

SPØRSMÅL OG SVAR I SPØRRESKJEMA- UNDERSØKELSER

Eksperimenter i forbindelse med valg av:
ordlyd i spørsmålsstilling, benevnelse på svarskala
og
rekkefølge på spørsmål

av

INGE JAN HENJESAND



Februar 1996

Avhandling innlevert til Institutt for Markedsøkonomi ved
Norges Handelshøyskole som del av kravet til graden
dr. oecon

96w002312

303.622

438s

cls. 2

ISBN 82-405-0024-2

Forord

“Noen ganger er det all right”, heter det i et dikt av Odd Børretzen. Dette må være en av de gangene. Å skrive forordet i dr. avhandlingen er noe jeg har sett frem til lenge. Alt for lenge, er det mange som vil påstå. Uansett, nå er jeg i hvertfall ferdig, og det er mange som skal takkes for at avhandlingen endelig kom i havn.

Først og fremst vil jeg rette en stor takk til professor Sigurd Villads Troye og professor Kjell Grønhaug for god veiledning, interessante diskusjoner og nyttige kommentarer gjennom hele arbeidet med avhandlingen. Sigurd har i tillegg til å være en god hovedveileder, vært en god støtte i perioder der de innsamlete dataene viste seg fra sin mest gjenstridige side. Kjell har gjennom sin interesse for det jeg har holdt på med, bidratt med mange nyttige innspill, interessante diskusjoner og “ekspress-tilbakemelding” på det jeg har skrevet.

En annen person som har betydd mye for meg som stipendiat ved Institutt for Markedsøkonomi, og for arbeidet med dr. avhandlingen er Ansgar Pedersen. I tillegg til å være den direkte årsaken til et godt arbeidsmiljø på instituttet, har Ansgar lest og gitt innsiktsfulle kommentarer på store deler av avhandlingen.

Alle på Institutt for Markedsøkonomi fortjener en stor takk for et godt faglig og sosialt arbeidsmiljø. Spesielt vil jeg takke Einar Breivik for godt samarbeid. En spesiell takk rettes også til Olav Kvitastein som fortsatt har svar på et hvilket som helst spørsmål

En rekke personer har blitt plaget med forespørsler om å lese deler av, eller hele avhandlingen. Jeg vil spesielt takke Anita Kåveland for grundig korrekturlesning, og for nyttige korrigeringer på min omstendelige skrivemåte. Dersom observante lesere mot

formodning skulle finne omstendelige setninger i det endelige verket, er dette helt og holdent mitt ansvar.

Datamaterialet som danner grunnlag for testing av hypotesene som er fremsatt i avhandlingen, er samlet inn i forbindelse med gjennomføring av gjesteundersøkelser for Inter Nor hotellene. Jeg vil benytte denne anledningen til å rette en spesiell takk til de hotellene som deltok i ferieundersøkelsen sommeren 1992.

En forutsetning for å kombinere forskning med mer konsulentpregete aktiviteter, er å ha en finansiering som gjør det mulig å "grave seg ned i" spennende faglige problemstillinger. Forskningen som denne avhandlingen er et resultat av, er i stor grad finansiert av Nærings- og energidepartementet (NOE), Norges Forskningsråd (NFR) og Statens Nærings- og Distriktsutviklingsfond (SND). Gjennom, på ulike måter, å ha bidratt med finansiering til forskningsprogrammet: Satsing i Reiseliv (SiR) ved Stiftelsen for Samfunns- og næringslivsforskning (SNF), har disse gjort det mulig å kombinere teori- og metodeutvikling med fremskaffelse av beslutningsrelevant informasjon for oppdragsgivere i reiselivsnæringen.

Bergen, februar 1996

Inge Jan Henjesand

==== Sammen drag ====

Er utvikling av spørreskjema en kunst eller en vitenskap? Dette er et spørsmål man med rette kan stille når et spørreskjema skal benyttes for innsamling av data. Formålet med denne avhandlingen er ikke å forsøke å besvare spørsmålet ovenfor. Vi har valgt å rette oppmerksomheten mot to valg man står overfor når et spørreskjema skal benyttes for måling av kunders produktvurderinger.

Det ene valget er knyttet til hvilken *instruksjon og vurderingsskala* som skal benyttes for vurderingsspørsmålene i spørreskjemaet. Når formålet med en spørreundersøkelse har vært å måle kunders produktvurderinger, er det særlig to ulike måter å stille spørsmål på, som både i litteraturen og i praksis har blitt brukt omhverandre. På den ene siden har man instruert respondenter til å vurdere hvorvidt produktet er bra eller dårlig, gjennom å benytte det vi har definert som *objektforankret instruksjon og vurderingsskala*. På den andre siden har det vært vanlig å spørre kundene i hvilken grad de er tilfreds eller utilfreds med produktet. I avhandlingen har vi definert denne tilnærmingen til måling av vurderingene som bruk av *subjektforankret instruksjon og vurderingsskala*.

Respondentenes produktkunnskap får stor betydning for responseffektene som fremkommer som følge av de metodologiske valgene vi har belyst i avhandlingen. For respondenter med lav produktkunnskap finner vi støtte for at det bør benyttes subjektforankrete instruksjoner og -vurderingsskaler i spørreskjemaet når man ønsker å måle vurderinger av spesifikke produktegenskaper. For disse respondentene er subjektforankrete vurderinger bedre egnet til å forklare totalvurderinger og atferdsintensjoner enn objektforankrete vurderinger. For respondenter med høy produktkunnskap finner vi at det motsatte er tilfellet. Når respondentenes produktkunnskap øker gir objekt-

forankrete vurderinger best prediksjon, både på totalvurderinger og atferdsintensjoner. Dersom det er mulig å finne ut noe om respondentenes produktkunnskap før et spørreskjema benyttes i en undersøkelse, bør disse funnene få implikasjoner for hvilken instruksjon og -vurderingsskala man velger å benytte i forbindelse med de spesifikke vurderingsspørsmålene i spørreskjemaet.

Et annet viktig funn er at det å benytte subjektforankrete instruksjoner og -vurderingsskalaer i forbindelse med spesifikke vurderingsspørsmål i et spørreskjema fører til økt relativ betydning av immaterielle produktegenskaper. Dette gjelder uavhengig av kunnskapen respondentene har om produktene vi ber dem vurdere. Dersom man i en undersøkelse skal måle kunders vurderinger av et produkt der service- eller tjenesteaspekter utgjør en vesentlig del, eller kjernen i produktet, bør disse funnene føre til at man velger å benytte subjektforankret instruksjon og -vurderingsskala i forbindelse med de spesifikke vurderingsspørsmålene i spørreskjemaet.

Det har vært vanlig å operasjonalisere kunders holdninger og produktvurderinger i et spørreskjema ved hjelp av multiattributtmodeller. Multiattributtmodeller blir vanligvis operasjonalisert ved hjelp av to sett med spørsmål, ett sett med viktighets spørsmål og ett sett med vurderingsspørsmål. Det andre valget vi har fokusert på i avhandlingen omhandler beslutninger vedrørende inkludering og alternativ plassering av viktighets spørsmålene i spørreskjemaet.

Kjennskap til respondentenes produktkunnskap får stor betydning, også når man står overfor en situasjon der man skal ta beslutning om inkludering og alternativ plassering av viktighetsspørsmål i et spørreskjema. For respondenter med lav produktkunnskap finner vi klare indikasjoner på at rapportering av viktighet før man foretar spesifikke produktvurderinger får betydning for vurderingene, mens dette ikke er tilfellet for respondenter med høy produktkunnskap. Hovedkonklusjonen vi kan trekke ut av disse funnene er at man bør unngå inkludering av viktighetsspørsmål i spørreskjemaet dersom man står overfor respondenter med lav produktkunnskap.

Et funn som leder oss til å konkludere med at man, uavhengig av kjennskap til respondentenes produktkunnskap, bør unngå å måle viktighet ved ulike produktegenskaper gjennom å inkludere viktighetsspørsmål i et spørreskjema, er et ønske om å måle spesifikke produktvurderinger på en måte som gir høyest mulig prediksjon på totalvurderinger. Funnene fra hypotesetestingen gir klare indikasjoner på en reduksjon i prediktiv evne når viktighetsspørsmål inkluderes i spørreskjemaet.

Innhold

| | |
|---------------------|-----|
| Forord..... | i |
| Sammendrag..... | iii |
| Innhold..... | v |
| Tabelloversikt..... | xi |
| Figuroversikt..... | xix |

DEL I INTRODUKSJON

| | |
|--|----|
| Kapittel 1 Introduksjon..... | 3 |
| 1.1 Problemstillinger..... | 6 |
| 1.1.1 Ordlyd i spørsmålsstilling og benevnelse på svarskala..... | 6 |
| 1.1.2 Spørsmålsrekkefølge..... | 12 |
| 1.1.3 Produktkunnskap..... | 14 |
| 1.1.4 Kvalitetsnivå på produktet som vurderes..... | 16 |
| 1.2 Oppsummering og implikasjoner..... | 18 |
| 1.3 Avhandlingens hoveddeler..... | 21 |

DEL II LITTERATURGJENNOMGANG

| | |
|---|----|
| Kapittel 2 Metodologiske perspektiver på svargivning i spørreundersøkelser..... | 25 |
| 2.1 Kilder til variasjon i svargivning i spørreundersøkelser..... | 27 |
| 2.1.1 En ramme for klassifisering av feilkilder i spørreundersøkelser..... | 27 |
| 2.1.2 Klassisk test teori..... | 29 |
| 2.1.3 Karakteristika ved stimuli - respondent - måling..... | 31 |

| | | |
|-------------------|--|-----------|
| 2.2 | Skalatyper i spørreundersøkelser..... | 34 |
| 2.3 | Empiriske studier vedrørende utforming av spørreskjema | 36 |
| 2.3.1 | Ordlyd i spørsmålsstilling..... | 37 |
| 2.3.2 | Svaralternativer..... | 40 |
| 2.3.2 | Spørsmålsrekkefølge..... | 45 |
| 2.5 | Oppsummering og implikasjoner..... | 48 |
| Kapittel 3 | Kognitive perspektiver på svargivning i spørreundersøkelser..... | 51 |
| 3.1 | Hvordan blir produktvurderinger foretatt?..... | 52 |
| 3.1.1 | Produktegenskaper..... | 54 |
| 3.1.2 | Holdepunkter (<i>cues</i>)..... | 56 |
| 3.1.3 | Egenskaper brukt i kundens vurderingsprosess..... | 57 |
| 3.1.4 | Multiattributtmodeller | 61 |
| 3.1.5 | Respondentens produktkunnskap..... | 64 |
| 3.1.6 | Produktkunnskap og holdningsdannelse..... | 65 |
| 3.1.7 | Kundetilfredshet..... | 67 |
| 3.1.8 | Produktkunnskap og kundetilfredshet..... | 68 |
| 3.2 | Kognitive oppgaver ved svargivning på spørsmål i et spørreskjema | 69 |
| 3.2.1 | Tolkning av spørsmålet..... | 70 |
| 3.2.2 | Dannelse av oppfatning | 71 |
| 3.2.3 | Utforming av respons..... | 73 |
| 3.2.4 | Redigering av respons..... | 74 |
| 3.3 | Objektforankrete- og subjektforankrete vurderinger..... | 74 |
| 3.3.1 | Hvordan er objektforankrete- og subjektforankrete instruksjoner og vurderingsskalaer forskjellige?..... | 75 |
| 3.3.2 | Hvilke effekter kan bruk av de to alternative responskalaene gi opphav til?..... | 76 |
| 3.3.2.1 | Implikasjoner for prediktiv validitet | 76 |
| 3.3.2.2 | Implikasjoner for konsistens i vurderinger | 80 |
| 3.3.2.3 | Implikasjoner for nivå på vurderinger..... | 81 |
| 3.3.2.4 | Implikasjoner for sikkerhet i vurderinger | 84 |
| 3.3.2.5 | Implikasjoner for relativ betydning av immaterielle- og materielle egenskaper | 84 |
| 3.4 | Multiattributtmodeller og spørsmålsrekkefølge | 86 |
| 3.4.1 | Hvorfor kan variasjoner i plassering av vurderingsspørsmål og viktighetsspørsmål i et spørreskjema føre til forskjeller i respondentenes svargivning?..... | 88 |
| 3.4.1.1 | Respondentens informasjonsbehandlingsprosess og konteksteffekter | 93 |
| 3.4.1.2 | Betydning av kontekst for tolking av spørsmål..... | 95 |
| 3.4.1.3 | Betydning av kontekst for fremhenting av informasjon og dannelse av oppfatning | 97 |
| 3.4.1.4 | Betydning av kontekst for vurderinger..... | 99 |
| 3.4.1.5 | Betydning av kontekst for responsseleksjon og redigering av respons..... | 100 |
| 3.4.2 | Betydning av produktkunnskap for plassering av viktighets-spørsmål og vurderingsspørsmål i spørreskjemaet..... | 101 |
| 3.4.3 | Hvilke effekter kan ulik plassering av viktighetsspørsmål og vurderingsspørsmål i spørreskjema gi opphav til?..... | 103 |

| | | |
|---------|--|-----|
| 3.4.3.1 | Implikasjoner for nivå på vurderinger og viktighetsspørsmål | 103 |
| 3.4.3.2 | Implikasjoner for sammenheng mellom vurderinger og viktighet | 107 |
| 3.4.3.3 | Implikasjoner for prediktiv validitet | 109 |
| 3.5 | Oppsummering og implikasjoner | 109 |

DEL III HYPOTESER OG UNDERSØKELSE SOPPLEGG *Les*

| | | |
|------------|---|-----|
| Kapittel 4 | Hypoteser | 113 |
| 4.1 | Modell | 113 |
| 4.2 | Hypoteser | 117 |
| 4.2.1 | Valg mellom objektforankret- og subjektforankret instruksjon og -vurderingsskala | 117 |
| 4.2.2 | Inkludering, og alternativ plassering av viktighetsspørsmål i spørreskjemaet | 127 |
| 4.3 | Oversikt over hypotesene | 132 |
| Kapittel 5 | Undersøkelsesopplegg | 139 |
| 5.1 | Krav til undersøkelsesopplegget | 139 |
| 5.2 | Metodologisk tilnærming | 144 |
| 5.3 | Valg av objekt/produkt for studien | 146 |
| 5.4 | Valg av subjekter/respondenter for studien | 147 |
| 5.5 | Beskrivelse av stimulimaterialet | 148 |
| 5.6 | Utvalg og praktisk gjennomføring | 149 |
| 5.7 | Oppsummering og implikasjoner | 152 |
| Kapittel 6 | Operasjonalisering | 155 |
| 6.1 | Uavhengige variabler | 155 |
| 6.1.1 | Instruksjon og vurderingsskala, og inkludering og alternativ plassering av viktighetsspørsmål | 156 |
| 6.2 | Avhengige variabler | 160 |
| 6.2.1 | Vurderinger av spesifikke egenskaper | 160 |
| 6.2.1.1 | Vurdering av samhandlingsegenskaper (immaterielle egenskaper) | 165 |
| 6.2.1.2 | Vurdering av strukturegenskaper (materielle egenskaper) | 167 |
| 6.2.1.3 | Vurdering av backstage- og kundeavhengige egenskaper | 169 |
| 6.2.2 | Viktighet av egenskapene som vurderes | 170 |
| 6.2.3 | Totalvurderinger og atferdsintensjoner | 171 |
| 6.2.4 | Sikkerhet i vurderinger | 174 |
| 6.2.5 | Relativ betydning av immaterielle- og materielle egenskaper | 175 |
| 6.3 | Modererende variabler | 176 |
| 6.3.1 | Respondentenes produktkunnskap | 177 |
| 6.3.2 | Kvalitetsnivå på produktet som vurderes | 180 |

DEL IV BESKRIVELSE AV DATAMATERIALET OG HYPOTSETESTING

| | | |
|-------------------|--|------------|
| Kapittel 7 | Beskrivelse av data og valg av statistiske analyseteknikker. | 183 |
| 7.1 | Manglende svargivning..... | 183 |
| 7.2 | Beskrivende statistikk..... | 186 |
| 7.3 | Valg av statistiske analyseteknikker | 189 |
| 7.3.1 | Test av forutsetninger for å gjennomføre variansanalyser | 191 |
| 7.3.1.1 | Forutsetning om normalitet i eksperiment- populasjonene | 192 |
| 7.3.1.2 | Forutsetning om homogenitet i varians..... | 193 |
| Kapittel 8 | Resultater av hypotesetesting..... | 197 |
| 8.1 | Objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer..... | 198 |
| 8.1.1 | Forskjeller i nivå på spesifikke vurderingsspørsmål | 198 |
| 8.1.2 | Forskjeller i konsistens for spesifikke vurderinger..... | 202 |
| 8.1.3 | Forskjeller i nivå på spesifikke vurderingsspørsmål - kvalitetsnivået på produktene som vurderes som modererende variabel..... | 206 |
| 8.1.4 | Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner | 212 |
| 8.1.5 | Forskjeller i sikkerhet ved individuelle vurderinger..... | 228 |
| 8.1.6 | Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner - respondentenes produktkunnskap som modererende variabel | 231 |
| 8.2 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet..... | 251 |
| 8.2.1 | Forskjeller i nivå på spesifikke vurderingsspørsmål | 251 |
| 8.2.2 | Forskjeller i nivå på spesifikke vurderingsspørsmål - kvalitetsnivået på produktene som vurderes som modererende variabel..... | 255 |
| 8.2.3 | Forskjeller i nivå på viktighetsspørsmål..... | 261 |
| 8.2.4 | Sammenheng mellom svargivning på viktighetsspørsmål og vurderingsspørsmål | 264 |
| 8.2.5 | Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner | 266 |
| 8.2.6 | Sammenheng mellom svargivning på viktighetsspørsmål og vurderingsspørsmål - respondentenes produktkunnskap som modererende variabel | 272 |
| 8.2.7 | Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner - respondentenes produktkunnskap som modererende variabel | 275 |
| 8.3. | Oppsummering av funn fra hypotesetestingen..... | 281 |
| 8.3.1 | Objektforankrete- og subjektforankrete instruksjoner og vurderingsskalaer..... | 281 |
| 8.3.2 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet..... | 286 |

DEL V DISKUSJON

| | | |
|---------------------------|---|-----|
| Kapittel 9 | Diskusjon og praktiske implikasjoner | 293 |
| 9.1 | Valg av subjektforankret- vs. objektforankret instruksjon og -vurderingsskala..... | 293 |
| 9.1.1 | Valg av instruksjon og vurderingsskala, og spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner..... | 295 |
| 9.1.2 | Valg av instruksjon og vurderingsskala, og relativ betydning av immaterielle produktetegenskaper..... | 299 |
| 9.1.3 | Valg av instruksjon og vurderingsskala, og nivå, konsistens og sikkerhet i vurderinger..... | 301 |
| 9.2 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet..... | 304 |
| 9.2.1 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, og sammenheng mellom viktighet og vurderinger..... | 306 |
| 9.2.2 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, og forklaring av totalvurdering og atferdsintensjoner..... | 308 |
| 9.2.3 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, og nivå på vurderinger og viktighets-spørsmål..... | 311 |
| Kapittel 10 | Begrensninger og implikasjoner for videre forskning..... | 317 |
| 10.1 | Valg av perspektiver for å belyse problemstillingene..... | 318 |
| 10.2 | Valg av undersøkelsesopplegg | 319 |
| 10.3 | Valg av vurderingsobjekter..... | 321 |
| 10.4 | Valg av operasjonaliseringer..... | 322 |
| 10.5 | Valg av utvalg og utvalgsprosedyrer | 324 |
| 10.6 | Valg av statistiske analysemetoder | 325 |
| Litteraturreferanser..... | | 327 |

APPENDIKS

| | | |
|-------------|---|-----|
| Appendiks A | Empiriske studier: ordlyd i spørsmålsstilling, svarskala og rekkefølge på spørsmål..... | 345 |
| Appendiks B | Eksempel på spørreskjema benyttet for innsamling av data..... | 355 |
| Appendiks C | Korrelasjonsmatrise og operasjonaliseringer..... | 365 |
| Appendiks D | Beskrivende statistikk..... | 369 |
| Appendiks E | Hypotesetesting: tabeller og figurer..... | 375 |

==== Tabelloversikt ====

| | | |
|------------|--|-----|
| Tabell 2.1 | Kilder til varians i psykometriske målinger..... | 31 |
| Tabell 3.1 | Variabler som påvirker størrelse og retning på konteksteffekter. | 94 |
| Tabell 4.1 | Hypoteser i tilknytning til valg mellom objektforankret- og subjektforankret instruksjon og vurderingsskala for spesifikke vurderings-spørsmål - Nivå og konsistens i vurderinger..... | 133 |
| Tabell 4.2 | Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderings spørsmål - Nivå på vurderinger, kvalitetsnivå på produktet som vurderes som modererende variabel | 133 |
| Tabell 4.3 | Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderings-spørsmål - Prediktiv validitet og sikkerhet i vurderinger | 134 |
| Tabell 4.4 | Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderings-spørsmål - Prediktiv validitet og sikkerhet i vurderinger, respondentenes produktkunnskap som modererende variabel..... | 135 |
| Tabell 4.5 | Hypoteser i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderings-spørsmål i et spørreskjema - Nivå på vurderinger, kvalitetsnivå på produktet som vurderes som modererende variabel | 136 |
| Tabell 4.6 | Hypoteser i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderings-spørsmål i et spørreskjema - Sammenheng mellom viktighet og vurderinger og prediktiv validitet..... | 136 |
| Tabell 4.7 | Hypoteser i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderings-spørsmål i et spørreskjema - Sammenheng mellom viktighet og vurderinger og prediktiv validitet, respondentenes produktkunnskap som modererende variabel | 137 |
| Tabell 5.1 | Sammenhenger mellom avhandlingens hypoteser og krav til informasjon..... | 143 |
| Tabell 5.2 | Oppnådd utvalg i de ulike eksperimentgruppene | 151 |
| Tabell 5.3 | Oppnådd utvalg for hotellene som deltok i undersøkelsene | 151 |

| | | |
|--------------|--|-----|
| Tabell 6.1 | Samhandlingsskalaen (immaterielle egenskaper) med spørsmål og statistiske egenskaper | 166 |
| Tabell 6.2 | Struktur skalaen (materielle egenskaper) med spørsmål og statistiske egenskaper | 168 |
| Tabell 6.3 | Totalvurderingsskalaen med spørsmål og statistiske egenskaper..... | 173 |
| Tabell 6.4 | Atferdsintensjonsskalaen med spørsmål og statistiske egenskaper..... | 174 |
| Tabell 7.1 | Manglende svargivning på enkeltspørsmål | 184 |
| Tabell 7.2 | ANOVA, summert antall manglende svar på enkeltspørsmål mellom eksperimentgruppene..... | 185 |
| Tabell 7.3 | Beskrivende statistikk for hele utvalget..... | 186 |
| Tabell 8.1 | T-test, samtlige 27 vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 201 |
| Tabell 8.2 | MANOVA, nivå på vurderinger av produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala. Rapportering av Pillais-T | 201 |
| Tabell 8.3 | Standardavvik for spesifikke vurderinger gruppert i produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 205 |
| Tabell 8.4 | Kendall's coefficient of concordance (W) for vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 206 |
| Tabell 8.5 | Gruppegjennomsnitt for spesifikke produktvurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - kvalitetsnivå på hotellene som vurderes som modererende variabel | 211 |
| Tabell 8.6a | ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 214 |
| Tabell 8.6b | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 216 |
| Tabell 8.7 | Standardiserte beta-koeffisienter og forklart varians for objektforankrete- og subjektforankrete totalvurderinger | 217 |
| Tabell 8.8 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 219 |
| Tabell 8.9 | Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner..... | 220 |
| Tabell 8.10a | ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 221 |
| Tabell 8.10b | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 223 |

| | | |
|--------------|---|-----|
| Tabell 8.11 | Standardiserte beta-koeffisienter og forklart varians for objektforankret- og subjektforankret totalvurdering, betydning av materielle- og immaterielle produktegenskaper | 224 |
| Tabell 8.12 | ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala..... | 226 |
| Tabell 8.13 | Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner, betydning av materielle- og immaterielle produkt-egenskaper | 227 |
| Tabell 8.14 | T-test, sikkerhet i vurderinger av de 27 spesifikke produkt-egenskapene foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala..... | 230 |
| Tabell 8.15 | MANOVA, sikkerhet i vurderinger av produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala. Rapportering av Pillais-T..... | 230 |
| Tabell 8.16a | ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap..... | 233 |
| Tabell 8.16b | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap..... | 235 |
| Tabell 8.17 | Standardiserte beta-koeffisienter og forklart varians for objektforankret- og subjektforankret totalvurdering, innenfor ulike nivåer av produktkunnskap | 236 |
| Tabell 8.18 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivåer av produktkunnskap..... | 238 |
| Tabell 8.19 | Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner, innenfor ulike nivåer av produktkunnskap..... | 239 |
| Tabell 8.20a | ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap..... | 241 |
| Tabell 8.20b | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap..... | 243 |
| Tabell 8.21 | Standardiserte beta-koeffisienter og forklart varians for objektforankret- og subjektforankret totalvurdering, betydning av materielle- og immaterielle produktegenskaper, innenfor ulike nivå av produktkunnskap | 244 |
| Tabell 8.22 | ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap..... | 246 |

| | | |
|--------------|--|-----|
| Tabell 8.23 | Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner, betydning av materielle- og immaterielle produkt-egenskaper innenfor ulike nivå av produktkunnskap..... | 247 |
| Tabell 8.24 | ANOVA, sikkerhet i vurderinger av de 27 spesifikke produkt-egenskapene foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, avhengig av respondentenes produktkunnskap..... | 249 |
| Tabell 8.25 | MANOVA, sikkerhet i vurdering av produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, avhengig av respondentenes produktkunnskap. Rapportering av Pillais-T..... | 250 |
| Tabell 8.26 | ANOVA, nivå på vurderinger foretatt i spørreskjema uten viktighets-spørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål..... | 254 |
| Tabell 8.27a | MANOVA, nivå på vurderinger gruppert i produktelementer, foretatt i spørreskjema uten viktighetsspørsmål, med viktighets-spørsmål før spesifikke vurderingsspørsmål og med viktighets-spørsmål etter spesifikke vurderingsspørsmål. Rapportering av Pillais-T... | 254 |
| Tabell 8.27b | MANOVA, nivå på vurderinger gruppert i produktelementer foretatt i spørreskjema uten viktighetsspørsmål, med viktighets-spørsmål før spesifikke vurderingsspørsmål og med viktighets-spørsmål etter spesifikke vurderingsspørsmål. Univariate F-tester | 255 |
| Tabell 8.28 | Gruppegjennomsnitt for vurderinger foretatt i spørreskjema uten viktighetsspørsmål og med viktighetsspørsmål før spesifikke vurderingsspørsmål, for hotellene i studien gruppert etter kvalitetsnivå | 260 |
| Tabell 8.29a | MANOVA, nivå på viktighet gruppert i produktelementer for vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. Rapportering av Pillais-T..... | 263 |
| Tabell 8.29b | MANOVA, nivå på viktighet gruppert i produktelementer for vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. Univariate F-tester | 263 |
| Tabell 8.30 | T-test, gjennomsnittskorrelasjon mellom viktighetsspørsmål og vurderingsspørsmål foretatt i spørreskjema med viktighetsspørsmål før vs. etter spesifikke vurderingsspørsmål..... | 265 |
| Tabell 8.31 | ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighets-spørsmål før spesifikke vurderingsspørsmål og med viktighets-spørsmål etter spesifikke vurderingsspørsmål..... | 268 |
| Tabell 8.32 | Standardiserte beta-koeffisienter og forklart varians for totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderings-spørsmål og med viktighetsspørsmål etter spesifikke vurderings-spørsmål..... | 270 |
| Tabell 8.33 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighets-spørsmål før spesifikke vurderingsspørsmål og med viktighets-spørsmål etter spesifikke vurderingsspørsmål..... | 271 |

| | | |
|-------------|--|-----|
| Tabell 8.34 | Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål..... | 272 |
| Tabell 8.35 | T-test, gjennomsnittskorrelasjon mellom viktighetsspørsmål og spesifikke vurderingsspørsmål foretatt i spørreskjema med viktighetsspørsmål før vs. etter spesifikke vurderingsspørsmål. Respondenter med lav produktkunnskap..... | 273 |
| Tabell 8.36 | T-test, gjennomsnittskorrelasjon mellom viktighetsspørsmål og spesifikke vurderingsspørsmål foretatt i spørreskjema med viktighetsspørsmål før vs. etter spesifikke vurderingsspørsmål. Respondenter med høy produktkunnskap..... | 274 |
| Tabell 8.37 | ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, innenfor ulike nivå av produktkunnskap..... | 276 |
| Tabell 8.38 | Standardiserte beta-koeffisienter og forklart varians for totalvurdering avhengig av tilstedeværelse og plassering av viktighetsspørsmål innenfor ulike nivå av produktkunnskap..... | 278 |
| Tabell 8.39 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, innenfor ulike nivå av produktkunnskap..... | 279 |
| Tabell 8.40 | Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner avhengig av tilstedeværelse og plassering av viktighetsspørsmål innenfor ulike nivå av produktkunnskap..... | 280 |
| Tabell 8.41 | Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Nivå på-, og konsistens i vurderinger..... | 281 |
| Tabell 8.42 | Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Nivå på vurderinger, kvalitetsnivået på produktene som vurderes som modererende variabel..... | 282 |
| Tabell 8.43 | Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete og subjektforankrete instruksjoner og -vurderingsskalaer - Prediktiv validitet og sikkerhet i vurdering..... | 283 |
| Tabell 8.44 | Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Prediktiv validitet og sikkerhet i vurdering, respondentenes produktkunnskap som modererende variabel..... | 284 |
| Tabell 8.45 | Oversikt over funn i tilknytning til eksplorative analyser vedrørende inkludering og plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Nivå på viktighetsspørsmål og vurderingsspørsmål..... | 286 |
| Tabell 8.46 | Oversikt over hypoteser og funn i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Nivå på vurderingsspørsmål, kvalitetsnivået på produktene som vurderes som modererende variabel..... | 287 |

| | | |
|-------------|--|-----|
| Tabell 8.47 | Oversikt over hypoteser og funn i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Sammenheng mellom viktighet og vurdering og prediktiv validitet | 288 |
| Tabell 8.48 | Oversikt over hypoteser og funn i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Sammenheng mellom viktighet og vurdering og prediktiv validitet, respondentenes produktkunnskap som modererende variabel | 289 |
| Appendiks: | | |
| Tabell A.1 | Empiriske studier: Problemstillinger i tilknytning til ordlyd i spørsmålsstilling i spørreundersøkelser | 346 |
| Tabell A.2 | Empiriske studier: Problemstillinger i tilknytning til svargivning i spørreundersøkelser | 347 |
| Tabell A.3 | Empiriske studier: Problemstillinger i tilknytning til rekkefølge på spørsmål i spørreundersøkelser | 351 |
| Tabell C.1 | Korrelasjonsmatrise for vurderingsspørsmål, totalvurderinger og atferdsintensjoner | 366 |
| Tabell C.2 | Backstage skalaen med spørsmål og statistiske egenskaper | 367 |
| Tabell C.3 | Kundeavhengig skalaen med spørsmål og statistiske egenskaper | 367 |
| Tabell D.1 | Beskrivende statistikk for hele utvalget | 370 |
| Tabell D.2 | Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med objektforankret instruksjon og -vurderingsskala | 371 |
| Tabell D.3 | Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med subjektforankret instruksjon og -vurderingsskala | 371 |
| Tabell D.4 | Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema uten viktighetsspørsmål | 372 |
| Tabell D.5 | Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med viktighetsspørsmål før spesifikke vurderingsspørsmål | 372 |
| Tabell D.6 | Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med viktighetsspørsmål etter spesifikke vurderingsspørsmål | 373 |
| Tabell D.7 | Resultater av tester for univariat homogenitet i varians for H1 - H9b..... | 373 |
| Tabell D.8 | Resultater av tester for univariat homogenitet i varians for H10a - H14b..... | 374 |
| Tabell E.1 | F-tester gjennomført separat for hvert hotell for høyere standardavvik knyttet til spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala..... | 376 |
| Tabell E.2 | ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP... | 376 |
| Tabell E.3 | ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP..... | 377 |
| Tabell E.4 | ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.. | 377 |
| Tabell E.5 | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP... | 378 |

| | | |
|-------------|--|-----|
| Tabell E.6 | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP..... | 378 |
| Tabell E.7 | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.. | 379 |
| Tabell E.8 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP..... | 379 |
| Tabell E.9 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP..... | 380 |
| Tabell E.10 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP..... | 380 |
| Tabell E.11 | ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP..... | 381 |
| Tabell E.12 | ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP..... | 381 |
| Tabell E.13 | ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP..... | 382 |
| Tabell E.14 | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP..... | 382 |
| Tabell E.15 | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP..... | 383 |
| Tabell E.16 | ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP..... | 383 |
| Tabell E.17 | ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP..... | 384 |
| Tabell E.18 | ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produktegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP..... | 384 |

| | | |
|-------------|--|-----|
| Tabell E.19 | ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP..... | 385 |
| Tabell E.20 | ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, LAV PRODUKTKUNNSKAP..... | 385 |
| Tabell E.21 | ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt med henholdsvis spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, MIDDELS PRODUKTKUNNSKAP..... | 386 |
| Tabell E.22 | ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt med henholdsvis spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderings-spørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, HØY PRODUKTKUNNSKAP..... | 386 |
| Tabell E.23 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, LAV PRODUKTKUNNSKAP..... | 387 |
| Tabell E.24 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, MIDDELS PRODUKTKUNNSKAP..... | 387 |
| Tabell E.25 | ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, HØY PRODUKTKUNNSKAP..... | 388 |

Figuroversikt

| | | |
|-----------|---|-----|
| Figur 1.1 | Skjematisk avbildning av problemstillingene som blir tatt opp i avhandlingen..... | 19 |
| Figur 2.1 | En modell av svargivning ved måling av produktvurderinger..... | 33 |
| Figur 2.2 | En funksjonell klassifisering av skaleringsprosedyrer..... | 35 |
| Figur 3.1 | Produksjonsfaktorer og produktelementer - et ledelsesperspektiv..... | 55 |
| Figur 3.2 | Vurdering av et produkt som en samling av iboende- og utvendige produktholdepunkter..... | 59 |
| Figur 3.3 | Produktet som en samling av iboende produktholdepunkter og utvendige produktholdepunkter (x_{ij}) , vs. en samling av egenskaper (s_k) | 60 |
| Figur 3.4 | <i>The unidimensionalist view of attitude</i> | 63 |
| Figur 3.5 | En skjematisk fremstilling av en modell for kunders produktvurderinger... | 63 |
| Figur 3.6 | Modell for informasjonsbehandling i en spørreundersøkelsessituasjon..... | 70 |
| Figur 3.7 | Forventede forskjeller mellom subjektforankret ("tilfreds-utifreds") og stimulusforankret ("bra-dårlig") instruksjon og -vurderingsskala..... | 83 |
| Figur 3.8 | En skjematisk fremstilling av en modell for betydningen av forutgående spørsmål for tolking av spørsmål i et spørreskjema..... | 96 |
| Figur 3.9 | Betydningen av produktkvalitet, og tilstedeværelse og plassering av viktighetsspørsmål for sammenhengen mellom assimilasjons- og kontrasteffekter og nivå på vurderinger..... | 106 |
| Figur 4.1 | Analysemodell..... | 116 |
| Figur 5.1 | Organisering-, og kombinasjoner av manipulerede variabler ved spørreskjemaforming..... | 148 |
| Figur 6.1 | Manipulering av instruksjon og vurderingsskala i spørreskjemaet | 156 |
| Figur 6.2 | Manipulering av rekkefølge på vurderingsspørsmål og viktighets-spørsmål i spørreskjemaet | 159 |
| Figur 6.3 | Formative og reflektive målemodeller..... | 162 |
| Figur 6.4 | Måling av viktighet ved egenskapene som vurderes..... | 170 |
| Figur 6.5 | Måling av totalvurderinger | 172 |

| | | |
|------------|--|-----|
| Figur 6.6 | Måling av sikkerhet i vurderinger for de individuelle egenskapene som vurderes..... | 175 |
| Figur 6.7 | Måling av erfaring og produktkunnskap | 179 |
| Figur 8.1 | Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret og subjektforankret instruksjon og -vurderingsskala | 199 |
| Figur 8.2 | Standardavvik for spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 204 |
| Figur 8.3a | Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - lavkvalitetshotell | 208 |
| Figur 8.3b | Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - hotell med gjennomsnittskvalitet | 209 |
| Figur 8.3c | Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - høykvalitetshotell..... | 210 |
| Figur 8.4 | Sikkerhet i vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala | 229 |
| Figur 8.5 | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål | 253 |
| Figur 8.6a | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål og med viktighetsspørsmål før spesifikke vurderingsspørsmål - lavkvalitetshotell | 257 |
| Figur 8.6b | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål og med viktighetsspørsmål før spesifikke vurderingsspørsmål - hotell med gjennomsnittskvalitet..... | 258 |
| Figur 8.6c | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål og med viktighetsspørsmål før spesifikke vurderingsspørsmål - høykvalitetshotell | 259 |
| Figur 8.7 | Nivå på viktighetsspørsmål plassert henholdsvis før-, og etter spesifikke vurderingsspørsmål i spørreskjemaet..... | 262 |
| Figur 9.1 | Valg av subjektforankret- vs. objektforankret instruksjon og vurderingsskala - oversikt over funn fra hypotesetestingen..... | 294 |
| Figur 9.2 | Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet - oversikt over funn fra hypotesetestingen..... | 305 |
| Appendiks: | | |
| Figur E.1a | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, lavkvalitetshotell | 388 |
| Figur E.1b | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, hotell med gjennomsnittskvalitet..... | 389 |
| Figur E.1c | Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, høykvalitetshotell..... | 389 |

DEL

I

INTRODUKSJON

Denne delen av avhandlingen inneholder introduksjonskapittelet, der bakgrunnen for den gjennomførte studien presenteres. I dette kapittelet blir en rekke spesifikke problemstillinger presisert og relatert til tidligere forskning.

1

Introduksjon

Når man har besluttet å ta i bruk et strukturert spørreskjema for innsamling av data, stilles man både som akademiker og praktiker overfor en rekke valg i forbindelse med innhold i -, og utforming av spørreskjemaet som skal benyttes i undersøkelsen. Følgende spørsmål kan f.eks. være nødvendige å få svar på; Hvordan skal spørsmålene stilles? Skal det benyttes forhåndsdefinerte svarkategorier? Hvordan skal eventuelle vurderingsskalaer se ut? Hvor mange svarkategorier skal det benyttes i en skala? Skal det være midtpunkt på skalaen? Hvilke benevnelser skal det være på svarkategoriene i skalaen? Skal det være benevnelse på alle svarkategoriene, eller bare på ytterpunktene? Hvor mange spørsmål skal stilles i spørreskjemaet? I hvilken rekkefølge skal spørsmålene stilles? Hvilket format, og hvilken layout skal det være på spørreskjemaet? Disse spørsmålene representerer eksempler på noen av de valgene man må ta stilling til ved utformingen av selve spørreskjemaet.

Til tross for en mengde negativ kritikk, har gjennomføring av spørreundersøkelser der det benyttes standardiserte spørreskjema, fortsatt et solid fotfeste som metode for innsamling av primærdata i samfunnsvitenskapelig forskning. Mye av kritikken mot spørreundersøkelsene har nettopp vært rettet mot problemer forbundet med utformingen av selve spørreskjemaet, og de spørsmålene som stilles i dette. "Som man spør, får man svar" blir det hevdet¹. Respondentenes evne og vilje til å gi de "rette" svarene i en undersøkelse, er nært knyttet til forskerens evne til å lage de "riktige" spørsmålene. Begge disse forholdene begrenser mulighetene man har for å hevde at data innsamlet i en spørreundersøkelse gir "sann" informasjon om de problemstillingene som studeres.

¹Dette er også tittelen på den eneste norske boken der det spesielt fokuseres på arbeid med spørreundersøkelser, se: Mordal (1989).

I denne avhandlingen vil vi sette søkelyset på noen av de valgene man står overfor ved utvikling av spørreskjemaet. Mer spesifikt vil vi fokusere på valg knyttet til ordlyd i spørsmålsstilling, benevnelse på svaralternativer og spørsmålsrekkefølge. Et viktig utgangspunkt for å velge dette hovedfokuset er at det særlig er to kilder til *respons-effekter* som har vært viet stor oppmerksomhet i metodologiske studier, og som har vist seg å ha stor betydning for respondentenes svargivning; nemlig ordlyd i spørsmålsstilling og benevnelse på svarskala (*question wording*) og rekkefølgen på spørsmålene i spørreskjemaet (*question order*),² se f.eks. Bradburn og Sudman (1981), Hippler, Schwarz, og Sudman (1987), Hogarth (1982), Schuman og Presser (1981), Schwarz og Sudman (1991), Sudman og Bradburn (1974), Sudman og Bradburn (1983) og Tanur (1991).

Utforming av enkeltspørsmål og rekkefølgen på spørsmålene i spørreskjemaet har vist seg å være avgjørende for validitet og reliabilitet ved data som er samlet inn i en spørreundersøkelse (se f.eks. Hippler, Schwarz et al. (1987), Schuman og Presser (1981), Sudman og Bradburn (1974) og Sudman og Bradburn (1983)). Endringer i ordlyd og spørsmålsrekkefølge har f.eks. vist seg å få betydning, både for univariate og marginale svarfordelinger (Schuman og Presser, 1981; Schwarz og Sudman, 1991), og for størrelse og retning på sammenhenger mellom enkeltspørsmål i spørreskjemaet (Schuman og Presser, 1981; Schwarz og Sudman, 1991). I tidligere forskning har man imidlertid viet liten oppmerksomhet til forskning omkring effekter på retning og størrelse på sammenhenger. Man har i hovedsak vært opptatt av effekter på univariate fordelinger (Hippler og Schwarz, 1987; Schuman og Presser, 1981; Sudman og Bradburn, 1974).

Tradisjonelt har man benyttet et rent metodologisk utgangspunkt for utvikling av retningslinjer som skal hjelpe forskeren til å foreta valg ved konstruksjon av spørreskjema. Formålet har her vært å redusere feilvarians i målingene. Utforming av selve spørreskjemaet kan imidlertid betraktes som kun en av flere kilder til variasjon i svargivning i spørreundersøkelser. I metodelitteraturen blir variasjon i svargivning generelt tilskrevet tre forhold: Egenskaper ved måleprosedyren, egenskaper ved respondentene og egenskaper ved stimuli, eller objektet som vurderes, (Cox, 1982; Cox, 1980; Guilford, 1954; Hippler, Schwarz, og Sudman, 1987; Torgerson, 1958). I undersøkelser der kunder blir bedt om å foreta vurderinger av et produkt gjennom å benytte et strukturert spørreskjema, antar man ofte implisitt at variasjoner i svargivning på

²For en oversikt: se f.eks. Bradburn og Sudman (1981); Hippler, Schwarz et al. (1987); Schuman og Presser (1981); Schwarz og Sudman (1991); Sudman og Bradburn (1974); Sudman og Bradburn (1983) og Tanur (1991). I tillegg vil det i kapittel 2 bli foretatt en grundigere gjennomgang av empiriske studier knyttet til *question wording* og *question order* effekter

vurderingsspørsmålene kun reflekterer *sanne* variasjoner i produktets ytelse³. Denne antakelsen representerer i de fleste sammenhenger et urealistisk idealtilfelle.

For det første ser man da nettopp bort fra det at forhold ved måleprosedyren, som f.eks. spørsmålsformulering, svarskala og rekkefølgen på spørsmålene som benyttes i spørreskjemaet, i tillegg til administrasjonsmåte og utvalgsprosedyrer, kan gi opphav til variasjoner i respondentenes svargivning.

For det andre kan de personene som foretar vurderinger ved å svare på spørsmålene i spørreskjemaet være forskjellige. Respondentene som blir bedt om å delta i en spørreundersøkelse kan ha forskjellige måter å bruke et produkt på, de kan ha forskjellige måter å foreta vurderinger på og de kan ha ulik evne og vilje til å svare på spørsmål i et spørreskjema. Ulik erfaring og produktkunnskap, ulike forventninger, responsstil og responssett⁴ er eksempler på kjennetegn ved respondentene som kan gi opphav til variasjoner i svargivning på vurderingsspørsmålene i spørreskjemaet.

I avhandlingen vil vi først og fremst fokusere på det vi ovenfor har beskrevet som egenskaper ved måleprosedyren. I forbindelse med valg av ordlyd i spørsmålsstilling og benevnelse på svarskaalen i spørreskjemaet, vil vi rette oppmerksomheten mot to alternative, og mye benyttede instruksjoner og vurderingsskalaer som benyttes i forbindelse med undersøkelser av kunders produktvurderinger. På den ene siden kan man instruere respondenter til å rapportere de følelsene kjøp og bruk av et produkt gir opphav til, f.eks. gjennom å benytte en "tilfreds-tilfreds" skala. På den andre siden kan respondentene instrueres til å foreta en mer objektiv vurdering av produktets godhet, f.eks. ved hjelp av en "bra-dårlig" skala. Som vi skal se senere i avhandlingen kan valg mellom disse to instruksjonene, som i utgangspunktet kan se like ut, og som i mange tilfeller benyttes om hverandre, få relativt stor betydning for respondentenes svargivning.

I studier av kunders produktvurderinger velger man ofte å beskrive produktet som bestående av et sett egenskaper ved hjelp av en multiattributtmodell⁵. En multiattributtmodell består vanligvis av en viktighetskomponent og en vurderingskomponent, se f.eks. Fishbein og Ajzen (1975), Grünert (1989), Lutz (1991) og Rosenberg (1956). Operasjonalisering av multiattributtmodeller ved hjelp av data innhentet via et spørre-

³Variasjoner i produktets ytelse = variasjoner som tilskrives egenskaper ved stimuli.

⁴Respons-sett (*response set*) oppstår når respondenten svarer på en bestemt måte for å fremstille seg selv på en måte som gir en bestemt image. Respons-stil (*response style*) oppstår når respondenten svarer på en bestemt måte uavhengig av spørsmål, f.eks. ved å følge et bestemt mønster ved utfylling av spørreskjemaet (*careless respondents*).

⁵For en grundigere beskrivelse av multiattributtmodeller, se kapittel 3

skjema, innebærer dermed inkludering av to sett spørsmål i spørreskjemaet; et sett spørsmål knyttet til respondentenes oppfatning av *viktigheten* av at en samling gitte produktegenskaper er til stede, og et sett knyttet til *vurderingene* av de samme produkt-egenskapene. Problemstillinger i forbindelse med rekkefølge på spørsmålene i spørreskjemaet, vil med dette som utgangspunkt bli avgrenset til effekter av å inkludere-, og plassere viktighetsspørsmål før vs. etter spesifikke vurderingss spørsmål.

I tillegg til perspektivene fra holdningsteori som er trukket frem ovenfor, i forbindelse med operasjonalisering av multiattributtmodeller, har teorier om kundetilfredshet vært et viktig utgangspunkt for å studere kunders produktvurderinger. En sentral antakelse i kundetilfredshetslitteraturen er at positive avvik mellom kundenes forventninger/normer og produktets ytelse fører til tilfredshet, mens negative avvik fører til utilfredshet (Oliver, 1981; Tse og Wilton, 1988; Yi, 1990). Produktets ytelse eller kvalitet kan på denne måten få betydning for hvordan respondenter foretar vurderinger. Respondentene som foretar produktvurderingene, og lar disse komme til uttrykk ved å svare på spørsmålene i spørreskjemaet, kan i tillegg være forskjellige med hensyn til hvilken erfaring de har med bruk av produktet. Erfaring kan føre til økt produktkunnskap som også kan få betydning for hvordan produktvurderinger blir foretatt. Vi ønsker å ta hensyn til dette i avhandlingen, og betrakter kvalitetsnivå, eller ytelse ved produktet som vurderes, og respondentenes produktkunnskap som forhold som kan moderere effektene av de metodologiske valgene vi har valgt å studere.

1.1 Problemstillinger

I det følgende vil vi begrunne og utdype en rekke konkrete problemstillinger i tilknytning til de metodologiske valgene vi ovenfor har introdusert vedrørende ordlyd i spørsmålsstilling, benevnelse på svarskala og rekkefølge på spørsmålene i et spørreskjema.

1.1.1 Ordlyd i spørsmålsstilling og benevnelse på svarskala

Et av de mange valgene som både akademikere og praktikere stilles overfor i studier av kunders holdninger til -, vurderinger av -, og tilfredshet med produkter de har kjøpt, er som nevnt hvordan spørsmålene skal stilles, og hvilke benevnelser som skal benyttes på svaralternativene på svarskalaen i spørreskjemaet. I spørreundersøkelser der formålet er

å finne ut noe om hvordan kunder vurderer et produkt synes det som om “bra-dårlig” og “tilfreds-tilfreds” skalaene har blitt benyttet om hverandre, se f.eks. Cadotte, Woodruff, og Jenkins (1987). Instruksjonen som blir gitt i spørsmålene i spørreskjemaet blir regnet som som et av de viktigste grunnlagene respondenten har for tolke hvilke oppgaver intervjueren forventer blir utført i forbindelse med svargivningen (Hippler, Schwarz, og Sudman, 1987). Vi ønsker å se nærmere på dette, og er spesielt interessert i å studere hvordan svargivning kan påvirkes om vi instruerer respondenter til å rapportere de følelsene kjøp og bruk av et produkt gir opphav til ved bruk av “tilfreds-tilfreds” skala, eller om vi alternativt instruerer dem til å foreta en mer “objektiv” vurdering av produktets godhet ved bruk av “bra-dårlig” skala. Disse to dimensjonene representerer ulike måter kunder kan bli bedt om å foreta produktvurderinger på, og er knyttet til høyst forskjellige utgangspunkt.

Det at selve spørsmålsformuleringen kan påvirke de svar som blir avgitt i en spørreundersøkelse er viet stor oppmerksomhet; se f.eks. Bradburn og Sudman (1981), Hippler, Schwarz, og Sudman (1987), Hogarth (1982), Schuman og Presser (1981), Schwarz og Sudman (1991), Sudman og Bradburn (1974), Sudman og Bradburn (1983) og Tanur (1991). I avhandlingens kapittel 2 vil det bli foretatt en grundig gjennomgang av en rekke empiriske studier der problemstillinger knyttet til ordlyd i spørsmålsformuleringer og benevnelse på svars skala er fokusert. Nedenfor gir vi derfor kun en summarisk oversikt over noen av de mest sentrale effektene forskjeller i spørsmålsstilling kan gi opphav til.

For det første kan endringer i spørsmålsformulering og benevnelse på svars skalaen endre det substansielle meningsinnholdet i et spørsmål. Variasjoner i spørsmålsformulering kan f.eks. føre til at man gjennom målingen fanger opp ulike sider ved et begrepsdomene⁶ (Hippler og Schwarz, 1987). Bruk av multiple indikatorer⁷ er nødvendig ved operasjonalisering av flerdimensjonale begreper for å fange opp flere dimensjoner. Selve spørsmålsformuleringen kan innvirke på i hvilken grad man makter å fange opp den enkelte dimensjon. F.eks. kan operasjonalisering av servicebegrepet innebære at man stiller spørsmål om servicepersonalets vennlighet, høflighet og profesjonalitet.

For det andre kan det samme spørsmålet bli oppfattet og tolket forskjellig av ulike respondenter. En forklaring på dette kan være bruk av kontekstspesifikt språk i spørsmålene som benyttes (Strack og Martin, 1987). Hvordan et bestemt spørsmål oppfattes

⁶Et begrepsdomene refererer til hva som omfattes av et teoretisk/konseptuelt begrep.

⁷Bruk av multiple indikatorer innebærer at man stiller flere spørsmål for å måle flere dimensjoner ved det samme teoretiske/konseptuelle begrepet.

og tolkes av respondentene kan således være avhengig av den kjennskapen disse har til den aktuelle konteksten.

For det tredje har man funnet at endringer i ordlyd i spørsmålsstilling kan påvirke respons ved at ulikt ladete assosiasjoner aktiveres hos respondenten. Inkludering av benevnelser som "kommuniststyre" i forbindelse med måling av amerikaneres holdninger til USA's militæraksjoner i Vietnam er et klassisk eksempel på dette (Schuman, Kalton, og Ludwig, 1983; Schuman og Presser, 1981). I spørsmål der man begrunnet USA's eventuelle deltakelse i militæraksjoner i Vietnam med formål å stoppe en *kommunistisk maktovertakelse*, ble det observert en fordobling av antall respondenter som støttet militæraksjoner og amerikanske troppeforsendelser, sammenlignet med spørsmål der *komunistisk maktovertakelse* begrunnelsen var utelatt.

Endelig kan endringer i ordlyd i spørsmålsstilling bidra til å påvirke respondentenes oppfatning av hva som er deres oppgave i forbindelse med svargivningen. Slike effekter kan oppstå selv for endringer som kan regnes for å være ikke-substansielle. Et eksempel på problemstillinger man har fokusert på i denne sammenheng er de forskjeller man oppnår i respondentenes svargivning ved å bruke ord som "forby" eller "tillate" i undersøkelser vedrørende holdninger til abort (Hippler og Schwarz, 1987). Funnene i disse studiene indikerer at respondentene generelt er mer tilbøyelige til å si seg enig i at noe bør forbys, enn at det samme bør tillates. I tillegg er de mer tilbøyelige til å si seg enig i at noe ikke bør tillates, enn at det ikke bør forbys. Man er altså mer tilbøyelig til å si seg enig i nektende utsagn.

Et kjennetegn ved den forskningen som er gjennomført med hensyn til ordlyd i spørsmålsstilling er at man i hovedsak har fokusert på effekter som kan få betydning for *reliabilitet* ved de dataene som er samlet inn. Man har altså vært interessert i *konsistens* i svargivning, det vil si hvorvidt resultater kan sammenlignes på tvers av undersøkelser dersom spørsmålene ikke er formulert på samme måte i de forskjellige undersøkelsene. Man har som nevnt vært spesielt opptatt av å studere effekter av bruk av ord som "rammer inn"⁸ spørsmålet og bidrar til at respondenten tolker det på en bestemt måte. Dette kan føre til at man får forskjellige svar avhengig av hvilken ordlyd man velger å benytte (Hippler og Schwarz, 1987).

Vi skal i avhandlingen rette oppmerksomheten mot et forhold som ikke har vært studert tidligere, og som i utgangspunktet kan betraktes som et forhold som faller inn under den forskningen som er gjennomført vedrørende ordlyd i spørsmålsstilling og benevnelse på

⁸Et eksempel på innramming (*framing*) av et spørsmål er f.eks bruk av *forby* i motsetning til *tillate* i undersøkelser av holdninger til abort.

svarskala. Respondenter kan på den ene siden bli bedt om å vurdere om et produkt er bra eller dårlig. "Bra-dårlig" skalaen representerer et eksempel på det vi har valgt å kalle *objektforankret*⁹ instruksjon og -vurderingsskala basert på skaleringslitteraturens definisjon på *stimulussentrert* skalering (Cox, 1982; Cox, 1980; Guilford, 1954; Hippler, Schwarz, og Sudman, 1987; Torgerson, 1958). Stimulussentrert skalering benyttes når man er interessert i å få respondenter til å skalere et objekt, eller vurdere et produkt som sådan. Ved bruk av stimulussentrert skalering styrer forskeren varians i respons mot objektet, eller produktet som vurderes.

På den andre siden kan respondentene bli bedt om å gi uttrykk for i hvilken grad de er tilfreds eller utilfreds med produktet. "Tilfreds-utilfreds" skalaen representerer da et eksempel på det vi velger å kalle *subjektforankret* instruksjon og -vurderingsskala med referanse til skaleringslitteraturens definisjon av *subjektsentrert* skalering (Cox, 1982; Cox, 1980; Guilford, 1954; Hippler, Schwarz, og Sudman, 1987; Torgerson, 1958). Subjektsentrert skalering benyttes når vi er interessert i respondentenes affektive eller følelsesmessige indre tilstander, eller holdninger overfor objektet. Ved bruk av subjektsentrert skalering styres variansen i respons mot respondentene som foretar vurderingene.

Når vi instruerer respondentene til å vurdere i hvilken grad et produkt oppfattes som bra eller dårlig ved bruk av en objektforankret skala, eller ber dem gi uttrykk for hvor tilfreds eller utilfreds de er med produktet ved bruk av en subjektforankret skala, vil dette medføre kognitive oppgaver som man må anta er vesentlig forskjellige. Når vi benytter objektforankret ("bra-dårlig") instruksjon og -vurderingsskala instruerer vi respondentene til å ta på seg en rolle som "dommere" som skal vurdere produktets beskaffenhet. Bruk av subjektforankret ("tilfreds-utilfreds") instruksjon og -vurderingsskala innebærer på den andre siden at vi ber respondentene gi uttrykk for sin tilfredshet med produktets beskaffenhet.

Problemstillinger i tilknytning til forskjeller i svargivning, som resultat av valg mellom de to instruksjonene og vurderingsskalaene, kan sees i relasjon til respondentens informasjonsbehandlingsprosess, og forskningen som er gjennomført med hensyn til hva respondenter oppfatter som sin oppgave når de skal svare på spørsmål i et spørreskjema (Hippler og Schwarz, 1987).

⁹Skaleringslitteraturen fokuserer på hvilke estimeringsteknikker og eksperimentelle design som er passende for subjektsentrert-, stimulussentrert og respons tilnærminger. Dette er ikke problemområder som blir tatt opp i denne avhandlingen. Formålet er heller å belyse hvilke forskjeller som oppstår når 1) subjekter instrueres til å være "dommere" som skal vurdere objekter (objektforankret tilnærming) og når 2) subjektene instrueres til å rapportere indre affektive tilstander som objektene gir opphav til (subjektforankret tilnærming). For å gjøre dette skillet eksplisitt bruker vi benevnelsene subjektforankrete- og objektforankrete instruksjoner og -vurderingsskalaer i avhandlingen.

Til tross for den utstrakte bruken av de to alternative instruksjonene og vurderingsskalaene, reflektert ved disse måtene å stille spørsmål på i undersøkelser av produktkvalitet, servicekvalitet og kundetilfredshet, der referansepunktet for vurderinger er henholdsvis produktet (objektforankret) og kunden/respondenten (subjektforankret), er forskjeller disse kan forårsake i svargivning, som nevnt viet liten oppmerksomhet i tidligere forskning. Mulige årsaker til dette kan være at det kan eksistere en oppfatning av at det ikke betyr noe hvilken av de to måtene å stille spørsmål på man velger, eller at man rett og slett ikke har tenkt over forskjellene.

Denne innledende litteraturgjennomgangen viser at det kan være grunn til å anta at bruk av de to typene instruksjon og vurderingsskala vi har beskrevet ovenfor, vil kunne gi opphav til systematiske forskjeller i svargivning på vurderingsspørsmålene i et spørreskjema, og at dette ikke bare bør betraktes som et problem knyttet til metodologiske valg. Vi argumenterer senere i avhandlingen¹⁰ for at effekter av ulik instruksjon og vurderingsskala også kan belyses med utgangspunkt i teorier om kundetilfredshet, holdningsteori og bidrag fra informasjonsbehandlingslitteraturen. Valg av ordlyd i spørsmålsformulering og svaralternativer er som vi har sett, et viktig aspekt ved gjennomføring av spørreundersøkelser. Vi reiser derfor følgende problemstilling som konkretiseres i testbare hypoteser i kapittel 4:

RQ1: Hvordan påvirker bruk av subjektforankret- vs. objektforankret instruksjon og -vurderingsskala respondentenes svargivning på vurderingsspørsmål i et spørreskjema?

I markedsføringslitteraturen har et produkt blitt definert som “alt som tilbys som tilfredsstillende behov og ønsker” (Kotler, 1984) og som “en samling av egenskaper” (Bagozzi, 1986). I økonomisk litteratur har det vært vanlig å beskrive et produkt med utgangspunkt i “egenskapene som bidrar til konsumentens nytte”¹¹ (Lancaster, 1966; Lancaster, 1971). Disse egenskapene kan være både materielle (fysiske) og immaterielle (f.eks. service) (Lefkoff-Hagius og Mason, 1990; Levitt, 1981; Shostack, 1977), iboende (*intrinsic*) og utvendige (*extrinsic*)¹² (Jacoby, Olson, og Haddock, 1971; Szybillo og Jacoby, 1974).

¹⁰Teoretiske perspektiver som kan bidra til å belyse forskjellene mellom objektforankrete- (“god-dårlig”) og subjektforankrete (“tilfreds-utilfreds”) vurderinger blir viet større oppmerksomhet i kapittel 3.

¹¹*Characteristics*

¹²Iboende (*intrinsic*) egenskaper er relatert til iboende holdepunkter (*intrinsic product cues*) som oftest er fysiske aspekter ved produktet. Dersom disse endres, endres også det fysiske produktet.

Utvendige (*extrinsic*) egenskaper er relatert til utvendige holdepunkter (*extrinsic product cues*). Dette er holdepunkter som har sitt opphav i markedsføringens virkemidler.

I tjenestemarkedsføringslitteraturen hevdes det at det er vanskeligere for en kunde å vurdere tjenester enn "fysiske" varer (Parasuraman, Zeithaml, og Berry, 1985; Uhl og Upah, 1983; Zeithaml, 1981; Zeithaml, Parasuraman, og Berry, 1985). Dette forklares blant annet med at et av karakteristikaene som kjennetegner en tjeneste er at den er immateriell (Eiglier et al., 1979; Levitt, 1972; Levitt, 1981; Lovelock, 1988; Shostack, 1977; Uhl og Upah, 1983; Zeithaml, Parasuraman, og Berry, 1985), eller at tjenester oppfattes som å ha immaterielt dominante egenskaper¹³ (Shostack, 1977).

Det at tjenester er immaterielle hevdes også å bidra til at de blir vanskeligere å kommunisere (Levitt, 1981) enn fysiske varer. Dette innebærer en antagelse om at konsumenter må benytte andre holdepunkter og prosesser når de vurderer tjenester, enn når de vurderer varer (Zeithaml, 1981). I litteraturen har man fokusert på forskjellene mellom varer og tjenester som årsaker til problemene ved vurdering. I stedet for å snakke om varer og tjenester, vil vi i denne avhandlingen ta utgangspunkt i at de fleste braer består av både vare- og tjenesteelementer. Vi betrakter altså service- eller tjenesteelementer som et av mange elementer i et produkt (Troye, 1990), og ønsker å fokusere på forskjeller i vurderingsprosesser for henholdsvis materielle og immaterielle egenskaper ved det samme produktet.

Når vi ber kunder vurdere materielle og immaterielle egenskaper ved produktet ved bruk av objektforankret ("bra-dårlig") instruksjon og -vurderingsskala vs. subjektforankret ("tilfreds-utilfreds") instruksjon og -vurderingsskala, antar vi at dette gir seg ulike utslag i svargivning. Blant annet vil vi senere i avhandlingen argumentere for at subjektforankrete vurderinger kan være lettere for respondentene å foreta enn objektforankrete vurderinger. Dette kan få implikasjoner for vurderinger foretatt ved bruk av de to instruksjonene og vurderingsskalaene i et spørreskjema. Med bakgrunn i denne innledende drøftingen vil vi reise følgende problemstilling:

RQ2: Hvordan påvirker bruk av subjektforankret- vs. objektforankret instruksjon og -vurderingsskala respondentenes vurderinger av henholdsvis immaterielle- og materielle produktegenskaper i et spørreskjema?

Problemstillingen ovenfor konkretiseres i kapittel 4 ved hjelp av et sett hypoteser.

¹³Det at et produkt har immaterielt dominante egenskaper innebærer at tjenesteegenskapene ved produktet dominerer.

1.1.2 Spørsmålsrekkefølge

Ved måling av kunders produktvurderinger vil det være av stor interesse for den som skal ta beslutninger om hvilke tiltak som skal settes i verk for å forbedre produktet, å vite noe om hvordan de ulike egenskapene som man ber kundene vurdere inngår i kundenes vurderingsprosess. En optimal økonomisk ressursbruk tilsier at de egenskapene som kundene legger mest vekt på i sine vurderinger bør prioriteres ved produktutviklingsbeslutninger. Dette er en viktig begrunnelse for at spørsmål som har som formål å avdekke hvor viktige, eller sentrale, de ulike produktegenskapene er for respondentens valg, ofte blir inkludert i spørreskjemaet.

I studier av holdninger, og av hvordan kunder foretar produktvurderinger har ulike varianter av multiattributtmodeller spilt en sentral rolle, se f.eks.: Fishbein og Ajzen (1975), Grünert (1989), Lutz (1991) og Lutz og Bettman (1977). Multiattributtmodeller, slik disse er benyttet ved måling av holdninger, består vanligvis av to komponenter; Rosenberg (1956) skiller mellom en viktighetskomponent (*value importance*) og en evalueringskomponent (*perceived instrumentality*), mens Fishbein (1963) skiller mellom en evalueringskomponent (*evaluative aspect*) og en antakelseskomponent (*belief*). Både Rosenbergs og Fishbeins modell faller inn under det man kaller "forventningsverdi" modeller¹⁴.

Viktighetskomponenten i Rosenbergs modell og evalueringskomponenten i Fishbeins modell sier noe om hvor viktig det er for respondenten at en spesiell egenskap ved objektet er til stede, mens evalueringskomponenten i Rosenbergs modell og antakelseskomponenten i Fishbeins modell reflekterer respondentens evaluering av de ulike egenskapene. Holdning til et objekt er i modellene forklart ved summen av produktet av viktighetskomponenter og evalueringskomponenter (Rosenberg, 1956), og summen av produktet av evalueringskomponenter og antakelseskomponenter (Fishbein, 1963)¹⁵. I tillegg finnes en rekke hybride modeller som ofte benyttes i praksis. Operasjonalisering av kunders produktvurderinger i en multiattributtmodell, ved hjelp av data innhentet via et spørreskjema, innebærer altså at vi stiller to sett av spørsmål, et vedrørende vektlegging av produktegenskaper, og et vedrørende vurdering av de samme egenskapene.

En anseelig mengde studier har vært gjennomført for å teste disse modellene. Et av formålene har f.eks. vært å komme frem til modeller og operasjonaliseringer av disse, som på en best mulig måte predikerer atferd, representert ved sannsynlighet for kjøp,

¹⁴*Expectancy-value models*

¹⁵Se f.eks. Grünert (1989) og Lutz (1991)

gjenkjøp og anbefaling av produktet¹⁶, se f.eks. Fishbein (1963), Fishbein (1972) og Fishbein og Ajzen (1975). Det har blitt reist spørsmål ved nytten av å be respondenter angi viktigheten av forskjellige egenskaper ved et produkt som svar på spørsmål i et spørreskjema, se f.eks. Huber og McCann (1982) og Lutz og Bettman (1977). Man har spesielt vært opptatt av hvorvidt slike viktighetsvekter skal måles ved hjelp av respondentenes selvrappotering i spørreskjemaet, eller om man skal estimere viktighetsvektene ved bruk av statistiske regresjonsteknikker, se f.eks.. Akaah og Korgaonkar (1983), Beckwith og Lehmann (1973), Jaccard, Brinberg, og Ackerman (1986), Mazis, Ahtola, og Klippel (1975) og Neslin (1981). Vi skal imidlertid la dette spørsmålet ligge, og retter oppmerksomheten mot en side ved bruk av viktighetskomponenter, eller vekter i spørreskjema som vi ikke kan se har vært studert tidligere.

Når multiattributtmodeller operasjonaliseres i et spørreskjema, og det inkluderes spørsmål for å måle viktigheten av at bestemte egenskaper ved produktet er tilstede, kan dette bidra til å påvirke respondentenes vurderinger av de samme egenskapene når disse er plassert etter viktighetsspørsmålene. På tilsvarende måte kan vurderinger av bestemte produktegenskaper tenkes å påvirke vektlegging av egenskapene dersom de spesifikke vurderingss spørsmålene stilles først i spørreskjemaet. Vi står altså overfor en problemstilling knyttet til spørsmålsrekkefølge.

Problemstillinger i forbindelse med rekkefølgen på spørsmålene i et spørreskjema er viet stor oppmerksomhet i tidligere forskning, se f.eks. Bradburn og Sudman (1981), Hippler, Schwarz, og Sudman (1987), Hogarth (1982), Schuman og Presser (1981), Schwarz og Sudman (1991), Sudman og Bradburn (1974), Sudman og Bradburn (1983) og Tanur (1991)¹⁷. Forutgående spørsmål i en spørreundersøkelse regnes ofte for den viktigste kilde til aktivering av informasjon som er relevant for å svare på etterfølgende spørsmål (Strack og Martin, 1987). Den psykologiske prosessen som oppstår i slike tilfeller kalles fremkalling (*priming*) (Fiske og Taylor, 1991; Strack og Martin, 1987).

Fremkalling kan bidra til aktivering av ulike typer informasjon som kan være relevant for respondentens svargivning (Strack og Martin, 1987). Et forutgående spørsmål i et spørreskjema vil på denne måten ha en fremkallingseffekt. En respondents tilgjengelighet (*accessibility*) (Fazio, 1989) til vurderinger av spesifikke egenskaper vil f.eks. øke etterhvert som egenskapene blir mer fremtredende (*salient*) (Bodenhausen og Wyer, 1987; Schuman og Presser, 1981).

¹⁶word of mouth

¹⁷I kapittel 2 gir vi en oversikt over en rekke empiriske studier der spørsmålsrekkefølge har vært tema.

I avhandlingen vil vi undersøke om viktighetsspørsmålene bør plasseres før eller etter de spesifikke vurderingsspørsmålene i spørreskjemaet. Som utgangspunkt for å belyse dette forholdet, velger vi å presisere følgende problemstilling som i kapittel 4 blir konkretisert ved hjelp av et sett testbare hypoteser

RQ3: *Hvordan påvirker inkludering og plassering av viktighets-spørsmål, før vs. etter spesifikke vurderingsspørsmål, i et spørreskjema svargivning ved operasjonalisering av multi-attributtmodeller?*

1.1.3 Produktkunnskap¹⁸

Produktkunnskap kan betraktes som et *ervert kjennetegn* ved personen som foretar vurderinger, og dermed som et av de forholdene som kan bidra til variasjoner i respondentenes svar på spørsmål i et spørreskjema (Cox, 1982; Cox, 1980; Guilford, 1954; Hippler, Schwarz, og Sudman, 1987; Torgerson, 1958). Erfaring med bruk av et produkt har blitt hevdet å føre til økt produktkunnskap (Alba og Hutchinson, 1987). Det at kundene er forskjellige med hensyn til den erfaringen de har med bruk av et produkt kan således få betydning for deres evne til å gjøre seg nytte av produktet.

I teoribidrag innenfor holdningsteori har man vist at erfaring med bruk av et produkt påvirker både holdningene man har til produktet, se f.eks. Fishbein og Ajzen (1975), og tilgjengeligheten¹⁹ til disse holdningene, se f.eks. Fazio (1989). Dette innebærer blant annet for det første at erfaring kan føre til læring (kunnskap), og for det andre at kunnskap kan gjøre det lettere å foreta vurderinger. I samsvar med dette argumenterer man for at kunder med høy produktkunnskap vil være bedre i stand til å bruke produktet på en "riktig" måte, noe som kan få betydning for deres vurderinger av-, og tilfredshet med produktet.

I konsumentatferdslitteraturen har det vært vanlig å dele kjøpsprosessen inn i fem trinn: (1) Problemerkjenning, (2) Informasjonssøking, (3) Evaluering av alternativer, (4)

¹⁸Litteraturgjennomgangen i forbindelse med denne avhandlingen har vist at det hersker relativt stor uklarhet i hvordan erfaring, kunnskap og ekspertbegrepet er definert og benyttet. Det samme gjelder hvordan begrepene er blitt målt empirisk. Som et resultat av dette vil alle begrepene bli benyttet i de teoretiske resonnementene vedrørende effekter av produktrelatert erfaring og produktkunnskap. Når vi på bakgrunn av teoridrøftelsene formulerer hypoteser i kapittel 4, vil vi imidlertid fokusere på produktkunnskapsbegrepet.

¹⁹Tilgjengelighet til holdninger innebærer at individet har holdninger lagret, og at disse kan hentes frem

Kjøpsbeslutning og (5) Aktiviteter etter kjøp, se f.eks. Engel og Blackwell (1982). Denne trinndelte modellen viser at kjøpsprosessen starter før kjøpet foretas, og slutter etter at selve kjøpet er avsluttet. I forbindelse med kjøpsprosessens *aktiviteter etter kjøp*, fokuseres det på *kundetilfredshet* og *kognitiv dissonans* som alternative resultater av de valg konsumenten har foretatt med hensyn til kjøpet (Engel og Blackwell, 1982). Effekter av erfaring og produktkunnskap på produktvurderinger etter kjøp og bruk av produktet er særlig behandlet i forbindelse med studier av kundetilfredshet (*customer satisfaction*) (Cadotte, Woodruff, og Jenkins, 1987; Mazursky og Geva, 1989; Oliver, 1980; Tse og Wilton, 1988; Westbrook, 1987; Yi, 1990). Kundetilfredshet kan oppfattes som et resultat av en vurderingsprosess der kunder antas å foreta en vurdering av produktet de har kjøpt og brukt i forhold til de forventninger de hadde før kjøpet.

Erfaring er også en sentral variabel i teoribidrag som tar for seg hvordan behandling av informasjon foregår i en informasjonsbehandlingsprosess. I atferdsforskning har man f.eks. studert hvordan ulik erfaring med-, eller kunnskap om et produkt kan påvirke en kjøpers beslutnings- eller kjøpsprosess (se f.eks. Alba og Hutchinson (1987), Alba og Marmorstein (1986), Bettman og Park (1980), Brucks (1985), Brucks (1986), Selnes (1986), Selnes og Troye (1989) og Sujan (1985)). Man har spesielt vært opptatt av betydningen av erfaring og produktkunnskap for hvilken type og mengde informasjon som brukes, og for hvilken informasjonsbehandlingsprosess individet benytter²⁰. Med bakgrunn i denne innledende litteraturgjennomgangen antar vi at erfaring med bruk av produktet, og økt produktkunnskap, kan påvirke vurderingsprosessene, og dermed også svargivning på vurderingsspørsmålene i et spørreskjema.

Tidligere forskning har vist at eksperter har høyere samsvar mellom vurderinger og atferd enn ikke-eksperter (Fazio, 1989). Man har også funnet at eksperter er mindre sensitive til forskjeller i undersøkelseskontekst når det gjelder sammenhengen mellom vurderinger og atferdsintensjoner enn ikke-eksperter (Hutchinson, 1983; Lynch, Chakravarti, og Mitra, 1991). Vi antar med dette som utgangspunkt at effektene vi forventer å finne som resultat av forskjellig bruk av instruksjon og vurderingsskala, vil bli moderert av respondentenes ekspertise og produktkunnskap.

En person som har brukt et produkt flere ganger vil sannsynligvis oftere ha foretatt både "bra-dårlig" og "tilfreds-utilfreds" slutninger knyttet til produktet. Vi antar derfor at individer med høy produktrelatert erfaring kan ha et bedre grunnlag for å foreta objektforankrete ("bra-dårlig") vurderinger enn individer med lav erfaring. Vi har med bakgrunn i denne innledende litteraturgjennomgangen, valgt å se nærmere på hvordan

²⁰Informasjonsbehandlingsprosesser behandles grundigere i kapittel 3.

erfaring med bruk av produktet kan virke inn på respondentenes vurderingsprosess, når det benyttes henholdsvis subjektforankrete- (“tilfreds-tilfreds”) og objektforankrete (“bra-dårlig”) instruksjoner og -vurderingsskalaer i forbindelse med vurderings-spørsmålene i spørreskjemaet. Vår fjerde problemstilling som konkretiseres i et sett test-bare hypoteser i kapittel 4 blir derfor:

RQ4: *Hvordan påvirker respondentenes produktkunnskap svar-givning på vurderings spørsmål ved bruk av subjekt-forankret- vs. objektforankret instruksjon og -vurderings-skala?*

Økt brukserfaring kan, på samme måte som forutgående spørsmål i spørreskjemaet, bidra til at holdninger og produktvurderinger blir mer tilgjengelige for respondentene, se f.eks. Fazio (1989). Vi antar derfor at produktkunnskap også kan bidra til å moderere effektene vi forventer ved å inkludere-, og plassere viktighetsspørsmål før vs. etter spesifikke vurderings spørsmål i spørreskjemaet. Som utgangspunkt for å belyse dette forholdet reiser vi følgende problemstilling.

RQ5: *Hvordan påvirker respondentenes produktkunnskap svar-givning ved plassering av viktighetsspørsmål, før vs. etter spesifikke vurderings spørsmål, i et spørreskjema ved operasjonalisering av multiattributtmodeller?*

1.1.4 Kvalitetsnivå på produktet som vurderes

Et sentralt perspektiv for studier av kundetilfredshet har vært *the disconfirmation paradigm* (Oliver, 1980; Churchill og Surprenant, 1982). I følge dette perspektivet kan tilfredshet relateres til størrelse og retning på avvik mellom forventninger og opplevd ytelse. Forventninger bekreftes (*confirmed*) når produktets ytelse er som forventet. Når produktets ytelse er dårligere enn forventet, blir forventningene negativt avkreftet (*negatively disconfirmed*), mens det motsatte er tilfellet når produktets ytelse er bedre enn forventet. I dette tilfellet blir forventningene positivt avkreftet (*positively disconfirmed*). Utilfredshet blir definert til å oppstå når individets forventninger blir negativt avkreftet. Forventninger reflekterer her “antatt ytelse”. I kundetilfredshets-litteraturen eksisterer det fire typer forventninger om et produkts antatte ytelse: ideal-

ytelse, forventet ytelse, minimums tolerert ytelse og ønskelig ytelse (Churchill og Surprenant, 1982). Tilfredshetsvurderinger kan med utgangspunkt i dette betraktes som en kategoriseringsprosess (Mervis og Rosch, 1981) der produktets ytelse eller kvalitet får betydning for vurderingene.

Assimilasjon-kontrasteori (Sherif og Hovland, 1961)²¹ har vært et viktig utgangspunkt for å forklare forklare avvik mellom et produkts ytelse og forventninger i litteratur om kundetilfredshet (Anderson, 1973; Oliver og DeSarbo, 1988). Sammenholdt med element fra dissonansteori (Festinger, 1957), innebærer perspektivene fra assimilasjon-kontrasteori at vurderinger vil assimileres mot forventningene når avvikene mellom opplevd produktytelse og forventninger er små. Det motsatte av slike assimilasjonseffekter oppstår når et individ forstørker avvikene mellom opplevd ytelse og forventninger når avvikene er store. Dette kalles kontrasteffekter. I kundetilfredshetslitteraturen (f.eks. Anderson (1973)) benyttes teori om kontrasteffekter for å beskrive, både hvordan produktytelse som er lavere enn forventet vil bli vurdert dårligere enn det den i virkeligheten er, og hvordan produktytelse som er høyere enn forventet vurderes høyere (se f.eks. Oliver (1988)).

Vi har tidligere i dette kapitlet definert tilfredshetsvurderinger som å reflektere indre følelsesmessige tilstander, mens "bra-dårlig" vurderinger ble definert til å representere respondentens mer objektive vurdering av et objekt. Det forventes her at respondentene opptrer som "dommere" og vurderer produkter uavhengig av egne forventninger og normer (Tse og Wilton, 1988). Bruk av "bra-dårlig" instruksjon og -vurderingsskala kan derfor føre til at respondentene føler at de ikke kan bruke egne normer og forventninger som basis, eller sammenligningsstandard for vurderinger. Tilfredshetsvurderinger kan også i større grad enn "bra-dårlig" vurderinger tenkes å involvere respondentene i vurderingene.

Dersom kundenes forventninger er konstante på tvers av produktalternativer innenfor en bestemt produktgruppe, er det sannsynlig at ytelsen for høykvalitetsprodukter vil overstige kundenes forventninger. Vi har tidligere beskrevet hvordan både høy- og lav kvalitet kan føre til kontrasteffekter som resultat av at høykvalitetsprodukter vil ha en ytelse som overstiger kundenes forventninger, mens lavkvalitetsprodukter vil ha en ytelse som ikke når opp til forventningene. Med et slikt utgangspunkt kan det tenkes at nivået på "tilfreds-utilfreds" og "bra-dårlig" mål vil være forskjellig avhengig av ytelsen, eller kvalitetsnivået på produktene som vurderes. Vi ønsker å studere dette og reiser følgende problemstilling som konkretiseres i testbare hypoteser i kapittel 4:

²¹ Som referert av Yi (1990).

RQ6: *Hvordan påvirker kvalitetsnivået på produktet som vurderes svargivning på vurderingsspørsmål ved bruk av subjektforankret- vs. objektforankret instruksjon og -vurderingsskala?*

Svargivning på viktighetsspørsmål kan aktivere sammenligningsstandarder eller “ideal-produkter” som vurderingene relateres til. Når viktighetsspørsmål inkluderes, og plasseres før de spesifikke vurderingsspørsmålene i et spørreskjema, kan slike sammenligningsstandarder i større grad bli benyttet når respondentene foretar vurderinger av produktet. Avhengig av kvaliteten på produktet respondenten har erfart, og blir bedt om å vurdere kan både assimilasjons- og kontrasteffekter oppstå. Vi antar derfor at kvalitetsnivået på produktene som vurderes også kan bidra til å moderere effektene vi forventer ved å inkludere-, og plassere viktighetsspørsmål før vs. etter spesifikke vurderingsspørsmål i spørreskjemaet. Som utgangspunkt for å belyse dette forholdet reiser vi følgende problemstilling.

RQ7: *Hvordan påvirker kvalitetsnivået på produktet som vurderes svargivning ved plassering av viktighetsspørsmål, før vs. etter spesifikke vurderingsspørsmål, i et spørreskjema ved operasjonalisering av multiattributtmodeller?*

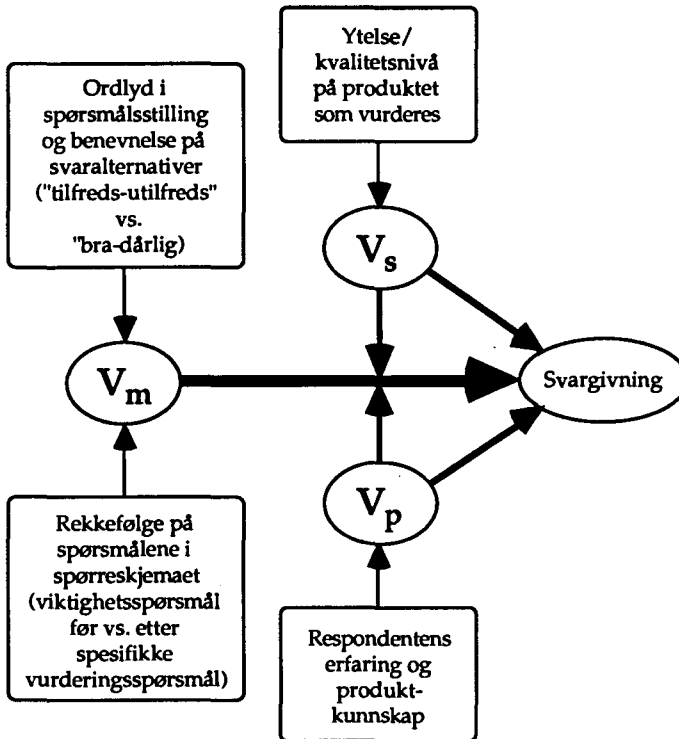
1.2 Oppsummering og implikasjoner

Som avslutning på dette introduksjonskapitlet vil vi illustrere de variablene vi fokuserer på i en modell som oppsummerer problemstillingene vi tar opp i avhandlingen. I figur 1.1 har vi derfor vist hvordan de tre hovedkildene til variasjoner i svargivning: egenskaper ved stimuli (V_S), egenskaper ved respondenten (V_P) og egenskaper ved måleprosedyren (V_M) kan relateres til problemstillingene vi har reist i dette kapitlet.

Hovedfokus i avhandlingen er variasjoner i svargivning som skyldes måleprosedyren (V_M), representert ved ordlyd i spørsmålsstilling, benevnelse på svarskala og spørsmålsrekkefølge. Måleprosedyren antas å ha en direkte effekt på svargivning. Når det gjelder varians som skyldes stimuli (V_S) har vi introdusert produktets ytelse eller kvalitetsnivå som et forhold som kan påvirke Egenskaper ved stimuli hevder vi altså vil

moderere effektene av skalavalg. Til slutt illustrerer modellen hvordan egenskaper ved personen som foretar vurderingene (V_p), representert ved respondentenes produktrelaterte erfaring og produktkunnskap, antas å ha en direkte effekt på svargivning, i tillegg til å moderere effektene av instruksjon og vurderingsskala og rekkefølge på spørsmålene i spørreskjemaet.

FIGUR 1.1
Skjematisk avbildning av problemstillingene som blir tatt opp i avhandlingen



Hensikten med å presentere denne modellen er å gi en oversikt over de forholdene vi ønsker å fokusere på i avhandlingen. Med utgangspunkt i litteraturgjennomgangen kan vi så langt bare hevde at de tre hovedkildene til variasjoner i svargivning kan gi opphav til forskjeller i respons. Vi har derfor ikke konkretisert hvilke effekter som kan forventes i modellen. I de to neste kapitlene vil vi belyse problemstillingene med utgangspunkt i relevante metodologiske og teoretiske bidrag, med formål å forsøke å predikere type effekter, samt retning og styrke på effektene.

Avhandlingen, som kan plasseres i grenseområdet mellom undersøkelsesmetode og teorier for individers informasjonsbehandling, har både et metodologisk og et teoretisk formål. Det metodologiske formålet er å *belyse konsekvenser av en rekke valg* man står overfor ved utvikling av spørreskjema til bruk for måling av kunders produktvurderinger. Mer spesifikt vil vi som nevnt, fokusere på ulike responseffekter, herunder type effekter, samt retning og styrke på effektene. Avhandlingen kan således betraktes som et bidrag i arbeidet med å forbedre teknikker for utvikling av spørreskjema og prosedyrer i forbindelse med måling av kunders produktvurderinger. Det teoretiske formålet er å forsøke å knytte sammen teoribidrag fra ulike områder innenfor kognisjon, informasjonsbehandling, og konsumentatferd for å forstå hvordan kunders produktvurderinger kan komme til uttrykk som svar på spørsmål i et spørreskjema. Teoribidragene blir her kun benyttet for å forstå *hvorfor*, og *hvordan* de ulike responseffektene fremkommer. De ulike teoribidragene vil ikke, i seg selv, bli gjort til gjenstand for teoritesting. I den utstrekning det er mulig, vil det imidlertid fremsettes *teori-baserte* testbare hypoteser knyttet til de metodologiske valgene. Disse presiseringene, med hensyn til det teoretiske og metodologiske formålet med avhandlingen, får en rekke implikasjoner, både for litteraturgjennomgangen vi foretar for å belyse problemstillingene, og for undersøkelsesopplegget vi velger for innsamling av data som grunnlag for testing av hypoteser.

På mange måter kan utfordringen vi står overfor i forbindelse med problemstillingene vi har reist i dette kapittelet beskrives som "metodologiske problemer som søker teoretiske forklaringer". En rekke teoretiske perspektiver vil være relevante for å belyse effekter av de metodologiske valgene. Som et felles utgangspunkt for å studere effekter av ordlyd i spørsmålsstilling, benevnelse på svaralternativer og spørsmålsrekkefølge vil det være naturlig å starte med litteratur som beskriver hvordan kunder foretar produktvurderinger, og hvordan disse kan tenkes å komme til uttrykk som svar på spørsmål i et spørreskjema. Med et slikt felles utgangspunkt vil de spesifikke problemstillingene deretter kreve ulike teoretiske tilnærminger for å forstå hvilke responseffekter som kan oppstå som følge av de metodologiske valgene. En implikasjon av dette er at litteraturgjennomgangen vil bli relativt omfattende.

De metodologiske valgene vi fokuserer på i avhandlingen, er valg man til enhver tid står overfor når et spørreskjema skal benyttes for innsamling av data. Når vi har presisert at formålet med avhandlingen ikke er å teste teoriene vi benytter som grunnlag for å fremsette hypoteser, men å fokusere på effektene som teoriene foreskriver, får dette implikasjoner for valg av undersøkelsesopplegg. På mange måter kan forskningsoppgaven vi står overfor oppfattes som et eksempel på det Calder, Phillips, og Tybout (1981) definerer som effektanvendelse (*effects application*). Et viktig utgangspunkt for

drøfting av undersøkelsesopplegg vil dermed være å sørge for at funn kan generaliseres til en realistisk situasjon. I kapittel 5 vil disse presiseringene danne grunnlag for en grundigere drøfting av undersøkelsesopplegg for den empiriske studien.

1.3 Avhandlingens hoveddeler

Vi startet dette kapitlet med en presentasjon av utgangspunktet for avhandlingen. Her ble det fokusert på en rekke valg man, både som akademiker og praktiker, må foreta i forbindelse med utforming av spørreskjema til bruk ved for innsamling av data i spørreundersøkelser. Disse valgene leder oss over til problemstillinger vedrørende ulike kilder til variasjon i svargivning i spørreskjemaundersøkelser. Hovedspørsmålene vi reiser for de empiriske delene av avhandlingen er hvordan bruk av objektforankret- vs. subjektforankret instruksjon og -vurderingsskala ved måling av kunders produktvurderinger kan påvirke respondentenes svargivning i et spørreskjema. I tillegg reiser vi spørsmål om inkludering og plassering av viktighetsspørsmål, før vs etter spesifikke vurderingss spørsmål i spørreskjemaet, kan påvirke svargivningen.

Avhandlingen omfatter 5 deler og 10 kapitler. I del II, som består av kapittel 2 og 3, gjennomgår vi relevante metodologiske- og teoretiske bidrag for å velge de sentrale begrepene og belyse sammenhengene mellom dem. I kapittel 2 redegjør vi for hvordan problemstillingene vi tar opp i avhandlingen kan belyses med utgangspunkt i sentrale bidrag fra måleteori og spørreundersøkelsesmetode. Vi tar her utgangspunkt i tre hovedkilder til variasjon i svargivning: Varians som skyldes måleprosedyren, varians som skyldes produktet eller objektet som vurderes og varians som skyldes respondentene. Kapittel 3 vil bli viet en drøfting av teoretiske perspektiver som kan belyse skillet mellom objektforankrete og subjektforankrete måleskalaer. I tillegg ser vi her også nærmere på teoretiske perspektiver som kan forklare effekter av inkludering-, og alternativ plassering av viktighetsspørsmål i spørreskjemaet.

Del III starter med presentasjon av en analysemodell med tilhørende hypoteser i kapittel 4. I kapittel 5 gjennomgår vi det metodologiske grunnlaget, og gjør rede for undersøkelsesopplegget som ble benyttet for de empiriske studiene vi har gjennomført for å teste hypotesene. Spørsmål knyttet til forskningsdesign blir her gjennomgått og diskutert. Hvordan de avhengige og uavhengige variablene operasjonaliseres i spørreskjemaet blir deretter gjort grundig rede for i kapittel 6. I dette kapitlet foretar vi en grundig drøfting av begrepsvaliditet.

I del IV foretar vi først, i kapittel 7, en beskrivelse av datamaterialet som ble samlet inn for å teste hypotesene. I tillegg til presentasjon av beskrivende statistikk drøftes valg av statistiske analyseteknikker, og det presenteres resultater fra tester som ble foretatt for å undersøke om dataene tilfredsstillende forutsetningene som ligger til grunn for gjennomføring av analysene. Etter denne innledende inspeksjonen av dataene gjennomføres testing av hypotesene i kapittel 8.

I del V, som består av kapittel 9 og 10, diskuterer vi resultatene fra hypotesetestingen. I forbindelse med diskusjonen i kapittel 9, fokuserer vi spesielt på studiens praktiske implikasjoner for utvikling av spørreskjema ved gjennomføring av spørreundersøkelser. Til slutt, i kapittel 10 drøfter vi begrensninger i den gjennomførte studien, og antyder implikasjoner for videre forskning.

DEL

II

LITTERATUR-
GJENNOMGANG

I denne delen av avhandlingen drøfter vi først problemstillingene som ble reist i kapittel 1, i forhold til metodologiske perspektiver i kapittel 2. Deretter, i kapittel 3, drøftes problemstillingene i forhold til kognitive perspektiver på svargivning i spørreundersøkelser.

2

Metodologiske perspektiver på svargivning i spørreundersøkelser

Når man, både som akademiker og praktiker, står overfor problemstillinger i tilknytning til metodologiske valg ved utvikling av et spørreskjema, er det naturlig å starte med å søke hjelp i metodelitteraturen. Dette gjelder også for problemstillingene vi reiste i foregående kapittel. I tillegg til en rekke generelle tommelfingerregler (se f.eks Churchill (1991)), er spørreskjema utvikling i metodelitteraturen nært knyttet til måling. Måling og måleproblemer er viet stor plass i teoriutvikling og forskning innenfor atferdsfagene, se f.eks Blalock (1982), Churchill (1979), Cox (1982), Ghiselli, Campbell, og Zedeck (1981), Hippler, Schwarz, og Sudman (1987), Nunnally (1978), Nunnally og Bernstein (1994), Sudman og Bradburn (1974), Sudman og Bradburn (1983) og Zeller og Carmines (1980). Her belyses blant annet ulike former for feilkilder og responseffekter som kan oppstå når spørreskjema benyttes for innsamling av data.

I dette kapittelet vil vi ta for oss en rekke sentrale og relevante bidrag som på forskjellig måte har fokusert på ulike feilkilder i spørreskjemaundersøkelser under overskriften *data kvalitet*. Følgende sitat, hentet fra Groves (1987), illustrerer det utgangspunktet denne forskningen gir for å belyse problemstillingene som blir tatt opp i avhandlingen:

Metodelitteraturen, med de ulike innfallsvinklene som benyttes for å studere forhold ved gjennomføring av undersøkelser med bruk av spørreskjema, kan ikke i seg selv sees på som en akademisk disiplin med et felles fagspråk, et felles sett av prinsipper for å evaluere nye ideer, og en velorganisert profesjonell faglig referansegruppe. Mangelen av en slik faglig organisering har ført til at forskning omkring gjennomføring av spørreundersøkelser med bruk av strukturerte spørreskjema er drevet frem av uavhengige og tildels ukoordinerte bidrag fra forskere innenfor fagområder som statistikk, psykologi og sosiologi. En slik kunnskapsgenerering med spredte og ad-hoc pregede koplinger mellom undersøkelsesmetodologi og ulike teoriretninger er grunnlaget for det man i dag mener å vite om hva som bidrar til undersøkelseskvalitet og datakvalitet (Groves 1987).

Status quo innenfor forskning omkring spørreundersøkelser, og bruk av spørreskjema gjenspeiler altså et sammensatt bilde der datakvalitet er definert som *graden av fravær av feilkilder i undersøkelsene*. Hvilke feilkilder det fokuseres på avhenger i vesentlig grad av *hva* som er formålet med å gjennomføre en bestemt spørreundersøkelse, men også ofte av *hvem* det er som står for gjennomføringen (Groves 1987). Vi finner f.eks et skille mellom forskere som gjennomfører spørreundersøkelser for å *beskrive* populasjoner og forskere som benytter data innsamlet gjennom spørreundersøkelser for å *teste* teorier. Videre finnes en rekke forskere som arbeider for å komme frem til empiriske estimater på de ulike feilkildene som kan oppstå i spørreundersøkelsene. Endelig finner vi forskere som ser som sin hovedoppgave å arbeide for å redusere eller eliminere disse feilkildene. Den delen av metodelitteraturen som spesielt tar for seg spørreundersøkelser, og utvikling av spørreskjema til bruk i disse, er sammensatt av bidrag skrevet av representanter fra alle disse gruppene, som hver for seg er opptatt av sin "favorittfeilkilde". Dette kan ha bidratt sterkt til den fragmenterte kunnskapsoppbyggingen som kjennetegner dette forskningsfeltet.

Innledningsvis i kapitlet gir vi først en oversikt over ulike perspektiver som alle fokuserer på kilder til variasjoner i svargivning i spørreundersøkelser. Felles for alle disse perspektivene er at den delen av variansen i respons på et spørsmål som ikke kan betegnes som *sann*, defineres som en *feilkilde* i undersøkelsene. Vi tar utgangspunkt i en klassifisering som definerer fire hovedtyper feilkilder: *dekningsfeil*, *utvalgsfeil*, *ikke-responsfeil* og *målefeil* (Alwin 1991; Groves 1987). Med dette utgangspunktet går vi så videre til *nyklassisk testteori* der observert varians ved en måling i hovedsak tilskrives tre komponenter: *sann varians*, *systematisk feilvariens* og *tilfeldig feilvariens* (Alwin 1991; Churchill 1979; Gerbing og Anderson 1988; Groves 1987; Nunnally 1978; Nunnally og Bernstein 1994; Zeller og Carmines 1980). I denne delen av kapitlet blir problemstillingene som ble reist i kapittel 1 relatert til de ulike feilkildene som behandles i metodelitteraturen.

Operasjonalisering og måling av teoretiske begreper har sitt utspring i den subjektive svargivningen til individene som svarer på spørsmål, f.eks ved å fylle ut et spørreskjema. Denne svargivningen kan være et resultat av mange typer påvirkning, som hver kan betraktes som en kilde til varians. Hovedkildene til varians kan imidlertid oppsummeres i tre kategorier: (1) karakteristika ved responsstimuli, (2) karakteristika ved respondenten og (3) karakteristika ved spørsmålene og måleprosedyren som blir benyttet (Cox 1982; Torgerson 1958). Med utgangspunkt i disse tre hovedkildene til varians i målinger, relateres problemstillingene fra kapittel 1 til skaleringslitteraturens skille mellom subjektssentret-, stimulusentret- og responskalaer (Cox 1982; Torgerson 1958). Denne tredelte, funksjonelle klassifiseringen har som utgangspunkt at

den, eller de, kildene til variasjon i svargivning som forskeren ønsker å ta hensyn til behandles som systematisk og sann varians.

Til slutt i kapittelet gir vi en oversikt over implikasjoner av en rekke sentrale og relevante empiriske studier, som i hovedsak er gjennomført med bakgrunn i de metodologiske perspektivene vi har presentert.

2.1 Kilder til variasjon i svargivning i spørreundersøkelser

I det følgende presenteres ulike perspektiver på måling der den totale variasjon i svargivningen på et spørsmål i et spørreskjema deles opp i delkomponenter.

2.1.1 En ramme for klassifisering av feilkilder i spørreundersøkelser

I empirisk forskning har man i lengre tid vært opptatt av *målingstrusler*, dvs. faktorer som kan føre til at de målinger som foretas blir feilaktige. Det har vært vanlig å dele målingstruslene inn i to hovedklasser, *feil som et resultat av manglende observasjoner* og *mdlefeil* (se f.eks. (Alwin 1991; Churchill 1991; Groves 1987).

Feil forbundet med manglende observasjoner, dvs. at en del av populasjonen ikke måles (dekningsfeil, ikke-responsfeil og utvalgsfeil), synes å være mest fokusert av forskere som er opptatt av å generalisere resultatene fra en bestemt undersøkelse til en veldefinert populasjon (f.eks. slik som i meningsmålinger).

Dekningsfeil oppstår når medlemmer av den definerte populasjonen ikke er med i utvalgsrammen. Denne feiltypen blir ofte ignorert av forskere som ikke har som formål å estimere karakteristika ved en gitt populasjon, men som fokuserer på sammenhenger mellom variabler. Oppmerksomheten omkring dekningsfeil er i hovedsak blitt rettet mot forsøk på å redusere disse, f.eks. gjennom å oppfordre til å foreta en bedre definisjon av utvalgsrammen som skal representere populasjonen en bestemt undersøkelse er rettet mot (Groves 1987).

Ikke-respons feil oppstår når det oppnådde utvalget i en undersøkelse ikke er dekkende for alle individer som er definert i utvalgsrammen. Tradisjonelt har responsrate i forbindelse med en undersøkelse blitt benyttet som en indikator på ikke-responsfeil (Groves 1987). I tillegg finnes det nå tilgjengelig modellverktøy for å estimere betydningen av ikke-respons. Slike estimater kan benyttes for å korrigere dataene for de gruppene som er underrepresentert i undersøkelsen.

Når vi f.eks. av praktiske eller økonomiske årsaker, ikke kan undersøke alle enheter i den populasjonen vi studerer, må vi trekke et utvalg av enheter og begrense observasjonene til disse. Ved slike undersøkelser oppstår ofte *utvalgsfeil* (*sampling-error*). Utvalgsfeil kan reduseres og kontrolleres for dersom enhetene som er med i utvalget trekkes ut ved "loddtrekning" (tilfeldig utvalg) slik at sannsynligheten for at en enhet skal komme med i utvalget er kjent. Dersom dette kravet kan imøtekommes, gir statistisk teori oss mulighet for å beregne størrelsen på de feil som kan oppstå når vi trekker konklusjoner om populasjonen på grunnlag av et utvalg. Under de samme betingelsene gjelder også at feilmarginene blir mindre, jo større andel av populasjonen som er med i utvalget og frafallet fordeler seg tilfeldig. Utvalgsfeilene vil først og fremst svekke en undersøkelses eksterne validitet.

Viktige grunner for at feilkilder kan oppstå, kan altså være at utvalgsrammen utelukker deler av populasjonen, eller at det kan oppstå selvseleksjon med hensyn til hvem som svarer, i form av både nektelses- og kontaktbortfall. I avhandlingen tar vi ikke hensyn til disse ovenfornevnte feilkildene. Disse inkluderes i den tilfeldige varians.

Den andre hovedklassen av målingstrusler kalles *målingsfeil* (*nonsampling-error*). Målingsfeil henspiller på at det er svakheter knyttet til selve målingen. Det kan f.eks. tenkes at de empiriske indikatorene vi har valgt dekker et område som ligger utenfor det teoretiske begrepet. Målingsfeilen vil da ligge i at vi egentlig måler et annet teoretisk begrep enn det vi tror vi måler. Målingsfeil kan også være knyttet til unøyaktigheter i måleinstrumentet. Responseeffekter er eksempel på en slik type målingsfeil som kan opptre i undersøkelser hvor det er fare for at forskeren ufrivillig påvirker undersøkelsesobjektene atferd på en måte som gir seg utslag i svargivningen.

Tradisjonelt har målingsfeil blitt knyttet til fire kilder: (1) Målingsfeil som skyldes intervjueren, (2) målingsfeil som skyldes respondenten, (3) målingsfeil som skyldes spørreskjemaet som benyttes, og (4) målingsfeil som skyldes administrasjon av data-innsamlingen, se f.eks. Cantril (1944) og Sudman og Bradburn (1974). I avhandlingen retter vi fokus mot denne typen feilkilder, og legger altså hovedvekt på hvordan forhold ved utforming av et spørreskjema kan påvirke svargivning.

2.1.2 Klassisk test teori

I følge klassisk test teori (*classical test theory*) blir enhver måling definert til å bestå av to komponenter: en *sann* komponent og en *feil* komponent (Churchill 1979; Nachmias og Nachmias 1985; Nunnally 1978; Nunnally og Bernstein 1994; Zeller og Carmines 1980). Observert score på et bestemt spørsmål for en gitt person kan dermed uttrykkes på følgende måte:

$$X = T + e$$

der X er den observerte score på spørsmålet i spørreskjemaet, T er den sanne score og e er størrelsen på feil i målingen.

I klassisk test-teori antar man at all målefeil er tilfeldig fordelt. Dette kan være en reell forutsetning i målinger foretatt i strengt utformet, eksperimentelle design. I undersøkelsesopplegg der data innsamles i utvalgsundersøkelser ved bruk av spørreskjema kan en slik forutsetning vise seg ikke å holde. For å ta hensyn til dette kan klassisk test-teori reformuleres slik at feilleddet brytes opp i *systematisk* og *tilfeldig* målefeil. Observert score kan dermed uttrykkes på følgende måte:

$$X = T + S + R$$

der X fortsatt er den observerte score på et spørsmål i et spørreskjema, T er den sanne score, mens e nå er brutt opp i systematisk feil, S , og tilfeldig feil, R (Churchill 1979; Zeller og Carmines 1980). Forskningsspørsmålene som blir tatt opp i avhandlingen er tradisjonelt blitt betraktet som forhold knyttet til systematisk feil i formelen ovenfor.

Effekter av ulik instruksjon og vurderingsskala for vurderingss spørsmål, og tilstedeværelse og ulik plassering av viktighetsspørsmål vil være systematiske i en situasjon der et tilfeldig utvalg respondenter blir bedt om å svare på spørsmål ved å fylle ut det samme spørreskjemaet. Systematisk feil kan også oppstå når respondentene tenderer til alltid å være enig med det han tror er intervjuerens meninger²², eller når det svares på en måte man oppfatter som å være i samsvar med det som ansees som sosialt akseptabelt (Sudman og Bradburn 1974).

Alwin (1991) detaljerer formelen ovenfor ytterligere ved å dele responsen på et bestemt spørsmål, y , på et gitt tidspunkt, t , inn i fem komponenter:

²²(*yeah-saying*)

1. Den "sanne" variabelen som måles, τ_t ²³
2. Målingsskjevhet eller andre forhold ved målingene som er konstante, f.eks på grunn av ordlyd på spørsmål, form på spørsmål eller kontekst²⁴. Denne komponenten kalles ω_t
3. Svakheter knyttet til feil ved begrepsdefinisjoner og/eller operasjonalisering. Dette vil da være andre begreper som måles med spørsmålet. Disse kalles η_t
4. Målingsfeil som er tilfeldig, ϵ_t
5. Målingsfeil som er tilfeldig med hensyn på τ_t , men som er korrelert med målingsfeil knyttet til målinger foretatt på ulike tidspunkt. Denne type målingsfeil kan betraktes som en stabil feilkomponent som er korrelert over tid og benevnes υ_t

Dette betyr at et gitt spørsmål, y , i en spørreundersøkelse med svarkategorier y_1, y_2, \dots, y_r , kan oppfattes som en funksjon av følgende komponenter:

$$y = \omega + \tau + \eta + \upsilon + \epsilon$$

der ω representerer målingsskjevhet, og de andre komponentene er kilder til varians²⁵. For forskningsspørsmålene som blir studert i avhandlingen vil dette bety at vi fokuserer på komponenten som kalles ω_t , og som er knyttet til målingsskjevhet eller andre forhold ved målingene som er konstante, f.eks på grunn av ordlyd på spørsmål, form på spørsmål eller kontekst (spørsmålsrekkefølge)

Som standard for å evaluere godheten av målinger benyttes ofte begrepene *reliabilitet* og *validitet*. En måling regnes som valid når forskjeller i observerte scorer kun reflekterer *sanne* forskjeller i de karakteristika eller egenskaper man forsøker å måle. En måling er reliabel i den grad uavhengige, men sammenlignbare mål på de samme karakteristika, egenskaper eller begreper er sammenfallende. Reliabilitet er altså definert som grad av konsistens i målingene. Formlene vi har presentert ovenfor danner utgangspunkt for reliabilitets- og validitetsdrøftinger i empiriske målinger. Reliabilitet blir da definert som andelen av ikke-tilfeldig varians, mens validitet defineres som den andel av variansen som de observerte score har felles med de sanne score (Zeller og Carmines 1980). Reliabilitet og validitet omfattes av begrepet *datakvalitet*. Problem-

²³Sann score, τ_t er her definert som i psykometrisk teori, og inkluderer muligheten for at det eksisterer objektive "sanne scorer" ved å la gjennomsnittsavvikene for disse bli absorbert i målingsskjevhet.

²⁴Konteksteffekter er i litteraturen definert på flere måter. I denne avhandlingen vil vi imidlertid følge (Schuman og Presser 1981) som definerer spørsmålsrekkefølge (*question order*) til å gjelde forhold mellom spørsmål, som inkluderer konteksteffekter (meningsoverføring) og sekvens effekter (mekaniske forhold).

²⁵Her er subskriften t som er knyttet til tidspunkt for målingen utelukket. Med unntak av ω og ϵ er alle komponentene i responsen korrelert over tid.

stillinger relatert til dem som blir tatt opp i denne avhandlingen, er i litteratur behandlet med utgangspunkt i metodologiske validitet- og reliabilitetsdisipliner f.eks Alwin (1991), Alwin og Krosnick (1991), Andrews (1984), Box-Cicchetti, Showalter, og Tyrer (1985), Jones og Lang (1982).

2.1.3 Karakteristika ved stimuli - respondent - måling

En alternativ måte å illustrere feilkilder i målinger er å ta utgangspunkt i hovedkildene til variansen på et gitt spørsmål. Hovedkildene til varians kan deles inn i tre kategorier: karakteristika ved responsstimuli, karakteristika ved respondenten og forhold ved målingen (Cox 1982). Denne inndelingen er illustrert i figur: 2.1.

TABELL 2.1
Kilder til varians i psykometriske målinger

| | |
|------|--|
| I. | Karakteristika ved stimuli |
| | A. Generelle |
| | B. Spesielle |
| II. | Karakteristika ved respondenten |
| III. | Forhold ved målingen |
| | A. Tidspunkt for datainnsamlingen |
| | B. Datainnsamlingsmedium |
| | C. Utforming av spørreskjema |
| | D. Spørsmålsform |
| | E. Andre kilder (f.eks lokalisering, tilgjengelig tid, osv.) |
| IV. | Interaksjoner |
| | A. Stimuli - Respondent |
| | B. Stimuli - Forhold ved målingen |
| | C. Respondenter - Forhold ved målingen |
| | D. Andre interaksjoner |
| V. | Feil |

Kilde: Cox (1982)

Ethvert datainnsamlingsproblem kan, med bakgrunn i fremstillingen i tabell 2.1, belyses med utgangspunkt i et variansanalyse perspektiv. Hver kilde til systematisk varians kan representeres ved en faktor (direkte effekt), eller som en interaksjon mellom faktorer. I et bestemt forskningsdesign vil vi vanligvis ha mulighet til å ta hensyn til, og kontrollere for bare et fåtall av de mulige varianskildene.

De tre hovedkildene til varians som ble presentert i tabell 2.1 kan illustreres med utgangspunkt i et eksempel der kvalitet ved hotell fra en hotellkjede skal vurderes av et utvalg respondenter som fyller ut et spørreskjema.

Karakteristika ved stimuli vil i vårt eksempel være karakteristika ved hotellene som vurderes. Generelle karakteristika (I.A.) er representert ved det totale gjennomsnitt for vurderinger av alle hotellene, og indikerer i hvilken grad de (som kjede) vurderes som kvalitetshoteller. Det totale gjennomsnittet for alle hotellene kan kalles "kjedenivå". De spesielle karakteristika (I.B.) vil være representert ved gjennomsnittet for hvert hotell. Disse gjennomsnittene vil indikere forskjeller mellom hotellene (*between stimuli variation*).

Karakteristika ved respondenten (II) reflekterer i hvilken grad respondentene skiller seg fra hverandre i sin vurdering av kvaliteten ved hotellene (*between respondent variation*). Her vil f.eks. ulik segmenttilhørighet og erfaringsbakgrunn kunne ha betydning.

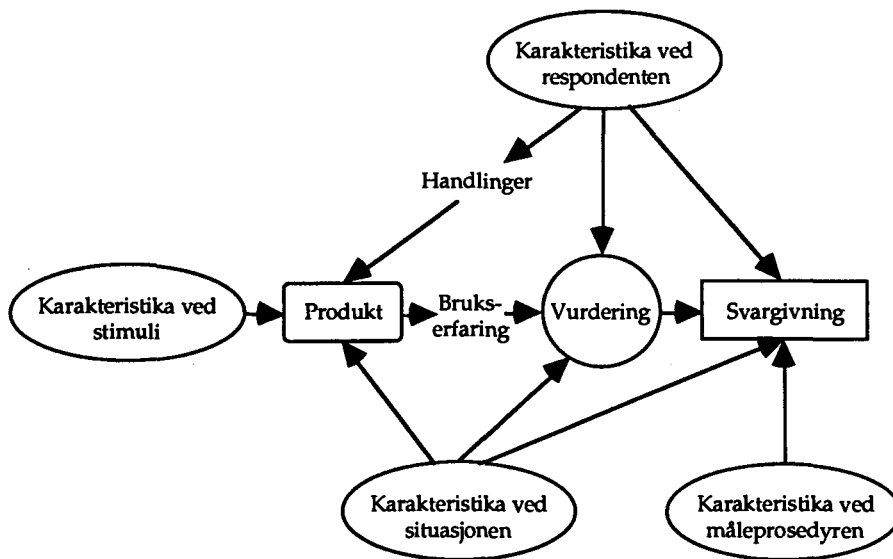
Forhold ved målingen (III) vil også være viktige kilder til variasjon i svargivning. For det første kan respons variere med hensyn til når dataene er innsamlet. Vurdering av kvalitet ved norske hoteller kan f.eks. variere med når på året undersøkelsen er foretatt. For det andre kan respons påvirkes av hvilket medium som benyttes i datainnsamlingen, f.eks. bruk av telefonintervju, postal spørreundersøkelse eller personlig intervju. For det tredje kan intervjueren påvirke respons og dermed være en kilde til varians. Disse tre første forholdene er knyttet til selve administrasjonen av undersøkelsen. For det fjerde kan utformingen av spørreskjemaet endre responsen på et bestemt spørsmål, f.eks. som resultat av haloeffekter²⁶ eller effekter av rekkefølge på spørsmålene. For det femte kan forhold ved selve spørsmålsstillingen påvirke respons gjennom antall svaralternativer, ordlyd i spørsmålsstilling og svaralternativer, format osv. Det er disse to siste forholdene som er hovedfokus for denne avhandlingen.

I tillegg til de direkte effektene, finnes det en rekke mulige *interaksjonseffekter*. Den viktigste interaksjonen er kanskje den mellom stimuli (I) og respondentene (II). Denne interaksjonen indikerer i hvilken grad ulike respondenter vurderer de ulike hotellene forskjellig. Interaksjonene mellom ulike forhold ved målingen (III) og enten stimuli (I), respondenter (II) eller høyere ordens interaksjoner kan også være kilder til varians. Disse blir imidlertid sjelden tatt hensyn til, og betraktes ofte som en del av feilledet (V), eller som den delen av variansen som forblir uforklart.

²⁶Haloeffekter oppstår når vurderinger av spesifikke egenskaper ved et objekt influeres av helhetsinntrykket.

Hvordan et spørreskjema for måling av kunders produktvurderinger blir utfylt, kan altså betraktes som et resultat av flere faktorer. I figur 2.1 har vi illustrert hvordan svargivning kan sees på som en funksjon av de tre hovedkildene til varians som ble presentert i figur 2.1: karakteristika ved responsstimuli, karakteristika ved respondenten og forhold ved målingen. I tillegg til disse tre kildene til variasjon i svargivning har vi utvidet modellen og tatt med situasjonsbestemte forhold (Steenkamp 1989; Troye 1992).

FIGUR 2.1
En modell av svargivning ved måling av produktvurderinger



Kilde: Troye (1992)

Forhold ved situasjonen er forhold som er utenfor både respondenten og den som foretar undersøkelsen sin kontroll, men som kan få betydning for svargivningen slik den kommer til uttrykk i spørreskjemaet.

Karakteristika ved responsstimuli kommer til uttrykk ved objektive, *sanne* variasjoner i produktets beskaffenhet. Disse bidrar til at produktet oppfattes som forskjellig fordi det eksisterer faktiske forskjeller i beskaffenheten ved produktet slik det leveres til samme kunde fra gang til gang, eller til forskjellige kunder over tid. Det at et produkt kan beskrives som bestående av både materielle og immaterielle egenskaper er et eksempel

på forhold ved responsstimuli som kan bidra til forskjeller i svargivning. I tillegg kan produktets faktiske ytelse, eller kvalitetsnivå bidra til variasjoner i svargivning.

Karakteristika ved respondenten gir seg utslag i det at kunder er forskjellige, f.eks. med hensyn på erfaring eller grad av produktkunnskap. Kundene kan også være forskjellige med hensyn til hvilke handlinger de utfører, eller hvordan de bruker produktet. Kundene kan videre være forskjellige i måten de foretar sine produktvurderinger. Forskjellige respondenter kan til slutt ha ulik måte å fylle ut et spørreskjema, f.eks. på grunn av responsstil eller responssett.

Karakteristika ved måleprosedyren er forhold som omfatter rekkefølge på spørsmålene, antall svaralternativer, ordlyd i spørsmålsstilling og svaralternativer, format osv. I tillegg vil forhold ved måling også omfatte administrasjonsmåten for undersøkelsen, f.eks. hvorvidt det benyttes personlig intervju, postal spørreundersøkelse eller telefonintervju.

Situasjonsbestemte forhold kan for det første påvirke produktets beskaffenhet direkte. Klimatiske forhold som f.eks. kulde og varme kan få et produkt til å endre karakter. Situasjonss bestemte forhold kan for det andre få konsekvenser for hvordan kundens vurderinger foretas. Vurderes f.eks. produktet forskjellig om det er kaldt, varmt, fuktig, travelt osv? Situasjonsbestemte forhold kan også direkte påvirke respons i spørreskjemaet. Tidspress under utfylling av spørreskjemaet er her et godt eksempel.

2.2 Skalatyper i spørreundersøkelser

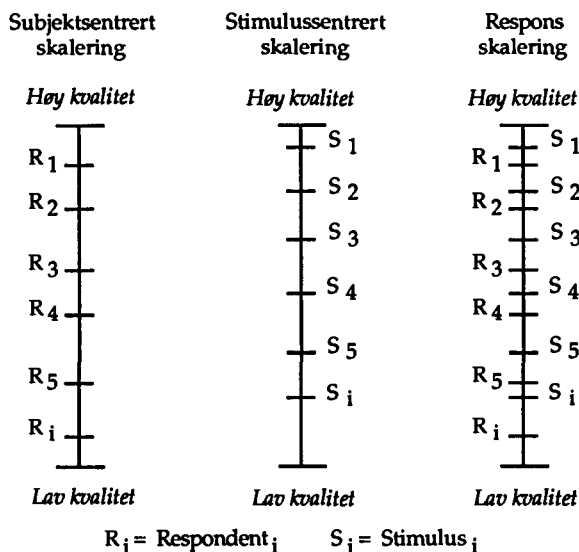
En av de mest brukte klassifiseringer av måleskalaer, både i markedsforskning og psykologi, er de som er utviklet av Torgerson (1958). Torgerson skiller mellom subjekt-sentrert-, stimulussentrert- og responsskalering. Denne funksjonelle klassifiseringen er basert på den-, eller de kildene til variasjon i svargivning, som forskeren ønsker å ta hensyn til og behandle som systematisk varians. Den systematiske variansen ønskes maksimert på bekostning av de gjenværende kildene til tilfeldig varians.

Ved bruk av subjekt-sentrert skalering gis respondentene et sett stimuli (*scale items*) som de skal vurdere langs en eller annen dimensjon (f.eks. "bra-dårlig" eller "tilfreds-tilfreds"). Man vil her være interessert i respondentenes vurderinger av hele utvalget av stimuli og har ingen interesse i vurderingene av de enkelte stimuli. Eksempler på bruk av subjekt-sentrert skalering finnes i holdningsskalaer. Ved subjekt-sentrert

skalering blir mellom-respondent-variens (*between-respondent variation*) betraktet som systematisk, mens forskjellene i hver respondents vurdering av stimuli betraktes som tilfeldig varians og målefeil (Cox, 1982). Fra vårt tidligere eksempel, med respondenter som skal vurdere kvaliteten ved hotell fra en hotellkjede, kan subjektssentrert skalering illustreres ved måling av respondentenes holdninger til hele utvalget av hotell. Vi er da ikke interessert i respondentenes vurdering av de enkelte hotellene.

Stimulussentrert skalering benyttes når vi ønsker å skalere stimuli og ikke er interessert i de systematiske forskjeller som kan eksistere mellom respondentene. Stimuli blir posisjonert på et kontinuum med utgangspunkt i den scoren de oppnår basert på respondentenes enighet (*concensus*). Den mest vanlige bruken av stimulussentrert skalering i markedsforskning består i at man presenterer et sett stimuli (eller vurderingsobjekter) for et stort antall *dommere*. Disse blir så bedt om å foreta vurderinger av de presenterte stimuli. Ved stimulussentrert skalering betraktes mellom-stimuli-variens (*between stimuli variation*) som systematisk, og forskjeller mellom respondentenes vurderinger av stimuli blir behandlet som tilfeldig varians og målefeil (Cox, 1982). I vårt hotelleksempel vil stimulussentrert skalering bety at vi skalerer de ti hotellene uten å være interessert i systematiske forskjeller som kan eksistere mellom respondenter.

FIGUR 2.2
En funksjonell klassifisering av skaleringsprosedyrer



Kilde: Cox (1982)

Når vi både er interessert i systematiske forskjeller mellom stimuli og mellom respondenter benyttes responstilnærming til skalering. Her blir responstilnærmingen et kontinuum der både de vurderte stimuli og respondentene plasseres samtidig (Cox, 1982). Responstilnærmingen har mye til felles med en to-veis variansanalyse (Torgerson, 1958) der de to hovedeffektene, stimuli og respondenter, blir behandlet som kilder til systematisk varians, mens deres interaksjon med andre kilder til ukontrollert varians betraktes som tilfeldig varians og målefeil (Cox, 1982). For vårt hotell-eksempel vil en responstilnærming bety at vi er interessert i systematiske forskjeller både mellom hoteller og respondenter. Dette vil f.eks være en aktuell innfallsvinkel dersom vi skal finne ut om det finnes ulike segmenter av respondenter som vurderer hotellene forskjellig.

Figur 2.3 illustrerer de tre skaleringsprosedyrene. For den subjektssentrerte skaleringen viser figuren hvordan respondentene fordeler seg langs skalaen som måler kvalitet. Den stimulus-sentrerte skalaen plasserer subjektene langs kvalitetskontinuumet. Til slutt er responstilnærmingen illustrert der både respondenter og subjekter skaleres sammen.

Vi hevder i denne avhandlingen at bruk av “bra-dårlig” instruksjon og -vurderingsskala reflekterer det vi har valgt å kalle objektforankret tilnærming til skalering, mens “tilfreds-tilfreds” instruksjon og -vurderingsskala reflekterer subjektforankret tilnærming²⁷. Instruksjonen vi gir respondentene når vi ber dem bruke den ene eller den andre vurderingsskalaen innebærer ulike kognitive oppgaver som kan føre til forskjeller i svargivning på vurderingsspørsmål i et spørreskjema. I kapittel 3 vil vi se nærmere på en rekke teoretiske perspektiver som kan hjelpe oss til å forstå disse forskjellene.

2.3 Empiriske studier vedrørende utforming av spørreskjema

Hovedformålet med avhandlingen er altså å undersøke om det innebærer noen forskjell om vi ber kundene gi uttrykk for om de er tilfreds eller utilfreds med produktet ved bruk av subjektforankret instruksjon og -vurderingsskala, eller om vi ber dem vurdere om produktet er bra eller dårlig ved bruk av objektforankret instruksjon og -vurderingsskala. I tillegg vil vi undersøke hvilke effekter plassering av viktighetsspørsmål og vurderingsspørsmål kan gi opphav til når vi ber kunder vurdere ulike egenskaper ved et

²⁷Se kapittel 3 for en grundigere begrunnelse

produkt. For å belyse disse forholdene er det nærliggende å søke hjelp i metodelitteraturen under temaet *responseeffekter*.

Vi vil nå gi en kort oversikt over en rekke studier som er gjennomført for å studere ulike typer responseeffekter som kan skyldes utforming av spørreskjema. Oversikten er organisert etter overskriftene: ordlyd i spørsmålsstilling, svaralternativer og rekkefølge på spørsmål. Under overskriften ordlyd i spørsmålsstilling tar vi for oss studier der man har fokusert på ulike måter å stille spørsmål for å måle det samme underliggende begrepet. Som vi skal se kan dette få stor betydning for respondentenes svargivning. Benevnelsene som benyttes i tilknytning til verdiene på vurderingsskalaen ved operasjonalisering av vurderinger i et spørreskjema, må i tillegg til selve spørsmålsstillingen og instruksjonen som ligger i denne, betraktes som en viktig informasjonskilde for respondentenes tolking av de oppgavene det forventes at de utfører ved svargivningen. Vi tar derfor for oss en rekke studier der forhold ved svaralternativene som benyttes i spørreskjemaet er fokusert. Til slutt gjennomgår vi studier der temaet har vært effekter av ulik rekkefølge på spørsmålene i spørreskjemaet.

Litteraturgjennomgangen som oppsummeres i avsnittene nedenfor gir ingen retningslinjer for beslutninger i forbindelse med problemstillingene vi tar opp i avhandlingen. En målrettet leser anbefales derfor å hoppe over resten av dette kapittelet. Litteraturgjennomgangen gir imidlertid en god illustrasjon av bredden i studier der relaterte problemstillinger er fokusert. Litteraturgjennomgangen viser også at de fleste studiene typisk er ad-hoc preget, og mangler en felles teoretisk forankring. Til tross for dette representerer studiene det naturlige utgangspunktet for å lete etter svar på de spørsmålene vi har reist i kapittel 1. Vi presenterer derfor en kort oppsummering av studiene nedenfor. Studiene er organisert i kronologisk rekkefølge. Dette er gjort for å få frem et bilde av en utvikling som viser en økning i bruk av kognitiv teori for å forstå effektene som studeres. I appendiks A presenteres tabeller som gir oversikt over hovedfunnene i de empiriske artiklene.

2.3.1 Ordlyd i spørsmålsstilling

I tidligere forskning har man vist at selv små endringer av ordlyd i spørsmålsstillingen kan føre til store forskjeller i respons, se f.eks Alwin og Krosnick (1985), Andrews (1984), Cantril (1944), Krosnick (1989), Miller (1984), Petty, Rennie, og Cacioppo (1987), Rasinski (1989), Schuman og Presser (1981), Sudman og Bradburn (1974) og Sudman og Bradburn (1983). Alle disse bidragene kan på mange måter karakteriseres

som metodologiske bidrag i den forstand at de betrakter forhold ved ordlyd i spørsmålsstilling som *tekniske problemer* i tilknytning til begreper som datakvalitet, reliabilitet og validitet.

I en studie utført av Andrews (1984) var formålet å undersøke hvilken effekt lengde på introduksjon til et spørsmål, og lengden på selve spørsmålet (antall ord) har på datakvalitet. I tillegg undersøkte man effekter av å benytte absolutte vs. komparative vurderingsspørsmål i et spørreskjema. Hovedfunnene fra studien viste at spørsmålstillinger med en introduksjon på 16 til 24 ord, og med etterfølgende spørsmål med mer enn 16 ord gir høyest validitet, lavest residualeffekt og lavest metodeeffekt. I tillegg resulterte studien i anbefaling om bruk av komparative-, fremfor absolutte spørsmålsformuleringer, da disse gir mer presise svar. Ved bruk av komparative spørsmål blir det etablert forankringspunkter som vurderingene foretas i forhold til.

Miller (1984) sammenlignet to ulike måter å stille 7-punktsskala holdningsspørsmål på i en nasjonal helseundersøkelse. En gruppe respondenter ble bedt om å svare på holdningsspørsmålene i ett trinn (*one step scale*) ved å velge ett tall, mens en annen gruppe først ble instruert til å gi en generell, og deretter en spesifikk vurdering på tallskalen (*two step scale*). Miller fant ingen forskjeller mellom gruppene med hensyn på frekvensfordelinger. *One step* skalaen viste mindre positiv responskjevhet, lavere andel manglende svar og høyere interkorrelasjon mellom spørsmålene.

Alwin, et al. (1985) gjennomførte en studie der de sammenlignet bruk av vurderingsskalaer og rangeringsskalaer ved måling av foreldres verdsetting av barns egenskaper. Vurderingsskalene og rangeringsskalene ble funnet å være likestilte mhp. å få frem relativ viktighet for verdivalg, men ulike mhp. latent struktur og korrelasjon med relevante prediksjonsvariabler. Rangeringsskaler ble funnet å gi best prediksjon.

Petty, Rennier, og Cacioppo (1987) undersøkte effekten av å presentere holdningsspørsmål i påstandsformat (*assertion*) vs. spørreformat (*interrogation*). Funnene indikerer en sterkere polarisering i svargivning ved bruk av spørreformat enn ved bruk av påstandsformat. I tillegg ble det observert at respondentene i studien benyttet en grundigere informasjonsbehandling ved bruk av spørreformat enn ved bruk av påstandsformat på holdningsspørsmålene. Forfatterne konkluderer med å anbefale bruk av spørreformat ved måling av holdninger.

Rasinski (1989) studerte effekter av ordlyd i spørsmålsstilling for måling av vurderinger av offentlig forbruk innenfor ulike politikkområder. Undersøkelsene viste signifikante effekter av endringer i ordlyd i spørsmålsstilling i løpet av de tre årene studien foregikk.

Krosnick (1989) gjennomførte en studie for å undersøke effekten av å benytte begrepene: *støtte* (*support*), *akseptabel* (*acceptable*) eller *foretrekke* (*favor*) ved måling av holdninger til rettssystemet i USA, og fant signifikante effekter av de manipulererte forskjellene i ordlyd i spørsmålsstilling. Bruk av *akseptabel* og *foretrekker* alternativene viste seg å gi høyere positiv svargivning enn ved bruk av *støtte* alternativet (*support*).

Litteraturgjennomgangen indikerer, som nevnt i innledningen, en lite systematisk, og ad-hoc preget tilnærming ved studier av problemstillinger knyttet til ordlyd i spørsmålsstilling ved gjennomføring av spørreundersøkelser. Studiene som er gjennomført gir ingen retningslinjer for valgene vi fokuserer på i de spesifikke problemstillingene i denne avhandlingen.

Problemstillingene som er tatt opp forankres i liten grad til substansielle teoretiske perspektiver, og kan beskrives ved å trekke paralleller til *Stimulus - Organisme - Respons* (S-O-R) modellen innenfor psykologifaget. Et kjennetegn ved studiene som er gjennomført er at man har observert forskjeller i svargivning, respons, ved endringer i spørsmålsstillingen i spørreskjemaet, stimulus. Man har vært mindre opptatt av å forklare *hvorfor* og *hvordan* endringene i spørsmålsstilling påvirker respondentenes svargivning. En slik tilnærming vil kreve at man "åpner opp organisme elementer" i S-O-R modellen og studerer respondentens informasjonsbehandlingsprosess i forbindelse med svargivningen.

I følge Hippler & Schwarz (1987) må endringer i spørsmålsstilling alltid betraktes som substansielle fordi de kan bidra til å endre meningsinnholdet i et bestemt spørsmål på en av tre måter: (i) henvisning til ulike sider ved et emne, (ii) aktivering av begreper som påvirker respondentens tolking av spørsmålet eller (iii) stille respondenten overfor forskjellige kognitive oppgaver ved svargivningen. I kapittel 3 vil vi bygge videre på dette når vi forsøker å anvende ulike kognitive perspektiver for å belyse problemstillingene vi tar opp i avhandlingen.

Relatert til problemstillingene som blir tatt opp i denne avhandlingen kan vi så langt konkludere med at endringer i ordlyd vil føre til endringer i vurderingsspørsmålenes substansielle meningsinnhold. Vi forventer at det vil bety noe annet for kundene om vi spør dem om de er tilfreds, eller om de synes produktet er bra. For det andre kan det samme spørsmålet tolkes forskjellig av ulike respondenter. Kunder med høy produktkunnskap hevder vi lettere kan vurdere hvorvidt produktet er bra eller dårlig enn kunder med lavere produktkunnskap. For det tredje kan endringer i ordlyd påvirke respons gjennom ulikt ladede assosiasjoner. Det å bli spurt om man er utilfreds kan oppfattes som et sterkere negativt ladet utsagn for respondenten enn det å bli spurt om noe er

dårlig. Til slutt kan endringer i ordlyd påvirke respondentenes oppfatning av hva som er deres oppgave i forbindelse med svargivningen ved utfylling av spørreskjemaet. Vi antar at det innebærer andre kognitive oppgaver for kundene å vurdere i hvilken grad de er tilfreds eller utilfreds med et produkt, enn det å bli bedt om å ta stilling til om det samme produktet er bra eller dårlig. Alle disse forholdene kan bidra til å forklare hvorfor vi skal kunne forvente forskjeller i respons avhengig av om vi spør respondentene om produktet er bra eller dårlig, ved hjelp av objektforankret instruksjon, eller om vi benytter subjektforankret instruksjon i spørsmålsstillingen, og be dem ta stilling til om de er tilfreds eller utilfreds med produktet²⁸.

2.3.2 Svaralternativer

Når respondentene skal svare på spørsmålene i et spørreskjema kan dette i hovedsak gjøres på to måter: bruk av åpne eller lukkede spørsmål. Fordeler og ulemper ved disse har vært gjenstand for utstrakt forskning, se f.eks (Sudman og Bradburn, 1983), og vi skal ikke komme nærmere inn på dette her. Vi vil heller rette oppmerksomheten mot studier der problemstillinger knyttet til ulike typer lukkede responsalternativer har vært tatt opp.

Wildt og Mazis, 1978) undersøkte betydningen av å bruke benevnelse på alle verdier på en vurderingsskala vs. benevnelse bare på endepunktene på skalaen. Utgangspunktet for studien er at svargivning på en skala hevdes å være en funksjon av både benevnelse på skalaverdier og plassering av benevnelsen på skalaen. Det finnes ingen signifikante forskjeller mellom skalaer med tilnærmet intervall benevnelser på samtlige skalaverdier, og skaler med benevnelse bare på endepunktene.

Dixon, Bobo, og Stevick (1984) sammenlignet to *Likert-type*²⁹ skaler der den ene har med benevnelse på alle skalaverdier, mens den andre har benevnelse bare på endepunktene. Skalaene med benevnelse kun på endepunktene viste seg å gi større varians enn skalaene med benevnelse på alle skalaverdiene.

Boote (1981) studerte i hvilken grad forskjeller i reliabilitet ved psykografiske skalaer kan tilskrives variasjoner i antall verdier på skalaen og forskjell i forankring, manipulert

²⁸I kapittel 3 vil vi "åpne opp organisme elementet" i S-O-R modellen for å forsøke å forklare hvorfor og hvordan de observerte responseffektene fremkommer.

²⁹Bruk av Likert-skala innebærer at man ber respondentene ta stilling til i hvilken grad de er enig eller uenig i påstander formulert av intervjueren

ved benevnelsene på skalaverdiene. Korrelasjon mellom vurderinger foretatt på to ulike tidspunkter viste seg ikke å gi forskjell i reliabilitet basert på antall verdier på skalaen, mens skalaer med benevnelse på samtlige skalaverdier ser ut til å være mer reliable enn skalaer med benevnelse kun på endepunktene. Konsistens i vurderinger (målt ved test-retest) gir høyere reliabilitet for 5-punktsskalaer enn for 7-punktsskalaer. Det finnes imidlertid ingen forskjell i reliabilitet mellom skalaer med benevnelse på alle skalaverdier og skalaer med benevnelse kun på endepunktene. Vurderinger av egenskaper som oppfattes som ekstremt viktige og ekstremt uviktige for respondentene, oppnår høyere reliabilitet enn egenskaper som oppfattes som middels viktige. Gjennomsnittsvurderingene viser seg å være konsistent lavere for skalaer med benevnelse på alle verdiene enn for skalaer med benevnelse kun på endepunktene.

Cox III (1980) foretok en litteraturgjennomgang over forskning som er gjennomført vedrørende optimalt antall responsalternativer på en skala. Hovedkonklusjonen er at det ikke finnes et "riktig" antall responsalternativer. Litteraturgjennomgangen oppsummeres med følgende konklusjoner og anbefalinger: (i) skalaer med 2-3 svaralternativer gir lite informasjon og frustrerer respondentene, (ii) skalaer med mer enn 9 svaralternativer gir små og marginale forbedringer, (iii) skalaer med midtpunkt anbefales der respondentene kan antas å ha en nøytral posisjon, (iv) forståelig instruksjon gjennom spørsmålsstillingen og forståelige benevnelse på svaralternativene er nødvendig, (v) for subjektentrert skalering anbefales det at man benytter opp til 5 svaralternativer, mens det for stimulusentrert skalering anbefales å benytte opp til 9 svaralternativer.

Neumann og Neumann (1981) sammenlignet seks vurderingsskalaer med ulik lengde (2, 3, 4, 5, 7 og 10 skalaverdier) for måling av studenters holdninger til det å motta instruksjon fra lærere. Neumann (1983) foretok en sammenligning av de samme vurderingsskalaene med tre ulike lengder (3, 5 og 7 verdier). I begge studiene fant man at alle svarfordelingene var skjeve i negativ retning. Ved bruk av de lange skalaene unngikk respondentene å benytte ekstremverdiene. Korrelasjon mellom vurderinger viste seg videre å øke med økende skalalengde. Eta-koeffisienter (grad av reduksjon i varians) økte også med økende skalalengde. Forfatterne avslutter med å anbefale bruk av 5-punktsskalaer ved gjennomføring av holdningsundersøkelser.

Cicchetti, Showalter, og Tyrer (1985) gjennomførte en studie med formål å undersøke hvorvidt *interrater*-reliabilitet for en klinisk måleskala er avhengig av antall kategorier eller verdier på skalaen. Reliabilitet i målingene viste seg å øke jevnt opp til bruk av skalaer med 7 skalaverdier, og flater deretter ut (selv om antall skalaverdier økes til 100).

Alwin og Krosnick (1991) studerte effekter av karakteristika ved spørsmålsstilling og kjennetegn ved respondentene på reliabilitet i målinger i spørreskjemaundersøkelser. Følgende funn ble rapportert: (i) Reliabilitet øker når antall responsalternativer på en vurderingsskala øker fra to til sju alternativer og flater deretter ut. For *enig-uenig* skalaer finnes ikke signifikante forskjeller i reliabilitet mellom skalaer med to verdier og skaler med fem verdier, (ii) bruk av midtpunkt på skalaen reduserer reliabilitet, men kan fungere som forankringspunkt på lange skalaer (f.eks. mer enn sju skalaverdier), (iii) benevnelse på alle skalaverdier gir høyere reliabilitet enn skalaer med benevnelse bare på endepunktene, og (iv) bruk av “vet ikke” alternativ reduserer reliabilitet ved måleskalaen.

Andrews (1984) undersøkte effekter av antall svarverdier/kategorier på måleskalaen, bruk av “vet ikke” alternativ og lengde på spørsmålsbatteri³⁰ på datakvalitet. Datakvalitet, representert ved reliabilitet, viste seg å øke med økende antall skalaverdier. Bruk av eksplisitt “vet ikke” alternativ viser seg å gi høyere datakvalitet. Datakvalitet reduseres med økende antall spørsmål i spørsmålsbatteriet, selv om det er fordeler forbundet med å benytte en felles introduksjon og svarskala knyttet til spørsmålsbatteriet.

Poe et al. (1988) fokuserte på effekten av å inkludere “vet ikke” alternativ på responsraten ved faktaspørsmål i en postal survey. Selv-administrert spørreskjema uten “vet ikke” alternativ anbefales benyttet fordi: (i) man oppnår høyere responsrate på enkeltspørsmål, (ii) det blir mindre behov for å korrigere for manglende svar, og dermed lavere sannsynlighet for å introdusere dette som en feilkilde i dataene, (iii) bruk av “vet ikke” alternativ fører til få store substansielle endringer i behandling og tolking av dataene, (iv) det finnes ingen vesentlige forskjeller i responsfeilrate når “vet ikke” alternativ utelates, (v) utelatelse av “vet ikke” alternativer påvirker total responsrate og (vi) spørreskjemaene blir kortere og enklere å fylle ut.

Ayidiya og McClendon (1990) undersøkte om tidligere funn vedrørende rekkefølge på spørsmål, rekkefølge på svaralternativer, bruk av “vet ikke”-filter, midtpunkt på svarskala og føyelighet i personlige intervjuer undersøkelser også gjelder ved gjennomføring av postale spørreundersøkelser. *Recency* effekter oppstår ikke i postale skriftlige spørreundersøkelser. Det finnes en svakt signifikant *primacy* effekt der midtalternativet, “fortsette som før”, velges oftere der dette alternativet er plassert på midten av svarskaalen i motsetning til der dette er plassert til slutt i skalaen.

³⁰Antall *item*

Bishop (1987) rapporterer resultater fra en undersøkelse der formålet var å studere effekter av å inkludere vs. utelukke midtpunkt (*middle response alternative*) på en vurderingsskala. Følgende funn blir rapportert: (i) Det er signifikant større sannsynlighet for å velge midtpunktet når dette er eksplisitt tatt med som del av spørsmålet. (ii) Ved å nevne midtpunktalternativet i innledningen til spørsmålet, økes sannsynligheten for at dette velges, selv om midtpunktet ikke er gitt som et eksplisitt svaralternativ. (iii) Plassering av midtpunktalternativet (i midten eller til slutt) har betydning. (iv) Respondenter som velger midtpunktalternativet når dette er tilgjengelig, vil ikke nødvendigvis svare på spørsmålet på samme måte som andre respondenter når de blir tvunget til å velge side.

Bishop (1990) testet antakelsene om at responseeffekter knyttet til variasjoner i form på spørsmål, ordlyd i spørsmålsstilling eller kontekst vil være sterkere for de respondentene som er minst involvert i temaet det spørres om. Effekter av spørsmålsrekkefølge og responsalternativ rekkefølge ser ut til å være upåvirket av respondentenes grad av involvering i temaet for undersøkelsen. Responseeffekter knyttet til bruk av balansert argumentasjon (*counter arguments*), og midtalternativer rapporteres å være sterkere for respondenter med lavere engasjement i temaet det spørres om.

Krosnick og Alwin (1987) hadde som formål å benytte kognitiv teori for å forklare effekter av rekkefølge på responsalternativer i spørreundersøkelser ved måling av foreldres rangering av hvilke verdier det er viktig at barn har. Plassering av et responsalternativ blant de tre første i et batteri av alternativer, øker sannsynligheten for at det rangeres som et av de tre viktigste. Respondenter som er mindre kognitivt sofistikert blir i sterkere grad påvirket av endringer i rekkefølgen på responsalternativene.

Schwarz et al. (1991) studerte effekter av bruk av ulike tallverdier (+/+ eller -/+) på respondenters tolkning av benevnelse på endepunktene på en vurderingsskala. Respondenter hevdes å bruke tallverdiene på vurderingsskalaen for å gi mening til de ulike benevnelsene på skalen. Følgende funn rapporteres: (i) ved bruk av 0 - 10 skala benyttes en unipolar konseptualisering, dvs. i hvilken grad egenskapen som vurderes er tilstede, (ii) ved bruk av -5 - 5 skala benyttes en bipolar konseptualisering, dvs. om egenskapen er tilstede = 0, mens negative verdier indikerer fravær av egenskapene. Forfatterne kommer med følgende anbefalinger: (a) Ved bruk av semantisk differensial skalaer bør det benyttes skalaer med negative til positive tallverdier. (b) Ved måling av intensitet/tilstedeværelse av en bestemt egenskap bør det benyttes skala med tallverdier fra 0 til positive skalaverdier. (c) Valg av skala bør generelt baseres på kunnskaper om responsfordelinger for de relevante svarene man forventer.

Schwarz et al. (1985) undersøkte effekter av rekkevidde (*range*) på responskategorier i lukkede svarformater i spørsmål for rapportering av en bestemt atferd (daglig bruk av TV) for vurdering av hvor tilfreds man er med den fritiden man har til disposisjon. Det finnes indikasjoner på at respondenter som blir presentert en skala med smal kategorilengde velger svaralternativer på midten av skalaen, mens respondenter som blir presentert en skala med vid kategorilengde velger den første kategorien i skalaen. Respondenter som presenteres en skala med vid kategorilengde rapporterte at de er mer tilfreds med fritiden enn respondenter som er presentert en skala med smal kategorilengde. Forfatterne anbefaler bruk av åpne svarkategorier ved måling av atferdsfrekvenser og bruksfrekvenser

En rekke studier knyttet til svarskalaer og svaralternativene som blir gitt i tilknytning til spørsmål, i tillegg til de vi har referert ovenfor, er gjennomgått og referert av Hippler og Schwarz (1987). Når det gjelder valg mellom ensidige eller tosidige responsalternativer anbefales det å bruke tosidige (balanserte) alternativer. Eksplisitt inkludering av midtpunkt på skalaen vil bidra til en økning i antall respondenter som vil benytte denne svarkategorien. Rekkefølge på svaralternativene får også betydning for respons. Alternativer som kommer tidlig, legges mer vekt på og kan etablere en referanseramme som de øvrige vurderes etter. Når det gjelder bruk av skala finnes det indikasjoner på at fem til ni svarkategorier med benevnelse kun på endepunktene fungerer best. Benevnelse på endepunktene og eventuelle benevnelser på de øvrige verdiene på skalaen har videre vist seg å påvirke respons og fordeling av verdiene på svarskalaen.

Også studiene knyttet til problemstillinger i forbindelse med svaralternativene som benyttes i spørreskjemaet, indikerer manglende eller ad-hoc preget teoretisk forankring. Det er imidlertid tre forhold som til en viss grad kan sies å være studert på en systematisk måte. Disse er problemstillinger knyttet til lengde på svarskala, om man skal benytte benevnelse på samtlige skalaverdier eller kun på endepunktene og om man skal inkludere "vet ikke" som et eksplisitt svaralternativ i tilknytning til måleskalaen.

Litteraturgjennomgangen gir ingen retningslinjer for å ta beslutninger om når man bør bruke objektforankrede benevnelser, f.eks "bra-dårlig", og når man bør bruke subjektforankrede benevnelser, f.eks "tilfreds-utilfreds", på svarskaalen som benyttes³¹. Vi har til tross for dette, valgt å ta med gjennomgangen ovenfor for å gi et komplett bilde av

³¹Selv om benevnelse på svarskaalen kan betraktes som et selvstendig element ved utforming av spørsmål og svaralternativer i et spørreskjema, vil vi i kapittel 3 betrakte benevnelser på svarskaalene som en del av instruksjonen som blir gitt respondenten i spørsmålet. Ordlyd i spørsmålsstilling og benevnelse på svarskaale behandles derfor samlet når vi senere trekker inn teoretiske perspektiver for å belyse problemstillingene knyttet til disse forholdene.

hvilke problemstillinger som er fokusert når man har studert forhold ved svaralternativene som benyttes i spørreskjema.

2.3.2 Spørsmålsrekkefølge

Rekkefølgen på spørsmålene i et spørreskjema kan påvirke respondentenes svargivning. Generelle tommelfingerregler³² som ofte trekkes frem er at de første spørsmålene i et spørreskjema bør være enkle og bidra til å vekke respondentens interesse, mens vanskelige, og mer "truende" spørsmål bør plasseres mot slutten av spørreskjemaet (Sudman og Bradburn, 1974). I tillegg til disse rent motiverende effektene av spørsmålsrekkefølge må vi også ta hensyn til mer komplekse kognitive effekter som har med selve prosessen som ligger bak det å svare på spørsmålene å gjøre. En lang rekke studier er gjennomført for å undersøke effekter av alternativ rekkefølge på spørsmål i et spørreskjema.

McClendon og O'Brien (1988) undersøkte effekten av å plassere et sett spesifikke tilfredshetsspørsmål før vs. etter et generelt tilfredshetsspørsmål. Ingen signifikant effekt av spørsmålsrekkefølge ble funnet for gjennomsnittsscorene på de spesifikke spørsmålene. Det ble rapportert signifikante effekter av spørsmålsrekkefølge på prediksjonsevnen til de spesifikke spørsmålene, når de spesifikke spørsmålene inngår som uavhengige variabler i en regresjonsanalyse, og de generelle spørsmålene defineres som den avhengige variabelen.

Tourangeau, Rasinski, & Bradburn (1991) hadde som formål å finne ut om det er slik at dersom et generelt tilfredshetsspørsmål følger etter et mer spesifikt spørsmål om tilfredshet med ekteskapet, vil respondentene tolke det generelle spørsmålet til å gjelde forhold som kommer *i tillegg til* forhold ved ekteskapet. Dette kalles "subtraksjonshypotesen" Spørreskjemaversjonen der det generelle spørsmålet eksplisitt hadde med "i tillegg til ekteskapet", gav respons som var lik versjonen der dette ikke var tatt med og der det generelle spørsmålet kom etter det spesifikke. Versjonen der det generelle spørsmålet eksplisitt hadde med "inkludert ekteskapet", gav respons som var lik versjonen der dette ikke var tatt med og der det generelle spørsmålet kom før det spesifikke. Subtraksjonshypotesen ble altså ikke støttet.

³²Se f.eks. *Procedure for developing a questionnaire* i Churchill (1991)

Schwarz, Strack, og Mai (1991) gjennomførte en studie for å undersøke om effekter av rekkefølge på generelle og spesifikke spørsmål om tilfredshet med livssituasjonen påvirkes av i hvilken grad spørsmålene oppfattes som å høre sammen (*conversational logic*). Funnene fra studien indikerer at respondentene benytter den sist aktiverte informasjonen når de svarer på det generelle spørsmålet. Inkludering av flere spesifikke spørsmål øker tilgjengeligheten til et variert sett av potensielt relevant informasjon, og reduserer dermed sammenhengen mellom et spesifikt spørsmål (relasjonstfredshet) og det generelle spørsmålet. Effekten av konversasjonsnorm, operasjonalisert ved felles introduksjon til de spesifikke og generelle spørsmålene, bidrar ikke til økning i sammenhengen mellom de spesifikke spørsmålene og det generelle spørsmålet. Respondenter som rapporterte en høy relasjonstfredshet, rapporterte også en høyere generell tilfredshet med livssituasjonen sett under ett.

Benton og Daly (1991) studerte effekter av rekkefølge på spesifikke- vs. generelle spørsmål i en undersøkelse av kvalitet ved offentlige tjenester. De ble ikke funnet noen effekt av rekkefølge på spørsmålene. Svargivningen varierte imidlertid som en funksjon av utdanning, der effektene av spørsmålsrekkefølge ble observert å være større for respondenter med lav utdanning

Schul og Schiff (1993) undersøkte effekten av rekkefølge på generelle- og spesifikke tilfredshetsspørsmål ved måling av kunders tilfredshet med serviceorganisasjoner. Korrelasjoner mellom generell tilfredshet og attributtspesifikke tilfredshetsspørsmål ble funnet å være høyere når de attributtspesifikke spørsmålene kommer først i spørreskjemaet. Når generell tilfredshetsspørsmål er plassert etter de attributtspesifikke tilfredshetsspørsmålene i spørreskjemaet økes tilgjengeligheten til både positiv og negativ informasjon. Når de generelle tilfredshetsspørsmålene kommer først blir negativ informasjon mer tilgjengelig enn positiv informasjon.

Jones og Lang (1982) fokuserte blant annet på rekkefølge på sammenligningsspørsmål og attributtvurderingsspørsmål ved respondenters vurderinger av sitt eget nabolag. Pretesting indikerer at attributtvurderingsspørsmål oppfattes som lettere å svare på enn sammenligningsspørsmål. Spørreskjema der attributtvurderingsspørsmål er plassert før sammenligningsspørsmål gir høyere reliabilitet men ingen effekt på validitet. I tillegg rapporteres en positiv effekt på responsrate i forhold til spørreskjema der sammenligningsspørsmålene stilles først.

Schuman, Kalton, og Ludwig (1983) hadde som formål å undersøke om kontekst-effekter kan reduseres ved å legge inn nøytrale spørsmål mellom spørsmål som man

antar, eller vet, vil påvirke hverandre når det gjelder svargivning³³. Studien gir indikasjoner på at det å innføre nøytrale mellomspørsmål ikke bidrar til å redusere konteksteffektene.

Crespi og Morris (1984) studerte effekten av ulik spørsmålsrekkefølge ved meningsmålinger i sammenheng med senats- og guvernørvalg, og fant motsatte preferanser for to ulike guvernørkandidater ved bruk av alternativ rekkefølge på spørsmålene.

Tourangeau et al. (1989) foretok en studie for å undersøke om overføringseffekter kan generaliseres over ulike holdningstemaer. Det ble fokusert på effekter av bruk av "buffer"-spørsmål mellom kontekst- og hovedspørsmål, i tillegg til effekter av holdningskonflikter og grad av sentralitet i holdningene. Faktorer som finnes å påvirke betydning av kontekst er følgende: Respondenter med blandete holdninger, eller med holdninger i konflikt, men der holdningene samtidig er sentrale, er spesielt påvirkelige av konteksteffektene. Tilstedeværelse av nøytrale/irrelevante "buffer"-spørsmål mellom kontekst- og hovedspørsmålene gav ingen klare effekter. Konteksteffektene var heller ikke avhengig av kjennskap til temaet eller svargivningen på kontekstspørsmålene. Det samme gjelder grad av enighet med kontekstspørsmålene.

Bickart et al. (1992) studerte hvordan kontekstuelle faktorer, som spørreskjemaformat og spørsmålsrekkefølge, kan påvirke respondenters selv-rapportering når de blir bedt om å vurdere ektefellen. Den avhengige variabelen var vektlegging av selvrapport ved vurdering av andre (*proxy*). Når selvrapportering og *proxy*-vurdering er skilt fra hverandre med en serie spørsmål, og når likheter mellom ektefellene er aktivert i et spørsmål som kommer før *proxy*-vurderingen, tillegges mer vekt til selvrapportering. Vektlegging på selvrapport reduseres når selvrapport og ektefelle (*proxy*) rapportering er sekvensiell, og likheter mellom ektefellene er aktivert.

Selv om studiene vi har gjennomgått ovenfor til en viss grad bygger på kognitive aspekter ved svargivning, er det først i de senere studiene at man eksplisitt drøfter forhold ved respondentens informasjonsbehandlingsprosess som forklaring på effekter av ulik rekkefølge på spørsmålene som stilles. Studiene har i hovedsak fokusert på forhold knyttet til plassering av generelle- vs. spesifikke vurderingss spørsmål.

Generelt kan oppsummere gjennomgangen ovenfor med å si at forutgående spørsmål kan bidra til å rette respondentenes oppmerksomhet mot en bestemt type informasjon,

³³Bruk av *filler questions*

og dermed øker denne informasjonens tilgjengelighet³⁴ ved tolking og svargivning på etterfølgende spørsmål i en undersøkelse (Schwarz og Sudman, 1991). Respondentene vil derfor med større sannsynlighet bruke denne informasjonen når de foretar vurderinger senere i spørreskjemaet³⁵ (Hastie, 1987). Forutgående spørsmål kan videre øke aktualiteten av et bestemt tema, noe som kan få betydning for svargivning på spørsmål vedrørende dette temaet senere i spørreskjemaet (Hippler og Schwarz, 1987). Til slutt kan forutgående spørsmål aktivere en bestemt norm som danner grunnlag for vurderinger som respondenten blir bedt om å foreta senere i spørreskjemaet (Hippler og Schwarz, 1987; Schuman og Presser, 1981).

2.5 Oppsummering og implikasjoner

Formålet med dette kapitlet har vært å etablere en ramme for problemstillingene fra kapittel 1, i hovedsak basert på det vi har definert som metodologiske tilnærminger. Vi startet kapitlet med å gi en generell oversikt over mulige innfallsvinkler for å forklare variasjoner i svargivning i spørreskjemaundersøkelser. Med utgangspunkt i disse presenterte vi en modell for å forstå kilder til variasjon i svargivning i undersøkelser, der vi måler kunders produktvurderinger ved hjelp av et spørreskjema i en utvalgsundersøkelse. Videre introduserte vi tre sentrale tilnærminger til skalering i markedsforskning og i psykologi: subjektentrert-, stimulusentrert- og respons tilnærming. I skaleringslitteraturen (Cox, 1982; Guilford, 1954; Torgerson, 1958) er det påpekt at to av innfallsvinklene til skalering, subjektentrert vs. stimulusentrert, korresponderer med ulike forskningsoppgaver og impliserer bruk av ulike skaleringsteknikker. En ren stimulusentrert skalering impliserer at formålet er å skalere objekter, og innebærer at respondentene betraktes som replikasjoner. En ren subjektentrert skalering impliserer at formålet er å skalere individer. Skaleringslitteraturen omhandler hvilke estimerings-teknikker og eksperimentelle undersøkelsesopplegg som bør benyttes for subjektentrert-, stimulusentrert- og responstilnærming. Dette er imidlertid ikke hovedfokus for denne avhandlingen. Vårt formål er å peke på hvilke forskjeller vi kan forvente i svargivning når a) subjekter blir instruert til å være "dommere" og vurdere objekter (objektforankret instruksjon), og b) subjekter blir instruert til å gi uttrykk for den indre følelsesmessige tilstand som objektene (eller andre faktorer) gir opphav til (subjektforankret instruksjon).

³⁴Spørsmålsrekkefølge og tilgjengelighet (*accessibility*) til informasjon vil bli grundigere behandlet i kapittel 3.

³⁵Hvordan informasjonen benyttes drøftes nærmere i kapittel 3.

Problemstillinger knyttet til rekkefølgen på spørsmålene i spørreskjemaet ble også drøftet i dette kapittelet. Til slutt gikk vi gjennom en rekke empiriske studier der problemstillinger vedrørende ordlyd i spørsmålsstilling, svaralternativer og spørsmålsrekkefølge har blitt studert.

I forhold til de problemstillingene vi reiste i kapittel 1, har litteraturen vi har gått gjennom så langt indikert følgende: (i) Vi kan forvente å finne forskjeller i svargivning ved valg av ulik ordlyd i spørsmålsstilling og benevnelse på svaralternativer, og (ii) vi kan også forvente at det betyr noe i hvilken rekkefølge vi stiller spørsmålene i spørreskjemaet. Litteraturen sier imidlertid lite om retning og styrke på disse effektene.

Gjennomgangen av de empiriske studiene gir altså ingen konkrete retningslinjer for valg mellom subjektforankrede vs. objektforankrede instruksjoner i et spørreskjema. Vi får heller ikke hjelp til å velge benevnelse på svarskalaen knyttet til disse to forskjellige instruksjonene. Studiene gir til slutt ingen konkrete svar på spørsmålet om hvilken plassering av spesifikke vurderingsspmål og viktighetsspmål som bør benyttes ved operasjonalisering av multiattributtmodeller. Vi har benevnt de empiriske studiene vi har gjennomgått som "metodologiske studier". Denne benevnelsen er benyttet fordi studiene er kjennetegnet av at man har hatt et utgangspunkt der formålet i hovedsak har vært å knytte metodologiske valg ved utvikling av spørreskjema til reliabilitets- og validitetsbetrakninger. Man har vært mer opptatt av å undersøke hvilke effekter de ulike valgene kan gi opphav til, enn å forklare hvorfor effektene oppstår. Hvorvidt resultatene kan generaliseres til å gjelde i andre situasjoner og for mer generelle problemstillinger kan diskuteres. Litteraturgjennomgangen vi har foretatt støtter de påstandene som er fremsatt av Groves (1987), og som vi startet dette kapittelet med. De empiriske studiene vi har gjennomgått, knyttet til gjennomføring av spørreundersøkelser der det benyttes strukturerte spørreskjema, ser ut til å være drevet frem av både uavhengige, og tildels ukoordinerte bidrag fra forskere innenfor fagområder som statistikk, psykologi og sosiologi. En slik kunnskapsgenerering, med spredte og ad-hoc pregete koplinger mellom undersøkelsesmetodologi og ulike teoriretninger, gir få generelle retningslinjer for metodologiske valg. I neste kapittel vil vi derfor belyse problemstillingene vi reiste i kapittel 1 med utgangspunkt i teoretiske perspektiver i hovedsak hentet fra kognisjonslitteraturen og informasjonsbehandlingslitteraturen.

3

Kognitive perspektiver på svargivning i spørreundersøkelser

I kapittel 2 relaterte vi beslutninger i tilknytning til utforming av enkeltspørsmål, svaralternativer og rekkefølge på spørsmål i et spørreskjema til litteratur der utgangspunktet i hovedsak har vært rent metodologisk. Fokus i denne litteraturen er på måleteoretiske perspektiver og tekniske aspekter knyttet til datakvalitet, herunder begreper som reliabilitet og validitet. I metodelitteraturen har man i hovedsak forklart variasjoner i svargivning som responseffekter (se f.eks. Hippler og Schwarz (1987), Schuman og Presser (1981) og Sudman og Bradburn (1974)).

Vi konkluderte med at metodelitteraturen gir få retningslinjer som grunnlag for å velge mellom alternativ instruksjon og vurderingsskala for vurderingsspørsmål, og alternativ plassering av viktighetsspørsmål i spørreskjemaet. Ved bruk av metodologiske og måleteoretiske tilnærminger kan vi i de fleste tilfeller bare påstå at ulik ordlyd og rekkefølge på spørsmålene kan gi forskjeller i respondentenes svargivning. Vi kan i liten grad si noe om retning og styrke på eventuelle forskjeller. I dette kapittelet vil vi derfor bygge videre på de perspektivene vi introduserte i kapittel 2, og forsøke å presisere retning og styrke på responseeffektene ved hjelp av relevante teoribidrag. Formålet her er ikke å teste teorier, men heller å anvende ulike teoretiske perspektiver for å forstå, og predikere hvorfor, og hvordan de ulike effektene fremkommer.

Selv om spørsmål i forbindelse med ordlyd i spørsmålsstilling og svaralternativer og spørsmålsrekkefølge ved utvikling av spørreskjema har generell interesse, fokuserer vi i avhandlingen spesielt på måling av kunders produktvurderinger ved hjelp av spørsmål i et spørreskjema. Basert på antakelsen om at kunder virkelig foretar vurderinger av produkter (Olshavsky og Granbois, 1979; Troye, 1990), gjør vi derfor innledningsvis nærmere rede for hvordan slike produktvurderinger kan tenkes å bli foretatt. I dette

avsnittet drøfter vi også hvordan litteratur om holdninger, erfaring, produktkunnskap og kundetilfredshet kan relateres til problemstillingene vi reiste i kapittel 1.

Når vi ber respondenter foreta vurderinger av produkter, eller andre objekter i en spørreundersøkelse pålegger vi dem ulike informasjonsbehandlingsoppgaver (Clark og Schober, 1991; Strack og Martin, 1987; Tourangeau og Rasinski, 1988). Vi knytter derfor svargivning på vurderingsspørsmålene i spørreskjemaet til en generell modell som illustrerer informasjonsbehandlingsprosessen en respondent kan tenkes å gå gjennom i en spørreundersøkelsessituasjon. I følge denne modellen må respondenten starte med å foreta en *tolkning* for å forstå meningsholdet i hvert enkelt spørsmål i spørreskjemaet. Deretter må respondenten *danne* seg en *oppfatning*, eller gjøre seg opp en mening om temaet det spørres om i spørsmålet. Videre må denne oppfatningen *innpasses* i spørreskjemaets responsformat. Til slutt kan det avgitte svaret bli *endret* som en funksjon av ulike motiver som kan oppstå i undersøkelsessituasjonen. Ulike konsekvenser av de metodologiske valgene vi fokuserer på i problemstillingene, som vi tar opp i avhandlingen, vil så langt det lar seg gjøre bli relatert til trinnene i denne informasjonsbehandlingsprosessen.

Etter denne generelle innledningen setter vi søkelys på to ulike dimensjoner i kunders produktvurderinger, objektforankrete vurderinger (f.eks. "bra-dårlig") og subjektforankrete vurderinger (f.eks. "tilfreds-utilfreds"). Formålet er her å forsøke å belyse: 1) Hvordan disse dimensjonene er forskjellige, og 2) Hvilke effekter bruk av objektforankret-, vs. subjektforankret instruksjon og -vurderingsskala kan gi opphav til ved måling av kundenes produktvurderinger i et spørreskjema.

Til slutt retter vi oppmerksomheten mot problemstillinger knyttet til inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet. Med bakgrunn i teoretiske perspektiver vil vi forsøke å forklare 1) Hvorfor variasjoner i plassering av viktighets-spørsmål og vurderingsspørsmål i spørreskjema kan gi forskjeller i respondentenes svargivning, og 2) Hvilke effekter disse forskjellene kan gi opphav til ved måling av kundenes produktvurderinger.

3.1 Hvordan blir produktvurderinger foretatt?

Vurderingsoppgavene en kunde står overfor før og etter kjøp av et produkt vil være fundamentalt forskjellige. Disse forskjellene kan skyldes informasjonsgrunnlaget kunden har i de to ulike situasjonene. Før kjøp av bil kan det f.eks. tenkes at en kunde

oppfatter holdepunkter som brede dekk, langt panser og hekkspoiler. Dette kan gi grunnlag for å anta at bilen har gode kjøreegenskaper, og at den derfor er sikker i trafikken. Denne typen slutninger kan kalles *prospektiv inferens* for å indikere at prosessen går fra holdepunkter via egenskaper til funksjoner (Troye, 1990). Etter kjøpet vil informasjonsgrunnlaget være basert på kundens erfaring med bruk av produktet. Prosessen vil dermed være snudd på hodet. Kunden har brukt produktet og erfart hvilke funksjoner det har. Oppgaven blir nå å tilskrive årsaker til disse funksjonene. Årsakene kan beskrives som de egenskapene som kunden tillegger produktet, f.eks. "bilen hadde gode kjøreegenskaper". Disse egenskapene må igjen ha sin årsak i bestemte holdepunkter som er innebygget i produktet fra markedsførers side, f.eks. brede dekk. Denne prosessen kan kalles *retrospektiv inferens* for å indikere at slutningene går fra funksjoner via egenskaper til holdepunkter (Troye, 1990).

Inferensprosessen er bare en av transformasjonene som kunden utfører for å komme frem til en totalvurdering av produktet (Stokmans, 1991; Troye, 1990). Den andre prosessen er selve evalueringsprosessen, eller vurderingsprosessen. Vurderinger innebærer at det knyttes verdier til egenskapene ved et objekt. Verdier kan knyttes både til holdepunkter (f.eks. brede dekk er bra), og til de oppfattede egenskapene (f.eks. gode kjøreegenskaper er bra) (Troye, 1990). Vi skal nå se nærmere på hvordan kunden kombinerer vurderinger av holdepunkter og egenskaper for å komme frem til en totalvurdering av produktet.

Produktvurderinger innebærer at kunnskap om et produkt omsettes til en vurdering av produktet. Basert på antakelser eller kunnskap om hvordan et produkt er, skal det foretas en vurdering av hvorvidt produktet er bra eller dårlig (Troye, 1990). Cox (1967) beskriver denne vurderingsprosessen ved bruk av *the sorting rule model*. I følge Cox innebærer produktvurderinger en prosess bestående av ni trinn: (1) *Etablering av egenskapskategorier*, f.eks. kan egenskapen, kjøreegenskaper, ved vurdering av en bil, være representert ved kategoriene "gode" eller "dårlige" kjøreegenskaper. (2) *Tillegging av verdi og konsekvenser* til egenskapskategoriene, f.eks. hvor viktig er kjøreegenskaper? (3) *Etablere kriterier* for å plassere et produkt i en spesifikk egenskapskategori. Kriterier vil i denne sammenheng bety det samme som holdepunkter for vurderingen (*cues*). (4) *Etablere kriteriumkategorier og sammenhenger* mellom holdepunkter for vurdering og egenskapene som vurderes. (5) *Tillegge prediktiv verdi* til hver kriterium-egenskap-sammenheng. Prediktiv verdi reflekterer sannsynligheten for at et holdepunkt som benyttes i vurderingen har sammenheng med en bestemt egenskap. (6) *Etablere (confidence) sikkerhetsverdi* for hvorvidt hvert holdepunkt kan plasseres i en bestemt kategori. Sikkerhetsverdi reflekterer her sikkerhet i kategoriseringen. (7) *Plassering* av produktet i en egenskapskategori, f.eks. "kjøreegenskapene er gode". (8)

Kombinering av vurderinger av egenskaper for å komme frem til en totalvurdering. (9) Vurdere *gyldigheten* av vurderingen som er foretatt. Denne modellen skiller mellom “holdepunkter for vurdering” og “egenskaper som vurderes”, og beskriver dermed produktvurderinger som en prosess der en totalvurdering er et resultat av ulike delprosesser. Vi kommer tilbake til denne vurderingsprosessen senere i kapittelet. Først vil vi se nærmere på skillet mellom holdepunkter og egenskaper som grunnlag for å foreta en beskrivelse av produktegenskapene som vurderes.

3.1.1 Produktegenskaper

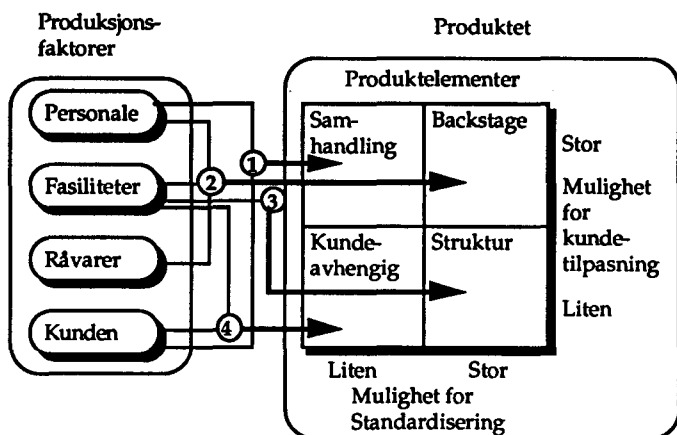
En vanlig definisjon på et produkt i markedsføringssammenheng er: *“en samling av egenskaper, eller attributter, tilbudt den ene parten, kjøper, av den andre parten, selger”* (Bagozzi, 1986). De tilbudte egenskapene kan være både materielle, fysiske egenskaper og immaterielle egenskaper som man ikke “kan ta og føle på” (*tangible* og *intangible* egenskaper). I økonomisk teori defineres et produkt som et sett av egenskaper eller *characteristics*, der tilstedeværelsen av bestemte egenskaper blir antatt å være relatert til kundens oppfattede nytte ved et spesifikt alternativ (Lancaster, 1966; Lancaster, 1971). I økonomisk teori om konsumenters atferd antar man videre at konsumenter maksimerer sin nytte når de skal bestemme hvilket produkt de skal kjøpe. Med “nytte” menes i hvilken grad et produkt tilfredsstiller konsumentens behov (Lancaster, 1966; Lancaster, 1971). Konsumenten avleder sin nytte av et produkts egenskaper, og ikke av produktet som sådan. Lancaster (1971) definerer egenskaper som objektive fysiske aspekter eller *characteristics* ved produktet. Basert på erkjennelsen av at et produkts egenskaper ikke nødvendigvis er fysiske aspekter ved produktet har følgende definisjon blitt foreslått for de egenskaper som vurderes (Grünert, 1989): *“Ethvert aspekt ved produktet, eller bruken av produktet, som kan brukes for å sammenligne ulike produktalternativer. Hvert alternativ kan (men trenger ikke) karakteriseres ved alle aspekter, dvs bruk av et aspekt er ikke uttømmende”*.

Fra et ledelsesperspektiv kan et produkts egenskaper sees på som et resultat av produksjonen. Kundene på sin side oppfatter egenskaper som bidrar til deres nytte av produktet (Lancaster, 1966; Lancaster, 1971). Vi skal nå se nærmere på forskjellene mellom disse to måtene å beskrive et produkt på.

Lederen som skal ta beslutninger om hvordan et produkt skal utformes, settes sammen, forbedres og markedsføres trenger en operasjonell definisjon, eller konseptualisering av produktet som tilbys kundene. En måte å beskrive produktet fra et slikt ledelses-

perspektiv kan bestå i at vi tar utgangspunkt i de ulike produksjonsfaktorene eller *inputs* som må til for å produsere produktet (Troye, 1990).

FIGUR 3.1
Produksjonsfaktorer og produktelementer - et ledelsesperspektiv



Kilde: Troye (1990)

Figur 3.1 viser hvordan fire ulike typer produksjonsfaktorer: råmaterialer, personale, fasiliteter (produksjonsanlegg), og kunden selv, kan kombineres eller interagere med hverandre i en prosess som gir opphav til følgende fire produktelementer (Troye, 1990; Troye, 1989): 1) *Samhandlingselementer*, omfatter personkontakten mellom kunden og personalet i bedriften. Disse elementene består av immaterielle produkttegenskaper. 2) *Backstageelementene*, er de sidene ved bedriftens produkt som er et resultat av aktiviteter som ikke skjer i direkte kontakt med kunden. Disse elementene består av egenskaper som ofte er materielle selv om de er høyst avhengig av personalet (f.eks. mat tilberedt i restauranter). 3) *Strukturelementene*, omfatter bygninger, fasiliteter og produksjonsutstyr. Tilstedeværelse, snarere enn bestemte aktiviteter fra kundens side er her nødvendig for å gi nytte. Elementene omfatter de mer materielle aspekter ved produktet som "design", "komfort" og "størrelse". 4) *Kundeavhengige elementer*, omfatter den egeninnsatsen som er nødvendig for at kunden skal kunne anskaffe, bruke og gjøre seg nytte av et gode. Disse elementene bringer inn Toffler's "prosumenterperspektiv" (Kotler, 1986; Toffler, 1980). Kunden er også "produsent" ettersom de fleste goder må viderefredes av kunden før de kan gi nytte. Det forutsettes aktivitet fra kundens side. Produktets egenskaper kan i liten grad oppleves bare ved å betrakte de strukturelle elementene, men tilskrives på grunnlag av kundens bruk av disse (Troye, 1990; Troye, 1989).

Modellen over kan anvendes både for produkter som tradisjonelt klassifiseres som tjenester, og for varer eller fysiske produkter. I tjenestemarkedsføringslitteraturen finner vi lignende innfallsvinkler til beskrivelse av tjenester (Levitt, 1981; Shostack, 1977). Grad av materielle egenskaper (*degree of tangibility*) blir her benyttet for å klassifisere produkter som varer eller tjenester. Produkter med høy grad av materielle egenskaper klassifiseres som fysiske varer, mens tjenester innehar en høy grad av immaterielle egenskaper. På denne måten kan alle produkter, det være seg varer eller tjenester, plasseres langs et kontinuum med et materielt dominant- og et immaterielt dominant³⁶ endepunkt.

Modellen i tablå 3.1 har vi kalt en *ledelsesmodell*. En slik modell kan være nyttig når formålet er å bestemme hva som må gjøres for å forbedre produktet. Med utgangspunkt i målinger av kundenes vurdering av egenskaper ved de fire produktelementene, kan modellen brukes for å relatere vurderte egenskaper ved produktet til innsatsfaktorene som inngår i produksjonen. Modellen ivaretar dermed et vitenskapsteoretisk krav om kontrollvaliditet (Zaltman, Pinson, og Angelmar, 1972)³⁷.

3.1.2 Holdepunkter (*cues*)

Ved utvikling av nye produkter er utvelgelse, og bruk av holdepunkter en viktig del av produktutforming og posisjonering (Stokmans, 1991). Følgende spørsmål er sentrale i denne sammenheng (Urban og Hauser, 1980): Hvilke funksjoner skal produktet utføre? Hvilke funksjoner skal oppfattes som kjernen i produktet? Hvordan skal produktet posisjoneres? Det gjøres altså et skille mellom selve produktet, kjennetegnet ved de to første spørsmålene, og produktet slik det markedsføres, beskrevet ved det tredje spørsmålet.

Selve produktet kalles ofte kjerneproduktet (Bagozzi, 1986; Kotler, 1984), og omfatter "holdepunkter som er *iboende* i produktet"³⁸ og som oftest er fysiske aspekter (Jacoby, Olsson, og Haddock, 1971; Olson, 1977; Olson og Jacoby, 1972). Dersom man endrer iboende holdepunkter, endrer man også det fysiske produktet. Iboende holdepunkter kan deles inn i: 1) Funksjonelle holdepunkter, som er relatert til den forventede nytten av produktet. 2) Strukturelle holdepunkter, som omfatter måten funksjonelle holde-

³⁶Eksempler på produkter som reflekterer endepunktene kan være *salt* for det materielt dominante endepunktet og *frisørjenester* for det immaterielt dominante endepunktet.

³⁷Kontrollvaliditet kan f.eks. for en definisjon av kvalitet bety at kvalitet må defineres på en slik måte at man har mulighet til å gjøre noe for å påvirke kvaliteten.

³⁸*Intrinsic product cues*

punkter realiseres i produktet, f.eks. størrelse, form, farge, materiale, osv. 3) Estetiske holdepunkter, som refererer til produktets design og bidrar til å skape et visuelt attraktivt produkt (Wind, 1982).

Det markedsførte produktet (Bagozzi, 1986; Kotler, 1984) består av det fysiske produktet utvidet med holdepunkter som har sitt opphav i markedsføringens virkemidler. Disse kalles "*utvendige* holdepunkter"³⁹ (Jacoby, Olson, og Haddock, 1971; Olson, 1977; Olson og Jacoby, 1972). Utvendige holdepunkter kan deles inn i: 1) holdepunkter som er avledet av virkemidler i markedsføringen uten å tilhøre kjerneproduktet, f.eks. merkenavn, innpakning, garantier og *service*, og 2) holdepunkter avledet fra de øvrige virkemidlene i markedsføring: pris, promosjon og distribusjon.

Andre måter å beskrive holdepunkter slik de benyttes av tilbyderer av et produkt, er i hvilken grad holdepunkter er *objektive* eller *subjektive* (Stokmans, 1991). Objektive holdepunkter kan presenteres på en slik måte at omtrent alle tolker dem på samme måte, f.eks. fysiske og kjemiske aspekter ved produktet. Objektive holdepunkter er vanligvis iboende holdepunkter. Subjektive holdepunkter kan ikke beskrives på en slik måte at de tolkes på samme måte av alle. Eksempler på subjektive holdepunkter er butikken produktet selges i, innhold i reklamen osv. Subjektive holdepunkter er dermed vanligvis utvendige holdepunkter. Distinksjonen mellom objektive- og subjektive holdepunkter er relatert til måten konsumenten oppfatter produktets holdepunkter på. De oppfattede produktholdepunktene danner egenskaper som brukes av kunden ved vurdering og valg mellom produkter (Anderson, 1981; Stokmans, 1991). Hvordan disse egenskapene fremkommer drøftes i neste avsnitt.

3.1.3 Egenskaper brukt i kundens vurderingsprosess

Forskjellen mellom objektive og subjektive holdepunkter, slik disse inngår i individets vurderingsprosess, kan relateres til transformasjonen som kobler et produkts holdepunkter (direkte observerbare karakteristika) til egenskapene ved produktet (de oppfattede motpartene til produktets holdepunkter). Dersom denne transformasjonen kan beskrives ved en enkel algebraisk funksjon (monoton og tilnærmet en-til-en) kan produktets holdepunkter betraktes som objektive. Dette gjelder de fleste iboende produktholdepunkter. Dersom den algebraiske funksjonen er kompleks (flere holde-

³⁹*Extrinsic product cues*

punkter har en påvirkning på de oppfattede egenskapene) kan holdepunkter kalles subjektive. De fleste utvendige produktholdepunkter er subjektive (Stokmans, 1991).

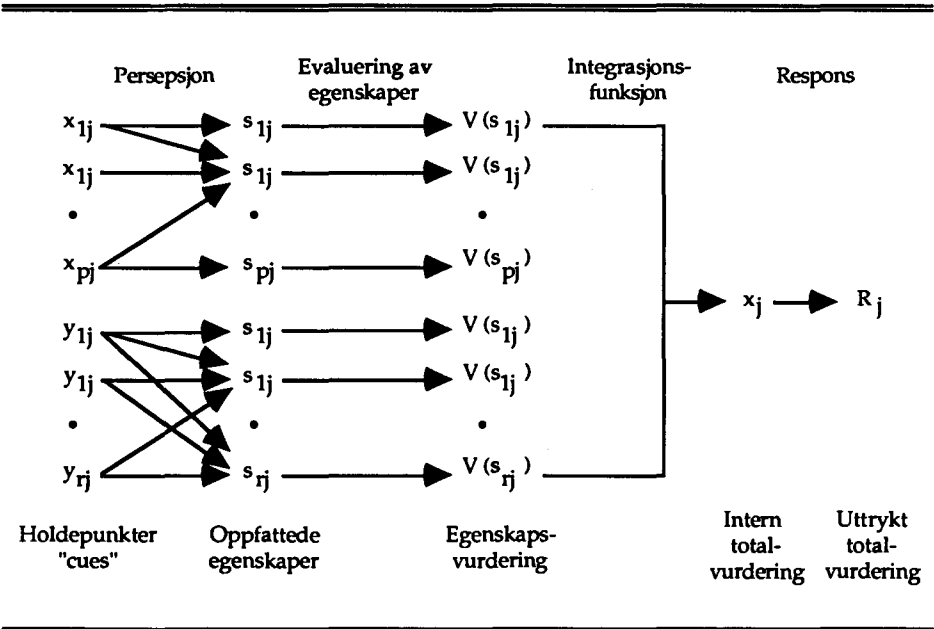
Cox (1967) hevder at ethvert holdepunkt som bidrar til å gi informasjon om produktet kan beskrives ved to faktorer; prediktiv- og sikkerhetsverdi⁴⁰. *Prediktiv verdi* reflekterer i hvilken grad konsumenten oppfatter eller antar at et holdepunkt (f.eks. pris) er relatert til, eller gir indikasjoner om egenskaper ved produktet (f.eks. kvalitet). *Sikkerhetsverdi* beskriver i hvilken grad en konsument er sikker på sine evner til å oppfatte og kategorisere et bestemt holdepunkt. Dersom konsumenten i sine vurderinger har tilgang til både iboende holdepunkter og utvendige holdepunkter med ekvivalent sikkerhetsverdi, vil iboende holdepunkter, ha større sannsynlighet for å bli tilegnet, og dermed ha sterkere innflytelse på produktvurderingene (Olson, 1977; Olson og Jacoby, 1972; Szybillo og Jacoby, 1974). Dette begrunnes med bakgrunn i at iboende produktholdepunkter vil ha høyere prediktiv verdi enn utvendige produktholdepunkter.

Basert på sammenhengene vi har beskrevet mellom produktholdepunkter og produkt-egenskaper, kan vi argumentere for at immaterielle og materielle egenskaper ved et produkt vil bli tillagt ulik betydning i produktvurderingene. De samme implikasjonene kan utledes av tjenestemarkedsføringslitteraturen, der det argumenteres for bruk av forskjellige vurderingsprosesser for varer og tjenester (Zeithaml, 1980). Zeithaml (1980) benytter skillet mellom *search qualities*, *experience qualities* (Nelson, 1974) og *credence qualities* (Darby og Karni, 1973) for å forklare hvordan vurderingsprosessene er forskjellige. *Search qualities* er her definert som egenskaper ved et produkt som en konsument kan vurdere (gjennom inspeksjon) før produktet kjøpes. Eksempler på *search qualities* er egenskaper som farge og design. *Experience qualities* defineres som egenskaper som kun kan vurderes etter kjøp og bruk av produktet. Eksempler er her smak, holdbarhet og tilfredshet. *Credence qualities* er egenskaper som kunden kan finne vanskelig å vurdere selv etter kjøp og bruk av produktet. Som eksempler på produkter der *credence qualities* dominerer, nevnes blindtarmoperasjoner og service på biler. Zeithaml (1980) beskriver videre tjenester som å inneha et større innslag av *credence qualities* og *experience qualities*, enn fysiske varer der *search qualities* dominerer. Dette hevdes å føre til at tjenester blir vanskeligere å vurdere enn fysiske varer. Som resultat av dette "tvinges" kundene til å benytte andre holdepunkter og vurderingsprosesser ved vurdering av tjenester.

⁴⁰*Confidence value*

Med utgangspunkt i produktmodellen vi beskrev i figur 3.1, der vi betrakter samhandling/service (tjenesteegenskaper) som en del av alle produkter, kan tjenestemarkedsføringslitteraturen benyttes for å forstå hvordan egenskaper ved samhandling vurderes forskjellig fra øvrige egenskaper ved produktet. De immaterielle samhandlingsegenskapene vil med dette utgangspunktet være vanskeligere for kundene å vurdere, enn de øvrige produktegenskapene. Vanskelighetene med å vurdere samhandlingsegenskaper antas å "tvinge" kundene til å legge mer vekt på de andre egenskapene ved vurdering av produktet. Vurderingsforskjellene vi har diskutert ovenfor vil senere i dette kapitlet danne utgangspunkt for fokusering på den relative betydningen av materielle (samhandling) egenskaper og immaterielle produktegenskaper for respondenters totalvurderinger og atferdsintensjoner.

FIGUR 3.2
Vurdering av et produkt som en samling av iboende- og utvendige produktholdepunkter



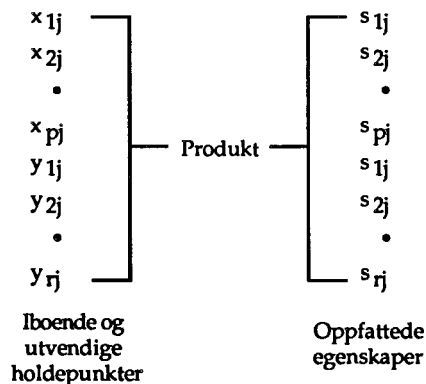
Kilde: Stokmans (1991)

Persepsjonsprosessen er bare en av transformasjonene som kunden går gjennom for å komme frem til en totalvurdering av produktet (Anderson, 1981; Stokmans, 1991). I figur 3.2 ovenfor, er det gjort et skille mellom iboende produktholdepunkter (x_{ij}) og utvendige produktholdepunkter (y_{ij}) ved produktet. Sammenhengen mellom iboende produktholdepunkter og utvendige produktholdepunkter, og de korresponderende oppfattede egenskapene er ikke nødvendigvis et en-til-en forhold. Relasjonen mellom et iboende produktholdepunkt og dets oppfattede motpart er sannsynligvis mindre

kompleks enn for utvendige produktholdepunkter. Konsumenten danner altså oppfatninger eller meninger om produktet basert på holdepunkter (x_{ij}) og (y_{ij}) som benyttes for å vurdere produktet (Alpert, 1971).

Den neste transformasjonen i konsumentens vurdering er å bestemme hvor viktig det er at produktet innehar bestemte egenskaper, og å foreta en vurdering av de valgte egenskapene ($V(s_{ij})$). Til slutt foretas en totalvurdering av produktet ved å integrere de vurderte egenskapene (r_{ij}) (Anderson, 1981; Stokmans, 1991). Beskrivelsen av et produkt som bestående av iboende-, og utvendige produktholdepunkter, eller materielle- og immaterielle produkt egenskaper, er vanligvis brukt av tilbyderen for å gi en detaljert beskrivelse av produktet slik det markedsføres. Produkt egenskaper brukes av kunden som grunnlag for vurderinger av produktet med formål å velge de produkter som maksimerer nytte. Denne distinksjonen er forsøkt illustrert i figur 3.3 som illustrerer hvordan et produkt, fra et ledelsesperspektiv, kan konseptualiseres ved hjelp av iboende-, og utvendige produktholdepunkter, mens det fra en kundens ståsted kan bli oppfattet som å bestå av en sammensetning av ulike produkt egenskaper.

FIGUR 3.3
Produktet som en samling av iboende produktholdepunkter og utvendige produktholdepunkter (x_{ij}), vs. en samling av egenskaper (s_{kj})



Kilde: Stokmans (1991)

I neste avsnitt skal vi se nærmere på hvordan vurderingsprosessene vi har beskrevet ovenfor kommer til uttrykk i ulike typer *multiattributtmodeller*, en type modeller som er mye benyttet, både for å beskrive produkter i markedsføringsammenheng, og for å beskrive vurderingsobjekter ved konseptualisering og måling av holdninger.

3.1.4 Multiattributtmodeller

I markedsførings- og konsumentatferdslitteraturen blir produkter ofte beskrevet i forhold til "attributter", et begrep lånt fra forskning vedrørende holdninger og bruk av multiattributtmodeller (Lutz, 1991). I en forbrukssituasjon er det vanlig å karakterisere et attributt som "*ethvert aspekt ved et produkt, eller bruken av et produkt, som kan brukes ved sammenligning av produktalternativer*" (Grünert, 1989). Bruk av multiattributtmodeller impliserer at konsumenter vurderer forskjellige aspekter ved en vare eller tjeneste. Forskjellige multiattributtmodeller har vært foreslått i litteraturen⁴¹. Det teoretiske rasjonale for disse modellene varierer. Hvordan attributter eller egenskaper blir tatt hensyn til i disse modellene varierer også. To former for multiattributtmodeller har stått sentralt i teorier innenfor konsumentatferd; Fishbein og Rosenbergs modeller (Lutz og Bettman, 1977).

Fishbeins multiattributtmodell (Fishbein, 1963; Lutz og Bettman, 1977) er basert på læringsteori. Modellen kan illustreres ved følgende ligning:

$$A_o = \sum_{i=1}^n b_i a_i$$

der A_o er holdning overfor et objekt; b_i er antakelsen (subjektiv sannsynlighet) om at objekt, o , innehar egenskap i ; a_i er vurderingen (f.eks. bra-dårlig) av egenskap i ; og n er antall fremtredende eller vurderte egenskaper. Affekt eller holdning overfor objektet fremkommer som en multiplikativ summering av vurderte egenskaper veiet med antakelsene om sannsynligheten for at objektet innehar egenskapene som vurderes (Fishbein, 1963; Lutz og Bettman, 1977).

Rosenbergs modell (Lutz og Bettman, 1977; Rosenberg, 1956) bygger på konsistensteori for å forklare hvordan holdninger endres. Ifølge konsistensteori vil et individ ha holdninger og handle i samsvar med sine verdier. Modellen kan illustreres på følgende måte:

$$A_o = \sum_{i=1}^m PI_i VI_i$$

der A_o er holdning overfor et objekt; PI_i er den oppfattede instrumentalitet objekt o har i å føre til blokkering eller oppnåelse av en bestemt verdi, i ; VI_i er viktigheten av verdi, i , uttrykt som grad av tilfredshet eller utilfredshet med hensyn til verdien; og m er antall fremtredende verdier. Rosenbergs modell beskriver kognitiv-affektiv konsistens. I

⁴¹For en oversikt, se Grünert (1989) og Lutz & Bettman (1977).

markedsforskning har man anvendt denne modellen og måleprosedyrene den impliserer, men man har da substituert verdier med produktegenskaper. Affekt eller holdning overfor objektet fremkommer som en multiplikativ summering av vurderte egenskaper ved et produkt veiet med viktigheten av egenskapenes tilstedeværelse (Lutz og Bettman, 1977; Rosenberg, 1956).

I markedsforskning har ulike hybride multiattributtmodeller blitt benyttet. Lutz (1977) samler disse under fellesbetegnelsen *the intuitive approach*. Standard multiattributtmodeller har alle en komponent for vektning av egenskaper. Disse er kalt *weights*, *salience* eller *importance* ved anvendelser av modellene i markedsforskning.

Av dette følger at alle attributter, eller egenskaper ikke nødvendigvis tillegges den samme vektlegging eller viktighet av konsumenten ved produktvurderinger. De egenskaper som vektlegges mest, f.eks. egenskaper som er kritiske i kundens vurderinger, er ofte karakterisert som fremtredende⁴² egenskaper. Dette er egenskaper som er dominerende ved vurdering av produkter. Vurdering av egenskaper multiplisert med sine respektive vekter antas kombinert av kunden ved bruk av en vurderingsregel⁴³ slik at totalvurderingen fremkommer.

Basert på de to modellene vi presenterte ovenfor kan individuelle egenskaper ved produktet kobles til atferd gjennom en totalvurdering. Holdningsteori kan her være til hjelp for å forstå hvordan dette skjer. Holdninger blir vanligvis definert som: "*a learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object*" (Fishbein og Ajzen, 1975). To ulike konseptualiseringer av holdninger og holdningsdannelse er benyttet i litteraturen: *the tripartite view of attitude* og *the unidimensionalist view of attitude* (Lutz, 1991).

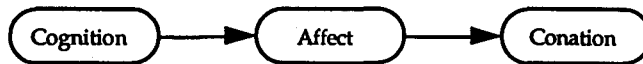
I følge *the tripartite view of attitude* betraktes både kognitive, affektive og atferdsmessige komponenter som en del av holdningsbegrepet. *The unidimensionalist view of attitude* skiller kognitive, affektive og atferdsmessige komponenter fra hverandre og definerer *beliefs* eller antakelser som kognitive årsaksfaktorer til holdninger mens intensjoner er de direkte kausale atferdsmessige konsekvenser, eller effekter av holdningene. Dette betyr at dersom en konsument lærer noe om et nytt produkt (f.eks. gjennom reklame), vil denne kunnskapen i form av antakelser lede til en holdning overfor produktet, som i neste omgang kan lede til en intensjon om å kjøpe, eller ikke kjøpe produktet (Lutz, 1991).

⁴²For en mer detaljert gjennomgang av *salience* i produktvurderinger, se Fiske & Taylor (1991).

⁴³For en gjennomgang av ulike vurderingsregler se f.eks. Lynch (1985).

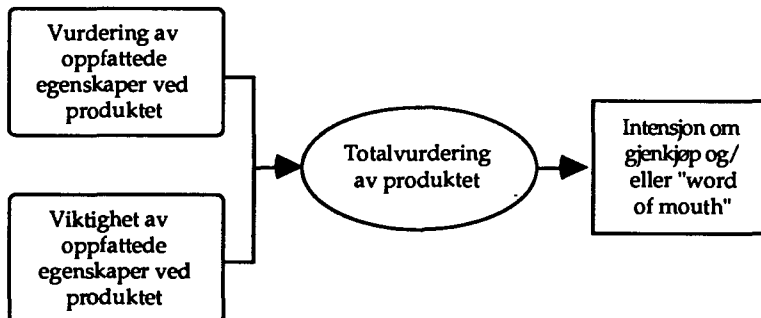
I figur 3.4 har vi gitt en skjematisk fremstilling av *the unidimensionalist view of attitude* som er utgangspunktet for Fishbeins *theory of reasoned action* modell (Fishbein og Ajzen, 1975; Grünert, 1989; Lutz, 1991; Lutz og Bettman, 1977).

FIGUR 3.4
The unidimensionalist view of attitude



I følge *the unidimensionalist view of attitude* skilles holdningskomponenten ut i en egen affektkomponent der holdning defineres som grad av affekt⁴⁴ (positiv eller negativ) overfor et objekt. *Cognition* er i forhold til Fishbein og Rosenbergs modeller representert ved antakelser, vurdering av egenskaper, styrke på antakelser og viktighetsvekter (Fishbein og Ajzen, 1975; Grünert, 1989; Lutz, 1991; Lutz og Bettman, 1977). *Conation* komponenten i *the unidimensionalist view of attitude* representerer individets intensjoner om å utføre en bestemt atferd overfor objektet, eller observert atferd. Basert på holdningsteori og informasjonsintegrasjonsteori er en skjematisk modell for kunders produktvurderinger fremstilt i figur 3.5.

FIGUR 3.5
En skjematisk fremstilling av en modell for kunders produktvurderinger



Modellen ovenfor viser hvordan totalvurdering av et produkt kan betraktes som et resultat av en transformasjon av kundens vektlegging og vurdering av oppfattede produkt egenskaper. Dette samsvarer med *the unidimensionalist view of attitude* slik

⁴⁴Positiv eller negativ innstilling overfor et objekt.

dette er illustrert i figur 3.4. I figur 3.5 har vi i tillegg inkludert intensjoner om gjenkjøp og/eller *word of mouth* som to mulige effekter av kundens totalvurdering. Vurderingsmodellen kan uttrykkes matematisk på følgende måte:

$$TV_p = \sum_{i=1}^n E_i V_i$$

der TV_p er totalvurderingen av produktet; E_i er vurdering av oppfattet egenskap i ved produkt p ; V_i er viktigheten av egenskap i ; og n er antall fremtredende egenskaper.

3.1.5 Respondentens produktkunnskap

Kundens erfaring med bruk av et produkt hevdes å føre til økt produktkunnskap (Alba og Hutchinson, 1987; Alba og Marmorstein, 1986) som både kan påvirke produktvurderingene (Alba og Hutchinson, 1987; Alba og Marmorstein, 1986; Bettman og Park, 1980; Sujan, 1985) og respons i spørreskjemaet (Hippler, Schwarz, og Sudman, 1987). Tradisjonelt har produktkunnskap blitt behandlet som et endimensjonalt begrep, og referert til som *produktkjennskap (product familiarity)* eller som *tidligere erfaring (prior knowledge)* (Alba og Hutchinson, 1987).

Man antar at konsumenter har en viss mengde erfaring med, eller informasjon om produkter som grunnlag for vurderinger. En måte å se konsumenters produktkunnskap på, er å betrakte kunnskapsbegrepet som bestående av to komponenter: *kjennskap* og *ekspertise* (Alba og Hutchinson, 1987; Jacoby et al., 1983). Kjennskap defineres her som "antall produktrelaterte erfaringer som har blitt akkumulert av konsumenten", mens ekspertise er definert som evne til å utføre produktrelaterte oppgaver på en vellykket måte. Produktrelaterte erfaringer fremkommer f.eks. gjennom eksponering for reklame, informasjonssøking, interaksjon med selgere, valg og beslutningstaking, kjøp og bruk av produkter i ulike situasjoner. Ekspertise reflekteres både i konsumentens *kognitive strukturer*, som f.eks. i antakelser om egenskaper ved produktet, og i *kognitive prosesser*, som f.eks. i beslutningsregler for atferd på bakgrunn av antakelsene. Både kognitive strukturer og kognitive prosesser er nødvendige for å kunne utføre produktrelaterte oppgaver på en vellykket måte. Økt produktkjennskap antas å føre til økt ekspertise (Alba og Hutchinson, 1987).

Alba og Hutchinson (1987) hevder at fem ulike aspekter ved ekspertise kan forbedres etterhvert som produktkjennskapen øker. For det første kan repetisjoner av bruken av et produkt forbedre vurderingsytelsen gjennom en reduksjon i den kognitive kapasiteten

som kreves for å utføre vurderingsoppgaven. I enkelte tilfeller kan gjentatt bruk til og med føre til at vurderingsytelsen blir automatisert. For det andre kan de kognitive strukturene som brukes for å differensiere produktet bli mer raffinert, mer komplette og mer korrekte etterhvert som produktkjennskapen øker. For det tredje kan evnen til å analysere informasjon, og isolere den delen av informasjonen som er viktigst og mest relevant forbedres. For det fjerde kan evnen til å generere presis kunnskap på basis av gitt informasjon forbedres. Til slutt kan evne til å huske produktinformasjon bli forbedret etterhvert som produktkjennskapen øker. I de neste to avsnittene skal vi se hvordan disse forholdene kan få betydning for hvordan kunder foretar produktvurderinger.

Kjennskap til produktet, eller produktrelatert erfaring, er altså en sentral dimensjon i kunnskapsbegrepet. Mer spesifikt kan produktrelatert erfaring betraktes som en nødvendig, men ikke tilstrekkelig betingelse for økt produktkunnskap (Rao og Monroe, 1988). Med dette som utgangspunkt har en konsumenters produktrelaterte erfaring ofte blitt benyttet som mål på konsumentens produktkunnskap (f.eks. Bettman og Park (1980)). Dette er imidlertid ikke uproblematisk. Selnes og Grønhaug (1986) hevder f.eks. at produktkunnskap kan utvikles gjennom informasjonssøking og bruk av informasjon, og på denne måten være uavhengig av brukserfaring. Brukserfaring behøver heller ikke alltid føre til at produktkunnskapen øker. Brukserfaring kan f.eks. medføre økt bruk av tommelfingerregler.

Litteraturgjennomgangen i forbindelse med denne avhandlingen har vist at det hersker relativt stor uklarhet om hvordan erfaring, kunnskap og ekspertbegrepet er definert og benyttet. Det samme gjelder hvordan begrepene er blitt målt empirisk. Som et resultat av dette vil alle begrepene bli benyttet i de teoretiske resonnementene knyttet til effekter av produktrelatert erfaring og produktkunnskap. Når vi på bakgrunn av teori- drøftelsene formulerer hypoteser i kapittel 4, vil vi imidlertid fokusere på produktkunnskapbegrepet.

3.1.6 Produktkunnskap og holdningsdannelse

Tidligere i dette kapitlet benyttet vi teorier om holdningsdannelse for å belyse hvordan kunder foretar produktvurderinger. I *the unidimensionalist view of attitude* blir det gjort et skille mellom kognitive, affektive og atferdsmessige komponenter. Antakelser eller vurderinger av egenskaper ved et produkt og vektlegging av egenskapene utgjør individets kognitive strukturer, mens prosessen der de kognitive strukturene omsettes i

en totalvurdering og holdning overfor produktet representerer de kognitive prosessene. Erfaring med bruk av produktet og produktkunnskap kan påvirke både kognitive strukturer (Alba og Hutchinson, 1987) og kognitive prosesser (Alba og Hutchinson, 1987; Sujan, 1985). Kognitiv struktur er i konsumentforskning referert til som den faktiske kunnskap konsumentene har om produkter, og måten denne kunnskapen er organisert. Kognitive strukturer har som funksjon å hjelpe individet med å differensiere varer og tjenester på en måte som er formålstjenlig for å kunne ta beslutninger. Faktisk produktkunnskap antas å øke etterhvert som den produktrelaterte erfaringen øker (Alba og Hutchinson, 1987).

I konsumentforskning har man vært opptatt av hvordan kategoristrukturer styrer akkumulering av kunnskap. Individer antas å foreta naturlige inndelinger av objekter i kategorier for å forenkle informasjonsbehandlingen (se f.eks. Rosch (1988) og Sujan, (1985)). Produktrelatert erfaring kan bidra til at kundene blir i stand til å diskriminere bedre mellom produktkategorier, f.eks. pga. bedre utviklede subkategorier (Alba og Hutchinson, 1987). Dette kan bety at kunder med høy erfaring er bedre i stand til å skille mellom ulike egenskaper ved et produkt. Produktrelatert erfaring forventes også å bidra til at et produkt kan beskrives på en mer komplett måte. "Ekspertene" vil med andre ord kunne ha flere produktegenskaper representert i sine kognitive strukturer. Til slutt kan produktrelatert erfaring føre til en dypere og mindre overfladisk struktur (Alba og Hutchinson, 1987).

For at holdninger skal virke styrende for atferd, må de være tilgjengelige⁴⁵ (Fazio, 1989). Økt brukserfaring og produktkunnskap kan bidra til at holdninger og produktvurderinger blir mer tilgjengelige. Holdninger eller produktvurderinger basert på erfaring vil også med større sannsynlighet påvirke atferd (Fazio, 1989; Zanna, Olson, og Fazio, 1980).

Ett av problemene man står overfor ved måling av holdninger og vurderinger i spørreundersøkelser, er at individer med dårlig utviklede, svake og i noen tilfeller ikke eksisterende holdninger eller vurderinger, ofte svarer på en måte som indikerer negative eller positive holdninger eller vurderinger. Når disse individene inkluderes som respondenter kan dette føre til en underestimert styrke på sammenhengen mellom vurderinger, holdninger og atferdsintensjoner (Sample og Warland, 1973). For å øke reliabilitet og validitet ved måling av holdninger og vurderinger, er det mulig å kontrollere for disse effektene ved å inkludere modererende variabler i undersøkelsene

⁴⁵*Attitude accessibility*

(Ghiselli, 1963). En slik moderende variabel som ofte inkluderes ved måling av holdninger er sikkerhet i vurderinger (*confidence*).

Sikkerhet i vurderinger på holdningsspørsmål kan være avhengig av i hvilken grad individet har utviklet holdninger overfor det objektet som det blir bedt om å vurdere (Fishbein og Ajzen, 1975; Sample og Warland, 1973). Hvorvidt holdninger er utviklet og eksisterer vil være avhengig av erfaringen respondentene har (Antil, 1983; Raju, 1977). Erfaring med bruk av et produkt vil derfor være avgjørende for hvor sikre individene vil være på sine holdninger og i sine vurderinger (Fazio, 1989). Sikkerhet i vurderinger vil også være positivt relatert til atferdsintensjoner (Howard og Sheth, 1969). Varer og tjenester varierer i grad av kompleksitet. Den oppfattede vanskeligheten med å lære om, vurdere, og sammenligne varer og tjenester kan variere på tvers av klasser av varer og tjenester. Erfaring kan resultere i læringsprosesser som reduserer oppfatningen av produktets kompleksitet, og dermed de kognitive oppgavene som behøves for å foreta en vurdering (se f.eks. Rosch (1988)).

I dette avsnittet har vi forsøkt å vise hvordan multiattributtmodeller og teoribidrag fra informasjonsintegrasjonsteori og holdningsteori kan benyttes for å forstå hvordan kunder foretar produktvurderinger. I tillegg drøftet vi hvordan produktrelatert erfaring og produktkunnskap kan få betydning for vurderingene. I kapittel 1 ble problemstillinger vedrørende kunders produktvurderinger, etter at de har kjøpt og brukt et produkt, relatert til kjøpsprosessens "aktiviteter etter kjøp". Et sentralt resultat av kundens etterkjøpsaktiviteter er tilfredshet (Engel og Blackwell, 1982). I neste avsnitt skal vi se nærmere på hvordan teorier og modeller om kundetilfredshet kan benyttes for å beskrive og forstå kunders produktvurderinger.

3.1.7 Kundetilfredshet

Kundetilfredshet har blitt et sentralt begrep i moderne markedsføringsteori, og i markedsføringspraksis. Dette har ført til en anseelig mengde forskning på dette området de siste tiårene⁴⁶. To forskjellige definisjoner på kundetilfredshet har blitt fremsatt. På den ene siden er kundetilfredshet definert som et resultat av den erfaringen konsum av produktet gir opphav til (Yi, 1990). Oliver (1981) definerer f.eks. kundetilfredshet som "den resulterende psykologiske tilstanden som oppstår når følelsene omkring uoppfylte

⁴⁶For en oversikt, se Yi (1990).

forventninger kobles med de følelser og forventninger konsumenten hadde før konsumerfaringen”.

Kundetilfredshet er også definert som en *prosess* (Yi, 1990). Tse og Wilton (1988) beskriver f.eks. kundetilfredshet som “konsumentens respons på en evaluering av avvik mellom forventninger (eller norm for ytelse) og produktets ytelse slik denne oppfattes etter kjøp og bruk. I det følgende velger vi å fokusere på *prosessorienterte* definisjoner på kundetilfredshet da disse legger mer vekt på de oppfatningsmessige (*perseptuelle*), vurderingsmessige (*evaluative*) og psykologiske prosessene som kombineres for å generere kundetilfredshet (Yi, 1990).

Et sentralt paradigme i studier av kundetilfredshet har vært *the disconfirmation paradigm* (Oliver, 1980; Churchill og Surprenant, 1982). Tilfredshet blir her relatert til størrelsen og retningen på avviket mellom opplevelse og forventninger. Forventninger blir bekreftet (*confirmed*) når produktet har en ytelse som er forventet. Forventninger blir negativt avkrefte (*negatively disconfirmed*) når produktet har en ytelse som er dårligere enn forventet. Til slutt blir forventninger positivt avkrefte (*positively disconfirmed*) når produktet har en ytelse som er bedre enn forventet. Utilfredshet oppstår når individets forventninger blir negativt avkrefte. Forventninger reflekterer her “antatt ytelse”. I kundetilfredshetslitteraturen eksisterer det fire typer forventninger om et produkts antatte ytelse: idealytelse, forventet ytelse, minimums tolerert ytelse og ønskelig ytelse (Churchill og Surprenant, 1982). Tilfredshetsvurderinger kan med utgangspunkt i dette betraktes som en kategoriseringsprosess (Mervis og Rosch, 1981).

3.1.8 Produktkunnskap og kundetilfredshet

Forventninger, *disconfirmation*, opplevd ytelse og eksisterende holdninger er forhold man har funnet påvirker kundetilfredshet (Alba og Hutchinson, 1987). Erfaring med bruk av et produkt kan påvirke både normer og de forventninger kunden har til produktets ytelse (Cadotte, Woodruff, og Jenkins, 1987; Mazursky og Geva, 1989; Oliver, 1980; Tse og Wilton, 1988; Woodruff, Cadotte, og Jenkins, 1983) En person med høy erfaring og produktkunnskap vil f.eks. ha “riktigere forventninger” til produktet. Forventninger kan både omfatte hvilke egenskaper som forventes å være til stede, og det forventete nivået på egenskapenes ytelse. “Riktige forventninger” vil bidra til å redusere avvik (*disconfirmation*) mellom forventninger til- og opplevd ytelse ved produktet når konsumenten vet hvilke egenskaper som kan forventes å være til stede, og hvilket nivå det vil være på disse.

I denne avhandlingen fokuserer vi både på teorier og modeller for holdningsdannelse og kundetilfredshet som grunnlag for å forstå hvordan produktvurderinger blir foretatt. Kundens tilfredshet med et produkt har av enkelte forskere blitt hevdet å være det samme som holdninger til produktet (Yi, 1990). Mål på tilfredshet betraktes i denne sammenhengen som "postkonsum holdningsmål"⁴⁷ (LaTour og Peat, 1979; Yi, 1990).

Andre hevder at kundetilfredshet er forskjellig fra holdninger, se f.eks. Yi (1990). Det sentrale argumentet for en slik påstand er at tilfredshet er relatert til *disconfirmation*, som er konsumspesifikk og en funksjon av et overraskelsesmoment, mens holdninger er kundens vedvarende affekt overfor et objekt uten innslag av overraskelsesmomenter. Holdningsmålinger er derfor mer generelle i forhold til et spesifikt produkt eller en spesifikk brukserfaring, og mindre situasjonsorienterte enn målinger av kundetilfredshet (Oliver, 1981; Yi, 1990). Videre hevdes det at holdningsmålinger mangler et element der man kan sammenligne forventninger og opplevelse. Det forventes på bakgrunn av dette at tilfredshet og holdninger avviker spesielt når forventninger og opplevelse avviker (*disconfirmation*) (Oliver, 1981; Yi, 1990).

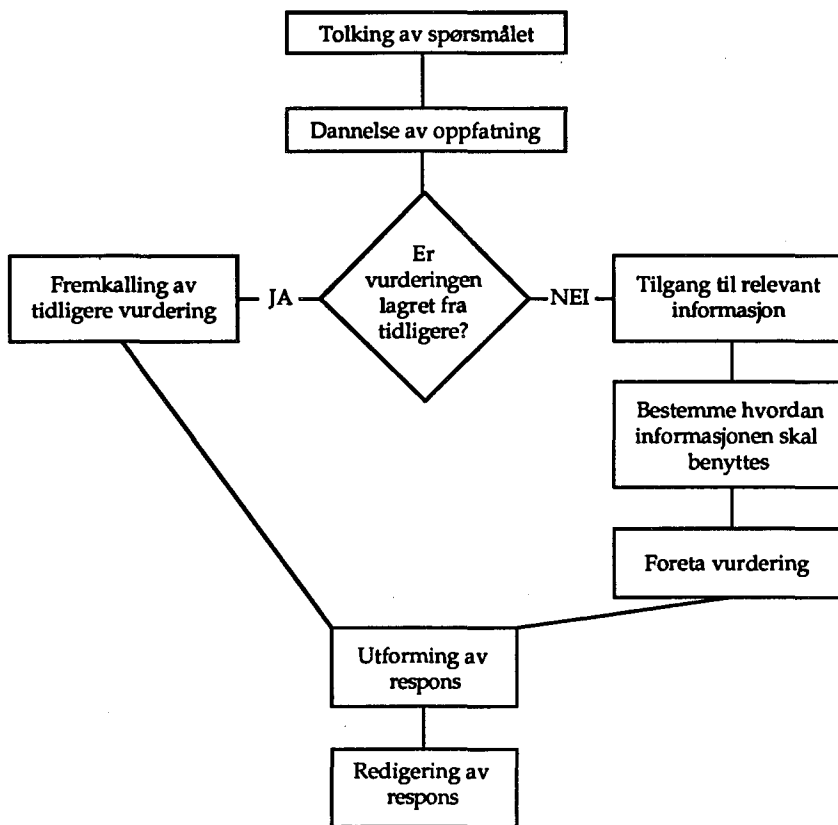
3.2 Kognitive oppgaver ved svargivning på spørsmål i et spørreskjema

Ved svargivning på spørsmål i et spørreskjema antas respondenten å utføre en sekvens av ulike kognitive oppgaver (Clark og Schober, 1991; Strack og Martin, 1987; Tourangeau og Rasinski, 1988). Først må respondenten foreta en *tolkning* for å forstå meningsholdet i hvert enkeltspørsmål i spørreskjemaet. Deretter må respondenten *danne seg en oppfatning* eller gjøre seg opp en mening om temaet det spørres om i spørsmålet. Videre må denne oppfatningen *innpasses* i spørreskjemaets responsformat. Til slutt kan det avgitte svaret bli *endret* som en funksjon av ulike motiver som kan oppstå i undersøkelsessituasjonen.

I dette avsnittet skal vi se nærmere på denne prosessen, og fokusere på forhold ved de kognitive oppgavene i tilknytning til de metodologiske valgene som er tema for avhandlingen. Sekvensene av kognitive oppgaver ved svargivning på vurderingss spørsmål i et spørreskjema er illustrert i en modell for respondentens informasjonsbehandling i en spørreundersøkelsessituasjon i figur 3.6 på neste side.

⁴⁷Mål på de holdninger et individ har til et produkt etter at produktet er kjøpt og brukt.

FIGUR 3.6
Modell for informasjonsbehandling i en spørreundersøkelsessituasjon



Kilde: Strack, et al. (1987)

3.2.1 Tolkning av spørsmålet

Det første trinnet i en respondents informasjonsbehandleingsprosess ved svargivning i en spørreundersøkelse, er å forstå hva som er meningsinnholdet i et bestemt spørsmål. Dette innebærer for det første at respondenten må *forstå* hva ordene som er benyttet i spørsmålet faktisk betyr. For det andre må ordbruken i spørsmålet kunne relateres til den aktuelle sammenhengen spørsmålet er knyttet til (Strack og Martin, 1987).

Forskning har vist at det er to typer prosesser som kan påvirke tolking av informasjon (Strack og Martin, 1987; Tourangeau og Rasinski, 1988). Den ene typen er passive og automatiske prosesser som oppstår utenfor respondentens kontroll, f.eks. når en holdning eller vurdering er lett tilgjengelig. Ved andre tilfeller kan det være vanskelig for respondentene å lokalisere de aktuelle vurderingene. Den andre typen prosesser innebærer derfor aktiv informasjonsbehandling fra respondentens side, og kan sammenlignes med problemløsning. Senere i dette kapittelet skal vi se hvordan instruksjonene som ligger i objektforankrete- og subjektforankrete vurderingsspmårl kan tolkes ulikt, og dermed stille en respondent overfor en situasjon der en av disse prosessene aktiveres.

Både automatiske og passive vurderingsprosesser kan fremkalles ved å svare på et tidligere spørsmål i spørreskjemaet. Svar på et forutgående spørsmål i spørreskjemaet kan på denne måten ha en fremkallingseffekt (*priming*)⁴⁸, fordi eksponering for et tema i et forutgående spørsmål kan øke sannsynligheten for at det samme, eller et lignende tema aktiveres i hukommelsen ved respondentens svargivning på et etterfølgende spørsmål (Bodenhausen og Wyer, 1987; Strack og Martin, 1987; Tourangeau og Rasinski, 1988). Eksponering overfor kontekst kan også føre til økt tilgjengelighet til begreper, og dermed bidra til tolkning av eventuelle tvetydige spørsmål senere i spørreskjemaet (Strack og Martin, 1987).

Tre ulike informasjonskilder kan være til hjelp for en respondents tolking av et spesifikt spørsmål (Clark og Schober, 1991; Strack og Martin, 1987). For det første kan selve ordlyden og innholdet i spørsmålet gi tilstrekkelig informasjon til at respondenten kan benytte sine generelle kunnskaper for å tolke meningsinnholdet i spørsmålet. For det andre kan responsskalaen som benyttes i forbindelse med et spørsmål bidra til å forenkle tolkning av spørsmålet. For det tredje kan forutgående spørsmål være til hjelp for å forstå et bestemt spørsmål som kommer senere i spørreskjemaet.

3.2.2 Dannelse av oppfatning

Etter at respondentene har tolket og forstått et spørsmål, slik dette er formulert og fremstilt i et spørreskjema, må de gjøre seg opp en oppfatning om det som det blir spurt etter. Ofte blir en slik oppfatning hentet frem fra hukommelsen. Ulike forhold kan forklare sannsynligheten for at dette skjer. For det første kan spørsmål som oppfattes

⁴⁸*Priming* av informasjon innebærer at informasjonen "fremkalles fra minnet og gjøres mer tilgjengelig".

som viktige være lettere å danne en oppfatning om enn spørsmål som oppfattes som mindre viktige. For det andre kan det være lettere for respondentene å danne seg en oppfatning om spørsmål vedrørende forhold som skal resultere i handling. Endelig kan en oppfatning lettere hentes frem fra hukommelsen dersom respondenten tidligere er blitt spurt om et tilsvarende emne (Strack og Martin, 1987).

I mange tilfeller har respondenten ikke tidligere vurdert et bestemt emne slik det er fremstilt gjennom spørsmål i et spørreskjema. I slike tilfeller må vurderinger eller oppfatninger dannes samtidig med at intervjuet foregår. Hvor lett det er å fremkalle informasjon vil virke bestemmende for hvilken informasjon som blir benyttet for å foreta vurderingene.

De viktigste faktorene som har vist seg å ha betydning for tilgjengelighet til informasjon er *frekvens* og *nærhet i tid* til tidligere aktivering (Bodenhausen og Wyer, 1987; Strack og Martin, 1987; Tourangeau og Rasinski, 1988). Hvor ofte en vurdering har blitt foretatt, i tillegg til hvor lenge det er siden vurderingen sist ble foretatt, får altså betydning for hvor lett det vil være for en respondent å danne seg en oppfatning i en undersøkelsessituasjon.

Fremtredende (*salient*) eller fremkallt (*primed*) informasjon (Fiske og Taylor, 1991; Strack og Martin, 1987) kan føre til at vurderinger påvirkes både i positiv og negativ retning. Forutgående spørsmål i et spørreskjema betraktes ofte som den viktigste faktor som bidrar til å aktivere informasjon som er relevant ved besvarelse av senere spørsmål.

I avhandlingen forkuserer vi på problemstillinger knyttet til plassering av henholdsvis viktighetsspørsmål og vurderingss spørsmål i spørreskjemaet. Vi antar på den ene siden at viktighetsspørsmålene, dersom disse plasseres før spesifikke vurderingss spørsmål, aktiverer informasjon som kan påvirke svargivningen på de etterfølgende vurderingss spørsmålene. På den andre siden kan svargivning på vurderingss spørsmålene, dersom disse plasseres før viktighetsspørsmålene, påvirke etterfølgende rapportering av viktighet knyttet til egenskapene som er vurdert. Oppfattet diagnoseverdi (Feldman og Lynch, 1988) ved spørsmålene som stilles først i spørreskjemaet, enten dette er vurderingss spørsmål eller viktighetsspørsmål, vil være bestemmede for i hvilken grad svargivning på de etterfølgende spørsmålene påvirkes.

De spesifikke problemstillingene i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingss spørsmål i spørreskjemaet vil senere bli drøftet i relasjon til litteratur om konteksteffekter og rekkefølgeeffekter i spørreundersøkelser. Her vil vi kun gi en generell oversikt som viser hvordan tidligere spørsmål, og respondentenes

svar på disse, kan bidra til å fremkalle (*prime*) forskjellige typer informasjon (Strack og Martin, 1987).

For det første kan slik fremkalling av informasjon føre til at et annet eksemplar av kategorien det er spørsmål om aktiveres i respondentens hukommelse. Vurderinger kan i slike tilfeller bli basert på-, og forankret i vurderinger av dette andre eksemplaret. For det andre kan fremkalling av en bestemt type informasjon føre til at respondenten aktiverer spesielle egenskaper eller kjennetegn ved objektet det spørres om. For det tredje kan fremkallingen etablere en normativ standard som påvirker etterfølgende vurderinger. For det fjerde kan fremkalling av en bestemt type informasjon aktivere tidligere atferd og dermed påvirke respons. Til slutt kan fremkalling av en bestemt type informasjon bringe respondenten i en spesiell sinnstillstand som kan påvirke aktivering av informasjon og dermed respons (se f.eks. Heide (1990)).

3.2.3 Utforming av respons

I de fleste tilfeller må respondentens vurderinger i forbindelse med spørsmål i et spørreskjema rapporteres i et format som er gitt av forskeren. Respondenten tvinges her til å benytte de svarkategorier som er gitt i spørreskjemaet, og kan ikke bruke de kategorier som kommer frem spontant, og som kanskje virker mer naturlige. Respons på en svarskala kan også influeres av respondentens forsøk på å tilpasse seg det han tror vil representere en gjennomsnittsrespons (Schwarz og Hippler, 1987). Slike forankringseffekter opptrer særlig i forbindelse med målinger av holdninger og vurderinger.

Forankringseffektene kan påvirkes av responsskalaen som benyttes. F.eks. antar vi at "bra-dårlig" og "tilfreds-utilfreds" benevnelse knyttet til endepunktene på svarskalaen vil medføre en forankring på vurderingene som er forskjellig. Ved måling av kundetilfredshet og produktvurderinger har man i tillegg funnet at svarfordelingene vanligvis blir svært skjeve i positiv retning (Aiello, Czepiel, og Rosenberg, 1977; Peterson og Wilson, 1992; Westbrook, 1980). Dette kan skyldes at respondentene benytter et forankringspunkt som ligger nær det mest positive svaralternativet. Lengde på en måleskala, representert ved antall svaralternativer eller skalaverdier, kan også bidra til at respondentene benytter forankringspunkter nær ekstremverdiene⁴⁹ (Heide, 1990; Strack og Martin, 1987).

⁴⁹Morten Heide eksperimenterte med ulik lengde på svarskalaen i sin doktoravhandling og valgte å benytte en 11 punktskala (-5 - 5) for å redusere *ceiling effekter*.

3.2.4 Redigering av respons

Respons på spørsmål i en spørreundersøkelse kan ofte betraktes som del av en sosial interaksjon mellom intervjuer og respondent (Tourangeau og Rasinski, 1988). Sosiale normer kan derfor påvirke respondentens svargivning i intervjusituasjonen. Effekter av sosial påvirkning antas å være mindre dersom svargivningen oppfattes som å være anonym. En lang rekke eksperimenter har vært gjennomført for å studere effekter av såkalt "sosialt ønskelige svaralternativer", se f.eks. (Hippler, Schwarz, og Sudman, 1987; Schuman og Presser, 1981; Sudman og Bradburn, 1974). Resultater fra disse undersøkelsene har gitt indikasjoner på at respondenters rapportering av tilfredshet vil være skjevt fordelt mot den positive enden av svarskaalen. Dette forklares blant annet ved at de som er tilfreds i større grad vil svare på spørsmålene, og at skjevheten derfor kan skyldes utvalgsfeil.

Med utgangspunkt i vår gjennomgang av hvordan individer foretar produktvurderinger, og kognitive aspekter knyttet til respondenters svargivning ved måling av vurderingene i en spørreundersøkelse, vil vi nå se nærmere på de konkrete problemstillingene i tilknytning til ordlyd i spørsmålsstilling, benevnelse på svarskala og spørsmålsrekkefølge som vi fokuserer på i avhandlingen.

3.3 Objektforankrete- og subjektforankrete vurderinger

I kapittel 1 gjorde vi et skille mellom objektforankrete- ("bra-dårlig"), og subjektforankrete ("tilfreds-tilfreds") instruksjoner og benevnelser på måleskalaene i tilknytning til spørsmål vedrørende kunders produktvurderinger. Dette skillet baserte vi blant annet på skaleringslitteraturens beskrivelse av stimulusentrert- og subjektentrert skalering (Cox, 1982; Guilford, 1954; Hippler, Schwarz, og Sudman, 1987; Torgerson, 1958).

Formålet med dette avsnittet er todelt. Først vil vi forsøke å belyse: 1) Hvordan objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer er forskjellige, og deretter 2) Hvilke effekter bruk av objektforankrete-, vs. subjektforankrete instruksjoner og -vurderingsskalaer kan gi opphav til ved måling av kundenes produktvurderinger.

3.3.1 Hvordan er objektforankrete- og subjektforankrete instruksjoner og vurderingsskalaer forskjellige?

Med utgangspunkt i skaleringslitteraturen benevner vi, som tidligere nevnt “bra-dårlig” skalaen som *objektforankret*. Med en slik tilnærming betrakter vi respondenten som en “dommer” som blir bedt om å vurdere produktet. “Tilfreds-utilfreds” skalaen kaller vi *subjektforankret*. Når vi spør kunden om han/hun er tilfreds eller utilfreds forventer vi en rapportering av en indre følelsesmessig tilstand. De ulike instruksjonene som ligger i subjekt- og objektforankrete skalaer antar vi påfører respondenten kognitive oppgaver som er forskjellige og som derfor resulterer i forskjeller i svargivning.

I skaleringslitteraturen påpekes det at de to skaleringsmåtene er relatert til forskjellige forskningsformål, og at dette også krever ulike skaleringsteknikker. Ved bruk av en stimulussenteret tilnærming er formålet å skalere objekter, mens respondentene betraktes som replikasjoner. Subjektsentrert skalering impliserer, på den andre siden en skalering av individer. Forskjeller i respons attribueres til individuelle forskjeller og objektene betraktes som replikasjoner. I enkelte tilfeller er vi interessert i å tilskrive forskjeller i respons både til objekter og subjekter. I skaleringslitteraturen kalles dette responstilnærming.

Skaleringslitteraturen fokuserer på hvilke estimeringsteknikker og eksperimentelle design som er passende for subjektsentrert-, stimulussentrert- og responstilnærming. Dette er ikke problemområder som blir tatt opp i denne avhandlingen. Formålet er heller å belyse hvilke forskjeller som oppstår når 1) subjekter instrueres til å være “dommere” som skal vurdere objekter (“bra-dårlig” skala), og når 2) subjektene instrueres til å rapportere hvilke indre følelsesmessige tilstander objektene gir opphav til (“tilfreds-utilfreds” skala). For å gjøre dette skillet eksplisitt benevner vi altså “bra-dårlig” skalaen *objektforankret* og “tilfreds-utilfreds” skalaen *subjektforankret*.

Selv om valg av den ene skalatypen fremfor den andre kan ha teoretiske begrunnelser, finner vi i litteraturen ingen teoretiske rasjonaler for å benytte subjektforankret instruksjon og -vurderingsskala fremfor objektforankret instruksjon og vurderingsskala. Subjektforankret skala virker som et naturlig valg i forbindelse med forskning i tilknytning til kundetilfredshet. Likevel ser vi at objektforankrete skalaer og subjektforankrete skalaer brukes om hverandre nettopp ved måling av kundetilfredshet, se f.eks. (Cadotte, Woodruff, og Jenkins, 1987; Swan og Trawick, 1981).

I tilfeller der det ikke finnes teoretiske rasjonaler for å velge den ene skalaen fremfor den andre, kan likevel en objektforankret tilnærming i form av “bra-dårlig” vurderinger

være å foretrekke. Ved etterkjøpsvurderinger vil f.eks. vurderinger foretatt av de som har brukt et produkt være direkte sammenlignbare med vurderinger foretatt av de som ikke har brukt produktet. Det er mulig å be både brukere og ikke-brukere av et produkt å vurdere hvor godt eller dårlig produktet er. Dette vil ikke være tilfelle for “tilfreds-tilfreds” skalaen. For denne vil det bare være meningsfylt å spørre nåværende eller tidligere brukere. Bruk av “tilfreds-tilfreds” instruksjon og -vurderingsskala innebærer altså at man må spørre kunder som har brukt produktet.

3.3.2 Hvilke effekter kan bruk av de to alternative respons-skalaene gi opphav til?

Forskjellen mellom objektforankret- og subjektforankret instruksjon og -vurderingsskala kan i utgangspunktet virke uvesentlig, og begge tilnærmingene har som nevnt blitt brukt som alternativ ved måling av produktvurderinger. Objektforankrete vurderingsskalaer har blitt brukt for å måle førkjøpsvurderinger, mens både subjektforankrete- og objektforankrete vurderingsskalaer har blitt benyttet for måling av produktvurderinger etter kjøp og bruk av produktet (Oliver, 1980; Tse og Wilton, 1988). Det finnes få eksempler på systematiske forsøk på å sammenligne objektforankrete- (“bra-dårlig”) og subjektforankrete (“tilfreds-tilfreds”) instruksjoner og -vurderingsskalaer basert på teoretiske forskjeller og responsmønster. Det finnes imidlertid empirisk grunnlag (Abelson et al., 1982; Breckler og Wiggins, 1989) for å påstå at subjektforankrete- og objektforankrete instruksjoner og -vurderingsskalaer vil lede til viktige forskjeller som gjør at de bør behandles forskjellig. Forskjellene fremkommer både for univariate responsmønstre og i forhold til sammenhenger mellom variabler.

3.3.2.1 Implikasjoner for prediktiv validitet

Breckler & Wiggins (1989) rapporterer at instruksjoner som vektlegger egenskaper ved et objekt (f.eks. “å gi blod er godt-dårlig”) gir svar som er forskjellige fra de svar som fremkommer når respondentene blir bedt om å rapportere indre følelsesmessige tilstander (f.eks. “å gi blod gjør at jeg føler meg bra-dårlig”). I en studie av Abelson & al. (1982) ble respondentene instruert til å tilskrive personlighetstrekk til politikere (*trait scales*), samtidig som de ble bedt om å rapportere de følelsene som politikerne gav opphav til (*affect scales*). Resultatene fra denne studien viste at instruksjonene i skalaene var forskjellige på flere måter.

For det første var reliabilitet, representert ved Cronbachs alpha, for positivt ladete spørsmål gjennomgående høyere for personlighetstrekkskalaer (*trait scales*) enn for skaler der respondentene ble instruert til å rapportere følelser (*affect scales*). Andre analyser av konsistens med hensyn til respondentenes svargivning bidro også til at man konkluderte med at affektive responser var mer motstridende enn personlighetsvurderinger. Det ble også funnet at affektive skalaer predikerte totalvurdering, representert ved intensjon og preferanse mål, bedre enn personlighetstrekkskalaer. Breckler & Wiggins (1989) studie gir dermed indikasjoner på at subjektforankrete vurderingsskalaer (f.eks. tilfredshetskalaer) ikke bare har responsmønstre som er forskjellige fra objektforankrete vurderingsskalaer. Subjektforankrete vurderingsskalaer gir også bedre prediksjoner på totalvurderinger og atferdsmål enn objektforankrete vurderingsskalaer. Studien gir imidlertid ingen teoretiske forklaringer på funnene.

En forklaring på hvorfor man skal kunne forvente at subjektforankrete instruksjoner og -vurderingsskaler skal gi bedre prediksjoner på totalvurdering og atferdsmål enn objektforankrete instruksjoner og -vurderingsskalaer, er at objektforankrete vurderingsskalaer kan måle tilstander ved vurderinger, eller vurderingsprosesser som er forskjellige fra de som måles med subjektforankrete vurderingsskalaer. I den grad subjektforankrete vurderingsskalaer måler vurderinger som i større grad er basert på *direkte erfaring* (Fazio, 1989; Fazio og Zanna, 1981; Rust, Zahorik, og Keiningham, 1994), er mer *fremtredende (salient)* (Fiske og Taylor, 1991), mer *tilgjengelige* ("accessible") (Fazio, 1989) og representerer *nedlagt engasjement (vested interest)* (Raden, 1985) kan man forvente at disse gir bedre prediksjoner på atferd og intensjoner om atferd.

Med unntak av arbeidene til Abelson & al. (1982) og Breckler og Wiggins (1989) har studiene vi har referert ovenfor foretatt sammenligninger av de samme vurderingsmålene på tvers av subjekter og situasjoner med ulike verdier på de modererende faktorene (f.eks. tilgjengelighet). Vi antar imidlertid at de to vurderingsskalaene måler ulike vurderingstilstander karakterisert ved forskjeller som f.eks. fremtredenhets og tilgjengelighet. Alternativt kan vi anta at instruksjonene som blir gitt ved bruk av de to vurderingsskalaene resulterer i vurderingsprosesser som gir opphav til slike forskjeller.

Med bakgrunn i litteraturgjennomgangen antar vi at subjektforankrete vurderinger er lettere *tilgjengelige*, og at de derfor er sterkere relatert til totalvurderinger og atferdsintensjoner enn objektforankrete vurderinger. I tillegg antar vi at subjektforankrete vurderinger er sterkere knyttet til individet og individets verdisystem. Dette vil bidra til de samme effektene i forhold til totalvurderinger og atferdsintensjoner.

For å forstå hvorfor subjektforankrete (“tilfreds-tilfreds”) vurderinger kan være mer tilgjengelige enn objektforankrete (“bra-dårlig”) vurderinger kan følgende forklaring være nyttig⁵⁰. Ved førkjøpsvurderinger må kunden dra slutninger på bakgrunn av holdepunkter, og hvilke konsekvenser disse kan gi opphav til (Steenkamp, 1989). Kjøperen har ikke tilgang til indre følelsesmessige vurderinger før kjøp og bruk av et produkt. I denne situasjonen må vurderinger av produktet baseres på informasjon eller inspeksjon. Ved etterkjøpsvurderinger kan vurderingene baseres på den erfarte eller opplevde ytelsen til produktet, og den tilfredshet som denne resulterer i (Troye, 1990).

Forskjeller mellom førkjøpsvurderinger og etterkjøpsvurderinger kan også relateres til skillet mellom *search qualities* og *experience qualities* (Nelson, 1974). *Search qualities* er her definert som egenskaper ved et produkt som en konsument kan vurdere (gjennom inspeksjon) før produktet kjøpes, mens *experience qualities* defineres som kvalitet som bestemmes etter kjøp og bruk av produktet.

Når respondenter blir bedt om å vurdere et objekt basert på episodisk erfaring, blir de bedt om å foreta slutninger som går ut over denne erfaringen (Abelson et al., 1982). Å gi uttrykk for sin egen tilfredshet medfører vurderinger av typen: “det jeg føler nå er godt for meg”, mens “bra-dårlig” vurderinger kan tolkes til å inneholde vurderinger av om produktet er godt slik dette oppleves over tid, og for en serie av kunder og bruks-situasjoner. Å gi uttrykk for hvorvidt noe er godt eller dårlig krever en generalisering fra én eller flere situasjoner der produktet er brukt. Dette krever slutninger fra noe sikkert og situasjonsspesifikt (“jeg er tilfreds”) til noe mindre sikkert og mer generelt (“det må være bra fordi jeg er tilfreds”). Et utsagn som “tilfredsheten jeg fikk fra produktet” trenger ikke være relatert til hvor godt produktet er, eller er antatt å være.

Det kan med bakgrunn i dette argumenteres for at objektforankrete vurderinger er indirekte og generelle, mens subjektforankrete tilfredshetsmål basert på spesifikke erfaringer er mer direkte og spesielle. Det vil også være rimelig å anta at tilfredshet, som en indre følelsesmessig tilstand relatert til en spesifikk konkret erfaring (Rust, Zahorik, og Keiningham, 1994), vil være mer direkte og lettere tilgjengelig enn mer generelle vurderinger av objektet. Respondenter som svarer på spørsmål i et spørreskjema etter at de har kjøpt og brukt produktet vil stå overfor en lettere oppgave når de blir bedt om å gi uttrykk for en indre følelsesmessig tilstand (f.eks. tilfredshet). Dersom dette er riktig vil tilfredshetsbenevnelse på svarskaalen gjøre vurderingene mer tilgjengelige enn “bra-dårlig” skalaer, og derfor sterkere relatert til totalvurderinger og intensjoner og atferd.

⁵⁰Se også drøfting av *prospektiv- og retrospektiv inferens* i kapittel 3.1.

Subjektforankrete vurderingsskalaer kan være mer selvrelatert og selvrelevant (Higgins og Bargh, 1987) enn objektforankrete vurderingsskalaer. I *self-awareness* teori (Duval & Wicklund, 1972)⁵¹ skiller man mellom den oppmerksomhet individet retter mot seg selv, og oppmerksomheten som rettes mot eksterne faktorer. Oppmerksomhet rettet mot seg selv innebærer aktivering av kunnskap som er kognitivt mer tilgjengelig (Schwarz, 1990), og vil derfor lede til sterkere konsistens mellom holdninger og atferd (Zanna, Olson, og Fazio, 1980). Indirekte støtte for påstanden om at bruk av subjektforankrete instruksjoner og -vurderingsskalaer kan føre til målinger som gir sterkere konsistens i sammenhenger mellom vurdering og atferd og mellom vurdering og intensjon om atferd kan derfor også finnes i *self-awareness* teori .

I litteraturgjennomgangen som er foretatt i forbindelse med denne avhandlingen finner vi ikke empirisk belegg for at ulike instruksjoner gitt i responsskalaen i et spørreskjema kan lede respondenten til å fokusere oppmerksomhet mot seg selv. Det virker likevel rimelig å anta at subjektforankrete tilfredshetsskalaer kan føre til at respondenten retter oppmerksomheten mot seg selv, mens objektforankrete, "bra-dårlig" skalaer kan lede oppmerksomheten mot eksterne faktorer. Bruk av tilfredshetsskalaer gir en eksplisitt instruksjon til respondenten om å rette oppmerksomheten mot indre tilstander (f.eks. "er jeg tilfreds?"). Instruksjonen i en "bra-dårlig" skala leder respondentens oppmerksomhet mot eksterne faktorer utenfor respondenten (f.eks. "er det bra eller dårlig?"). *Self-awareness* teori kan på denne måten gi støtte for påstanden om at tilfredshetsmål gir sterkere konsistens mellom vurderinger, totalvurderinger og atferd enn mindre selvfokuserte "bra-dårlig" vurderinger.

Forskjeller mellom objektforankrete- og subjektforankrete instruksjoner og vurderingsskalaer kan også sees i lys av distinksjonen mellom holdning til et objekt og holdning til handling overfor objektet (Fishbein og Ajzen, 1975). Måling av holdning til handling overfor et objekt er funnet å være sterkere relatert til atferdsintensjoner enn holdninger overfor objektet (Fishbein, 1972; Fishbein og Ajzen, 1975), fordi holdning til handling overfor et objekt er mer spesifikk. Objektforankrete ("bra-dårlig") vurderinger kan til en viss grad representere en objektiv, generell vurdering av et produkt, mens subjektforankrete tilfredshetsvurderinger impliserer en mer spesifikk vurdering av konsekvenser av å ha kjøpt og brukt produktet (handling overfor produktet). Vi antar derfor at subjektforankrete ("tilfreds-utilfreds") vurderinger er sterkere relatert til atferdsintensjoner enn objektforankrete ("bra-dårlig") vurderinger.

⁵¹ Som referert av Wilson (1990).

Tidligere forskning har vist at økt produktkunnskap fører til sterkere sammenheng mellom vurdering av egenskaper ved et produkt og totalvurdering av produktet. Dette kan skyldes at vurderinger som er basert på erfaring og produktkunnskap vil være mer *tilgjengelige* (Fazio, 1989; Fazio og Zanna, 1981). Forskning har også vist at respondenter med høy erfaring og produktkunnskap er mindre sensitive til forskningskonteksten enn respondenter med lav erfaring (Lynch, Chakravarti, og Mitra, 1991).

Vi antar at erfaring og produktkunnskap også kan sees i sammenheng med vurderinger basert på objektforankret- og subjektforankret instruksjon og måleskala. Gjennom erfaring med bruk av et produkt vil respondenten oftere ha foretatt både "tilfreds-tilfreds" og "bra-dårlig" slutninger. Respondenter med høy erfaring med bruk av et produkt vil derfor kunne ha et bedre fundament for sine objektforankrete "bra-dårlig" vurderinger enn respondenter med lav erfaring. Det kan med bakgrunn i dette antas at respondenter med høy erfaring og produktkunnskap vil være i bedre stand til å foreta objektforankrete vurderinger. I den grad "bra-dårlig" vurderinger er mer tilgjengelige, og lettere å foreta for respondenter med høy erfaring enn for respondenter med lav erfaring, vil også sammenhengen mellom spesifikke vurderinger og atferdsintensjoner være mindre sensitiv for valg av benevnelse på responsskalaen for respondenter med høy erfaring enn for respondenter med lav erfaring.

3.3.2.2 Implikasjoner for konsistens i vurderinger

Objektforankrete ("bra-dårlig") instruksjoner og -vurderingsskalaer kan oppfattes som å gi respondentene instruksjon om å basere sine vurderinger på objektive eller intersubjektive standarder (Troye, 1990) i den grad slike standarder eksisterer, eller er kjent for respondenten. Subjektforankrete ("tilfreds-tilfreds") instruksjoner og -vurderingsskalaer, antar vi, er nærmere knyttet til respondentens verdier og selvopfatning. Dette kan føre til at forskjeller i preferanser og smak, i mindre grad vil gi seg utslag i variasjoner i svargivning ved bruk av objektforankrete vurderingsskalaer enn ved bruk av subjektforankrete vurderingsskalaer. Vi forventer med andre ord høyere grad av konsistens i objektforankrete vurderinger enn i subjektforankrete vurderinger. En slik påstand bygger imidlertid på en antakelse om at objektive eller intersubjektive standarder er kjent for respondenten.

Dersom objektive eller intersubjektive standarder ikke eksisterer, kan instruksjonen som ligger i en objektforankret vurderingsskala medføre usikkerhet i vurderingene som kan redusere konsistensen. For enkelte produkter (f.eks. tjenester) er det sannsynlig at

kundene anvender de samme standardene ved sine produktvurderinger, samtidig som de kan være uvitende om andre kunders vurderingsstandarder. I slike situasjoner kan konsistensen mellom respondentene være lavere for objektforankrete vurderinger enn for subjektforankrete vurderinger.

I kundetilfredshetslitteraturen er det vanlig å betrakte tilfredshet som et resultat av opplevd produktytelse i forhold til en standard eller forventning, se f.eks. (Tse og Wilton, 1988; Yi, 1990). Vi antar at slike prosesser blir aktivert når tilfredshetsmål benyttes i et spørreskjema, mens dette ikke vil være tilfellet når “bra-dårlig” mål benyttes. Dette taler også for at bruk av objektforankret (“bra-dårlig”) instruksjon og -vurderingsskala vil føre til mer konsistente vurderinger enn bruk av subjektforankrete (“tilfreds-utilfreds”) instruksjoner og -vurderingsskalaer.

3.3.2.3 Implikasjoner for nivå på vurderinger

Subjektforankrete (“tilfreds-utilfreds”) instruksjoner og -vurderingsskalaer og objektforankrete (“bra-dårlig”) instruksjoner og -vurderingsskalaer blir, som tidligere nevnt, ofte benyttet om hverandre ved måling av kunders produktvurderinger. Dette kan være et resultat av at man ikke forventer at gjennomsnittsvurderinger ved bruk av de to skalaene skal være systematisk forskjellige. Imidlertid vil det være situasjoner der ytelsen til det vurderte produktet enten overstiger, eller er lavere enn forventningene kundene har. Vi antar at dette kan få implikasjoner for gjennomsnittsvurderinger ved bruk av de to instruksjonene og vurderingsskalaene.

I følge “confirmation-disconfirmation” perspektivet er det forventet at positive avvik mellom kundenes forventninger eller normer og produktets ytelse fører til tilfredshet, mens negative avvik fører til utilfredshet (Oliver, 1981; Tse og Wilton, 1988; Yi, 1990). Assimilasjon-kontrastteori⁵² (Sherif og Hovland, 1961)⁵³ har vært et viktig utgangspunkt for å forklare disse avvikene i litteratur om kundetilfredshet (Anderson, 1973; Oliver og DeSarbo, 1988).

Sammenholdt med elementer fra dissonansteori (Festinger, 1957), innebærer perspektivene fra assimilasjon-kontrastteori at individer vil være motstridende til å erkjenne avvik mellom forventninger og opplevelse. Vurderinger vil med et slikt

⁵²For en grundigere gjennomgang av assimilasjon- og kontrast effekter, se drøftingen i tilknytning til effekter av rekkefølge på viktighetsspørsmål og vurderingsskjemaer i kapittel 3.4.

⁵³Som referert av Yi (1990).

utgangspunkt assimileres mot forventningene når avvikene mellom den opplevde produktytelsen og forventningene er små. Det motsatte av slike assimilasjonseffekter oppstår når et individ forstørrer avvikene mellom opplevd ytelse og forventninger når disse er store. Dette kalles kontrasteffekter. I kundetilfredshetslitteraturen (Anderson, 1973) benyttes teori om kontrasteffekter for å beskrive både hvordan produktytelse som er lavere enn forventet vil bli vurdert dårligere enn det den i virkeligheten er, og hvordan produktytelse som er høyere enn forventet vurderes høyere (Oliver, 1988). Kontrasteffekter hevdes på denne måten å kreve større grad av involvering fra respondentens side (Yi, 1990).

Vi har tidligere argumentert for at subjektforankrete instruksjoner og -vurderingsskalaer reflekterer indre følelsesmessige tilstander, mens objektforankrete instruksjoner og -vurderingsskalaer representerer respondentens mer objektive vurdering av et objekt. Det kan derfor være en rimelig antakelse at kontrasteffekter spesielt vil oppstå i situasjoner der det benyttes subjektforankrete skalaer.

Ved bruk av objektforankrete skalaer er det forventet at respondentene opptrer som "dommere", og vurderer produkter uavhengig av forventninger og personlige normer når de skal gi uttrykk for den tilfredsstillende produktet gir opphav til (Tse og Wilton, 1988). "Bra-dårlig" instruksjon og -vurderingsskala kan derfor føre til at respondentene føler at de ikke kan bruke egne normer og forventninger som basis eller sammenligningsstandard for vurderinger.

Antakelsen om at objektforankrete skalaer gir vurderinger som er uavhengige av respondentens egne standarder, normer og forventninger er belyst av Tse & Wilton (1988). I deres studie ble både forventet- og opplevd ytelse målt ved en "very poor-very good" skala, mens tilfredshet ble målt ved "very dissatisfied-very satisfied" skala.

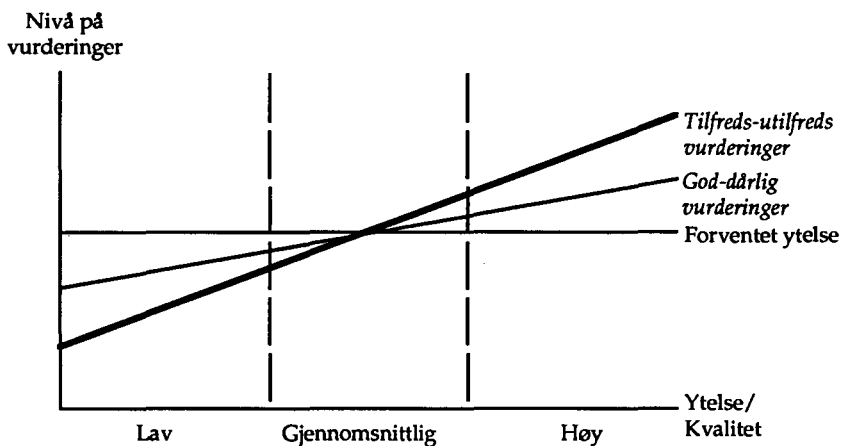
Dersom kundenes forventninger er konstante på tvers av produktalternativer innenfor en bestemt produktgruppe, er det sannsynlig at ytelsen for høykvalitetsprodukter vil overstige kundenes forventninger. Forventninger som er konstante på tvers av produktalternativer kan f.eks. oppstå i situasjoner der informasjonen kundene får gjelder alle produktene i produktgruppen. Dette kan være tilfellet i situasjoner med kjedesamarbeid bestående av felles profilering og markedskommunikasjon. Et annet eksempel på en situasjon der forventninger kan være konstante på tvers av produktalternativer, er når en bedrift lanserer ulike produktalternativer under samme familiermerke.

Vi har tidligere beskrevet hvordan både høy- og lav kvalitet vil føre til kontrasteffekter som resultat av at høykvalitetsprodukter vil ha en ytelse som overstiger kundenes

forventninger, mens lavkvalitetsprodukter vil ha en ytelse som ikke når opp til forventningene. Vi har også argumentert for at subjektforankrete instruksjoner og -vurderingsskalaer i større grad enn objektforankrete instruksjoner og -vurderingsskalaer involverer kunden i vurderingen.

Med dette som utgangspunkt kan vi forvente at nivået på “tilfreds-utilfreds” mål for høykvalitetsprodukter vil ligge over “bra-dårlig” mål. På tilsvarende måte kan vi forvente at nivået på tilfredshetsmål for lavkvalitetsprodukter vil ligge under “bra-dårlig” mål. For gjennomsnittsprodukter vil vi ikke forvente forskjeller mellom skalaene. Disse sammenhengene er illustrert i figur 3.7 nedenfor.

FIGUR 3.7
Forventede forskjeller mellom subjektforankret- (“tilfreds-utilfreds”) og stimulusforankret (“bra-dårlig”) instruksjon og -vurderingsskala



De samme resultatene kan forventes med utgangspunkt i forskning som har vist at “tilfreds-utilfreds” vurderinger tenderer mot å være svært skjeve i positiv retning (Peterson og Wilson, 1992). “Bra-dårlig” vurderinger kan være mer konservative enn “tilfreds-utilfreds” vurderinger. Å gi en generell vurdering av om et produkt er godt eller dårlig basert på spesifikke erfaringer, kan betraktes som en mer utfordrende oppgave som fører til at respondenten forsøker å gi en konservativ “bra-dårlig” vurdering. Dette vil føre til at stigningen på “bra-dårlig” kurven vil være slakkere enn for “tilfreds-utilfreds” kurven, og vi vil få et mønster som illustrert i figur 3.7.

3.3.2.4 Implikasjoner for sikkerhet i vurderinger

Ved bruk av subjektforankret instruksjon og -vurderingsskala instrueres respondentene til å gi uttrykk for en indre følelsesmessig tilstand som sier noe om hvor tilfreds de er med et objekt. Vi har tidligere argumentert for at indre følelsesmessige tilstander vil være mer tilgjengelige (*accessible*) (Fazio, 1989) for respondenten når han/hun skal foreta produktvurderinger. Vi har tidligere drøftet mulige sammenhenger mellom respondentens erfaring og produktkunnskap, tilgjengelighet til vurderinger og sikkerhet i vurderingene (*confidence*). Sikkerhet i vurderinger har i hovedsak blitt betraktet som en faktor som modererer sammenhengen mellom vurderinger av et objekt og holdninger og intensjoner om atferd overfor objektet (Sample og Warland, 1973). Sikkerhet i vurderinger har også vært funnet å være positivt relatert til atferdsintensjoner (Howard og Sheth, 1969).

I avhandlingen fokuserer vi ikke på sikkerhet i vurderinger som en modererende variabel. Vi argumenterer derimot for at respondentenes sikkerhet i vurdering av spesifikke produktegenskaper vil kunne påvirkes av instruksjonen og vurderingsskalaen som benyttes i spørreskjemaet. Mer spesifikt påstår vi at respondenter som foretar subjektforankrete vurderinger av et produkt vil være sikrere på at vurderingene er "riktige" enn respondenter som foretar objektforankrete vurderinger. Sikkerhet i vurderinger kan på denne måten betraktes som en effekt av valg mellom objektforankrete-, og subjektforankrete instruksjoner og -vurderingsskalaer.

3.3.2.5 Implikasjoner for relativ betydning av immaterielle- og materielle egenskaper ⁵⁴

I tjenestemarkedsføringslitteraturen blir det argumentert for at immaterielle samhandlingsegenskaper vil være vanskeligere for kundene å vurdere, enn andre egenskaper ved et produkt (Zeithaml, 1981). Disse vanskelighetene med å vurdere samhandlingsegenskaper antas å "tvinge" kundene til å legge mer vekt på de øvrige egenskapene ved vurdering av produktet.

I markedsforskning har man vært opptatt av kunders produktvurderinger, uttrykt både som sammenligningsvurderinger og som preferansevurderinger. Preferansevurderinger er generelt blitt betraktet som subjektive, individspesifikke og heterogene på tvers av

⁵⁴Relativ betydning er definert på følgende måte: Totalvurdering = f (vurdering av materielle egenskaper, vurdering av immaterielle egenskaper).

konsumenter, mens sammenligningsvurderinger betraktes som objektive og homogene (Lefkoff-Hagius og Mason, 1990). Med bakgrunn i at en objektforankret (“bra-dårlig”) instruksjon og -vurderingsskala kan oppfattes som en instruksjon til respondentene om å basere sine vurderinger på objektive, intersubjektive standarder, kan det argumenteres for at objektforankrete vurderinger kan relateres til sammenligningsvurderinger. På tilsvarende måte kan subjektforankrete (“tilfreds-tilfreds”) vurderinger relateres til preferansevurderinger, da disse vil være mer individspesifikke og subjektive, og gir uttrykk for individuelle indre følelsesmessige tilstander.

I en studie av Lefkoff-Hagius & Mason (1990) blir det foretatt en sammenligning av preferansevurderinger og sammenligningsvurderinger, der det fokuseres på forskjellene i betydning av materielle og immaterielle egenskaper ved et produkt ved bruk av de to måtene å foreta vurderinger på. Det argumenteres i studien for at konsumenter benytter hierarkiske informasjonsbehandlingsprosesser når de blir bedt om å foreta sammenligningsvurderinger (Johnson, 1988)⁵⁵. Konsumenten vurderer først konkrete materielle egenskaper ved produktet og fortsetter så med de mer abstrakte immaterielle egenskapene. Med utgangspunkt i dette hierarkiske informasjonsbehandlingsperspektivet hevdes det at når en konsument blir presentert et produkt med både immaterielle og materielle egenskaper, vil de materielle egenskapene vies mer oppmerksomhet, og disse blir derfor viktigere i sammenligningsvurderinger enn i preferansevurderinger. Lefkoff-Hagius & Mason (1990) finner støtte for disse påstandene i sin studie.

Vi har tidligere argumentert vi for at iboende holdepunkter (*intrinsic cues*) er relatert til materielle egenskaper, mens utvendige holdepunkter (*extrinsic cues*) er relatert til immaterielle egenskaper. Begrunnelsen for dette finner vi hos Cox (1967) som hevder at ethvert holdepunkt (*cue*) som bidrar til å gi informasjon om produktet kan beskrives ved to faktorer; prediktiv- og *confidence* verdi.

Prediktiv verdi reflekterer i hvilken grad konsumenten oppfatter eller antar at et generisk holdepunkt (f.eks. pris) er relatert til, eller gir indikasjoner om egenskaper ved produktet (f.eks. kvalitet). *Confidence* verdi beskriver i hvilken grad en konsument er sikker (*confident*) på sine evner til å oppfatte og kategorisere et bestemt holdepunkt. Olson (1977) og Szybillo og Jacoby (1974) hevder videre at iboende holdepunkter vil ha sterkere påvirkning på produktvurderingene enn utvendige holdepunkter. Dette forklares med at iboende holdepunkter har høyere prediktiv verdi enn utvendige holdepunkter.

⁵⁵Som referert av Lefkoff-Hagius & Mason (1990).

Basert på sammenligningene vi gjorde mellom preferansevurderinger og subjektforankrete vurderinger, og mellom sammenligningsvurderinger og objektforankrete vurderinger, kan det argumenteres for at materielle egenskaper vil være viktigere, og bety mer for respondentens totalvurdering og atferdsintensjoner når det benyttes objektforankret, "bra-dårlig", instruksjon og -vurderingsskala. Ved bruk av subjektforankret, "tilfreds-tilfreds", instruksjon og -vurderingsskala vil forskjellen i betydning mellom materielle og immaterielle egenskaper være mindre. De samme effektene antar vi vil fremkomme med bakgrunn i forskjellene mellom iboende og utvendige holdepunkter, der iboende holdepunkter som ble relatert til materielle egenskaper har høyere predikativ verdi enn utvendige holdepunkter.

3.4 Multiattributtmodeller og spørsmålsrekkefølge

Det at rekkefølgen på spørsmålene i et spørreskjema kan være en kilde til variasjoner i respondentenes svargivning er viet stor oppmerksomhet i litteraturen, se f.eks. Hippler, Schwarz, og Sudman (1987), Schuman og Presser (1981), Schwarz og Sudman (1991) og Sudman og Bradburn (1974)⁵⁶. Uttrykket "rekkefølgeeffekt" refererer til det fenomenet som oppstår når ulik rekkefølge på spørsmål (eller responsalternativer) påvirker respondentenes svargivning på en systematisk måte (Schuman og Presser, 1981). Et spørsmål, A, kan således oppnå forskjellig svargivning dersom det stilles før spørsmål B, sammenlignet med en situasjon der rekkefølgen på de to spørsmålene er motsatt. I litteraturen hevdes det at slike effekter av spørsmålsrekkefølge er betinget av at forskeren har kontroll over rekkefølgen spørsmålene blir presentert, slik dette er tilfellet i et personlig intervju og i et telefonintervju. Dersom dette ikke er tilfellet, som når respondentene selv fyller ut et spørreskjema, antar man at plassering av spørsmålene i spørreskjemaet vil ha mindre betydning som forklaring på variasjoner i svargivning (Ayidiya og McClendon, 1990; Strack, 1991). Disse forbeholdene om rekkefølgeeffekter er basert på en antakelse om at respondenter som svarer på spørsmål ved selv å fylle ut et spørreskjema, ikke nødvendigvis svarer på spørsmålene i den rekkefølgen disse er presentert i spørreskjemaet. I denne avhandlingen vil vi imidlertid studere rekkefølgeeffekter som vi antar også vil oppstå når respondenter svarer på spørsmål ved å fylle ut et spørreskjema.

⁵⁶Se også appendix A.3 som gir en oversikt over en rekke studier der spørsmålsrekkefølge har vært fokusert.

På mange måter kan gjennomføring av et intervju i en spørreundersøkelse sammenlignes med en konversasjon mellom to individer, intervjuer og respondent, der intervjuet består av spørsmål og svar knyttet til et bestemt tema. Selv om respondenten svarer på spørsmålene ved å fylle ut et spørreskjema kan denne sammenligningen gi nyttige perspektiver. Mange av de prinsippene som styrer vanlig og mindre strukturert konversasjon mellom to individer kan gjøres gjeldende også i intervjusituasjonen. To prinsipper har fått særlig betydning for å forstå konversasjon (Grice, 1975); *maxim of relevance*, som innebærer at innholdet i et spørsmål i en konversasjon refererer til innholdet i et forutgående spørsmål, og *maxim of quantity*, som innebærer et prinsipp om å være informativ, og å unngå å tilføre redundant informasjon til samtalen. Det første prinsippet kan føre til at svargivningen på et spørsmål påvirkes i en bestemt retning avhengig av implikasjonene av et forutgående spørsmål. Det andre prinsippet kan ha motsatt effekt, og føre til at respondenten unngår å benytte informasjonen som ligger i et spørsmål som grunnlag for å tolke et etterfølgende spørsmål. Dette kan være tilfelle der respondenten forsøker å være informativ og søker å unngå å gi overflødig informasjon.

Operasjonalisering av multiattributtmodeller, se f.eks. Fishbein og Ajzen (1975), Grünert (1989) og Lutz (1991), innebærer at forskeren stiller to sett av spørsmål, et vedrørende vektlegging av egenskaper ved et objekt, eller styrke på antakelser om objektet, og et vedrørende vurderinger av de samme egenskapene eller antakelsene. Problemstillinger knyttet til bruk av viktighetsspørsmål ved operasjonalisering av multiattributtmodeller i spørreskjema har vært viet stor oppmerksomhet i tidligere forskning. Man har spesielt vært opptatt av hvorvidt viktighet skal måles ved hjelp av respondentenes selvrappotering i spørreskjemaet, eller om man skal estimere de vurderte egenskapenes relative viktighet ved bruk av statistiske teknikker, se f.eks. Akaah og Korgaonkar (1983), Beckwith og Lehmann (1973), Jaccard, Brinberg, og Ackerman (1986), Mazis, Ahtola, og Klippel (1975) og Neslin (1981). I tillegg til viktighetsspørsmål og vurderingssspørsmål, inkluderes ofte totalvurderingssspørsmål og spørsmål som fanger opp respondentenes atferdsintensjoner i spørreskjemaet.

Formålet med studiene som er gjennomført i forbindelse med bruk av multiattributtmodeller har i hovedsak vært å utvikle måleprosedyrer for å måle respondentenes vurdering av spesifikke produktegenskaper på en måte som gir best mulige prediksjoner på totalvurderinger, atferdsintensjoner, atferd og preferanser. Resultatene er tvetydige, og det eksisterer fortsatt ulike oppfatninger av hvorvidt selvrappotert vektlegging av produktegenskaper bør inkluderes i spørreskjemaet eller ikke. Imidlertid blir selvrappotering av viktighet ofte benyttet i praksis, både i kvalitetsundersøkelser (Parasuraman, Zeithaml, og Berry, 1988) og i holdningsundersøkelser (Lutz, 1991).

Når man har bestemt seg for å inkludere spørsmål for å kartlegge viktighet ved de spesifikke egenskapene som man ber respondentene vurdere, i tillegg til spesifikke vurderingss spørsmål i spørreskjemaet, antar vi at variasjoner i plasseringen av disse spørsmålene kan gi seg utslag i respondentenes svargivning. I dette avsnittet søker vi først å belyse hvorfor ulik plassering av viktighetsspørsmål i forhold til spesifikke vurderingss spørsmål i spørreskjemaet kan gi seg utslag i forskjeller i respondentenes svargivning. Etter en drøfting av mulige årsaker til forskjellene vil vi se nærmere på de effektene som spørsmålsrekkefølgen kan gi opphav til.

3.4.1 Hvorfor kan variasjoner i plassering av vurderingss spørsmål og viktighetsspørsmål i et spørreskjema føre til forskjeller i respondentenes svargivning?

Problemstillinger knyttet til bruk av viktighetsspørsmål kan inndeles i to hovedkategorier. 1) Plassering av viktighetsspørsmål før spesifikke vurderingss spørsmål i spørreskjemaet, og 2) Plassering av viktighetsspørsmål etter spesifikke vurderingss spørsmål i spørreskjemaet. I dette avsnittet vil vi behandle disse samlet, da vi antar at de samme teoretiske perspektivene kan gjøres gjeldende i begge tilfellene.

Ved plassering av viktighetsspørsmålene før de spesifikke vurderingss spørsmålene blir respondentene først bedt om å gi uttrykk for hvor viktig det er at ulike egenskaper ved et produkt er tilstede. Etter at de har tillagt viktighet til produktegenskapene blir de så bedt om å foreta vurderinger av produktet på basis av de samme egenskapene. Når viktighetsspørsmålene plasseres etter de spesifikke vurderingss spørsmålene, blir respondentene først bedt om å vurdere en rekke produktegenskaper, for deretter å rapportere hvor mye de ulike egenskapene betyr for deres totalvurdering. Uavhengig av rekkefølgen på viktighetsspørsmål og spesifikke vurderingss spørsmål blir respondentene til slutt bedt om å foreta en helhetsvurdering, eller en totalvurdering, av produktet. I tillegg inkluderes ofte spørsmål om respondentenes atferdsintensjoner i spørreskjemaet.

Et viktig og relevant spørsmål i forbindelse med problemstillingene vi har reist ovenfor er hvorvidt spesifikke vurderinger, viktighet, totalvurderinger og atferdsintensjoner kan bli skapt av måleprosedyren dersom disse ikke allerede eksisterer i respondentens langtidshukommelse. Spørsmål i et spørreskjema antas å aktivere midlertidige kunnskapsstrukturer hos respondenten (Schwarz og Bless, 1992). Disse midlertidig aktiverte kunnskapsstrukturene kan ha stor påvirkning, både på vurderinger av objekter og på relatert atferd. Hvilke kunnskapsstrukturer som blir midlertidig aktivert vil være en

funksjon av: (1) holdepunkter i omgivelsene som retter respondentens oppmerksomhet mot spesifikke egenskaper ved vurderingsobjektet, (2) fremkalling av informasjon (*priming*) og (3) individuelle forskjeller, som f.eks. produktkunnskap (Feldman og Lynch, 1988). En viktig implikasjon av disse forholdene for gjennomføring av spørreundersøkelser der vi ønsker å måle holdninger, vurderinger, totalvurderinger, intensjoner og atferd, er at selve prosessen ved måling av disse variablene retter oppmerksomhet, både mot deler av respondentens eksisterende kunnskap og mot deler av vurderingsobjektet.

En tradisjonell oppfatning innenfor fagområder som sosialpsykologi, organisasjonspsykologi og konsumentpsykologi er at målinger av individers antakelser, holdninger og intensjoner kun reflekterer antakelser, holdninger og intensjoner som allerede er tilstede, og som er robuste i respondenters hukommelse. I tillegg antar man at atferd er under bevisst kontroll av atferdsintensjoner. Videre betraktes antakelser, holdninger, intensjoner og atferd som spontant aktivert og som del av et kausalt hierarki der antakelser integreres i totalvurderinger som påvirker intensjoner. Intensjonene antas deretter å påvirke faktisk atferd direkte. Implikasjonene av disse forholdene er at begreper som antakelser, holdninger, intensjoner og atferd behandles som kontekst-uavhengige (Feldman og Lynch, 1988). Relatert til problemstillingene vi har reist i tilknytning til måling av viktighet, vurderinger, totalvurderinger og atferdsintensjoner vil dette innebære at vi står overfor begreper som er eksisterende i respondentenes hukommelse, og som ikke påvirkes av måleprosedyren.

Et alternativt syn på disse tradisjonelle antakelsene fremsettes av Feldman og Lynch (1988). Utgangspunktet for det som her blir kalt *selvgenerert validitet*, er at kun et fåtall av de antakelser om objekter, holdninger og intensjoner som måles, enten i et laboratorieeksperiment eller i en surveyundersøkelse, allerede eksisterer og er robuste i respondentenes langtidshukommelse. I samsvar med et prinsipp om individers kognitive økonomi (Wyer og Srull, 1986), antar man at kognitive ressurser ikke benyttes for å utvikle vurderinger, antakelser eller intensjoner dersom det ikke eksisterer spesifikke motiver for dette. Videre hevdes det at kun en liten del av de responser som er lagret vil være tilgjengelige når disse måles f.eks. ved hjelp av et spørreskjema. Dette antar man imidlertid ikke vil hindre respondenter i å svare på spørsmålene i spørreskjemaet. Dersom det ønskete svaret ikke kan fremhentes direkte fra hukommelsen, kan et svar likevel lett bli gitt med utgangspunkt i andre tilgjengelige kilder. Til slutt kan både fremhentete og beregnete responser bli modifisert av undersøkelseskonteksten. Disse alternative perspektivene på måling av antakelser, holdninger, intensjoner og atferd representerer et viktig utgangspunkt for å belyse problemstillingene vi har tatt opp i tilknytning til inkludering og alternativ plassering av

viktighetsspørsmål i et spørreskjema. Perspektivene impliserer at selve målingen kan påvirke respondentenes rapportering av viktighet, vurderinger, totalvurderinger og atferdsintensjoner. I tillegg kan plassering av viktighetsspørsmålene påvirke respondentenes svargivning, både på viktighetsspørsmålene og vurderingss spørsmålene i spørreskjemaet.

I spørreskjemalitteraturen, f.eks. Hippler, Schwarz og Sudman (1987), Schuman og Presser (1981), Schwarz og Sudman (1991) og Sudman og Bradburn (1974), er effekter av spørsmålsrekkefølge studert under betegnelsen konteksteffekter. Man har spesielt fokusert på hvordan tidligere spørsmål kan påvirke svargivning på etterfølgende spørsmål. Innenfor markedsforskning har man f.eks. lenge vært oppmerksom på at måling av holdninger er kontekstavhengig, og at svargivning på holdningsspørsmålene dermed påvirkes av tidligere spørsmål. Imidlertid finnes det lite kunnskap om når slike konteksteffekter kan oppstå. I tillegg har det vært vanskelig å predikere hvilken retning effektene kan ha dersom de oppstår (Schwarz og Bless, 1992).

To forskjellige utgangspunkt er benyttet for å studere konteksteffekter; et metodologisk utgangspunkt, og et mer teoretisk utgangspunkt der respondentens informasjonsbehandlingsoppgaver ved utfylling av et spørreskjema har dannet basis for studiene. I studier av rekkefølgeeffekter med bakgrunn i metodologiske perspektiver har utgangspunktet vært et ønske om å finne ut om et spørsmål, A, blir besvart forskjellig avhengig av om dette blir stilt før eller etter ett annet spørsmål, B (Strack, 1991)⁵⁷.

Dersom viktighetsspørsmål blir inkludert, og plasseres før spesifikke vurderingss spørsmål i spørreskjemaet, kan det tenkes at de spesifikke vurderingene påvirkes på en måte som gjør dem forskjellige fra en situasjon der viktighetsspørsmålene er utelatt fra spørreskjemaet. Ved plassering av viktighetsspørsmål etter at respondentene er bedt om å foreta vurderinger av de spesifikke produktegenskapene i spørreskjemaet skal man i utgangspunktet kunne anta at viktighetsspørsmålene ikke vil påvirke de spesifikke vurderingene. Denne plasseringen av viktighetsspørsmålene skal ideelt sett gi samme resultat for kundenes vurderinger av spesifikke produktegenskaper som spørreskjema der viktighetsspørsmål er utelatt. I dette tilfellet antar vi imidlertid at vurderinger av de spesifikke produktegenskapene kan påvirke svargivning på de etterfølgende viktighets spørsmålene.

⁵⁷Med utgangspunkt i et kausalitetsperspektiv vil det ikke være rekkefølgen på spørsmålene som påvirker svaret på spørsmål A, men det forutgående spørsmålet, B, som har en effekt på svargivningen, og ikke omvendt. På denne måten kan "rekkefølgeeffekten" på mange måter betraktes som en "spørsmåls effekt", en effekt av et spørsmål som enten er stilt eller ikke stilt.

I tillegg til at plassering av viktighetsspørsmål og spesifikke vurderingss spørsmål kan påvirke respondentenes svargivning på disse spørsmålene som beskrevet ovenfor, kan det også tenkes at plassering av viktighetsspørsmål og spesifikke vurderingss spørsmål vil påvirke respondentenes svargivning i forbindelse med etterfølgende totalvurderinger og atferdsintensjoner. Når viktighetsspørsmål plasseres etter de spesifikke vurderingss spørsmålene kan viktighetsspørsmålene få større betydning for totalvurderinger og atferdsintensjoner. Vi står altså overfor en situasjon der rekkefølgeeffektene omfatter tre blokker av spørsmål, der plasseringen av de to første spørsmålsblokkene kan påvirke svargivningen, både på spørsmålene i disse spørsmålsblokkene (viktighetsspørsmålene og vurderingss spørsmålene), og på totalvurderinger og atferdsintensjoner som representerer den tredje spørsmålsblokken. I avhandlingen ønsker vi å studere begge disse effektene. Vi vil derfor senere i dette kapitlet diskutere betydningen av spørsmålsrekkefølge for nivå, både på viktighetsspørsmål og spesifikke vurderingss spørsmål. I tillegg vil se nærmere på sammenhengene mellom viktighets spørsmål og vurderingss spørsmål avhengig av hvor disse er plassert i forhold til hverandre i et spørreskjema. Til slutt vil betydningen av spørsmålsrekkefølge bli drøftet med utgangspunkt i de spesifikke vurderingens evne til å forklare totalvurderinger og atferdsintensjoner⁵⁸.

Spørsmåls effekter er i litteraturen klassifisert med bakgrunn i retning på effektene (Strack, 1991). Denne klassifiseringen er basert på en antakelse om at individer som blir bedt om å foreta en vurdering av et objekt først må hente en kognitiv representasjon av objektet frem fra minnet. Deretter må de bestemme seg for en standard som vurderingsobjektet skal sammenlignes med, og vurderes i forhold til. Både fremhenting av den kognitive representasjonen av objektet, og etableringen av vurderingsstandard er kontekst avhengig (Schwarz og Bless, 1992). Individer antas ikke å hente frem og bruke all kunnskap knyttet til hverken vurderingsobjektet eller sammenligningsstandard for å foreta en vurdering. Istedenfor benyttes den delen av informasjonen som er mest tilgjengelig når vurderingen blir foretatt. Dette medfører at den midlertidige representasjonen av vurderingsobjektet, og konstruksjonen av sammenligningsstandard som benyttes inneholder informasjon som er kronologisk tilgjengelig, dvs tilstede i respondentens hukommelse, og dermed uavhengig av kontekstuell påvirkning. Den midlertidige representasjonen av vurderingsobjektet inneholder også informasjon som er midlertidig tilgjengelig, og et resultat av kontekstuell påvirkning (Schwarz og Bless, 1992). Forskjeller knyttet til kronologisk tilgjengelighet av informasjon kan sees på som en refleksjon av karakteristika ved respondenten, som f.eks. erfaring og produktkunnskap, mens forskjeller knyttet til midlertidig tilgjengelighet av informasjon

⁵⁸På denne måten vil betydningen av plassering av viktighetsspørsmål og vurderingss spørsmål for nivået på totalvurderinger og atferdsintensjoner bli studert.

primært har sin årsak i forhold ved spørreskjemaet som benyttes for måling av respondentenes vurderinger. Den mest vanlige, og viktigste konteksteffekten oppstår når informasjon som har blitt benyttet for å svare på et tidligere spørsmål aktiveres, er relevant og påvirker svargivningen på etterfølgende spørsmål. I forhold til problemstillingene som blir tatt opp i avhandlingen er dette en sannsynlig årsak til forskjeller i svargivning som resultat av ulik plassering av viktighetsspørsmål og vurderings-spørsmål i spørreskjemaet.

To forskjellige benevnelser blir benyttet for å beskrive prosessene som antas å lede til konteksteffektene. Schwarz, Strack og Mai (1991) og Lynch, Chakravatri og Mitra (1991) er representanter for forskere som benytter benevnelserne assimilasjon og kontrast, mens Tourangeau og Rasinski (1988) og Bickart (1993) representerer forskere som benytter benevnelserne carryover og backfire for å beskrive de samme prosessene. Vi har tidligere i avhandlingen benyttet benevnelserne assimilasjons- og kontrasteffekter i forbindelse med betydningen av instruksjon og vurderingsskala for nivå på vurderinger når kvaliteten på produktene er forskjellig. Basert på et ønske om konsistent begrepsbruk gjennom hele avhandlingen, velger vi derfor også å benytte benevnelserne assimilasjon og kontrast når vi nedenfor beskriver disse prosessene mer i detalj, og drøfter hvorvidt perspektivene de representerer er relevante for å forstå hvorfor ulik plassering av viktighetsspørsmål og vurderingss spørsmål i et spørreskjema kan gi opphav til forskjeller i svargivning.

Svargivning på et vurderingss spørsmål i en intervjusituasjon innebærer at respondentene har aktivert informasjon. Hvordan informasjonen som aktiviseres påvirker vurderingene avhenger av hvordan den aktiverte informasjonen kategoriseres av respondenten. Informasjon som er inkludert i (den midlertidige) representasjonen individene har av objekt-kategorien vurderingsobjektet tilhører, kan resultere i en assimilasjonseffekt. I en spørreundersøkelsessituasjon vil en slik effekt kunne oppstå dersom respondenter svarer på en måte som er konsistent med antakelser, vurderinger eller kunnskap aktivert og gjort tilgjengelig gjennom svargivning på et tidligere spørsmål i spørreskjemaet. Det kan f.eks. tenkes at produktgenskaper som oppfattes som viktige, og som blir gjort mer tilgjengelige gjennom svargivning på viktighetsspørsmål fører til mer positive eller negative vurderinger når disse plasseres etter viktighetsspørsmålene i spørreskjemaet. Dette kan skyldes respondenters ønske om å virke konsistente i svargivning.

Den motsatte effekten, som kalles kontrasteffekt, kan oppstå dersom informasjon som aktiveres gjennom tidligere spørsmål ekskluderes fra den kognitive representasjonen respondentene har av vurderingsobjektet. I dette tilfellet vil påvirkningen av det forutgående spørsmålet gå i motsatt retning (Schwarz og Bless, 1992; Tourangeau, 1991).

Kontrasteffekter kan være et resultat av to ulike mekanismer: På den ene siden kan en kontrasteffekt oppstå som et resultat av en subtraksjon av informasjon fra objekt-representasjonen. På den andre siden kan en kontrasteffekt oppstå dersom informasjon som ekskluderes fra objektrepresentasjonene benyttes for å etablere en standard for sammenligning. Svargivning på viktighetsspørsmål kan her føre til at respondentene aktiverer en standard for vurderinger som tilsvarer et "ideal produkt". Dersom dette er tilfellet kan det tenkes at etterfølgende vurderinger vil være mindre positive på grunn av forankringen som benyttes for vurderingene. Dette betinger imidlertid at produktet som respondeneten blir bedt om å vurdere, kvalitetsmessig ligger langt fra den aktiverte sammenligningsstandard, eller det aktiverte idealproduktet.

Forskning i tilknytning til assimilasjons- og kontrasteffekter har i hovedsak fokusert på implikasjoner av plassering av generelle vs. spesifikke spørsmål. Dersom innholdet i et spørsmål fremkaller respons som er dominerende i positiv eller negativ retning, kan valøren på den dominerende responsen bli benyttet som et kriterium for assimilasjons- eller kontrasteffekter. Et eksempel på dette finnes i en studie av (Ottati et al., 1989). Her fant man at respondenter uttrykte en mer positiv holdning til generell ytringsfrihet dersom et tidligere spørsmål refererte til et positivt evaluert objekt, f.eks. "the American Civil Liberties Union" (assimilasjon), enn om det tidligere spørsmålet refererte til objekt med negativ valør, f.eks. "the American Nazi Party (kontrast). I dette tilfellet er det selve valøren på det første spørsmålet som gir opphav til effektene.

Dersom svargivning på et forutgående spørsmål ikke er påvirket av innholdet, eller valøren på spørsmålet vil det ikke eksistere noen dominant respons, og den individuelle reaksjonen eller svargivningen på det første spørsmålet må benyttes som grunnlag for assimilasjon eller kontrast. Eksempel på dette finnes hos Schuman, Presser, og Ludwig (1981). Her fant man at respondenter som hadde uttrykt positiv holdning til legalisering av abort, f.eks. dersom det er sannsynlig at et barn kan være misdannet (et spesifikt spørsmål), før de svarte på et generelt spørsmål om legalisering av abort, var mindre positive til legalisering av abort (kontrast), enn respondenter som svarte positivt på det spesifikke spørsmålet når dette kom etter det generelle (assimilasjon).

3.4.1.1 Respondentens informasjonsbehandlingsprosess og konteksteffekter

Forskningen som er drøftet ovenfor gir oss få retningslinjer for å si noe om hvorfor ulik plassering av viktighetsspørsmål og spesifikke vurderingssspørsmål i et spørreskjema skal gi opphav til forskjeller i respondentenes svargivning. En alternativ innfallsvinkel

for å forklare effekter av spørsmålsrekkefølge, er derfor å ta utgangspunkt i den modellen som ble presentert tidligere i dette kapitlet, og som illustrerer de ulike psykologiske informasjonsbehandlingsoppgavene respondenter står overfor ved svargivningen. Tourangeau og Rasinski (1988) drøfter ulike konteksteffekter som kan oppstå i holdningsundersøkelser nettopp med utgangspunkt i denne modellen. I artikkelen kartlegges variabler som antas å påvirke både retning og styrke på effektene. Drøftingen oppsummeres i tabell 3.1.

TABELL 3.1
Variabler som påvirker størrelse og retning på konteksteffekter

| Variabel | Effekt |
|---|---|
| <u>Variabler som påvirker tolking</u> | |
| • Kjennskap til tema | Ukjente tema er mottakelige for assimilasjonseffekter |
| • Tilgjengelighet til vurderinger | Ikke-tilgjengelige vurderinger er mer mottakelige for assimilasjonseffekter |
| <u>Variabler som påvirker fremhenting av informasjon/dannelse av oppfatning</u> | |
| • Blandete eller konfliktfylte vurderinger | Respondenter med blandete vurderinger er mer mottakelige for assimilasjonseffekter ved fremkalling av informasjon |
| • Ekspertise og involvering i tema | “Ekspertes” og involverte respondenter er i mindre grad påvirkelige av assimilasjonseffekter |
| • Spørsmålsform (direkte oppfatning eller relatert vurdering) | Relaterte vurderingsspmåler er mer på-påvirkelige av assimilasjonseffekter |
| • Hvor opplagt konteksten er | Opplagte kontekstsituasjoner neddiskonteres |
| • Dybde i informasjonsbehandlingen | Bruk av selvadministrert spørreskjema og gjennomføring av intervjuer i langsomt tempo kan redusere konteksteffekter |
| <u>Variabler som påvirker vurderingene</u> | |
| • Komplexitet i vurderingene | Mangesidige tema er påvirkelige av assimilasjonseffekter knyttet til vurderingene. Sammenligningsvurderinger er påvirkelige av kontrasteffekter |
| • Sammenlignbarhet og ekstremitet i kontekstspørsmålene | Ekstreme eller ulike forankringer gir opphav til kontrasteffekter |
| <u>Variabler som påvirker responsseleksjon/redigering av respons</u> | |
| • Karakteristika ved forankring (benevnelse på svarskala, midtpunkt vs. ekstremverdier) | Benevnelse på alle svaralternativer /skala-verdier kan bidra til å redusere forankringseffekter ved responsstadiet. Ekstreme forankringer gir opphav til kontrasteffekter |
| • Hvor fremtredende sammenhengen mellom spørsmålene er | Økt fremtredenhets av relasjon mellom spørsmålene kan øke konsistenseffektene for involverte respondenter og moderere effektene for ikke-involverte respondenter |

Kilde: Strack (1991) og Tourangeau, et al. (1988)

3.4.1.2 Betydning av kontekst for tolking av spørsmål

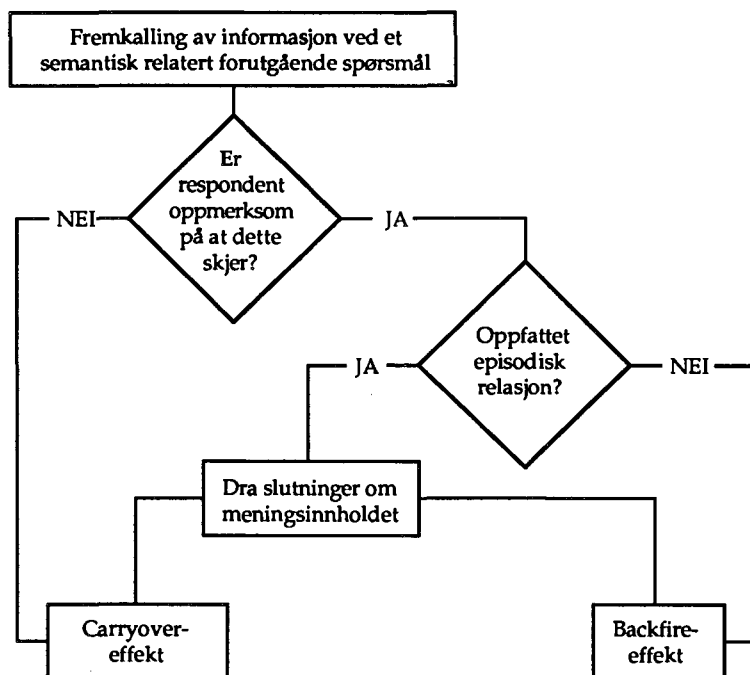
Når vi fokuserer på det første trinnet i prosessen, *tolking*, vil en viktig konsekvens være at det å svare på et spørsmål fører til aktivering av informasjon. Forutgående spørsmål regnes derfor ofte som den viktigste kilde til aktivering av informasjon som er relevant for tolkning og svargivning på etterfølgende spørsmål (Strack og Martin, 1987). Den psykologiske prosessen som oppstår i slike tilfeller kalles fremkalling (*priming*) (Fiske og Taylor, 1991; Strack og Martin, 1987). Fremkalling kan bidra til aktivering av ulike typer informasjon som kan være relevant for respondentens svargivning (Strack og Martin, 1987). Et forutgående spørsmål vil på denne måten ha en fremkallingseffekt der en respondents tilgjengelighet (*accessibility*) (Fazio, 1989) til vurderinger av spesifikke egenskaper antas å øke etterhvert som egenskapene blir mer fremtredende (*salient*) (Bodenhausen og Wyer, 1987; Schuman og Presser, 1981). Informasjon som blir mer fremtredende kan bidra til at en bestemt respons gjøres mer tilgjengelig (Fiske og Taylor, 1991), eller mer attraktiv gjennom en bevissthetsøkende prosess, aktivert av tidligere spørsmål i spørreskjemaet (Schuman, Kalton, og Ludwig, 1983; Schuman og Presser, 1981; Schwarz og Sudman, 1991).

Forutgående spørsmål i et spørreskjema kan altså ha to ulike funksjoner i forbindelse med respondenters tolking av etterfølgende spørsmål. For det første kan forutgående spørsmål ha en aktiveringsfunksjon som innebærer en automatisk aktivering av informasjon på en slik måte at bestemte begreper lettere fremkalles og benyttes ved senere vurderinger. For det andre kan et forutgående spørsmål ha en informasjonsfunksjon ved at det gir informasjon som basis for respondentenes tolking av meningsinnholdet i et etterfølgende spørsmål. Informasjonsfunksjonen betinger at respondenten er *oppmerksom* på at et relatert spørsmål er stillt tidligere, og at dette fører til fremkalling av informasjon. I tillegg til en forutsetning om at respondenten må være oppmerksom på informasjonen som ligger i det forutgående spørsmålet, må de to spørsmålene oppfattes som *episodisk relatert*, dvs. spørsmålene må oppfattes som å høre sammen. Relatert til problemstillingene i avhandlingen antar vi med bakgrunn i dette at plassering av viktighetsspørsmål før de spesifikke vurderingss spørsmålene kan implisere både aktivering av informasjon og hjelp til tolking av de etterfølgende vurderingene.

Fremkallingseffekter er direkte relatert til assimilasjons- og kontrasteffekter (Schuman, Kalton, og Ludwig, 1983; Schuman og Presser, 1981; Schwarz og Sudman, 1991). For å være i stand til å predikere type effekt og retning på effekten er det imidlertid nødvendig å forstå de undeliggende psykologiske prosessene knyttet til tolking av spørsmål. I figur 3.8 har vi illustrert disse prosessene i en modell slik vi har gjort rede for dem i dette avsnittet.

Modellen impliserer at påvirkning fra forutgående spørsmål i en spørreundersøkelses-situasjon kan forstås som fremkalling av informasjon. Fremkallingspåvirkningen kan som tidligere nevnt ha både en aktiveringsfunksjon og en informasjonsfunksjon (Strack, 1991). Aktiveringsfunksjonen impliserer økt tilgjengelighet til aktivert informasjon, Dette kan skje automatisk, og krever ikke at respondentene er klar over at fremkallingen skjer. Denne prosessen vil resultere i en assimilasjonseffekt på vurderingene. Informasjonsfunksjonen krever at respondentene er klar over at fremkalling skjer. I tillegg kreves det at det oppfattes en episodisk sammenheng mellom de to spørsmålene på en måte som gjør at de oppfattes å tilhøre samme kontekst i intervjuet.

FIGUR 3.8
 En skjematisk fremstilling av en modell for betydningen av forutgående spørsmål for tolking av spørsmål i et spørreskjema



Kilde: Strack (1991)

I en situasjon der disse kravene er tilfredsstillt, vil resultatet i form av enten assimilasjons- eller kontrasteffekt være avhengig av om respondenten benytter informasjonen som ligger i det forutgående spørsmålet. Dersom dette er tilfellet vil svargivning på det etterfølgende spørsmålet være påvirket av implikasjonene fra det forutgående spørsmålet og vi vil få en assimilasjonseffekt. Dersom respondentene i

motsatt fall ikke benytter informasjonen fra det forutgående spørsmålet kan vi oppleve en kontrasteffekt fordi respondentene vil tendere mot å ville gi ny informasjon og unngå å virke redundant i sin svargivning. Relatert til problemstillingene som blir tatt opp i avhandlingen er det sannsynlig at plassering av viktighetsspørsmål før spesifikke vurderingsspørsmål i spørreskjemaet vil implisere en assimilasjonseffekt for tolking av de etterfølgende vurderingsspørsmålene. Dersom viktighetsspørsmålene plasseres etter vurderingsspørsmålene forventer vi, basert på diskusjonen ovenfor, at vurderingsspørsmålene har en assimilasjonseffekt på tolking av viktighetsspørsmålene. Både viktighetsspørsmål og vurderingsspørsmål antar vi altså vil aktivere informasjon som er relevant for tolking av-, og svargivning på etterfølgende vurderingsspørsmål og viktighetsspørsmål.

Tourangeau og Rasinski (1988) fremhever respondentenes kjennskap til et bestemt tema og deres tilgjengelighet til holdninger og vurderinger som to forhold som kan påvirke tolking av spørsmål, og dermed også om assimilasjons- eller kontrasteffekter vil oppstå. Kjennskap til et bestemt tema, og tilgjengelighet til holdninger og vurderinger kan påvirke i hvilken grad vurderinger som er relevante for å svare på et spørsmål aktiveres gjennom tidligere spørsmål i spørreskjemaet. Dersom temaet det spørres om i et spørsmål er kjent for respondentene, kan det tenkes at en vurderingsstruktur aktiveres automatisk. Konteksteffekter i tilknytning til tema som ikke er kjent for respondentene vil vanligvis tendere i retning av assimilasjonseffekter. Dette impliserer at respondenter med god kjennskap til objektet de blir bedt om å vurdere vil være mindre påvirkelige av plassering av viktighetsspørsmål og vurderingsspørsmål i spørreskjemaet. Respondenter med dårligere kjennskap til vurderingsobjektet vil i større grad påvirkes av plasseringen av viktighetsspørsmålene i spørreskjemaet.

3.4.1.3 Betydning av kontekst for fremhenting av informasjon og dannelse av oppfatning

Gjennom *fremhenting av informasjon, og dannelse av oppfatning* komponenten i svar-givningsprosessen danner respondentene en oppfatning av hva de mener om temaet, eller objektet det spørres om. Dette innebærer fremhenting av relevant informasjon. Slik informasjon kan f.eks. bestå av antakelser, følelser eller eksisterende vurderinger. Hvilken informasjon som benyttes for å svare på et bestemt spørsmål kan bli påvirket av forhold som ordlyd i spørsmålsstilling, type vurdering, instruksjonen som blir gitt

respondenten, tilgjengelighet til tidligere vurderinger, og tidligere spørsmål i spørreskjemaet (Tourangeau, 1991).

Rekkefølge på spørsmål i et spørreskjema kan altså påvirke respondentenes dannelse av oppfatning, eller hva som betraktes som svar på et bestemt spørsmål. Dette kan skje gjennom fremkalling av informasjon, eller *priming*, ved hjelp av et forutgående spørsmål. Strack og Martin (1987) drøfter fem ulike effekter som fremkalling av informasjon gjennom forutgående spørsmål kan ha. For det første kan fremkalling av informasjon føre til at et annet eksemplar av kategorien det er spørsmål om i den aktuelle undersøkelsen aktiveres i respondentens minne. Vurderingene kan i slike tilfeller bli basert på vurderinger av dette andre eksemplaret. For det andre kan fremkalling av informasjon føre til at respondentene aktiverer spesielle egenskaper eller kjennetegn ved objektet det spørres om. For det tredje kan fremkalling av informasjon etablere en normativ standard som påvirker etterfølgende vurderinger. For det fjerde kan fremkalling av informasjon påvirke svargivning gjennom aktivering av tidligere atferd.

I forhold til problemstillingene som tas opp i avhandlingen antar vi at plassering av viktighets spørsmål før spesifikke vurderingss spørsmål i spørreskjemaet kan resultere i alle disse effektene, og påvirke respondentenes fremkalling av informasjon og dannelse av oppfatning, og dermed svargivning på vurderingss spørsmålene. På samme måte antar vi at svargivning på vurderingss spørsmålene kan påvirke respondentenes oppfatning av viktigheten ved de ulike produkt egenskapene dersom viktighets spørsmålene plasseres etter vurderingss spørsmålene i spørreskjemaet. Det er imidlertid vanskelig å si noe om retning og styrke på effektene i de spesifikke tilfellene bare på bakgrunn av diskusjonen ovenfor. Tourangeau og Rasinski (1988) diskuterer en rekke forhold som kan påvirke fremhenting av relevant informasjon for dannelse av oppfatning. Respondentenes kjennskap til tema og tilgjengelighet til vurderinger trekkes frem som variabler som i tillegg til å ha betydning for tolkning, også kan få betydning for fremhenting av informasjon og dannelse av oppfatning. Andre forhold av betydning er hvorvidt respondentene har blandete vurderinger av objektet, grad av ekspertise eller produkt kunnskap, hvilken spørsmålsform som benyttes og dybde i informasjonsbehandlingen ved svargivning. Det argumenteres for at respondenter med blandete vurderinger av objektet og respondenter med lav produkt kunnskap⁵⁹ er mer påvirkelige av svargivning på tidligere spørsmål, og at det her vil oppstå en assimilasjonseffekt. Det samme vil være tilfelle dersom respondentene blir bedt om å foreta relaterte vurderinger.

⁵⁹Betydningen av produkt kunnskap vil bli grundigere drøftet senere i kapittelet.

3.4.1.4 Betydning av kontekst for vurderinger

Forutgående spørsmål kan føre til fremkalling av tidligere vurderinger gjennom å gjøre bestemte aspekter ved vurderingsobjektet mer tilgjengelige. Dersom disse aspektene er relevante for svargivning på etterfølgende spørsmål vil de bli benyttet av respondentene (Tourangeau, 1991). Assimilasjonseffekter i forbindelse med vurderinger kan på denne måten oppstå når standardene som benyttes for å svare på et spørsmål også benyttes for å svare på etterfølgende spørsmål. Gjennom å gjøre en bestemt dimensjon, norm eller standard tilgjengelig, antar vi at forutgående viktighetsspørsmål i et spørreskjema kan påvirke respondentenes vurderinger, på samme måte som forutgående vurderings-spørsmål kan påvirke svargivningen på viktighetsspørsmålene.

Kontekst, eller rekkefølge på spørsmål, kan altså også påvirke hvordan vurderingen foretas, hvilke dimensjoner som inngår og hvilke kriterier eller sammenligningsstandarder som benyttes. På denne måten er konteksteffekter i tilknytning til spesifikke vurderinger relatert til hvordan antakelser eller kunnskap benyttes i vurderingsprosessen, mens konteksteffekter i forbindelse med fremhenting av informasjon og dannelse av oppfatning refererer til hvilke antakelser, eller hvilke "kunnskapsbiter" som inngår i produktvurderingene. I praksis kan det imidlertid være vanskelig å skille disse to typene effekter fra hverandre. Forutgående viktighetsspørsmål i et spørreskjema kan f.eks. etablere en standard som etterfølgende vurderinger relateres til. I slike tilfeller har ikke konteksten bare påvirket kva som fremkalles, f.eks. vurderingsstandarden, men også hvordan denne benyttes i vurderingene, f.eks. lavere vurderinger av produkter som oppfattes som dårligere enn den aktiverte vurderingsstandarden (idealet).

Assimilasjonseffekter i tilknytning til vurderingsstadiet i informasjonsbehandlingsprosessen oppstår vanligvis når vurderingsstandarden representerer et gjennomsnitt eller et midtpunkt (Tourangeau og Rasinski, 1988). Dersom vurderingsstandarden reflekterer ekstreme standarder for sammenligning vil kontrasteffekter kunne oppstå. I relasjon til problemstillingen om plassering av viktighetsspørsmål og vurderings-spørsmål i et spørreskjema som vi tar opp i denne avhandlingen kan både assimilasjons- og kontrasteffekter oppstå i vurderingsstadiet. Tourangeau og Rasinski (1988) fremhever to forhold som kan virke bestemmende for konteksteffekter i dette stadiet i informasjonsbehandlingsprosessen. For det første antar man at enkle vurderinger vil være lite påvirkelige av kontekst. Det motsatte vil være tilfelle dersom kriteriene for en vurdering er uklare, eller når vurderingen er komparativ, og en eksplisitt standard for sammenligning mangler. Her vil vurderingsprosessen være mindre reliabel og mer påvirkelig av kontekstuell fremkaltte vurderingskriterier, og forankringer. Basert på disse forholdene antar vi at enkle vurderinger vil være mindre påvirkelige av svar-

givning på viktighetsspørsmål dersom disse er plassert før de spesifikke vurderings-
spørsmålene i spørreskjemaet.

Det andre forholdet som trekkes frem for å forklare hvorvidt konteksteffekter vil oppstå i vurderingsstadiet er kjennetegn ved kontekstspørsmålene. Dersom en vurdering er basert på en sammenligning med et tidligere spørsmål, eller en annen forankring som resultat av kontekst, vil størrelse og retning på konteksteffekten avhenge av slike karakteristika ved forankringen som grad av ekstremitet og likhet med vurderingsobjektet. Ekstreme standarder for sammenligning vil som tidligere nevnt mest sannsynlig føre til kontrasteffekter på vurderingene, mens forankringer som representerer moderate verdier (som f.eks. gjennomsnittet for en kategori eller midtpunktet på en vurderingsskala) vil føre til assimilasjonseffekter.

Relatert til problemstillingene i avhandlingen, vil vi med drøftingen ovenfor som utgangspunkt forvente assimilasjonseffekter på de spesifikke vurderingene når viktighets-
spørsmål plasseres før disse i spørreskjemaet, og vurderingsobjektet eller produktet som respondentene blir bedt å vurdere oppfattes som et høykvalitetsprodukt. Produkter med høy kvalitet innenfor en kategori antar vi i mindre grad vil avvike fra mulige standarder for vurdering som kan bli aktivert gjennom svargivning på viktighetsspørsmål. Dersom produktet som vurderes i motsatt fall oppfattes som dårlig vil den samme vurderingsstandarden bli mer ekstrem, og kontrasteffekter vil være mest sannsynlig når viktighets-
spørsmål inkluderes før de spesifikke vurderingssørsmålene i spørreskjemaet.

3.4.1.5 Betydning av kontekst for responsseleksjon og redigering av respons

Det siste trinnet i respondentenes informasjonsbehandlingsprosess ved svargivning på et vurderingssørsmål er å *velge blant responsalternativene* som er gitt i spørreskjemaet. Dette innebærer at respondentene på den ene siden velger hvilket responsalternativ som representerer deres vurdering. På den andre siden antas respondenter å *rapportere og redigere* sine svar på et spørsmål for å unngå "uønsket respons", eller for å gjøre dem konsistente med svar på tidligere spørsmål i spørreskjemaet (Tourangeau og Rasinski, 1988). Det mest vanlige er altså assimilasjonseffekter ved valg av responsalternativ og redigering av respons. Et eksempel på en slik assimilasjonseffekt i dette siste stadiet i informasjonsbehandlingsprosessen er når viktighetsspørsmål som er plassert før spesifikke vurderinger i spørreskjemaet fører til at viktige produkt-egenskaper vurderes mer positivt (eller negativt) enn om viktighetsspørsmålene ikke hadde vært stillt.

Dersom respondentene svarer på viktighetsspørsmål etter at de har foretatt en vurdering av et produkts spesifikke egenskaper kan assimilasjonseffekter oppstå. Assimilasjons-effekter kan her føre til at respondentene legger mer vekt på de egenskapene som vurderes ekstremt positivt eller ekstremt negativt. For å være konsistent i sin svar-givning kan en egenskap som av respondenten er vurdert som svært bra eller dårlig, kunne bli tillagt høyere vekt enn den ville blitt dersom vurderingen ikke hadde vært foretatt. Assimilasjonseffekter kan også oppstå når respondenten forsøker å redusere kognitiv dissonans (Festinger, 1957) ved å redusere viktigheten av egenskaper som er vurdert som dårlig.

3.4.2 Betydning av produktkunnskap for plassering av viktig-hetsspørsmål og vurderingssspørsmål i spørreskjemaet

Kontekstuelle forhold, som f.eks. spørsmål plassert tidligere i et spørreskjema, kan altså påvirke etterfølgende vurderinger ved å øke tilgjengeligheten til informasjon (Feldman og Lynch, 1988). Imidlertid garanterer ikke økt tilgjengelighet til informasjon at denne informasjonen vil bli benyttet ved senere vurderinger (Bickart, 1991). Et forhold som påvirker hvordan tilgjengelig informasjon blir brukt for å foreta en vurdering er respondentenes produktkunnskap. Alba og Hutchinson (1987) beskriver produkt-kunnskap som bestående av den erfaringen et individ har med bruk av et produkt (kjennskap), og evnen til å utføre produktrelaterte oppgaver på en vellykket måte (ekspertise). Økt kunnskap om en produktkategori kan resultere i flere effekter, inkludert: (a) forbedret evne til å huske produktinformasjon, (b) mer komplett og raffinert kognitiv struktur relatert til produktkategorien og (c) reduksjon i den kognitive anstrengelsen som er nødvendig for utføre produktrelaterte oppgaver. Vi skal nå se nærmere på implikasjoner av disse forholdene for rekkefølgeeffektene vi fokuserer på i avhandlingen.

Respondenter med høy produktkunnskap vil med større sannsynlighet fremhente informasjon fra hukommelsen enn respondenter med lav produktkunnskap (Srull, 1983). Fordi eksperter vil være bedre i stand til å fremhente en tidligere vurdering som grunnlag for å foreta en vurdering, antar man at de ikke baserer vurderingene på informasjon gjort tilgjengelig av kontekst på samme måte som noviser. Dette impliserer at eksperter vil være mindre påvirkelige av tidligere spørsmål i et spørreskjema. Noviser vil f.eks. med større sannsynlighet benytte informasjon gjort tilgjengelig gjennom tidligere spesifikke spørsmål for å svare på et senere og mer generelt spørsmål. Dette vil resultere i økt korrelasjon mellom spørsmålene når det spesifikke

spørsmålet stilles først. Ekspertene vil på den andre siden bli mindre påvirket av kontekst når de foretar vurderinger i fremhentingstadiet i responsprosessen (Bickart, 1991).

Produktkunnskap påvirker ikke bare individers evne til å fremhente diagnostisk informasjon fra hukommelsen, men også hvordan informasjonen blir benyttet. I vurderingsstadiet i responsprosessen antar man at eksperter vil være mer påvirket av tidligere spørsmål i spørreskjemaet enn noviser. Tidligere spørsmål vil mest sannsynlig aktivere antakelser som favoriserer den ene siden av et tema for eksperter, på grunn av at deres kognitive strukturer er bedre utviklet. Aktiverte antakelser vil dermed ha større sannsynlighet for å bli benyttet ved senere vurderinger (Tourangeau et al., 1989a, Tourangeau et al., 1989b). Forutgående spørsmål i spørreskjemaet vil i en slik situasjon med stor sannsynlighet føre til assimilasjonseffekter for eksperter, mens vi her ikke forventer konteksteffekter for vurderinger foretatt av noviser (Bickart, 1991). Ekspertene og noviser er forskjellige med hensyn til hvor stor kognitiv anstrengelse som behøves for å svare på spørsmål om produkter (Alba og Hutchinson, 1987). Dette kan ha to forskjellige implikasjoner for svargivningsprosessen i en spørreundersøkelsessituasjon.

For det første antas det at spørsmål som er vanskelige for respondentene å svare på fører til satisfisering, eller at respondentene gir raske og enkle svar i motsetning til nøyaktige svar (Krosnick og Alwin, 1987). Respondenter med lav kognitiv kompleksitet er mer påvirkelige av spørsmålsrekkefølge. Fordi noviser behøver mer kognitiv anstrengelse for å svare på spørsmål i et spørreskjema antar man også at disse med større sannsynlighet vil satisfiere. Dette vil resultere i bruk av aktivert informasjon og dermed en assimilasjonseffekt. Ekspertene behøver mindre kognitiv anstrengelse for å svare på spørsmål i et spørreskjema. Ekspertene forventes derfor å gi mer gjennomtenkte svar. Tourangeau og Rasinski (1988) argumenterer for at respondenter som gir gjennomtenkte svar med større sannsynlighet vil være oppmerksom på sine tidligere svar og vil aksessere relevansen av disse for senere svargivning.

Oppmerksomhet på tidligere svargivning kan være en viktig faktor i en respondents bruk av informasjon. Når en respondent er oppmerksom på sine tidligere svar vil bruk av informasjon relateres til diagnoseverdi (Feldman og Lynch, 1988), som reflekterer i hvilken grad informasjon gjør respondenten i stand til å gjennomføre en vurdering. Dersom produktkunnskap reduserer den kognitive anstrengelsen som behøves for å svare på et spørsmål i en spørreundersøkelse, vil økt oppmerksomhet mot tidligere ikke-diagnostiske responser føre til en kontrasteffekt, mens økt oppmerksomhet mot diagnostiske responser vil føre til en assimilasjonseffekt (Bickart, 1991).

Som oppsummering av diskusjonen ovenfor kan altså tre forskjellige prediksjoner være aktuelle i forbindelse med betydningen av produktkunnskap for effekter av spørsmålsrekkefølge: (a) noviserespondenter vil bli positivt påvirket av tidligere spørsmål i et spørreskjema, mens eksperter vil bli mindre påvirket som resultat av deres evne til også å fremhente og benytte andre infomasjonskilder, (b) eksperter vil bli positivt påvirket av tidligere spørsmål fordi tidligere spørsmål aktiverer antakelser som favoriserer den ene siden av et tema, eller (c) eksperter vil bli negativt påvirket av tidligere spørsmål fordi de er oppmerksom på sine svar og ønsker å unngå vurderinger som er farget av svarene på de tidligere spørsmålene. Disse tre prosessene er ikke uavhengige av hverandre. Dette impliserer at forskjellige resultater sier noe om den relative betydningen av de tre prosessene for å bestemme effektene av spørsmålsrekkefølge. Resultatene kan også indikere på hvilket trinn i svargivningsprosessen effektene er lokalisert. (Tourangeau og Rasinski, 1988).

3.4.3 Hvilke effekter kan ulik plassering av viktighetsspørsmål og vurderingss spørsmål i spørreskjema gi opphav til?

Bruk av viktighetsspørsmål i vurderings- og holdningsmodeller har i hovedsak blitt studert med utgangspunkt i modellenes prediktive validitet (Akaah og Korgaonkar, 1983; Beckwith og Lehmann, 1973; Jaccard, Brinberg, og Ackerman, 1986; Mazis, Ahtola, og Klippel, 1975; Neslin, 1981). Prediktiv validitet er et uttrykk for hvor godt modellene predikerer totalvurderinger, holdninger, atferdsintensjoner, atferd eller preferanser. I dette avsnittet vil vi drøfte problemstillingene vi har tatt opp i forbindelse med plassering av viktighetsspørsmål i forhold til spesifikke vurderingss spørsmål i spørreskjema både i forhold til de spesifikke vurderingenenes evne til å forklare totalvurderinger og atferdsintensjoner, og i forhold til nivå på spesifikke vurderinger og viktighetsspørsmål. I tillegg vil vi se nærmere på sammenhenger mellom viktighets spørsmålene og vurderingss spørsmålene, avhengig av plassering i spørreskjemaet

3.4.3.1 Implikasjoner for nivå på vurderinger og viktighetsspørsmål

De fleste studiene som er gjennomført med hensyn på konteksteffekter og spørsmålsrekkefølge i forbindelse med spørreundersøkelser har fokusert på betydningen av kontekst for univariate svarfordelinger (Bickart et al., 1992; Bickart, 1991; Bickart, 1993; Hippler og Schwarz, 1987; Hippler, Schwarz og Sudman, 1987; Schuman,

Kalton og Ludwig, 1983; Schuman og Presser, 1981; Schuman, Presser og Ludwig, 1981; Schwarz og Bless, 1992; Schwarz, Strack og Mai, 1991; Schwarz og Sudman, 1991). Til tross for utstrakt forskning er det fortsatt en rekke uklarheter knyttet til når konteksteffekter vil oppstå, og når de oppstår hvilken retning og styrke effektene vil ha.

Tidligere i dette kapitlet har vi forsøkt å benytte en rekke bidrag der effekter av ulik rekkefølge på spørsmål i et spørreskjema drøftes i relasjon til assimilasjons- og kontrasteffekter, også betegnet *carryover-* og *backfireeffekter*. I relasjon til problemstillingene som blir tatt opp i avhandlingen definerte vi assimilasjonseffekter til å oppstå dersom svargivning på viktighetsspørsmål (eller vurderingsspørsmål dersom disse plasseres først) aktiverer informasjon som påvirker respondentenes vurdering av (eller tillegging av viktighet til) spesifikke produktenskaper, slik at disse vurderes på en måte som er konsistent med tilleggingen av viktighet til (eller vurderingen av) egenskapene.

I motsetning til assimilasjonseffekter der den aktiverte informasjonen inkluderes, og benyttes for svargivning på de etterfølgende spørsmålene, vil den aktiverte informasjonen, dersom en kontrasteffekt oppstår, ekskluderes og unngå å bli benyttet for svargivning på de etterfølgende spørsmålene. Kontrasteffekter ble i relasjon til våre problemstillinger definert til å oppstå dersom svargivning på viktighetsspørsmål aktiverer standarder for vurdering som er svært forskjellige fra det produktet som respondentene blir bedt om å vurdere i spørreskjemaet. Når vurderingene plasseres før viktighetsspørsmålene i spørreskjemaet vil de samme psykologiske prosessene som inntre når viktighetsspørsmålene plasseres før vurderingene kunne oppstå. Assimilasjonseffekter kan oppstå dersom svar på viktighetsspørsmålene påvirkes av respons på vurderingsspørsmålene på en slik måte at viktighetsspørsmålene samsvarer med vurderingene⁶⁰. I motsetning til det som er tilfelle ved assimilasjonseffekter vil kontrasteffekter føre til større forskjeller i svargivning som resultat av spørsmålsrekkefølgen. Når viktighetsspørsmål plasseres før de spesifikke vurderingene i spørreskjemaet, vil vektleggingen som et resultat av assimilasjons- og kontrasteffekter kunne påvirke vurderingene. Det motsatte vil være tilfelle dersom vurdering av spesifikke egenskaper foretas før man tillegger egenskapene vekt gjennom selvrapportert viktighet.

Assimilasjons- og kontrasteffekter vil i første rekke kunne påvirke nivået på viktighets-spørsmål og vurderingsspørsmål. Som vi har drøftet tidligere i dette kapitlet er det mange forhold på ulike nivå i respondentenes informasjonsbehandlingsprosess som kan

⁶⁰Dette kan f.eks. være tilfelle ved forsøk på å redusere kognitiv dissonans. Egenskaper som er vurdert som dårlig tillegges da lavere vekt.

føre til både assimilasjons- og kontrasteffekter. Det er derfor også her, som i mange tidligere studier, f.eks. Bickart et al. (1992), Bickart (1991), Bickart (1993), Hippler og Schwarz (1987), Hippler, Schwarz og Sudman (1987), Schuman, Kalton og Ludwig (1983), Schuman og Presser (1981), Schuman, Presser og Ludwig (1981), Schwarz og Bless (1992), Schwarz, Strack og Mai (1991) og Schwarz og Sudman (1991), vanskelig å predikere retning og styrke på generelle hypoteser i tilknytning til nivå på viktighetsspørsmål og vurderingsspørsmål, avhengig av plasseringen av disse i spørreskjemaet.

Vi har tidligere i dette kapittelet gjort et skille mellom høykvalitetsprodukter og lavkvalitetsprodukter ved drøfting av forskjeller i nivå på vurderinger foretatt med henholdsvis subjektforankrete- og objektforankrete instruksjoner og vurderingsskalaer. Dette skillet kan også hjelpe oss til å forstå og forklare forskjeller i svargivning på spesifikke vurderingsspørsmål som resultat av tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet.

I samsvar med et kausalitetsperspektiv antar vi at viktighetsspørsmål, og svargivning på disse, bare vil påvirke svargivning og nivå på spesifikke vurderingsspørsmål når viktighetsspørsmålene plasseres før de spesifikke vurderingene i spørreskjemaet. Denne antakelsen er basert på en forutsetning om at respondentene foretar en sekvensiell gjennomgang og utfylling av spørreskjemaet. Som vi har drøftet tidligere kan dette være en streng antakelse når respondentene svarer på spørsmålene ved selv å fylle ut et spørreskjema, og muligheten for å hoppe frem og tilbake i spørreskjemaet er tilstede.

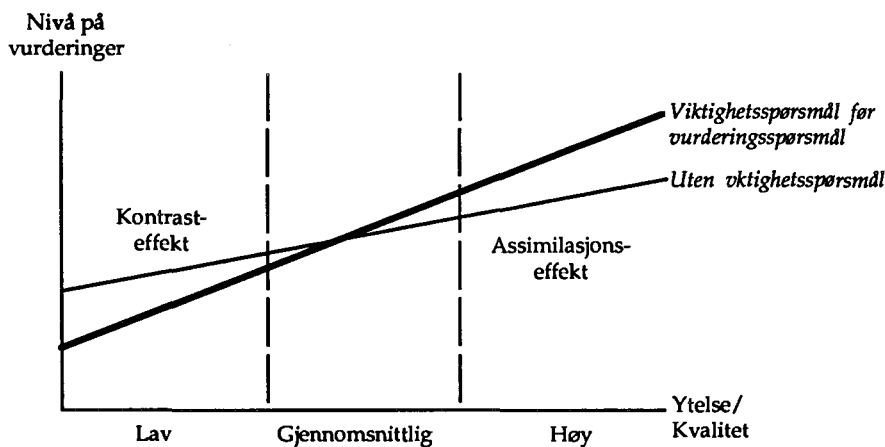
I litteraturen vi har benyttet for å belyse ulike konteksteffekter, og effekter av ulik rekkefølge på spørsmål i et spørreskjema, fokuseres betydningen av tidligere spørsmål i spørreskjemaet for aktivering av informasjon som kan påvirke svargivning på etterfølgende spørsmål. Vi har tidligere diskutert hvordan svargivning på viktighetsspørsmål kan aktivere sammenligningsstandarder eller "idealprodukter". Når viktighetsspørsmål inkluderes, og plasseres før de spesifikke vurderingsspørsmålene i et spørreskjema, kan disse sammenligningsstandardene i større grad bli benyttet når respondentene foretar vurderinger av produktet. Avhengig av kvaliteten på produktet respondenten har erfart, og blir bedt om å vurdere kan både assimilasjons- og kontrasteffekter oppstå.

Dersom respondenten blir bedt om å vurdere et høykvalitetsprodukt etter at en sammenligningsstandard, eller et "idealprodukt", er aktivert gjennom svargivning på viktighetsspørsmål, antar vi at en assimilasjonseffekt vil oppstå. Begrunnelsen for dette er at et høykvalitetsprodukt ytelsesmessig vil ligge nær idealproduktet eller sammenligningsstandard (Lynch, Chakravarti, og Mitra, 1991). For høykvalitetsprodukter,

antar vi med dette som utgangspunkt at nivået på vurderinger av spesifikke produkttegenskaper der respondentene først har tillagt egenskapene viktighet, vil overstige, eller være tilnærmet lik nivået på vurderinger foretatt i et spørreskjema der viktighetsspørsmålene er utelatt eller plassert etter de spesifikke vurderings-spørsmålene.

Når produktet respondenten har erfart og blir bedt om å vurdere kan kategoriseres som et lavkvalitetsprodukt, vil produktet ytelesemessig avvike fra eventuelle sammenligningsstandarder eller "idealprodukter" (Lynch, Chakravarti, og Mitra, 1991). Når viktighetsspørsmål inkluderes og blir plassert før de spesifikke vurderingsspørsmålene antar vi at slike sammenligningsstandardene eller "idealprodukter blir aktivert. For lavkvalitetsprodukter vil derfor mest sannsynlig en kontrasteffekt oppstå når viktighetsspørsmål plasseres før de spesifikke vurderingsspørsmålene i spørreskjemaet. Nivået på vurderinger av spesifikke produkttegenskaper vil her være lavere enn om viktighetsspørsmålene hadde vært utelatt eller plassert etter de spesifikke vurderings-spørsmålene.

FIGUR 3.9
 Betydningen av produktkvalitet, og tilstedeværelse og plassering av viktighetsspørsmål for sammenhengen mellom assimilasjons- og kontrasteffekter og nivå på vurderinger



Assimilasjons- og kontrasteffekter kan altså relateres til høykvalitetsprodukter og lavkvalitetsprodukter. Figur 3.9 illustrerer hvordan vi for lavkvalitetsprodukter forventer høyere nivå på spesifikke vurderinger dersom viktighetsspørsmål er utelatt eller plassert etter vurderingsspørsmålene i spørreskjemaet. Dette forklares med en kontrasteffekt. For høykvalitetsprodukter forventer vi høyere nivå på vurderinger

foretatt i et spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene. Forklaringen er her en assimilasjonseffekt.

3.4.3.2 Implikasjoner for sammenheng mellom vurderinger og viktighet

Forskjeller i nivå på viktighetsspørsmål og vurderingsspørsmål er ikke en nødvendig konsekvens av plasseringen av spørsmålene i et spørreskjema. Nivå, eller gjennomsnitt, representerer aggregerte størrelser som sier noe om svarmønster i et utvalg sett under ett. Når vi ønsker å studere forskjeller i sammenhengen mellom rapportert viktighet ved ulike produktegenskaper og vurdering av de samme egenskapene, avhengig av rekkefølgen på spørsmålene, må dette skje på individnivå. Feldman og Lynch (1988) fremhever effektene av tre forskjellige (men relaterte) faktorer for å forklare sannsynligheten for at respons i forbindelse med målinger av et begrep vil bli benyttet som basis for respons på et etterfølgende spørsmål. Disse faktorene er: (1) den oppfattede diagnoseverdien av den første vurderingen på den andre vurderingen, (2) tilgjengeligheten til den første vurderingen i hukommelsen og (3) tilgjengeligheten til alternative kilder for den andre vurderingen. Vi har tidligere beskrevet hvordan tilgjengelighet til informasjon aktivert av forutgående spørsmål i et spørreskjema kan påvirke svargivning på etterfølgende spørsmål. Vi vil nå se nærmere på hvordan diagnoseverdien som kan ligge i et forutgående spørsmål kan påvirke respons på etterfølgende spørsmål.

Oppfattet diagnoseverdi av en vurdering på en etterfølgende vurdering defineres som i hvilken grad en respondent oppfatter at svargivning på det første spørsmålet korrekt identifiserer hvordan det andre spørsmålet bør besvares (Feldman og Lynch, 1988). En mulig test for å avdekke diagnoseverdi ved viktighetsspørsmål og vurderingsspørsmål er å sammenligne korrelasjoner mellom disse. Dersom viktighetsspørsmålene og vurderingsspørsmålene har symmetrisk diagnoseverdi vil korrelasjonene mellom spørsmålene ikke påvirkes av rekkefølgen på spørsmålene i spørreskjemaet. Dette innebærer at diagnoseverdien av viktighet for vurderinger er lik diagnoseverdien av vurderinger for rapportering av viktighet. Dersom viktighetsspørsmål og vurderingsspørsmål har asymmetrisk diagnoseverdi, vil korrelasjon mellom viktighetsspørsmål og vurderingsspørsmål være avhengig av rekkefølgen på spørsmålene. Rapportering av viktighet vil ha diagnoseverdi for vurderinger dersom korrelasjonen mellom viktighets-spørsmål og vurderingsspørsmål er høyere når viktighetsspørsmålene er plassert før vurderingsspørsmålene i spørreskjemaet, enn når vurderingsspørsmålene er plassert før viktighetsspørsmålene.

Tidligere i dette kapittelet stilte vi et viktig og relevant spørsmål i forbindelse med rekkefølgeproblemstillingene vi har reist ovenfor: Er det slik at vurderinger, viktighet, totalvurderinger og atferdsintensjoner kan bli skapt av måleprosedyren dersom disse ikke allerede eksisterer i respondentens langtidshukommelse? Dersom vi svarer nei på spørsmålet ovenfor innebærer dette en antakelse om at viktighet, vurderinger, totalvurderinger og atferdsintensjoner er begreper som eksisterer i respondentenes langtidshukommelse, og som ikke påvirkes av måleprosedyren.

Et alternativt syn på disse tradisjonelle antakelsene er at selve målingen vil påvirke respondentenes rapportering av viktighet, vurderinger, totalvurderinger og atferdsintensjoner. Med et slikt utgangspunkt antar vi at plassering av viktighetsspørsmålene vil påvirke respondentenes svargivning, både på viktighetsspørsmålene og vurderingss spørsmålene i spørreskjemaet.

Hvilke effekter vi forventer når det gjelder sammenhenger mellom viktighetsspørsmål og vurderingss spørsmål vil altså være avhengig av hvilket perspektiv vi legger til grunn når det gjelder måling av teoretiske begreper. Et grunnleggende utgangspunkt for denne avhandlingen har vært en erkjennelse av at måling og måleprosedyren som benyttes kan påvirke respondentenes svargivning, og dermed måling av de teoretiske begrepene vi er opptatt av. Imidlertid kan det skilles mellom viktighet og vurderinger når det gjelder sannsynligheten for at begrepene er skapt av måling, eller om de eksisterer i respondentenes langtidshukommelse.

Det kan argumenteres for at viktighet knyttet til produktegenskaper er individspesifikk og knyttet til individuelle preferanser, mens vurderinger i større grad er knyttet til objektene som vurderes. Med dette som utgangspunkt antar vi at det er mer sannsynlig at vurderinger vil bli påvirket av respondentenes rapportering av viktighet, enn at rapportering av viktighet vil bli påvirket av vurderingene.

Vi har tidligere beskrevet hvordan respondentenes produktkunnskap kan moderere konteksteffekter i spørreundersøkelser. Produktkunnskap antar vi også vil ha en modererende effekt for de forskjellene vi har argumentert for når det gjelder sammenheng mellom viktighetsspørsmål og vurderingss spørsmål. Vi antar derfor at viktighets spørsmålene vil ha diagnoseverdi for vurderingss spørsmålene for respondenter med lav produktkunnskap. For respondenter med høy produktkunnskap antar vi at viktighets spørsmål og vurderingss spørsmål vil ha tilnærmet symmetrisk diagnoseverdi.

3.4.3.3 Implikasjoner for prediktiv validitet

Tidligere i dette kapittelet har vi argumentert for at plassering av viktighetsspørsmål før spesifikke vurderingss spørsmål i spørreskjemaet kan føre til både *priming* og *saliency* effekter. *Priming* og *saliency* effekter bidrar til å gjøre vurderingene mer tilgjengelige for respondentene. Tilgjengelighet (*accessibility*) påvirker holdningsdannelse og vurderinger (Fazio, 1989). Respondenter med holdninger eller vurderinger som er lett tilgjengelige vil ha en sterkere sammenheng mellom vurdering av spesifikke egenskaper og totalvurderinger og intensjoner om atferd (Fazio, 1989). Med bakgrunn i dette antar vi derfor at respondenter som benytter spørreskjema der viktighetsspørsmål er plassert før spesifikke vurderinger, vil ha sterkere sammenheng mellom vurderinger av spesifikke produkt egenskaper, totalvurdering og atferdsintensjoner enn respondenter som benytter spørreskjema der viktighetsspørsmålene er plassert etter de spesifikke vurderingss spørsmålene.

Vi har tidligere konkludert med at erfaring med bruk av et produkt og produktkunnskap, på lik linje med fremtredenhet og fremkalling av informasjon kan bidra til økt tilgjengelighet til vurderinger. Vi antar derfor at de effektene av spørsmålsrekkefølge for de spesifikke vurderingenes prediktive validitet som vi beskrev ovenfor, vil gi seg mindre utslag for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap.

3.5 Oppsummering og implikasjoner

I dette kapitlet har vi fokusert på respondenten som foretar produktvurderinger, og uttrykker disse gjennom svar på spørsmål i et spørreskjema. Informasjonsintegrasjonsteori, holdningsteori og bidrag fra kundetilfredshetslitteraturen ble benyttet for å belyse hvordan slike produktvurderinger kan tenkes å bli foretatt. For å illustrere hvordan kunder vurderer varer og tjenester drøftet vi ulike multiattributtmodeller hentet fra holdningsteori. I følge disse modellene er kundens totalvurdering en funksjon av vurderinger av oppfattede egenskaper ved produktet, veiet med viktigheten eller betydningen av disse egenskapene. Videre benyttet vi teorier for kundetilfredshet som et supplement til holdningsmodellene. Et sentralt paradigme i studier av kundetilfredshet er *the disconfirmation paradigm* der "tilfredshet" relateres til størrelsen, og retningen på avviket mellom opplevelse og forventninger. Både holdningsteori og

teorier for kundetilfredshet ble benyttet både for å forstå forskjeller mellom objektforankrete- og subjektforankrete vurderinger.

For å forstå hvordan produktvurderingene kommer til uttrykk gjennom svar på spørsmål i et spørreskjema, gjorde vi rede for en modell som setter søkelyset på hvilke kognitive informasjonsbehandlingsoppgaver et individ antas å utføre som respondent når han/hun svarer på disse spørsmålene. I følge denne modellen må respondenten først foreta en tolkning for å forstå meningsholdet i spørsmålet. Videre må det dannes en oppfatning av temaet det spørres om før det foretas en vurdering. Denne vurderingen må så passes inn i spørreskjemaets responsformat. Til slutt kan det avgitte svar bli endret som en funksjon av ulike motiver som kan oppstå i undersøkelsessituasjonen. Vi har i dette kapitlet benyttet denne modellen som utgangspunkt for på drøfte hvordan bruk av objektforankrete- og subjektforankrete instruksjoner pålegger respondentene ulike kognitive oppgaver. Vi har argumentert for at objektforankrete- og subjektforankrete vurderinger er forskjellige på mange måter. Med utgangspunkt i drøftingen av disse forskjellene, hevder vi at bruk av de to instruksjonene og vurderingsskalaene, ved måling av produktvurderinger, vil føre til forskjeller i egenskapsvurderingenes evne til å predikere kundens totalvurdering og atferdsintensjoner, forskjeller i konsistens i vurderingene, forskjeller i sikkerhet i vurderingene og forskjeller i nivå på vurderingene ved ulik kvalitet på produktet. Med bakgrunn i de teoretiske perspektivene har vi i tillegg argumentert for at de direkte effektene av valg mellom de to instruksjonene og vurderingsskalaene kan bli moderert av kundens produktkunnskap og erfaring med bruk av produktet.

Til slutt i kapitlet ble modellen for respondentenes informasjonsbehandlingsprosess ved svargivning på spørsmål i et spørreskjema benyttet for å drøfte effekter av inkludering og ulik plassering av viktighetsspørsmål i spørreskjemaet. Litteratur om konteksteffekter i spørreundersøkelser, *priming* og *salience* effekter ble her benyttet for å forsøke å forstå og forklare forskjeller, både for nivå på spesifikke vurderings-spørsmål og på viktighetsspørsmål, i tillegg til de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner. Vi argumenterte for at inkludering og ulik plassering av viktighetsspørsmål i spørreskjemaet vil få konsekvenser, både for de spesifikke vurderingenes evne til å forklare kundens totalvurderinger og atferdsintensjoner, for sammenhenger mellom viktighetsspørsmål og vurderingsspørsmål og for nivå på vurderingsspørsmål. Selv om vi antar at nivået på viktighetsspørsmålene, på samme måte som vurderingsspørsmålene, kan bli påvirket av plasseringen av disse, er vi på bakgrunn av litteraturgjennomgangen og teoridrøftelsene vi har foretatt, ikke i stand til å formulere hypoteser som sier noe om retning og styrke på nivåforskjeller for viktighetsspørsmålene.

DEL
III
HYPOTESER
OG
UNDERSØKELSES-
OPPLEGG

Denne delen av avhandlingen består av 3 kapitler. I kapittel 4 presenteres først en modell for de variabler og sammenhenger som vi ønsker å studere. Deretter fremsettes og drøftes en rekke hypoteser basert på analysemodellen. I kapittel 5 diskuteres metodologiske valg som grunnlag for den empiriske studien. Til slutt, i kapittel 6, redegjør vi for operasjonalisering av avhandlingens uavhengige, avhengige og modererende variabler.

4

Hypoteser

I dette kapitlet vil vi på nytt rette oppmerksomheten mot problemstillingene vi tok opp i kapittel 1. Basert på de metodologiske og teoretiske perspektivene vi drøftet i kapittel 2 og 3, vil vi nå presisere problemstillingene i form av hypoteser. Formålet med avhandlingen er ikke å teste de teoretiske perspektivene vi benytter, men heller forsøke å anvende teorier for å predikere effekter av bestemte valg man, både som akademiker og praktiker, står overfor ved gjennomføring av spørreundersøkelser. Med utgangspunkt i prediksjoner fra den teoretiske drøftingen, blir det formulert hypoteser som senere vil bli underlagt testing. Denne presiseringen av formålet vil gi føringer både for hvordan vi anvender de ulike teoretiske perspektivene i presisering av hypoteser, og for undersøkelsesopplegget som velges for å teste hypotesene.

Vi innleder kapitlet med å presentere en modell som illustrerer hvordan variablene og sammenhengene vi ønsker å studere, er relatert til problemstillingene vi tok opp i kapittel 1.

4.1 Modell

Vi oppsummerte drøftingen av problemstillingen vi reiste i kapittel 1, basert på måle-teoretiske og metodologiske innfallsvinkler, med følgende foreløpige konklusjoner: 1) Vi kan forvente å finne systematiske forskjeller i respondentenes svargivning ved bruk av ulik ordlyd i spørsmålsstilling og benevnelse på svarskala, og 2) Rekkefølgen på spørsmålene i et spørreskjema kan påvirke svargivning. Drøfting av problemstillingene

basert på de måleteoretiske og metodologiske perspektivene i kapittel 2, gir oss imidlertid ikke indikasjoner på hvilke effekter vi kan forvente å finne for de spesifikke problemstillingene som blir tatt opp i avhandlingen.

Dersom vi bare skulle ha tatt utgangspunkt i metodologiske og måleteoretiske perspektiver, noe som ved første øyekast synes å være det mest nærliggende, ville vi ikke ha vært i stand til å presisere problemstillingene fra kapittel 1 i presise hypoteser som sier noe om type effekter, samt retning og styrke på effektene. Med svært upresise hypoteser, hvis noen i det hele tatt, ville en naturlig konsekvens for videre arbeid vært en eksplorativ tilnærming, der formålet hadde vært å generere forklaringer.

Formålet med avhandlingen er som nevnt i innledningen å rette søkelyset mot bruk av spørreskjema for måling av kunders produktvurderinger som grunnlag for å si noe om produktets kvalitet. Tre begreper er sentrale i en slik sammenheng: produktet som vurderes, respondenten (kunden) som foretar vurderinger og måling av vurderingene. Med utgangspunkt i litteraturgjennomgangen og de teoretiske drøftelsene i kapittel 3, presiserer vi problemstillingene i hypoteser med bakgrunn i disse begrepene. Sentrale bidrag fra informasjonsbehandlingsteori, sosial kognisjonsteori, konsumentatferd, holdningsteori, kundetilfredshetsteori, og beslutningsteori er benyttet som grunnlag for å definere avhengige variabler som kan si noe om de effektene disse teoriperspektivene foreskriver. Hypotesene er derfor basert på implikasjoner av de teoretiske perspektivene vi har valgt.

Når formålet med en spørreundersøkelse har vært å måle kunders produktvurderinger, er det særlig to ulike måter å stille spørsmål på, som i litteraturen har blitt brukt omhverandre. På den ene siden har man instruert respondentene til å vurdere hvorvidt produktet er bra eller dårlig. På den andre siden har det vært vanlig å spørre kundene i hvilken grad de er tilfreds eller utilfreds med produktet. Basert på de teoretiske og metodologiske perspektivene vi har drøftet, antar vi at bruk av disse instruksjonene og vurderingsalternativene de impliserer, medfører ulike kognitive oppgaver for respondentene, og at de derfor fører til systematiske forskjeller i svargivning. Vi betrakter denne spesifikke problemstillingen, knyttet til valg av ordlyd i spørsmålsstilling og benevnelse på svarskala (*question wording*) som viktig, og ønsker å belyse hvilke responseffekter som kan oppstå ved valg av den ene måten å stille spørsmål på fremfor den andre.

Det har vært vanlig å operasjonalisere kunders holdninger og produktvurderinger i et spørreskjema ved hjelp av multiattributtmodeller. Multiattributtmodeller blir vanligvis operasjonalisert ved hjelp av to sett med spørsmål, ett sett med viktighets spørsmål og ett

sett med vurderingsspmårl. En lang rekke studier har vrt gjennomfrt for å forbedre multiattributtmodellenes prediktive validitet. Srlig har man vrt opptatt av hvilken betydning det har dersom viktighetsspmårl inkluderes i sprreskjemaet. Resultatene fra disse studiene er ikke entydige. Vi har valgt å fokusere på et problem i tilknytning til sprsmålsrekkeflge (*question order*) som vi ikke kan se har vrt viet oppmerksomhet i tidligere forskning. Når man har bestemt seg for å ta med viktighetsspmårl i sprreskjemaet kan disse enten plasseres fr eller etter vurderingene. Vi nsker altså å belyse responseeffekter som kan oppstå ved ulike plassering av viktighetsspmårlene.

Respondentens produktkunnskap har vrt viet stor oppmerksomhet i forbindelse med forskning vedrrende vurdering av produkter. Man har spesielt fokusert på betydningen produktkunnskap både for vurderingsstruktur og for vurderingsprosess. I avhandlingen betraktes produktkunnskap både som en variabel med direkte effekter på svargivning, og som en variabel som kan moderere effektene vi forventer som resultat av de spesifikke valgene vi fokuserer på i forbindelse med instruksjon og vurderingsskala. Produktkunnskap betraktes også som en variabel som kan bidra til å moderere effektene av sprsmålsrekkeflge.

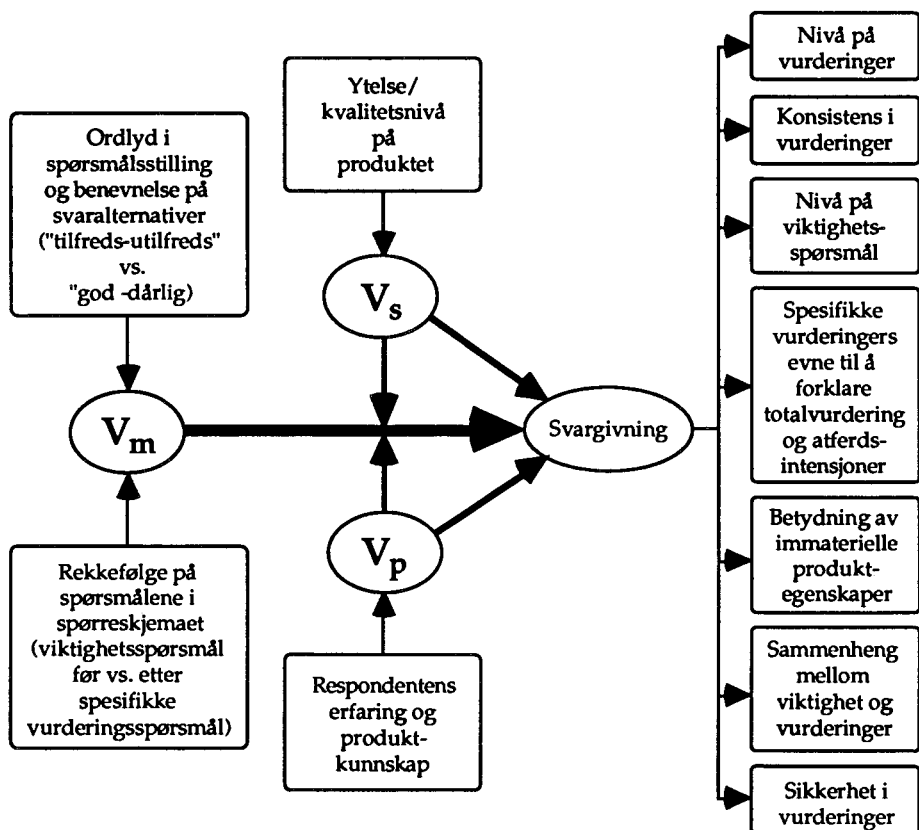
En sentral antakelse i kundetilfredshetslitteraturen er at positive avvik mellom kundenes forventninger/normer og produktets ytelse frer til tilfredshet, mens negative avvik frer til utilfredshet. Produktets ytelse eller kvalitet kan på denne måten få betydning for hvordan respondenter foretar vurderinger. Kvalitetsnivået på produktene som vurderes ble med dette som utgangspunkt, i kapittel 3, trukket frem som et forhold som kan moderere effektene av de metodologiske valgene vi nsker å studere. Med bakgrunn i litteraturgjennomgangen, kan modellen vi benyttet for å illustrere problemstillingene i kapittel 1, nå presiseres i en analysemodell der den avhengige variabelen er presisert i form av ulike effekter.

I Figur 4.1 har vi inkludert de tre hovedkildene til varians i respons: egenskaper ved produktet (V_s), egenskaper ved respondenten (V_p) og egenskaper ved måleprosedyren (V_m). Modellen viser i tillegg hvordan disse kan relateres til problemstillingene vi har tatt opp i dette kapittelet.

Hovedfokus legges på variasjoner i svargivning som skyldes måleprosedyren (V_m), representert ved ordlyd i sprsmålsstilling, benevnelse på svarskalaen og sprsmålsrekkeflge. Disse forholdene antas å ha direkte betydning for respondentenes svargivning. Når det gjelder varians som skyldes stimuli (V_s), har vi valgt å fokusere på produktets kvalitetsnivået eller ytelse. Basert på teoridrftingen antar vi kvalitetsnivået

på produktet, både har en direkte effekt og modererer effektene av de metodologiske valgene.

FIGUR 4.1
Analysemodell



Til slutt illustrerer modellen hvordan egenskaper ved personen som foretar vurderingene (V_p), representert ved erfaring og produktkunnskap, antas å moderere effektene av skalavalg, i tillegg til å ha en direkte effekt på svargivning. Respons-effektene er i modellen representert ved nivå på vurderinger, konsistens i vurderinger, sammenheng mellom viktighet og vurderinger, sammenheng mellom vurderinger av spesifikke egenskaper, totalvurdering, og intensjoner om atferd, sikkerhet i vurderinger og relativ betydning av immaterielle og materielle produkt egenskaper. Vi vil nå presisere de ulike effektene i testbare hypoteser.

4.2 Hypoteser

Hypotesene er organisert i to hovedblokker. I den første hovedblokken presenterer vi hypoteser som tar for seg effekter i tilknytning til valg mellom bruk av objektforankret- og subjektforankret instruksjon og -vurderingsskala. I den andre hovedblokken fokuseres problemstillingene vedrørende tilstedeværelse og plassering av viktighets-spørsmål i spørreskjemaet, ved operasjonalisering av multiattributtmodeller.

4.2.1 Valg mellom objektforankret- og subjektforankret instruksjon og -vurderingsskala

Ordlyd i spørsmålsstilling og svarskala er som nevnt viet relativt mye oppmerksomhet i tidligere forskning (Schuman og Presser, 1981; Sudman og Bradburn, 1983). Sammenligninger av benevnelsene “bra-dårlig” og “tilfreds-tilfreds”, slik de benyttes ved måling av kunders produktvurderinger, er imidlertid ikke foretatt tidligere. I kapittel 2 og 3 argumenterte vi for at “bra-dårlig” skalaen kan kalles objektforankret og “tilfreds-tilfreds” skalaen kan kalles subjektforankret. Dette skillet baserte vi på skaleringslitteraturens klassifisering av ulike kilder til varians (Cox, 1982; Guilford, 1954; Hippler, Schwarz, og Sudman, 1987; Torgerson, 1958). Basert på det faktum at subjektforankrete- (“tilfreds-tilfreds”) og objektforankrete (“bra-dårlig”) instruksjoner og -vurderingsskalaer blir benyttet om hverandre, både ved gjennomføring av praktiske markedsundersøkelser og i akademisk forskning (Cadotte, Woodruff, og Jenkins, 1987; Swan og Trawick, 1981), kan det tyde på at man ikke forventer forskjeller i nivå på vurderinger foretatt ved bruk av de to skalaene. Det finnes heller ikke støtte i litteraturen for at slike forskjeller skal eksistere. Med utgangspunkt i disse “ikke-teoretiske” betraktningene har vi valgt å formulere følgende generelle hypotese i tilknytning til nivå på vurderinger avhengig av hvilken instruksjon og vurderingsskala som benyttes i spørreskjemaet:

H1: *Nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer vil ikke avvike fra nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer⁶¹*

⁶¹Hypotese 1 kan betraktes som en null-hypotese i den forstand at vi ikke forventer forskjeller mellom eksperimentgruppene. Dette får konsekvenser for testlogikken vi benytter ved hypotesetestingen. Et signifikansnivå $\alpha < 0,05$ impliserer at rapportering av signifikante forskjeller mellom eksperiment-

I kapittel 3 argumenterte vi for at det kan eksistere situasjoner der objektive eller intersubjektive standarder som grunnlag for vurderinger eksisterer. Slike standarder kan f.eks. skapes av tilbyderer gjennom informasjon til markedet, som kan føre til at det etableres en felles norm for vurderinger. Bruk av objektforankret (“bra-dårlig”) instruksjon og -vurderingsskala inviterer respondentene til å basere sine vurderinger på objektive eller intersubjektive standarder, i den grad slike eksisterer. Ved bruk av subjektforankret (“tilfreds-tilfreds”) instruksjon og -vurderingsskala blir vurderingene knyttet nærmere respondentens indre følelser overfor sin egen erfaring med produktet, og vurderingene kan dermed bli mer subjektive. Bruk av objektforankrete instruksjoner og -vurderingsskalaer impliserer på bakgrunn av dette en høyere konsistens i respondentenes vurderinger enn bruk av subjektforankrete instruksjoner og -vurderingsskalaer, i situasjoner der objektive eller intersubjektive standarder eksisterer.

Når objektive eller intersubjektive standarder ikke eksisterer, kan bruk av objektforankrete instruksjoner og vurderingsskalaer medføre usikkerhet med hensyn til hvilken norm kundene skal bruke som grunnlag for vurderinger. Denne usikkerheten kan bidra til å redusere konsistensen i objektforankrete vurderinger. Det kan derfor være vanskelig å si noe om retning på forskjellene i konsistens mellom de to skalaene dersom man ikke vet hvorvidt objektive eller intersubjektive standarder eksisterer i markedet. Vi ønsker ikke å studere dette spørsmålet i denne studien, og velger derfor å formulere følgende hypotese i tilknytning til konsistens⁶² i vurderinger.

H2: *Bruk av objektforankrete instruksjoner og -vurderingsskalaer vil føre til høyere konsistens i respondentenes vurdering av spesifikke produktegenskaper enn bruk av subjektforankrete instruksjoner og -vurderingsskalaer*

Selv om vi i hypotese 1 har fremmet påstand om at det ikke vil være forskjeller i nivå på vurderinger foretatt med bakgrunn i objektforankret- og subjektforankret instruksjon og vurderingsskala argumenterte vi i kapittel 3 for at det kan finnes situasjoner der nivået på vurderinger ved bruk av de to instruksjonene og vurderingsskalaene vil være forskjellig. En slik situasjon kan oppstå når vurderinger omfatter produkter med forskjellig kvalitetsnivå.

grupper med liten sannsynlighet kan antas å skyldes tilfeldigheter. For testing av hypotese 1 vil eventuelle forskjeller mellom eksperimentgruppene derfor betraktes som tilfeldige dersom rapporterte p-verdier er høyere enn valgt signifikansnivå.

⁶²Konsistens i vurderinger reflekterer i hvilken grad respondentene er enige, og måles ved varians. Lav varians tilsvarer høy konsistens.

I følge assimilasjons-kontrasteori (Sherif og Hovland, 1961)⁶³ vil kunden forstørre avvik mellom opplevd ytelse og forventninger når disse er store. Dette kalles kontrasteffekter. Når avvikene mellom opplevd ytelse og forventning er små vil vurderingene tilnærmes forventet nivå. Yi (1990) påpeker i tillegg at kontrasteffekter innebærer involvering fra respondentenes side.

Vi har i kapittel 3 argumentert for at bruk av subjektforankret instruksjon og -vurderingsskala reflekterer indre følelsesmessige tilstander, mens bruk av objektforankret instruksjon og -vurderingsskala representerer respondentens mer objektive vurdering av et objekt. Vi har også argumentert for at bruk av subjektforankrete vurderingsskalaer kan resultere i mer involverte vurderinger. Det kan derfor hevdes at kontrasteffekter spesielt vil oppstå når det benyttes subjektforankrete instruksjoner og -vurderingsskalaer. Ved bruk av objektforankrete instruksjoner og -vurderingsskalaer antas respondentene å opptre som “dommere”, og vurdere produktet uavhengig av forventninger og personlige normer (Tse og Wilton, 1988). “Bra-dårlig” benevelser kan på denne måten føre til at respondentene føler at de ikke kan bruke egne normer og forventninger som basis for vurderinger.

Dersom kundenes forventninger er konstante på tvers av produktalternativer innenfor en bestemt produktgruppe, er det sannsynlig at ytelsen for høykvalitetsprodukter vil overstige kundenes forventninger. Forventninger som er konstante på tvers av produktalternativer kan f.eks. oppstå i situasjoner med kjedesamarbeid, felles markeds-kommunikasjon og profilering der informasjonen kundene får gjelder alle produktene i en produktgruppe.

Med bakgrunn i at både høy og lav kvalitet vil føre til kontrasteffekter (Yi, 1990), kan det hevdes at nivået på tilfredshetsmål for høykvalitetsprodukter vil ligge over “bra-dårlig” mål. Nivået på tilfredshetsmål for lavkvalitetsprodukter vil på den andre siden ligge under “bra-dårlig” mål. For gjennomsnittsprodukter vil det ikke være forskjeller mellom skalaene. Disse påstandene kan formuleres i hypoteser på følgende måte:

H3a: *For lavkvalitetsprodukter vil nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer overstige nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer*

⁶³Referert av Yi (1990)

H3b: *For produkter med kvalitet lik gjennomsnittet, vil nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer ikke avvike fra nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer*

H3c: *For høykvalitetsprodukter vil nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer overstige nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer*

Vi har i teoridrøftelsen argumentert for at bruk av de to instruksjonene og vurderingsskalaene pålegger respondentene forskjellige kognitive oppgaver. I kapittel 3 benyttet vi sentrale teoribidrag fra informasjonsintegrasjonsteori, holdningsteori og teori om kundetilfredshet, for å fremsette en modell for hvordan kunder kan tenkes å foreta produktvurderinger etter kjøp og bruk av produktet. Vi delte modellen inn i en kognitiv, en affektiv og en konativ komponent, se f.eks (Lutz, 1991).

Et sentralt problemområde innenfor teorier om holdningsdannelse og vurderinger har vært modellenes prediktive validitet. Med prediktiv validitet vil vi i denne avhandlingen forstå hvor godt vurdering av spesifikke egenskaper ved et produkt forklarer totalvurderinger og intensjoner om atferd eller faktisk atferd f.eks. gjenkjøp og “word of mouth”. Et sentralt forhold som kan få betydning for målingenes prediktive validitet i holdningsmodeller er hvorvidt holdningene er tilgjengelige for individet (Fazio, 1989). I kapittel 3 hevdet vi at en respondentens tilgjengelighet til sine holdninger og vurderinger ville være reflektert i bruk av vurderingsskala. Mer spesifikt antar vi at subjektforankrete vurderinger vil være mer tilgjengelige for respondenten enn objektforankrete vurderinger. Dette impliserer at vi forventer å finne høyere prediktiv validitet for “tilfreds-utilfreds” skalaer enn for “bra-dårlig”skalaer.

I kapittel 3 argumenterte vi i tillegg for at subjektforankrete vurderingsskalaer måler vurderinger som er mer fremtredende, basert på direkte erfaring, representerer høyere grad av engasjement og er mer rettet mot respondentenes selvoppfatning enn objektforankrete vurderingsskalaer. Disse forholdene taler også for at vi skal kunne forvente de samme effektene for prediktiv validitet.

Til slutt sammenlignet vi de to instruksjonene og vurderingsskalaene med Fishbein og Ajzen's (1975) skille mellom holdning til et objekt, og holdning til handling overfor objektet. Vi relaterte "bra-dårlig" vurderinger til *holdninger til et objekt*, mens "tilfreds-tilfreds" skalaen ble antatt å være nærmere *holdning til handling overfor objektet*. Også dette peker i retning av at subjektforankrete vurderinger vil ha høyere prediktiv validitet enn objektforankrete vurderinger. Basert på litteraturgjennomgangen har vi valgt å fremsette følgende hypoteser:

H4a: *Vurderinger av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer vil ha større evne til å forklare totalvurderinger enn vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer*

H4b: *Vurderinger av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer vil ha større evne til å forklare atferdsintensjoner enn vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer*

I kapittel 3 beskrev vi et produkt som et sett av egenskaper, der tilstedeværelsen av spesielle egenskaper blir antatt å være relatert til kundens oppfattede nytte ved et spesifikt alternativ (Lancaster, 1966; Lancaster, 1971). Ledelsen og kundene bruker produktegenskapene for ulike formål. Fra et ledelses perspektiv er egenskapene et resultat av produksjonen, mens kundene oppfatter egenskaper som bidrar til deres nytte av produktet (Lancaster, 1966; Lancaster, 1971). Vi gjorde et skille mellom egenskaper og holdepunkter for å indikere at holdepunkter er direkte observerbare, som f.eks. brede dekk og langt panser, mens egenskaper er det som oppfattes og erfares av kunden, f.eks. gode kjøreegenskaper (Steenkamp, 1989).

Egenskaper ved et produkt kan være både materielle (fysiske) og immaterielle (service) (Levitt, 1981; Shostack, 1977). I tjenestemarkedsføringslitteraturen fremheves det at tjenester er vanskeligere for kunden å vurdere enn fysiske varer. En av årsakene til dette hevdes å være at tjenester er immaterielle (Zeithaml, 1981). Materielle egenskaper ved et produkt ble i kapittel 3 relatert til iboende holdepunkter, mens immaterielle egen-

skaper ble utledet av utvendige holdepunkter (Olson, 1977). I følge Olson (1977) og Szybillo og Jacoby (1974) vil iboende holdepunkter ha større sannsynlighet for å bli tilegnet av konsumenten, og dersom benyttet, ha større påvirkning på produktvurderinger enn utvendige holdepunkter.

Dersom vi relaterer disse perspektivene til tjenestemarkedsføringens skille mellom materielle og immaterielle egenskaper, og de antatte vanskelighetene kunder har ved vurdering av immaterielle egenskaper, kan det hevdes at det vil bli lagt større vekt på materielle egenskaper enn immaterielle egenskaper ved produktvurderinger (Zeithaml, 1981).

Vi har tidligere argumentert for at det kan være lettere for en kunde å foreta "tilfreds-tilfreds" vurderinger enn "bra-dårlig" vurderinger. Dette ble begrunnet med at subjektforankrete vurderinger vil være mer tilgjengelige for kundene enn objektforankrete vurderinger. Videre ble det argumentert for at "bra-dårlig" vurderinger er generelle og indirekte, mens rapportering av indre følelsesmessige tilstander, som f.eks. grad av tilfredshet, er mer direkte og spesiell, og basert på spesifikke erfaringer.

Self-awareness teori gir også en viss støtte for at det kan være lettere for kunden å foreta "tilfreds-tilfreds" vurderinger enn "bra-dårlig" vurderinger. Fokusering av oppmerksomhet mot seg selv og egne følelser ved bruk av "tilfreds-tilfreds" mål (f.eks. "jeg er tilfreds") kan være en lettere oppgave for kunden å rapportere, enn om oppmerksomheten blir styrt mot eksterne faktorer ved bruk av "bra-dårlig" mål (f.eks. "er det bra eller dårlig?").

Basert på sammenligningene vi gjorde mellom preferansevurderinger og subjektforankrete vurderinger, og mellom sammenligningsvurderinger og objektforankrete vurderinger i kapittel 3, argumenterte vi til slutt for at materielle egenskaper vil være viktigere og bety mer for kundens totalvurdering og atferdsintensjoner ved bruk av objektforankret instruksjon og -vurderingsskala. Når det benyttes subjektforankret instruksjon og -vurderingsskala vil forskjellen i betydning mellom materielle- og immaterielle egenskaper være mindre.

Med utgangspunkt i perspektivene fra teoridrøftelsen hevder vi at de antatte problemene kundene har ved vurdering av immaterielle egenskaper vil være mindre fremtredende ved bruk av subjektforankrete vurderingsskalaer. Dette vil gi seg utslag i den prediktive verdi de immaterielle egenskapene vil ha for kundens totalvurdering. I hypoteseform kan dette uttrykkes på følgende måte:

H5a: *Når det benyttes subjektforankrete instruksjoner og -vurderingsskalaer i tilknytning til spesifikke vurderings-spørsmål i et spørreskjema, vil immaterielle egenskaper ved produktet ha sterkere relativ betydning for prediksjon av totalvurderinger enn ved bruk av objektforankrete instruksjoner og -vurderingsskalaer*

H5b: *Når det benyttes subjektforankrete instruksjoner og -vurderingsskalaer i tilknytning til spesifikke vurderings-spørsmål i et spørreskjema, vil immaterielle egenskaper ved produktet ha sterkere relativ betydning for prediksjon av atferdsintensjoner enn ved bruk av objektforankrete instruksjoner og -vurderingsskalaer*

Et viktig problem man har vært opptatt av ved måling av holdninger er at individer med dårlig utviklede, svake og ikke eksisterende holdninger ofte svarer på en måte som indikerer enten negative eller positive holdninger (Sample og Warland, 1973). Når disse individene blir inkludert i utvalget i en undersøkelse, kan styrken i sammenhengene mellom holdninger og atferdsintensjoner bli underestimert. Vi har tidligere argumentert for at subjektforankrete vurderinger vil være mer tilgjengelige for respondenten enn objektforankrete vurderinger. Tilgjengelighet til holdninger vil videre være avhengig av om holdninger er utviklet. Sikkerhet i vurdering av et objekt vil dermed være avhengig av i hvilken grad et individ har utviklet holdninger overfor objektet. På bakgrunn av dette kan det argumenteres for at respondentene vil være sikrere i sine vurderinger når de benytter en subjektforankret vurderingsskala.

I kapittel 3 ble det også argumentert for at objektforankrete vurderinger er generelle og indirekte, mens subjektforankrete vurderinger er mer direkte og spesielle og basert på spesifikke erfaringer. Subjektforankrete vurderinger kan derfor, i mange tilfeller, representere en lettere oppgave for respondentene enn objektforankrete vurderinger. Dette argumentet kan gi støtte til påstanden om at respondentene vil være sikrere på rapportering av egen tilfredshet enn vurderinger av om et produkt er godt eller dårlig

Til slutt kan det finnes en viss støtte i påstanden om at kunder vil være sikrere i "tilfreds-utilfreds" vurderinger, enn i "bra-dårlig" vurderinger med bakgrunn i self-

awareness teori (Duval og Wicklund, 1972)⁶⁴. Ved fokusering av oppmerksomhet mot seg selv og sine følelser, kan det argumenteres for at kunden vil være sikrere i sine vurderinger enn om oppmerksomheten blir styrt mot eksterne faktorer. På bakgrunn av argumentasjonen fra den teoretiske drøftingen i kapittel 3, kan dermed følgende hypotese vedrørende respondentenes sikkerhet på vurderinger av spesifikke produkt-egenskaper formuleres:

H6: *Respondenter vil være sikrere på vurdering av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer enn vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer*

Kundens produktkunnskap og erfaring med bruk av produktet kan påvirke både vurderinger (Alba og Hutchinson, 1987; Bettman og Park, 1980; Sujan, 1985) og respons i spørreskjemaet (Hippler, Schwarz og Sudman, 1987). I kapittel 3 relaterte vi kundens produktkunnskap til produktvurderinger gjennom holdningsteori og teorier for kundetilfredshet. I holdningsteori gjøres det et skille mellom kognitive, affektive og atferdsmessige komponenter. Antakelser om-, eller vurderinger av egenskaper ved produktet, og tillegging av viktighet til egenskapene utgjør individets kognitive strukturer, mens prosessen der de kognitive strukturene omsettes i en totalvurdering og holdning overfor produktet representerer de kognitive prosessene. Vi benyttet disse perspektivene fra holdningslitteraturen for å bygge en modell for måling av kunders produktvurderer etter kjøp og bruk av produktet (se figur 3.5). Erfaring med bruk av produktet kan påvirke både kognitive strukturer (Alba og Hutchinson, 1987) og kognitive prosesser (Alba og Hutchinson, 1987; Sujan, 1985). Forskning knyttet til teorier om holdningsdannelse har vist at et individ må ha en viss *tilgjengelighet* til sine holdninger for at disse skal virke styrende for atferd (Fazio, 1989). Økt brukserfaring og produktkunnskap er sentrale forhold som kan bidra til at holdninger og produktvurderinger blir mer tilgjengelige. Holdninger eller produktvurderinger basert på erfaring, vil også med større sannsynlighet påvirke atferd (Fazio, 1989; Zanna, Olson, og Fazio, 1980).

Foran, i hypotese 4a og 4b, fremsatte vi en påstand om at bruk av subjektforankrete, ("tilfreds-utifreds") instruksjoner og -vurderingsskalaer i spørreskjemaet ved måling av kundenes produktvurderinger vil gi sterkere prediksjoner på totalvurdering og atferds-

⁶⁴Referert av (Wilson 1990)Wilson (1990)

intensjoner enn objektforankrete (“bra-dårlig”), instruksjoner og vurderingsskalaer. Denne hypotesen er blant annet begrunnet i at “tilfreds-tilfreds” vurderinger vil være mer tilgjengelige for kundene enn “bra-dårlig” vurderinger. I kapittel 3 argumenterte vi for at et individs tilgjengelighet til sine holdninger kan virke styrende for hvorvidt en bestemt holdning vil bli aktivert (Fazio, 1989). Økt brukserfaring og produktkunnskap kan bidra til å øke tilgjengeligheten til holdninger og produktvurderinger. Holdninger eller produktvurderinger basert på produktkunnskap vil også med større sannsynlighet påvirke atferd (Fazio, 1989; Zanna, Olson, og Fazio, 1980). Det kan derfor argumenteres for at erfaring med bruk av et produkt vil bidra til å redusere forskjellen i prediksjonsevne mellom subjektforankrete og objektforankrete vurderinger. Denne påstanden kan presiseres i følgende hypoteser:

H7a: Forskjellene mellom spesifikke vurderinger foretatt i spørreskjema med subjektforankrete- vs. objektforankrete instruksjoner og -vurderingsskalaer, når det gjelder de spesifikke vurderingenes evne til å forklare totalvurderinger, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

H7b: Forskjellene mellom spesifikke vurderinger foretatt i spørreskjema med subjektforankrete- vs. objektforankrete instruksjoner og -vurderingsskalaer, når det gjelder de spesifikke vurderingenes evne til å forklare atferdsintensjoner, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

Ovenfor, i hypotese 5a og 5b, fremmet vi påstander om at materielle egenskaper ved produktet vil ha sterkere betydning i kundenes vurderinger og ha større evne til å forklare totalvurderinger og atferdsintensjoner enn immaterielle egenskaper. Materielle egenskaper ved et produkt ble relatert til iboende holdepunkter, mens immaterielle egenskaper ble relatert til utvendige holdepunkter (Olson, 1977). Iboende holdepunkter vil ha større sannsynlighet for å bli tilegnet av konsumenten (Olson, 1977; Szybillo og Jacoby, 1974), og ha større påvirkning på produktvurderinger enn utvendige holdepunkter. Når det benyttes subjektforankret instruksjon og -vurderingsskala argumenterte vi i tillegg for at materielle- og immaterielle egenskaper vil ha tilnærmet lik betydning

for prediksjon av totalvurderinger og atferdsintensjoner. Hovedargumentet for denne antakelsen er at det generelt sett vil være lettere for respondentene å foreta "tilfreds-tilfreds" vurderinger. Produktkunnskap kan imidlertid gjøre det lettere for en respondent å foreta "bra-dårlig" vurderinger. Produktkunnskap kan dermed bidra til å redusere forskjellen i betydning av materielle- og immaterielle egenskaper for totalvurderinger og atferdsintensjoner når vi benytter en objektforankret ("bra-dårlig") måleskala. Støtte for en slik antakelse kan også finnes hos Olson (1977), der det argumenteres for at erfaring med bruk av et produkt vil påvirke den relative betydningen av iboende og utvendige holdepunkter for totalvurderinger av produktet. Basert på drøftingen av produktkunnskap som en modererende faktor kan dermed følgende hypoteser fremsettes:

H8a: Forskjellen mellom spesifikke produktvurderinger foretatt i spørreskjema med objektforankrete- vs. subjektforankrete instruksjoner og -vurderingsskalaer, når det gjelder den relative betydningen av materielle- og immaterielle egenskaper ved produktet for forklaring av totalvurderinger, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

H8b: Forskjellen mellom spesifikke produktvurderinger foretatt i spørreskjema med objektforankrete- vs. subjektforankrete instruksjoner og -vurderingsskalaer, når det gjelder den relative betydningen av materielle- og immaterielle egenskaper ved produktet for forklaring av atferdsintensjoner, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

I hypotese 6 hevdet vi at kunder som foretar subjektforankrete vurderinger vil være sikrere i sine vurderinger enn kunder som blir bedt om å foreta objektforankrete vurderinger. Som grunnlag for hypotesen ble det argumentert for at subjektforankrete vurderinger vil være mer tilgjengelige for kundene enn objektforankrete vurderinger, og at dette vil føre til økt sikkerhet i subjektforankrete vurderinger. Sikkerhet i vurdering av et objekt vil også være avhengig av i hvilken grad et individ har utviklet holdninger overfor objektet. Økt produktkunnskap kan være et forhold som kan bidra til å øke tilgjengeligheten til holdninger og produktvurderinger. Økt produktkunnskap kan dermed også være bestemmende for hvor sikre kundene vil være i sine vurderinger. I

kapittel 3 argumenterte vi for at respondenter med høy produktkunnskap lettere kan foreta objektforankrete vurderinger enn kunder med lav produktkunnskap. Det at kunder med høy produktkunnskap generelt antas å være sikrere i sine vurderinger av egenskaper ved produktet, enn kunder med lav produktkunnskap, kan bidra til å moderere effektene av skalavalg for sikkerhet i vurderinger. På bakgrunn av den teoretiske drøftingen, kan følgende hypoteser formuleres i forbindelse med produktkunnskap som modererende faktor for respondentenes sikkerhet i vurderinger.

H9a: *Respondenter med lav produktkunnskap vil være sikrere på sine vurderinger av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer enn vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer*

H9b: *Respondenter med høy produktkunnskap vil være tilnærmet like sikre på sine vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer som vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer*

4.2.2 Inkludering, og alternativ plassering av viktighetsspørsmål i spørreskjemaet

I metodelitteraturen, f.eks Hippler, Schwarz, og Sudman (1987) og Schuman og Presser (1981), er effekter av spørsmålsrekkefølge viet stor oppmerksomhet. Man har fokusert på hvordan forutgående spørsmål i en undersøkelse kan påvirke svargivning på etterfølgende spørsmål.

Effekter av spørsmålsrekkefølge kan relateres til *assimiasjons-* og *kontrasteffekter* (Schuman, Kalton og Ludwig, 1983; Schuman og Presser, 1981; Schuman, Presser og Ludwig, 1981)⁶⁵. Assimilasjonseffekter oppstår når respons på etterfølgende spørsmål bringes nærmere respons på tidligere spørsmål enn de ville gjort dersom de forutgående spørsmålene hadde vært utelatt. Man antar at respondenten har behov for å være

⁶⁵*Carryover-* og *backfireeffekter* (Tourangeau og Rasinski 1988, Bickart 1993)

konsistent i sin svargivning. I motsetning til assimilasjonseffekter vil kontrasteffekter føre til større forskjeller i svargivning som resultat av spørsmålsrekkefølge. Når viktighetsspørsmål plasseres før vurderingene i spørreskjemaet, kan svargivning på viktighetsspørsmålene, som et resultat av kontrast- og assimilasjonseffekter kunne påvirke vurderingene. Det motsatte kan være tilfelle dersom vurdering av spesifikke produktegenskaper foretas før man tillegger egenskapene vekt gjennom selvrapporterte viktighetsspørsmål. Assimilasjons- og kontrasteffekter kan i første rekke tenkes å påvirke nivået på viktighetsspørsmål og vurderingsspørsmål.

Til tross for den relativt grundige drøftingen av assimilasjons- og kontrasteffekter vi gjennomførte i kapittel 3, som grunnlag for å belyse effekter av spørsmålsrekkefølge, er det vanskelig å predikere retning på disse effektene. Vi velger derfor ikke å fremsette generelle hypoteser i tilknytning til nivå på vurderingsspørsmål og viktighetsspørsmål. I kapittel 8 vil vi imidlertid foreta sammenligninger av nivå både på vurderingsspørsmål og på viktighetsspørsmål, avhengig av hvor disse er plassert i spørreskjemaet. Disse eksplorative analysene vil danne grunnlag for diskusjon i kapittel 9.

I kapittel 3 introduserte vi kvalitetsnivå på produktet som vurderes, som modererende variabel for nivået på vurderinger, avhengig av om viktighetsspørsmål inkluderes, og i tilfelle hvor disse plasseres i forhold til spesifikke vurderingsspørsmål. Vi diskuterte hvordan svargivning på viktighetsspørsmål kan aktivere sammenligningsstandarder eller idealprodukter. Sammenligningsstandardene kan bli benyttet når respondentene foretar vurderinger av produktet. Mer spesifikt antar vi at en assimilasjonseffekt kan oppstå når en respondent blir bedt om å vurdere et høykvalitetsprodukt etter at en sammenligningsstandard, eller et idealprodukt, er aktivert gjennom svargivning på forutgående viktighetsspørsmål. En begrunnelse for dette er at et høykvalitetsprodukt ytelsesmessig vil ligge nær idealproduktet eller sammenligningsstandard. For høykvalitetsprodukter, antar vi at nivået på vurderinger av spesifikke produktegenskaper der respondentene først har tillagt egenskapene viktighet, vil overstige, eller være tilnærmet lik nivået på vurderinger foretatt i et spørreskjema der viktighetsspørsmålene er utelatt.

Når et produkt kan kategoriseres som et lavkvalitetsprodukt, vil dette ytelsesmessig avvike fra eventuelle sammenligningsstandarder eller idealprodukter. Når viktighets-spørsmål inkluderes og blir plassert før de spesifikke vurderingsspørsmålene antar vi at slike sammenligningsstandarder, eller idealprodukter blir aktivert. For lavkvalitetsprodukter vil derfor mest sannsynlig en kontrasteffekt oppstå når viktighetsspørsmål plasseres før de spesifikke vurderingsspørsmålene i spørreskjemaet. Nivået på vurderinger av spesifikke produktegenskaper antar vi her vil være lavere enn om viktig-

hetsspørsmålene hadde vært utelatt, eller plassert etter de spesifikke vurderings-spørsmålene i spørreskjemaet. Med dette som utgangspunkt kan vi nå formulere følgende hypoteser:

H10a: *For lavkvalitetsprodukter vil nivå på vurderinger av spesifikke produktegenskaper, foretatt i spørreskjema der viktighetsspørsmål er utelatt, overstige nivå på vurderinger av spesifikke produktegenskaper foretatt i spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene (kontrasteffekt)*

H10b: *For produkter med kvalitet lik gjennomsnittet, vil nivå på vurderinger av spesifikke produktegenskaper, foretatt i spørreskjema der viktighetsspørsmål er utelatt, ikke avvike fra nivå på vurderinger av spesifikke produktegenskaper foretatt i spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene*

H10c: *For høykvalitetsprodukter vil nivå på vurderinger av spesifikke produktegenskaper, foretatt i spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene, overstige eller være tilnærmet lik nivå på vurderinger av spesifikke produktegenskaper foretatt i spørreskjema der viktighetsspørsmål er utelatt (assimilasjonseffekt)*

Nivå, eller gjennomsnitt, representerer aggregerte størrelser som sier noe om svar-mønster i et utvalg respondenter sett under ett. Når vi ønsker å studere forskjeller i sammenhenger mellom rapportert viktighet ved ulike produktegenskaper, og vurdering av de samme egenskapene avhengig av rekkefølgen på spørsmålene, må dette skje på individnivå.

I kapittel 3 ble det argumentert for at viktighet knyttet til spesifikke produktegenskaper er individspesifikk og knyttet til individuelle preferanser, mens vurderinger i større grad er knyttet til objektene som vurderes. Med dette som utgangs-punkt antar vi at det er mer sannsynlig at vurderinger vil bli påvirket av respondentenes rapportering av

viktighet, enn at rapportering av viktighet vil bli påvirket av vurderingene. Rapportering av viktighet vil på denne måten ha større diagnoseverdi (Feldman og Lynch, 1988) enn vurderingene. Følgende hypotese kan dermed formuleres i tilknytning til sammenhenger mellom viktighet og vurderinger.

H11: *Sammenhengen mellom viktighet og vurderinger vil være sterkere når viktighetsspørsmål er plassert før spesifikke vurderingss spørsmål i spørreskjemaet, enn når viktighets spørsmål er plassert etter de spesifikke vurderings spørsmålene*

Vi har tidligere argumentert for at en respondents tilgjengelighet (Fazio, 1989) til vurderinger av egenskaper vil øke etterhvert som egenskapene blir mer fremtredende (Bodenhausen og Wyer, 1987; Schuman og Presser, 1981). Forutgående spørsmål regnes ofte som den viktigste kilde til aktivering av informasjon som er relevant for å svare på etterfølgende spørsmål (Strack, 1991; Strack og Martin, 1987). Den psykologiske prosessen som oppstår i slike tilfeller kalles *priming* eller fremkalling (Fiske og Taylor, 1991; Strack og Martin, 1987). Fremkalling kan bidra til aktivering av ulike typer informasjon som kan være relevant for respondentens svargivning (Strack, 1991; Strack og Martin, 1987). Når kundens tilgjengelighet til vurderingene øker, f.eks. gjennom svar på tidligere spørsmål i spørreskjemaet vil sammenhengene mellom vurderinger av egenskaper, totalvurdering og atferdsintensjoner forsterkes (Fazio, 1989). Vi velger å presisere effektene av spørsmålsrekkefølge for de spesifikke vurderingss spørsmålenes prediktive validitet i følgende hypoteser:

H12a: *Målinger der viktighetsspørsmål er plassert før spesifikke vurderingss spørsmål i spørreskjemaet vil gi bedre forklaring på totalvurderinger enn målinger der viktighetsspørsmålene er plassert etter de spesifikke vurderingss spørsmålene*

H12b: *Målinger der viktighetsspørsmål er plassert før spesifikke vurderingss spørsmål i spørreskjemaet vil gi bedre forklaring på atferdsintensjoner enn målinger der viktighets spørsmålene er plassert etter de spesifikke vurderings spørsmålene*

Respondenter med høy produktkunnskap antas å ha vurderinger som er mer tilgjengelige enn det som er tilfellet for kunder med lav produktkunnskap (Fazio, 1989). Dette fører til at påvirkningen fra tidligere spørsmål i et spørreskjema antas å være mindre for disse respondentene. Diagnoseverdien av viktighetsspørsmålene for vurderingene, representert ved sammenhengen mellom viktighet og vurderinger antar vi dermed vil være sterkere for respondenter med lav produktkunnskap.

Et forhold som styrker disse antakelsene er at respondenter med høy produktkunnskap også generelt sett er funnet å være mindre påvirkelige av kontekst enn respondenter med lav produktkunnskap (Feldman og Lynch, 1988). Med bakgrunn i dette har vi valgt å formulere følgende hypotese i forbindelse med betydningen av respondentenes produktkunnskap for plassering av viktighetsspørsmål og spesifikke vurderingssørsmål i et spørreskjema:

H13: *Forskjellen mellom målinger der viktighetsspørsmål er plassert før, vs. målinger der viktighetsspørsmål er plassert etter spesifikke vurderingssørsmål i et spørreskjema, når det gjelder sammenhengen mellom viktighetsspørsmål og vurderingssørsmål vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap*

Når kundens tilgjengelighet til vurderingene øker, f.eks. gjennom svar på tidligere spørsmål i spørreskjemaet vil fremkallingseffekter kunne oppstå og sammenhengene mellom vurdering av spesifikke egenskaper, totalvurdering og atferdsintensjoner forsterkes (Fazio, 1989). I hypotese 12a og 12b fremsatte vi påstander om at målinger der viktighetsspørsmålene er plassert før de spesifikke vurderingene i spørreskjemaet vil gi sterkere prediksjoner for totalvurderinger og atferdsintensjoner.

Fordi respondenter med høy produktkunnskap antas å ha vurderinger som er mer tilgjengelige, vil sammenhengen mellom vurderinger av spesifikke egenskaper, totalvurderinger og atferdsintensjoner, være sterkere for disse respondentene enn for respondenter med lav produktkunnskap.

På bakgrunn av teoridrøftelsene kan vi dermed formulere følgende hypotese vedrørende betydningen av respondentenes produktkunnskap for plassering av viktighetsspørsmål og spesifikke vurderingssørsmål i et spørreskjema:

H14a: Forskjellen mellom målinger der viktighetsspørsmål er plassert før, vs. målinger der viktighetsspørsmål er plassert etter spesifikke vurderingsspørsmål i et spørreskjema, når det gjelder de spesifikke vurderingsspørsmålenes evne til å forklare totalvurderinger, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

H14b: Forskjellen mellom målinger der viktighetsspørsmål er plassert før, vs. målinger der viktighetsspørsmål er plassert etter spesifikke vurderingsspørsmål i et spørreskjema, når det gjelder de spesifikke vurderingsspørsmålenes evne til å forklare atferdsintensjoner, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

4.3 Oversikt over hypotesene

Hypotesene ble organisert i to hovedblokker. I den første hovedblokken presenterte vi hypoteser som omhandler effekter i tilknytning til bruk av objektforankret- vs. subjektforankret instruksjon og -vurderingsskala. I den andre hovedblokken fokuserte vi på problemstillingene i tilknytning til tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet.

For å gi en bedre oversikt over hypotesene oppsummeres disse i tabellform nedenfor. I tabellene indikerer > og < retning på hypotesene, mens = betyr at vi ikke forventer forskjeller mellom eksperimentgruppene.

Først, i tabell 4.1 oppsummeres hypoteser som omhandler betydningen av valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer for nivå og konsistens for spesifikk vurderingsspørsmål. Som det fremgår av tabellen forventer vi ingen forskjeller i nivå på de spesifikke vurderingene. Når det gjelder konsistens i vurderingene forventer vi at objektforankrete vurderinger vil være mer konsistente enn subjektforankrete vurderinger.

TABELL 4.1

Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderingsspørsmål - Nivå og konsistens i vurderinger

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | | |
|----------|--------------------------|--------------------------------|---|-----------------|
| | | INSTRUKSJON OG VURDERINGSSKALA | | |
| H1 | Nivå på vurderinger | Subjektforankret | ≈ | Objektforankret |
| H2 | Konsistens i vurderinger | Subjektforankret | < | Objektforankret |

> < : retning på hypoteser, ≈ : ingen forskjell mellom eksperimentgruppene

Vi har argumentert for at kvalitetsnivået på produktene som vurderes kan påvirke nivå på vurderinger foretatt i spørreskjema der det benyttes henholdsvis subjektforankrete- og objektforankrete instruksjoner og vurderingsskalaer. I tabell 4.2 oppsummeres hypoteser som omhandler disse effektene.

TABELL 4.2

Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderingsspørsmål - Nivå på vurderinger, kvalitetsnivå på produktet som vurderes som modererende variabel

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | | |
|----------|--|--------------------------------|---|-----------------|
| | | INSTRUKSJON OG VURDERINGSSKALA | | |
| H3a | Nivå på vurderinger for lavkvalitetsprodukter | Subjektforankret | < | Objektforankret |
| H3b | Nivå på vurderinger for gjennomsnittsprodukter | Subjektforankret | ≈ | Objektforankret |
| H3c | Nivå på vurderinger for høykvalitetsprodukter | Subjektforankret | > | Objektforankret |

> < : retning på hypoteser, ≈ : ingen forskjell mellom eksperimentgruppene

Prediktiv validitet og sikkerhet i vurderinger er sentrale avhengige variabler som vi antar påvirkes av hvilken instruksjon og vurderingsskala som benyttes i forbindelse med de spesifikke vurderingsspørsmålene i et spørreskjema.

I tabell 4.3, oppsummeres først hypoteser som omhandler de spesifikke produktvurderingenes evne til å forklare respondentenes totalvurderinger og atferdsintensjoner. Deretter oppsummeres hypoteser for respondentenes sikkerhet i vurderinger. For disse hypotesene forventer vi høyere prediksjonsevne og høyere sikkerhet i vurderingene når det benyttes subjektforankrete instruksjoner og -vurderingsskalaer i forbindelse med de spesifikke vurderingsspørsmålene i spørreskjemaet

TABELL 4.3

Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderingsspørsmål - Prediktiv validitet og sikkerhet i vurderinger

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | |
|----------|---|--------------------------------|-------------------|
| | | INSTRUKSJON OG VURDERINGSSKALA | |
| H4a | Spesifikke vurderingers evne til å forklare totalvurdering | Subjektforankret | > Objektforankret |
| H4b | Spesifikke vurderingers evne til å forklare atferdsintensjoner | Subjektforankret | > Objektforankret |
| H5a | Immaterielle egenskapers betydning for forklaring av totalvurdering | Subjektforankret | > Objektforankret |
| H5b | Immaterielle egenskapers betydning for forklaring av atferdsintensjoner | Subjektforankret | > Objektforankret |
| H6 | Sikkerhet i vurderinger | Subjektforankret | > Objektforankret |

> < : retning på hypoteser, ≈ : ingen forskjell mellom eksperimentgruppene

Som det fremgår av tabell 4.3, er hypotesene formulert uten at vi tar hensyn til respondentenes produktkunnskap. I kapittel 3 ble nettopp produktkunnskap trukket frem som et sentralt forhold som kan moderere forskjellene vi forventer å finne når det benyttes subjektforankrete- og objektforankrete instruksjoner og -vurderingsskalaer i forbindelse med spesifikke vurderingsspørsmål i et spørreskjema.

Hypoteser der respondentenes produktkunnskap inkluderes som modererende variabel for effektene vi forventer som følge av valg mellom å benytte objektforankrete- og subjektforankrete- instruksjoner og -vurderingsskalaer, oppsummeres i tabell 4.4.

Først oppsummeres hypoteser som omhandler betydningen av produktkunnskap for de spesifikke vurderingers evne til å forklare respondentenes totalvurderinger og atferdsintensjoner. Deretter oppsummeres hypoteser der vi ser nærmere på den relative betydningen av immaterielle- og materielle produkttegnegenskaper, både for totalvurderinger og atferdsintensjoner innenfor ulike nivå av produktkunnskap. Til slutt inkluderes respondentenes produktkunnskap som modererende variabel for hypotesene som omhandler respondentenes sikkerhet i vurderingene de foretar ved å benytte de to instruksjonene og vurderingsskalaene vi fokuserer på i avhandlingen.

TABELL 4.4

Hypoteser i tilknytning til valg mellom objektforankret-og subjektforankret instruksjon og vurderingsskala for spesifikke vurderingsspørsmål - Prediktiv validitet og sikkerhet i vurderinger, respondentenes produktkunnskap som modererende variabel

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | | | |
|----------|--|--------------------------------|------------------|---|-----------------|
| | | INSTRUKSJON OG VURDERINGSSKALA | | | |
| H7a | Spesifikke vurderingers evne til å forklare totalvurdering: | Lav produktkunnskap | Subjektforankret | > | Objektforankret |
| | | Høy produktkunnskap | Subjektforankret | ≈ | Objektforankret |
| H7b | Spesifikke vurderingers evne til å forklare atferdsintensjoner: | Lav produktkunnskap | Subjektforankret | > | Objektforankret |
| | | Høy produktkunnskap | Subjektforankret | ≈ | Objektforankret |
| H8a | Immaterielle egenskapers betydning for forklaring av totalvurdering: | Lav produktkunnskap | Subjektforankret | > | Objektforankret |
| | | Høy produktkunnskap | Subjektforankret | ≈ | Objektforankret |
| H8b | Immaterielle egenskapers betydning for forklaring av atferdsintensjoner: | Lav produktkunnskap | Subjektforankret | > | Objektforankret |
| | | Høy produktkunnskap | Subjektforankret | ≈ | Objektforankret |
| H9a, H9b | Sikkerhet i vurderinger | Lav produktkunnskap | Subjektforankret | > | Objektforankret |
| | | Høy produktkunnskap | Subjektforankret | ≈ | Objektforankret |

> < : retning på hypoteser, ≈ : ingen forskjell mellom eksperimentgruppene

Den andre hovedblokken av hypoteser er knyttet til inkludering og alternativ plassering av viktighetsspørsmål i et spørreskjema. I tabell 4.5 har vi først oppsummert hypoteser som omhandler forskjeller i nivå på vurderinger, når kvalitetsnivået på produktet som vurderes betraktes som modererende variabel. Tabellen viser at vi forventer lavere nivå på vurderinger av lavkvalitetsprodukter når viktighetsspørsmål inkluderes og plasseres før de spesifikke vurderingsspørsmålene i spørreskjemaet. For høykvalitetsprodukter forventer vi høyere nivå på vurderingene når viktighetsspørsmål inkluderes i spørreskjemaet.

TABELL 4.5

Hypoteser i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderings-spørsmål i et spørreskjema - Nivå på vurderinger, kvalitetsnivå på produktet som vurderes som modererende variabel

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | |
|----------|--|---------------------------------|----------------------------|
| | | PLASSERING AV VIKTIGHETSPØRSMÅL | |
| H10a | Nivå på vurderinger for lavkvalitetsprodukter | Før spesifikke vurderinger | < Uten viktighets-spørsmål |
| H10b | Nivå på vurderinger for gjennomsnittsprodukter | Før spesifikke vurderinger | = Uten viktighets-spørsmål |
| H10c | Nivå på vurderinger for høykvalitetsprodukter | Før spesifikke vurderinger | > Uten viktighets-spørsmål |

> < : retning på hypoteser, = : ingen forskjell mellom eksperimentgruppene

Sammenhenger mellom viktighet og vurderinger og prediktiv validitet er sentrale avhengige variabler i hypoteser som omhandler inkludering og plassering av viktighets-spørsmål i et spørreskjema ved operasjonalisering av multiattributtmodeller. I tabell 4.6 nedenfor presenteres først hypoteser for sammenhenger mellom viktighetsspørsmål og vurderingss spørsmål. Deretter oppsummeres hypoteser som omhandler de spesifikke vurderingenes prediktive validitet.

TABELL 4.6

Hypoteser i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderings-spørsmål i et spørreskjema - Sammenheng mellom viktighet og vurderinger og prediktiv validitet

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | |
|----------|---|---------------------------------|--------------------------------|
| | | PLASSERING AV VIKTIGHETSPØRSMÅL | |
| H11 | Sammenheng mellom viktighet og vurdering | Før spesifikke vurderinger | > Etter spesifikke vurderinger |
| H12a | Spesifikke vurderingers evne til å forklare total-vurdering | Før spesifikke vurderinger | > Etter spesifikke vurderinger |
| H12b | Spesifikke vurderingers evne til å forklare atferds-intensjoner | Før spesifikke vurderinger | > Etter spesifikke vurderinger |

> < : retning på hypoteser, = : ingen forskjell mellom eksperimentgruppene

Produktkunnskap er også trukket frem som et viktig kjennetegn ved respondentene som kan bidra til å moderere effektene vi forventer, som resultat av å inkludere og plassere viktighetsspørsmålene, før vs. etter vurderingss spørsmålene i spørreskjemaet, både for

sammenhenger mellom viktighet og vurderinger, og for de spesifikke vurderingenes prediktive validitet. Til slutt, i tabell 4.7, oppsummeres hypoteser der respondentenes produktkunnskap inkluderes som en modererende variabel.

TABELL 4.7

Hypoteser i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderings-spørsmål i et spørreskjema - Sammenheng mellom viktighet og vurderinger og prediktiv validitet, respondentenes produktkunnskap som modererende variabel

| HYPOTESE | AVHENGIG VARIABEL | UAVHENGIG VARIABEL | |
|----------|---|----------------------------------|--------------------------------|
| | | PLASSERING AV VIKTIGHETSSPØRSMÅL | |
| H13 | Sammenheng mellom viktighet og vurdering: | | |
| | Lav produktkunnskap | Før spesifikke vurderinger | > Etter spesifikke vurderinger |
| | Høy produktkunnskap | Før spesifikke vurderinger | ≈ Etter spesifikke vurderinger |
| H14a | Spesifikke vurderingers evne til å forklare totalvurdering: | | |
| | Lav produktkunnskap | Før spesifikke vurderinger | > Etter spesifikke vurderinger |
| | Høy produktkunnskap | Før spesifikke vurderinger | ≈ Etter spesifikke vurderinger |
| H14b | Spesifikke vurderingers evne til å forklare atferdsintensjoner: | | |
| | Lav produktkunnskap | Før spesifikke vurderinger | > Etter spesifikke vurderinger |
| | Høy produktkunnskap | Før spesifikke vurderinger | ≈ Etter spesifikke vurderinger |

> < : retning på hypoteser, ≈ : ingen forskjell mellom eksperimentgruppene

5

Undersøkelsesopplegg

I dette kapittelet vil vi beskrive det metodologiske opplegget som ble benyttet for innsamling av data som grunnlag for å teste hypotesene vi formulerte i kapittel 4. Innledningsvis vil vi imidlertid først foreta en drøfting av sentrale krav som ble stillt til undersøkelsesopplegget, for at dette skal gi et informasjonsgrunnlag som hypotesene kan testes ut fra. Deretter beskriver vi den metodologiske tilnærmingen som ble benyttet for innsamling av data. Til slutt gir vi en beskrivelse av datamaterialet som ble samlet inn.

5.1 Krav til undersøkelsesopplegget

Utforming av forskningsdesign eller undersøkelsesopplegg innebærer en spesifisering av regler og prosedyrer for å innhente den ønskede informasjonen. Hensikten med en slik forhånds-utforming av undersøkelsesopplegget, er innenfor de rammer vi står overfor, å sikre så pålitelig informasjon som mulig for å belyse problemstillingene (Grønhaug, 1985). Undersøkelsesopplegget må altså utformes på bakgrunn av de problemstillinger vi har, og de hypoteser som vi ønsker å teste. Ved utforming av undersøkelsesopplegg må vi i tillegg ta hensyn til hvilket gyldighetsområde vi ønsker at resultatene skal ha. Valg av undersøkelsesopplegg blir dermed styrt av ulike hensyn til validitet.

Validitet angir hvorvidt man har målt hva man har til hensikt å måle, altså om innsamlete data er relevante for problemstillingene som er formulert (Cook og Campbell,

1979). Det har vært vanlig å skille mellom fire hovedformer for validitet (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979): 1) Ekstern validitet, 2) Intern validitet, 3) Statistisk konklusjonsvaliditet og 4) Begrepsvaliditet. Det kan være vanskelig å tilgodese de ulike validitetskravene like mye i en og samme studie. Ulike validitetshensyn vil ofte gå på bekostning av hverandre. Det er derfor viktig å være eksplisitt i forhold til hvilke validitetskrav vi ønsker å tilgodese (Cook og Campbell, 1979).

Formålet med oppgaven er å *anvende teoretiske perspektiver* som grunnlag for å forklare implikasjoner av to valg man som forsker står overfor når formålet er å måle kunders produktvurderinger ved hjelp av et spørreskjema. Vi skal altså ikke teste de teoretiske perspektivene i seg selv, men ønsker å finne ut om implikasjonene de teoretiske perspektivene foreskriver for valg av ordlyd i spørsmålsstilling, benevnelse på svarskala og spørsmålsrekkefølge inntreffer. Forskningsoppgaven vi står overfor kan oppfattes som et eksempel på det Calder, Phillips, og Tybout (1981) definerer som effektanvendelse (*effects application*). Et viktig utgangspunkt for drøfting av undersøkelsesopplegg vil dermed være å sørge for at funn kan generaliseres til en realistisk situasjon.

I våre hypoteser er vi i hovedsak interessert i å teste effektene av disse valgene, samt undersøke hvordan forhold som respondentenes produktkunnskap og kvalitetsnivået på produktene som vurderes kan moderere hovedeffektene. Vi er dermed opptatt av å fastslå *årsak-virkningsforhold* mellom utforming av spørreskjema og de respons-effektene dette kan gi opphav til. Dette innebærer at kausaltitet og intern validitet må stå sentralt i studien. Et kausalt undersøkelsesopplegg vil derfor ideelt sett peke seg ut som et naturlig valg. For testing av kausale sammenhenger er det generelt tre krav som må tilfredsstilles (Cook og Campbell, 1979; Nachmias og Nachmias, 1985; Stinchcombe, 1968):

- 1) *Det må være samvariasjon mellom de avhengige og de uavhengige variablene (Er det f.eks samvariasjon mellom endringer i ordlyd og spørsmålsrekkefølge og endringer i respons?).*
- 2) *Spuriøse sammenhenger, dvs. sammenhenger som skyldes andre uavhengige variabler enn de vi studerer, må kunne utelukkes (Finnes det andre årsaker enn ordlyd og spørsmålsrekkefølge som kan være årsak til eventuelle forskjeller i respons mellom eksperimentgrupper?).*
- 3) *Vi må kunne fastslå at de uavhengige variablene opptrer før de avhengige variablene i tid (Kommer årsaken, ordlyd og spørsmålsrekkefølge, foran virkningen i respons i tid?).*

Disse kravene blir best tilfredsstillt ved bruk av et *klassisk eksperimentdesign* (Cook og Campbell, 1979; Kidder og Judd, 1986; Nachmias og Nachmias, 1985). Klassiske eksperimentdesign består generelt av to sammenlignbare grupper, en eksperimentgruppe og en kontrollgruppe. Disse gruppene er like med unntak av at eksperimentgruppen, i motsetning til testgruppen, utsettes for et *treatment*, en behandling, som representerer en uavhengig variabel. Tilordning av subjekter til eksperimentgruppen og testgruppen skjer tilfeldig (*randomisert*) slik at spuriøse effekter kan utelukkes. For å måle effekter på de avhengige variablene blir det foretatt målinger av disse på to tidspunkter; før og etter at subjektene i eksperimentgruppen har blitt utsatt for behandlingen som representerer den uavhengige variabelen i eksperimentsituasjonen.

For at en undersøkelse skal ha *ekstern validitet* må resultatene kunne generaliseres til også å ha gyldighet utenfor den spesifikke undersøkelsessituasjonen. Det har vært vanlig å skille mellom to former for generalisering (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979). På den ene siden kan man være interessert i å generalisere resultater, eller effekter i en spesifikk undersøkelse til å gjelde for en spesifikk reell situasjon. Et eksempel på en problemstilling der vi ønsker å generalisere effekter, er når vi ønsker å finne ut hvordan feriegjester vurderer et bestemt hotell. Vi ønsker da å generalisere vurderingene som ble foretatt av de gjestene som er med i utvalget, til å gjelde hele populasjonen av feriegjester som bor på det bestemte hotellet i ferien. Den andre formen for generalisering er basert på et ønske om å generalisere forklaringer. Formålet er nå å anvende teori for å forklare resultater ut over undersøkelsessituasjonen. Et eksempel på generalisering av forklaringer er når man som forsker ønsker å finne ut om en bestemt teori også gjelder utover den situasjon og kontekst testingen blir foretatt i. Data innsamlet blant feriegjester ved et bestemt hotell kan f.eks benyttes for å teste teoretiske antakelser om sammenhenger mellom kognitive, affektive og konative komponenter i vurderingsmodeller.

Intern validitet omhandler i hvilken grad observerte sammenhenger kan betraktes som kausale (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979). Denne validitetsformen stiller spørsmål om den eksperimentelle *treatment* variabelen har en effekt på den avhengige variabelen. Ved mangel på intern validitet, eksisterer det alternative forklaringer på sammenhengen mellom årsaksvariabelen og effektvariabelen (Mitchell, 1985).

Begrepsvaliditet er et spørsmål om godheten av de operasjonelle mål vi har valgt for de teoretiske begrepene i studien (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979; Nachmias og Nachmias, 1985). Begrepsvaliditet vil bli grundigere drøftet i kapittel 6.

Statistisk konklusjonsvaliditet tar for seg hvorvidt det er statistisk grunnlag for å trekke de konklusjoner som blir presentert (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979; Nachmias og Nachmias, 1985). Mangel på reliabilitet i målinger og feilaktig bruk av statistiske tester og analyseteknikker representerer trusler mot statistisk konklusjonsvaliditet. Statistisk konklusjonsvaliditet er en betingelse for å si noe om både intern validitet og ekstern validitet. I kapittel 7 vil vi drøfte hvorvidt de innsamlede dataene tilfredsstillende forutsetningene som ligger til grunn for de analysemetodene som benyttes for å teste hypotesene.

I mange tilfeller er det vanskelig eller umulig å oppnå randomisering. Kravene til intern validitet vil da bli vanskelig å oppfylle. I slike tilfeller benyttes *kvasi-eksperimentelle design* (Cook og Campbell, 1979; Kidder og Judd, 1986; Nachmias og Nachmias, 1985). kvasi-eksperiment er en eksperimentell tilnærming der vi har *treatments*, mål på utfall og eksperimentgrupper, men der vi mangler randomisert tilordning.

I metodelitteraturen har det vært diskutert hvordan man bør prioritere mellom validitetskravene (McGrath og Brinberg, 1983). Ved teoriforskning og teorianvendelse blir det anbefalt at man i prioritert rekkefølge legger vekt på intern validitet, begrepsvaliditet, statistisk konklusjonsvaliditet og ekstern validitet (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979; Nachmias og Nachmias, 1985). Et alternativt perspektiv presenteres av Lynch som hevder at alle validitetstyper er like sentrale, og at det ikke nødvendigvis er noen konflikt mellom dem (Lynch, 1982; Lynch, 1983).

Relatert til problemstillingene vi har tatt opp i avhandlingen, vil de to ovenfornevnte formene for generalisering få stor betydning, både for valg av undersøkelsesopplegg og for prioritering mellom de ulike validitetskravene. Vi ønsker på den ene siden å generalisere effekter av metodologiske valg ved utvikling av spørreskjema til en reell situasjon der spørreskjema benyttes. Dette innebærer at datainnsamlingen må foretas i en reell setting. På den andre siden vil det være viktig å sikre at responseffektene vi studerer bare skyldes forskjeller i valg av ordlyd i spørsmålsstilling og benevnelse på svars skala, og variasjoner i plassering av viktighetsspørsmålene i spørreskjemaet. Vi må i tillegg være i stand til å kontrollere for de antatt modererende effektene av kvalitetsnivået på produktet som vurderes og respondentens produktkunnskap. Til slutt må de data som samles inn være av en slik art, og inneholde tilstrekkelig varians til at vi kan teste hypotesene. Den foreskrevne prioriteringen mellom validitetskrav (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982; Cook og Campbell, 1979; Nachmias og Nachmias, 1985) vil derfor gjelde for denne studien.

TABELL 5.1

Sammenhenger mellom avhandlingens hypoteser og krav til informasjon

| HYPOTESE: | KRAV TIL INFORMASJON: |
|---|--|
| H 1. Sammenheng mellom ordlyd i spørsmålsstilling og nivå på vurderinger ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt |
| H 2. Sammenheng mellom ordlyd i spørsmålsstilling og varians på vurderinger ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt |
| H 3a. Sammenheng mellom ordlyd i spørsmålsstilling og nivå på vurderinger ved ulik kvalitet på produktet ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt og mål på totalvurdering for produkter av ulik kvalitet |
| H 4a. Sammenheng mellom ordlyd i spørsmålsstilling og evne til å predikere totalvurdering og atferd, ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt og mål på totalvurdering og atferdsintensjoner |
| H 5a. Sammenheng mellom ordlyd i spørsmålsstilling ved måling av kunders produktvurdering, egenskaper ved produktet og evne til å predikere totalvurdering og atferdsintensjoner | Ulike kunders "bra-dårlig" og tilfreds-utifreds" vurderinger av et produkt bestående av immaterielle og materielle egenskaper og mål på totalvurdering og atferdsintensjoner |
| H 6. Sammenheng mellom ordlyd i spørsmålsstilling og sikkerhet i vurderinger ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt og mål på sikkerhet for de samme vurderingene |
| H 7a. Sammenheng mellom ordlyd i spørsmålsstilling, kundens produktkunnskap, og evne til å predikere totalvurdering og atferdsintensjoner ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt, mål på kundens produktkunnskap og mål på totalvurdering og atferdsintensjoner |
| H 8a. Sammenheng mellom ordlyd i spørsmålsstilling, kundens produktkunnskap, egenskaper ved produktet og evne til å predikere totalvurdering og atferdsintensjoner ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt bestående av materielle og immaterielle egenskaper, mål på kundenes produktkunnskap og mål på totalvurdering og atferdsintensjoner for det samme produktet |
| H 9a. Sammenheng mellom ordlyd i spørsmålsstilling, kundens produktkunnskap og sikkerhet i vurderingene ved måling av kunders produktvurderinger | Ulike kunders "bra-dårlig" og "tilfreds-utifreds" vurderinger av et produkt, mål på kundens produktkunnskap og mål på sikkerhet i de samme vurderingene |
| H 10a. Sammenheng mellom inkludering og plassering av viktighetsspørsmål vs. vurderingsspørsmål og nivå på vurderingsspørsmål ved ulik kvalitet på produktet | Ulike kunders vurderinger av produkter med ulik kvalitet i spørreskjema uten viktighetsspørsmål, og med ulik plassering av viktighetsspørsmål og mål på nivå på vurderingsspørsmål |
| H 11. Sammenheng mellom viktighet og vurderinger avhengig av plassering av viktighetsspørsmål og vurderingsspørsmål | Ulike kunders vurderinger og tillegging av viktighet til egenskaper ved et produkt i et spørreskjema med ulik plassering av viktighetsspørsmål og vurderingsspørsmål |
| H 12a. Sammenheng mellom plassering av viktighetsspørsmål vs. vurderingsspørsmål og evne til å predikere totalvurdering og atferd ved måling av kunders produktvurderinger | Ulike kunders vurderinger av et produkt i et spørreskjema med ulik plassering av viktighetsspørsmål vs. vurderingsspørsmål og mål på totalvurdering og atferdsintensjoner |
| H 13a. Sammenheng mellom viktighet og vurderinger avhengig av plassering av viktighetsspørsmål og vurderingsspørsmål for respondenter med ulik produktkunnskap | Ulike kunders vurderinger og tillegging av viktighet til egenskaper ved et produkt i et spørreskjema med ulik plassering av viktighetsspørsmål og vurderingsspørsmål og ulik produktkunnskap |
| H 14a. Sammenheng mellom plassering av viktighetsspørsmål vs. vurderingsspørsmål ved måling av kunders produktvurderinger, kundens produktkunnskap og evne til å predikere totalvurdering og atferdsintensjoner | Ulike kunders vurderinger av et produkt i et spørreskjema med ulik plassering av viktighetsspørsmål og vurderingsspørsmål, kundens produktkunnskap og mål på totalvurdering og atferdsintensjoner |

Sammenhenger mellom hypotesene og kravene vi må stille til informasjonen er nærmere spesifisert i Tabell 5.1. Studiens krav til informasjon innebærer at vi må utforme et undersøkelsesopplegg som sikrer varians i variablene som inngår i hypotesene, og som i tillegg kan sikre at responseffektene er ekte. Det kreves det altså en studie med variasjon i følgende variabler:

- *Ordlyd i spørsmålsstilling og svaralternativer.* Vi trenger to sett med spørreskjema. Ett sett der respondentene blir bedt om å vurdere produktet som “bra-dårlig”, og ett sett der de blir bedt om å gi uttrykk for i hvilken grad de er “tilfreds-tilfreds” med produktet
- *Totalvurdering og atferdsintensjoner*
- *Nivå på vurderinger og viktighetsspørsmål*
- *Mål for kundenes sikkerhet i sine vurderinger*
- *Rekkefølge på viktighetsspørsmål og vurderinger i spørreskjemaet.* Vi trenger tre sett med spørreskjema. Ett sett der viktighetsspørsmålene plasseres før vurderingene, ett sett der viktighetsspørsmålene plasseres etter vurderingene og ett sett uten viktighetsspørsmål⁶⁶.
- *Egenskaper ved produktet.* Vi trenger et produkt som inneholder både materielle og immaterielle egenskaper
- *Mål for kundenes produktkunnskap.* Vi trenger kunder med både høy og lav produktkunnskap

5.2 Metodologisk tilnærming

Siktemålene med studien, slik det kommer til uttrykk gjennom hypotesene, er å undersøke implikasjoner av valg mellom subjektforankrete- (“tilfreds-tilfreds”) og objektforankrete (“bra-dårlig”) instruksjoner og -vurderingsskalaer ved måling av kunders produktvurderinger etter at de har kjøpt og brukt produktet. Videre ønsker vi å studere implikasjoner av å inkludere og variere plassering av viktighetsspørsmål, ved operasjonalisering av multiattributtmodeller i et spørreskjema. I tillegg til de direkte implikasjonene av disse valgene, ønsker vi å studere antatt modererende effekter av respondentens produktkunnskap og kvalitetsnivået på produktet som vurderes.

Faktorielle design er i litteraturen beskrevet både som eksperimentelle undersøkelsesopplegg (Nachmias og Nachmias, 1985), og som kvasi-eksperimentelle undersøkelses-

⁶⁶Spørreskjema der vektene utelates vil bli benyttet som kontrollgruppe.

opplegg (Cook og Campbell, 1979). Faktorielle design er en form for eksperiment design der man kan ta hensyn til både direkte effekter og indirekte (interaksjons) effekter. Et faktorielt design vil derfor være å foretrekke for å teste hypotesene vi fremsatte i kapittel 4.

Undersøkelser der vi studerer effekter av mer enn én uavhengig variabel krever et stort antall eksperimentgrupper som utsettes for ulike behandlinger (*treatments*). Ordlyd i spørsmålsstilling og svaralternativer, og spørsmålsrekkefølge representerer uavhengige variabler som kan manipuleres gjennom å endre på spørreskjemaet. Et faktorielt design som tar hensyn til våre problemstillinger vil kreve to sett spørreskjema med ulik ordlyd i spørsmålsstilling, og tre sett med ulik manipulering av rekkefølgen på viktighets-spørsmål og vurderingsspørsmål. Vi har altså et (2x3) faktorielt design. Ved beskrivelse av stimulumaterialet senere i dette kapittelet, vil vi foreta en nærmere redegjørelse for hvilke behandlinger (*treatments*) som ble benyttet i dette undersøkelsesopplegget.

I avhandlingen er vi opptatt av å belyse responseffekter i *survey-undersøkelser*. *Survey*-metoden kan både betraktes som et forskningsdesign, og som en datainnsamlings teknikk. Som forskningsdesign skiller *survey* seg fra andre undersøkelsesopplegg ved at de uavhengige variablene er naturlige fenomener, som ikke kan manipuleres slik tilfellet er ved bruk av eksperimentelle og kvasi-eksperimentelle opplegg. I våre problemstillinger og hypoteser opererer vi både med manipulerbare- og ikke-manipulerbare uavhengige variabler. Ordlyd i spørsmålsstilling og svaralternativer, og spørsmålsrekkefølge representerer som nevnt uavhengige variabler som kan manipuleres gjennom å endre på spørreskjemaet. Respondentenes produktkunnskap og kvalitetsnivået ved produktene som vurderes er uavhengig variabler⁶⁷ som ikke kan manipuleres, men som må kontrolleres for ved hjelp av målinger gjennom spørsmål i spørreskjemaet.

Survey som datainnsamlingsteknikk innebærer utdeling av spørreskjema til et tilfeldig utvalg av respondenter fra en kjent populasjon, som vi ønsker å uttale oss om. Når vi benytter “survey” som datainnsamlingsmetode for å skaffe data for å teste hypotesene vi fremsatte i kapittel 4, vil randomisering innebære at vi sørger for at det er tilfeldig hvem som får en bestemt type spørreskjema. Hvorvidt vi har lyktes med dette vil bli diskutert senere i dette kapittelet.

Med utgangspunkt i drøftingen av ulike designtilnærminger, har vi valgt å benytte et kvasi-eksperimentelt, faktorielt design der vi kan ta hensyn til både direkte effekter og

⁶⁷Respondentenes produktkunnskap og kvalitetsnivå på produktet som vurderes betraktes både som uavhengige variabler med direkte effekter, og som variabler som modererer effektene av ordlyd og spørsmålsrekkefølge.

indirekte effekter av de uavhengige variablene vi studerer. I tillegg ble det valgt å benytte en feltstudie som gjør det mulig å undersøke problemstillingene i sin naturlige setting. Feltstudier blir tradisjonelt oppfattet som å styrke resultatenes eksterne validitet (Cook og Campbell, 1979) Selv om hensynet til ekstern validitet ikke er det primære i vår sammenheng⁶⁸, gjør bruk av feltstudie det mulig å si noe om generaliserbarhet. Studien gjør videre bruk av tverrsnittsdata. På denne måten kan vi få inn data fra et forholdsvis stort utvalg, slik at dataene kan brukes for å undersøke sammenhenger mellom avhengige og uavhengige variabler, og multivariat dataanalyse kan benyttes for å teste hypotesene.

5.3 Valg av objekt/produkt for studien

I den empiriske studien har vi valgt å måle feriegjesters vurderinger av hotell i en større norsk hotellkjede, Inter Nor Hotels. Det er ulike årsaker til at nettopp hotell fra denne kjeden ble valgt som objekter for studien. Valg av objekt (produkt) vil i dette avsnittet bli begrunnet.

For det første representerer hotellproduktet en kompleks kombinasjon av materielle- og immaterielle egenskaper. Materielle egenskaper kan f.eks. være hotellrom, senger og mat. Et hotellopphold vil også inneholde mer immaterielle egenskaper, som f.eks. service. Ved vurdering av et hotell antas en hotellgjest å måtte vurdere, og tillegge viktighet til, et stort antall produktegenskaper for å komme frem til en totalvurdering. Hotellproduktet tilfredsstiller dermed vårt krav til et produkt med både materielle- og immaterielle egenskaper.

I hypotese 3a, 3b og 3c fremsatte vi påstander om at forskjellen i nivå mellom objektforankrete- og subjektforankrete vurderinger, vil være større for produkter med kvalitet over eller under gjennomsnittet enn for produkter med kvalitet lik gjennomsnittet. For at det skal være mulig å teste denne hypotesen trenger vi vurderinger av flere enn ett produkt. De samme kravene til objekter for studien følger av påstandene som ble fremsatt i hypotese 10a, 10b og 10c. Dette innebærer at vi trenger vurderinger av flere hotell fra Inter Nor kjeden.

I tillegg til disse faglige begrunnelsene for å velge hotell fra en hotellkjede som objekt, er det to rent praktiske årsaker som også var av betydning. For det første er mange av

⁶⁸Vi er ikke interessert i målinger som er representative for hotellgjester som sådan

variablene som vil bli benyttet i studien hentet fra tidligere hotellundersøkelser utført ved Stiftelsen for Samfunns og Næringslivsforskning (SNF). Dette har sikret en god pretesting av variablene, og stadige forbedringer er foretatt med hensyn på hvilke variabler som bør være med i spørreskjemaet. For det andre er studien ledd i et prosjektsamarbeid med Inter Nor Hotels og Reiselivsseksjonen i Næringsdepartementet som har bidratt med finansiell støtte til prosjektet. Samarbeidet med Inter Nor Hotels har også sikret et godt apparat for utdeling av spørreskjema og innsamling av store datamengder.

5.4 Valg av subjekter/respondenter for studien

I forrige avsnitt ble det argumentert for å velge hotell fra Inter Nor Hotels som objekt for studien. Subjekter vil derfor måtte bli gjester som bor på disse hotellene. Nærmere bestemt ble det bestemt å benytte feriegjester som bor på hotellene i løpet av sommerferien. I tillegg valgte vi å begrense utvalget av subjekter til norske feriegjester. Valg av subjekter og begrensningene vi har foretatt har flere årsaker.

For det første er formålet med avhandlingen å studere kunders vurdering av produkter etter kjøp og bruk. Dette innebærer at vi, som subjekter, må velge gjester som har brukserfaring med hotellene, da det er disse som har grunnlag for å foreta vurderinger. Vi måtte altså velge gjester som har overnattet en eller flere netter på et av de utvalgte hotellene.

Kundens produktkunnskap er som påpekt i den teoretiske drøftingen i kapittel 3 et forhold som kan påvirke produktvurderingene. Vi har derfor valgt å se nærmere på betydningen av produktkunnskap som modererende variabel. For å sikre at vi oppnår varians i gjestenes produktkunnskap og erfaring med bruk av hotell som overnattingsprodukt, valgte vi å begrense undersøkelsen til feriegjester. Vi utelukker med andre ord forretningsgjester og kurs/konferansegjester fra totalpopulasjonen av hotellgjester. Begrunnelsen for dette valget er at det sannsynligvis vil være større variasjoner i feriegjestenes produktkunnskap og erfaring med bruk av hotell.

Et annet argument for å begrense undersøkelsen til å gjelde kun ferie- og fritidsgjester er at disse vil representere en mer homogen gruppe enn om vi hadde inkludert gjester fra de øvrige segmentene. Et homogent utvalg er å foretrekke når formålet er å anvende teori fordi vi da kan redusere varians som kan skyldes ulike bakgrunnsvariabler (Calder, Phillips, og Tybout, 1981; Calder, Phillips, og Tybout, 1982). Ønske om et homogent

utvalg er også en av årsakene til at vi begrenset undersøkelsen til norske feriegjester på Inter Nor Hotellene. Vi har da mulighet for å utelukke variasjoner i respons som kan skyldes eventuelle kulturelle forskjeller mellom ulike nasjonaliteter.

En annen, mer praktisk begrunnelse for valg av norske feriegjester som subjekter, er mulighetene for å gjennomføre en undersøkelse innenfor kostnadmessige og tidsmessige rammer. Når utvalget skal deles inn i eksperimentgrupper som blir tilordnet forskjellige typer spørreskjema, vil det bli både tidkrevende og kostbart å håndtere gjennomføring av undersøkelsen med mange typer spørreskjema på forskjellige språk.

5.5 Beskrivelse av stimulimaterialet

I det faktorielle eksperimentdesignet vi skisserte tidligere i dette kapittelet ble to variabler manipulert: Ordlyd i spørsmålsstilling/benevnelse på svarskala, og rekkefølge på viktighetsspørsmål vs. vurderingsspørsmål i spørreskjemaet. Manipulering av ordlyd i spørsmålsstilling og benevnelse på svarskala ble foretatt ved å benytte to sett med spørreskjema. I det ene settet ble respondentene bedt om å foreta en objektforankret "bra-dårlig" vurdering av egenskaper ved hotellet. I det andre settet ble respondentene bedt om å gi uttrykk for i hvilken grad de er "tilfreds-tilfreds" med hotellet, basert på en subjektforankret vurdering av de samme egenskapene.

FIGUR 5.1
Organisering-, og kombinasjoner av manipuleerte variabler ved spørreskjemaforming

| Ordlyd i spørsmålsstilling, benevnelse på svarskala og spørsmålsrekkefølge | Viktighetsspørsmål før spesifikke vurderingsspørsmål | Viktighetsspørsmål etter spesifikke vurderingsspørsmål | Uten viktighetsspørsmål |
|--|--|--|-------------------------|
| Subjektforankret instruksjon og -vurderingsskala | 1 | 3 | 5 |
| Stimulusforankret instruksjon og -vurderingsskala | 2 | 4 | 6 |

Rekkefølge på spørsmål ble manipulert ved å benytte tre sett med spørreskjema. I det ene settet ble viktighetsspørsmålene plassert før de spesifikke vurderingsspørsmålene. I det andre ble viktighetsspørsmålene plassert etter de spesifikke vurderingsspørsmålene,

og i det tredje settet utelates viktighetsspørsmålene fra spørreskjemaet. Respondenter som er tilordnet spørreskjema der viktighetsspørsmålene ble utelatt, vil bli benyttet som kontrollgruppe i analysene.

Som illustrert i figur 5.1 ble det benyttet seks forskjellige typer spørreskjema i undersøkelsen. Disse ble utdelt på en slik måte at det er tilfeldig hvilken gjest som fikk en bestemt type spørreskjema. På denne måten ble det foretatt en randomisering av hvilke subjekter som fikk hvilken behandling (*treatment*). Dette gjør at vi kan anta at mulige forklaringsfaktorer som skyldes forskjeller ved respondentene, og som ikke er tatt med i undersøkelsen utligner hverandre. Produktkunnskap blir målt i spørreskjemaene.

5.6 Utvalg og praktisk gjennomføring

Spørsmål vedrørende utvalg er sentrale i undersøkelser der vi baserer generaliseringer på et utvalg av respondenter fra en bestemt populasjon. I denne avhandlingen er vi som tidligere nevnt ikke opptatt av å generalisere resultatene av de vurderingene som ble foretatt av et utvalg av hotellgjester til å gjelde for hele populasjonen av hotellgjester. Utvalget trenger dermed ikke være representativt for hotellgjester som har overnattet på de ulike hotellene. Når formålet er å generalisere forklaringer, er det imidlertid viktig at utvalget av hotellgjester ikke er så spesielt at de effektene vi ønsker å undersøke, blir et resultat av forskjeller mellom respondentene på andre variabler enn de vi har fremsatt som uavhengige variabler i hypotesene.

Dataene som danner grunnlag for hypotesetesting er innsamlet gjennom gjesteundersøkelser gjennomført for Inter Nor Hotellene sommeren 1992. Antall deltagende hotell i undersøkelsene var 10. Dataene ble samlet inn ved bruk av strukturerte spørreskjema med lukkede spørsmål. Valg av spørreskjema som innsamlingsinstrument er begrunnet i formålet med avhandlingen, som er å studere ulike responseffekter i "survey" undersøkelser. Problemstillingene og hypotesene er slik at spørreskjema med lukkede svaralternativer må benyttes. Bruk av spørreskjema gjorde det også mulig å nå et stort antall respondenter på en kostnadseffektiv måte. Ved utvikling av spørreskjemaene ble det lagt stor vekt på pilottesting og erfaringer fra tidligere undersøkelser⁶⁹. Utdelingen ble foretatt av resepsjonspersonalet på hotellet ved gjestens ankomst til hotellet og innsjekkning. For at det skulle være tilfeldig hvilke gjester som fikk utdelt spørreskjema, i tillegg til hvilken type skjema som ble delt ut, ble utdelingen organisert etter

⁶⁹Jfr. neste kapittel om operasjonalisering av variabler.

romnummer på hotellet. Disse ble trukket tilfeldig hver dag. Innsamling av spørreskjemaene skjedde enten ved at gjestene leverte de ferdig utfylte skjemaene i svarkonvolutt i resepsjonen ved avreise fra hotellet, eller ved at man benyttet en ferdig-frankert svarkonvolutt og sendte det utfylte spørreskjemaet i posten. For å øke responsraten ble respondentene gitt muligheten for å delta i en trekning av en gavesjekk til en verdi av kr. 5.000,-.

Når utdeling av spørreskjema overlates til ansatte i resepsjonen på hotellet er det en fare for at vi mister kontrollen med utdelingen av skjemaene. For å redusere feilkildene knyttet til utdelingsmåte ble det gitt en grundig informasjon til resepsjonspersonalet gjennom instruks for utdeling av spørreskjema. Spørreskjemaene ble i tillegg blandet i følgende rekkefølge før de ble utdelt på hotellene, slik at det var tilfeldig hvilken gjest som skulle få en bestemt type spørreskjema:

1. *Bra - dårlig instruksjon, uten viktighetsspørsmål*
2. *Tilfreds - utilfreds instruksjon, uten viktighetsspørsmål*
3. *Bra - dårlig instruksjon, med viktighetsspørsmål plassert før spesifikke vurderingsspørsmål, totalvurderingsspørsmål og atferdsintensjoner*
4. *Tilfreds - utilfreds instruksjon, med viktighetsspørsmål plassert før spesifikke vurderingsspørsmål, totalvurderingsspørsmål og atferdsintensjoner*
5. *Bra - dårlig instruksjon, med viktighetsspørsmål plassert etter spesifikke vurderingsspørsmål, men før totalvurderingsspørsmål og atferdsintensjoner*
6. *Tilfreds - utilfreds instruksjon, med viktighetsspørsmål plassert etter spesifikke vurderingsspørsmål, men før totalvurderingsspørsmål og atferdsintensjoner*

Gjennom å blande spørreskjemaene på denne måten sikret vi at det ble utdelt likt antall av alle spørreskjemavariantene, og at det var tilfeldig hvilke gjester som fikk en bestemt type spørreskjema.

Totalt ble det distribuert 3000 spørreskjema til hotellene. Av disse ble omtrent halvparten utdelt til gjestene. Ca. 1500 spørreskjema danner derfor grunnlag for beregning av responsrate. Totalt 629 spørreskjema ble utfylt og returnert. Ved kontroll av spørreskjemaene ble det avdekket 21 respondenter som systematisk hadde markert ekstremverdiene (+5) på samtlige vurderingsspørsmål. Det er umulig å vite om disse svarene reflekterer en "sann" vurdering, eller om ekstremverdiene ble markert uten at det ble tatt hensyn til de spesifikke spørsmålene. For å unngår potensielle skjevheter ved

analyse av dataene blir de 21 respondentene betraktet som *careless respondents* og fjernet fra utvalget. Effektivt utvalg for dataanalysene blir derfor 608 respondenter.

TABELL 5.2
Oppnådd utvalg i de ulike eksperimentgruppene

| EKSPERIMENTGRUPPE: | ANTALL RESPONDENTER: |
|--|-----------------------------|
| Subjektforankret instruksjon og -vurderingsskala, viktighetsspørsmål før spesifikke vurderingsspørsmål | 100 |
| Objektforankret instruksjon og -vurderingsskala, viktighetsspørsmål før spesifikke vurderingsspørsmål | 113 |
| Subjektforankret instruksjon og -vurderingsskala, viktighetsspørsmål etter spesifikke vurderingsspørsmål | 107 |
| Objektforankret instruksjon og -vurderingsskala, viktighetsspørsmål etter spesifikke vurderingsspørsmål | 84 |
| Subjektforankret instruksjon og -vurderingsskala, uten viktighetsspørsmål | 101 |
| Objektforankret instruksjon og -vurderingsskala, uten viktighetsspørsmål | 103 |
| Sum | 608 |

Tabell 5.2 illustrerer hvordan disse fordeler seg i de ulike eksperimentgruppene i det faktorielle designet. Gjennomsnittlig responsrate for undersøkelsen er ca 40 %. Dette er langt fra perfekt, men vesentlig høyere enn nedre grenser for responsrate i lignende studier, se f.eks Harris (1978) som rapporterer responsrater på 18%.

TABELL 5.3
Oppnådd utvalg for hotellene som deltok i undersøkelsene

| HOTELL: | ANTALL RESPONDENTER: |
|-----------------------------|-----------------------------|
| Solborg Hotell, Mandal | 64 |
| Hotell Røros, Røros | 175 |
| Grand Hotel, Larvik | 136 |
| Grand Hotel Olav, Trondheim | 50 |
| Hotel Scandinaviè, Ålesund | 26 |
| Ulstein Hotel, Ulsteinvik | 47 |
| Sunnfjord Hotel, Førde | 27 |
| Hotel Prinsen, Trondheim | 37 |
| Hotel Residence, Sandnes | 14 |
| Highland Hotel, Geilo | 23 |
| Grand Hotel, Kristiansund | 9 |
| Sum | 608 |

De oppnådde utvalgene i de forskjellige eksperimentgruppene er som vi ser i tabell 5. 2 svært jevne. Dette indikerer at utdelingen har fungert. Det viste seg imidlertid at man i enkelte hotell var flinkere til å få delt ut spørreskjemaene enn i andre hotell. Tabell 5.3 viser antall respondenter fordelt på hotellene som deltok i undersøkelsen. På bakgrunn av den relativt høye responsraten og de jevne cellestørrelsene i det faktorielle designet argumenterer vi så langt for at manglende svar (*non-respons*) som feilkilde i studien er svært liten.

Under drøftingen av undersøkelsesopplegget har vi lagt vekt på å begrunne de valg vi har foretatt. Som en avslutning på dette kapittelet skal vi peke på problemer og svakheter ved det valgte undersøkelsesopplegget.

Når vi skal tilordne ulike typer spørreskjema (*treatment*) til et randomisert utvalg av respondenter, må vi ha en kjent populasjon som vi trekker dette utvalget fra. Dersom formålet med den empiriske undersøkelsen hadde vært å finne ut hvordan feriegjester som gruppe vurderer hotellene måtte vi ideelt sett ha trukket et tilfeldig utvalg fra en kjent populasjon av feriegjester som bor på hotell. Dette vil ikke være mulig da en slik kjent populasjon ikke finnes.

Når formålet er å finne ut om de predikerte responseffektene er ekte, vil randomisering innebære at det må være tilfeldig hvem som får en bestemt type spørreskjema. De valgte hotellgjestene trenger dermed ikke være representative for hotellgjester som gruppe. Kravet er heller at det ikke må være en spesiell type gjester som får en bestemt type spørreskjema. I og med at vi har overlatt til hotellet og resepsjonsmedarbeiderne å foreta utdelingen av spørreskjemaene, har vi svak kontroll med hvordan denne foregår. Randomiseringen kan derfor bli svak.

En svakhet ved å bruke en populasjon som ikke er kjent, er at vi ikke har muligheter for å kontrollere hvem som ikke svarer på spørreskjemaene (*non-respons*). Selv om det blir foretatt en tilfeldig utdeling, kan svarene være et resultat av selvsleksjon⁷⁰. Det kan således være vanskelig å vite om det er gjester som er spesielt fornøyde eller spesielt misfornøyde som svarer. Når spørreskjemaene blir utdelt ved gjestens ankomst til hotellet, har vi i tillegg ingen kontroll med selve utfyllingen. Bruk av svarkonvolutt i tillegg muligheten for innlevering av spørreskjemaet i resepsjonen øker responsraten, men kan bidra til forskjeller i respons på grunn av at utfyllingen av spørreskjemaet kan skje på ulike tidspunkt.

⁷⁰Selvsleksjon vil bety at subjektene bestemmer selv om de vil være en del av utvalget eller ikke.

5.7 Oppsummering og implikasjoner

I dette kapittelet har vi beskrevet undersøkelsesopplegget som ble benyttet som grunnlag for å foreta en empirisk testing av hypotesene som ble formulert i kapittel 4. Med utgangspunkt i problemstillingene og hypotesene fremsatte vi innledningsvis en del sentrale krav til undersøkelsesopplegget. Valg av undersøkelsesopplegg har på denne måten blitt styrt av ulike hensyn til validitet. Med utgangspunkt i den teoretisk drøftingen har vi formulert hypoteser med årsak-virkning sammenhenger. Dette innebærer at krav til intern validitet blir sentrale. Videre krever hypotesene at vi har klart å fange opp teoretiske begreper, som f.eks. produktkunnskap, erfaring, egenskaper ved produktet og sikkerhet i vurderinger, på empiriplan. Drøfting av begrepsvaliditet vil bli foretatt i forbindelse med forslag til operasjonalisering av begrepene i neste kapittel. For å sikre at vi kan trekke konklusjoner på bakgrunn av resultatene fremsatte vi krav til statistisk konklusjonsvaliditet. Hvorvidt disse kan tilfredsstilles diskuteres i tilknytning til forutsetningene for de valgte analysemetodene i kapittel 7.

Basert på kravene til undersøkelsesopplegg har vi valgt å benytte et (2x3) faktorielt design der vi manipulerer ordlyd og spørsmålsrekkefølge. Som produkt valgte vi å benytte hotell fra en norsk hotellkjede. Subjekter for studien er norske feriegjester som valgte å benytte disse hotellene som overnattingssted i sin ferie. Data ble samlet inn ved hjelp av en *survey*-tilnærming.

6

Operasjonalisering

I dette kapittelet skal vi beskrive hvordan avhandlingens variabler er operasjonalisert i den empiriske studien. Utvikling av måleinstrument bygger i tillegg til litteraturgjennomgangen i kapittel 3, på resultater og erfaringer fra tidligere undersøkelser der en rekke av målene stadig har vært underlagt testing og forbedringer (Henjesand, 1991a, Henjesand, 1991b, Henjesand og Troye, 1991). Først beskrives uavhengige variabler i avsnitt 6.1. Deretter tar vi for oss de avhengige variablene i avsnitt 6.2. Til slutt drøftes måling av respondentenes produktkunnskap og kvalitetsnivået på produktene som vurderes som modererende variabler i avsnitt 6.3.

6.1 Uavhengige variabler

Avhandlingens uavhengige variabler er av to forskjellige typer. På den ene siden har vi uavhengige eksperimentvariabler som blir manipulert ved å endre på spørreskjemaet. Dette gjelder instruksjonen og vurderingsskalaen som blir benyttet i forbindelse med de spesifikke vurderingsspørsmålene, i tillegg til inkludering og alternativ plassering av viktighetsspørsmålene.

På den andre siden opererer vi med kundens produktkunnskap, og kvalitetsnivået på produktene som vurderes som ikke-manipulerte modererende variabler. I dette avsnittet gjør vi rede for de manipulerede uavhengige eksperimentvariablene. Operasjonalisering og måling av de modererende variablene drøftes i avsnitt 6.3

6.1.1 Instruksjon og vurderingsskala, og inkludering og alternativ plassering av viktighetsspørsmål

Vi velger å betrakte de ulike stimulikategoriene, representert ved forskjellige spørreskjema varianter, som uavhengige variabler, og behandler måling av vurderinger og viktighet som avhengige variabler⁷¹.

Som nevnt i foregående kapittel blir det benyttet seks forskjellige spørreskjema som hver representerer et eksperimentelt stimulus i et kvasieksperimentelt faktorielt undersøkelsesopplegg. I figur 5.2 viste vi en oversikt som illustrerer hvordan spørreskjemaene er forskjellige. Den ene hovedkategorien av spørreskjema har subjektforankret instruksjon og -vurderingsskala i forbindelse med de spesifikke vurderingsspørsmålene. Innenfor denne hovedkategorien vil inkludering og plassering av viktighetsspørsmål bli manipulert.

FIGUR 6.1
Manipulering av instruksjon og vurderingsskala i spørreskjemaet

Objektforankret instruksjon:

Nedenfor ber vi Dem gi en vurdering av det hotellet hvor De fikk utlevert dette spørreskjemaet. Hvor bra eller dårlig er dette hotellet når det gjelder de egenskapene som er nevnt nedenfor?

Sett en sirkel rundt tallet som uttrykker Deres svar.

Et eksempel:

| | <i>Svært dårlig</i> | | | | | | | | | | <i>Svært bra</i> | | | | | | | | | | | |
|-----------------------------|---------------------|----|----|----|----|---|---|---|---|---|------------------|--|--|--|--|--|--|--|--|--|--|--|
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | | | | | | | | | | | |
| 1. Informasjon om utflukter | | | | | | | | | | | | | | | | | | | | | | |

Denne gjesten har vurdert hotellets informasjon om utflukter som hverken bra eller dårlig

Subjektforankret instruksjon:

Nedenfor ber vi Dem gi en vurdering av det hotellet hvor De fikk utlevert dette spørreskjemaet. Hvor tilfreds eller utilfreds er De når det gjelder de egenskapene som er nevnt nedenfor?

Sett en sirkel rundt tallet som uttrykker Deres svar.

Et eksempel:

| | <i>Svært utilfreds</i> | | | | | | | | | | <i>Svært tilfreds</i> | | | | | | | | | | | |
|-----------------------------|------------------------|----|----|----|----|---|---|---|---|---|-----------------------|--|--|--|--|--|--|--|--|--|--|--|
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | | | | | | | | | | | |
| 1. Informasjon om utflukter | | | | | | | | | | | | | | | | | | | | | | |

Denne gjesten er hverken tilfreds eller utilfreds med hotellets informasjon om utflukter.

⁷¹Vurdering av spesifikke egenskaper ved produktet betraktes også som en uavhengig variabel i hypotesene som omhandler de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner

Den andre hovedkategorien av spørreskjema har objektforankret instruksjon og -vurderingsskala, og de samme manipulasjonene med spørsmålrekkefølge blir foretatt. Hvordan operasjonalisering av objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer ble manipulert i spørreskjemaene er illustrert i figur 6.1.

Skalaene som blir benyttet ved måling av både objektforankrete-, og subjektforankrete vurderinger er utviklet på bakgrunn av erfaringer fra tidligere undersøkelser (Henjesand, 1991a, Henjesand, 1991b). -5 til 5 skalaen er valgt med bakgrunn i ønsket om å redusere trusselen for *ceiling* effekter⁷² (Heide, 1990), og for å øke variansen i vurderingsmålene. For å øke sannsynligheten for at måleskalaene måler vurderinger på tilnærmet intervallnivå blir det, som anbefalt av Parasuraman (1986) og Wildt og Mazis (1978), kun benyttet benevnelse på endepunktene. Respondentene ble bedt om å sette sirkel rundt det tallet som representerer deres vurdering.

Gjennomgang av en rekke metodologiske studier⁷³ viser at det hersker relativt stor uenighet omkring hvor mange skalapunkter som bør benyttes på en måleskala. I noen sammenhenger anbefales 5 skalaverdier eller 7 skalaverdier. I andre studier anbefales bruk av 11 skalaverdier. Alle disse anbefalingene er basert på reliabilitetsbetraktninger. En mer kognitiv teoretisk basert diskusjon av optimalt antall responsalternativer på måleskalaer finnes hos Fazio (1990). Her diskuteres problemer forbundet med å benytte måleskalaer med mange skalaverdier. De kognitive oppgavene som pålegges en respondent ved bruk av 11-punktsskalaer er krevende, og det kan stilles spørsmål om man er i stand til å benytte så mange kategorier. Vi har likevel valgt å benytte 11-punktsskalaer i denne studien med begrunnelse i at denne skalalengden har vist seg å ha bedre statistiske egenskaper enn skalaer med et lavere antall skalaverdier.

Nivå og konsistens i vurderinger vil bli målt ved gjennomsnittsmål og standardavvik. Hvilke spesifikke egenskaper ved hotellproduktet respondentene blir bedt om å vurdere ved bruk av disse instruksjonene blir nærmere redegjort for i kapittel 6.2 der avhandlingens avhengige variabler beskrives.

I forbindelse med manipuleringene av instruksjonene som benyttes i spørreskjemaet er det relevant å diskutere problemer forbundet med mono-operasjonaliseringsskjjevhet⁷⁴ (Cook og Campbell, 1979). Mono-operasjonaliseringsskjjevhet kan oppstå når man i en

⁷²*Ceiling* effekter oppstår når måleskalaen ikke er fanger opp vurderinger som dekker hele utfallsrommet. Et fenomen kan godt være normalfordelt, men på grunn av "uheldig" valg av måleskala fanges ikke ytterpunktene opp av skalaen. F.eks vil en måleskala fra 100 -170 cm ikke være tilstrekkelig til å måle høyden på norske menn, som kan være normalfordelt.

⁷³Se appendix A

⁷⁴*Mono operation bias*

studie benytter kun en operasjonalisering av en mulig årsak, og/eller når man bare har ett mål på hver av de mulige effektvariablene.

I studien som er gjennomført i forbindelse med denne avhandlingen opererer vi med multiple mål på de avhengige variablene. Dette er ikke tilfelle for de uavhengige eksperimentvariablene, ordlyd i spørsmålsstilling og benevnelse på svarskala. Her har vi operasjonalisert subjektforankret instruksjon, og -vurderingsskala ved å instruere respondentene til å gi uttrykk for hvor tilfreds eller utilfreds de er med spesifikke egenskaper ved hotellproduktet. Objektforankrete instruksjon og -vurderingsskala er operasjonalisert gjennom å be respondentene vurdere hvor bra eller dårlig hotellet var på de samme spesifikke egenskapene. Vi står altså overfor en situasjon der mono-operasjonalisering kan representere en trussel mot begrepsvaliditet (Cook og Campbell, 1979). Mono-operasjonaliseringer kan både underrepresentere begrepet man ønsker å måle og tilføre irrelevans i målingene, noe som kan redusere begrepsvaliditet. I motsetning til mono-operasjonalisering av effektvariabler, der man lett kan inkludere ekstra mål i samme studie, innebærer det å unngå mono-operasjonalisering av årsaksvariablen ofte et større problem. Bruk av multiple mål på en årsaksvariabel vil innebære at man må øke antall manipulasjoner i et faktorielt design. En økning av antall manipulasjoner vil kreve svært store utvalg, eller medføre små utvalg innenfor hver celle i det faktorielle designet, dersom det totale utvalget ikke kan økes. Dette vil redusere statistisk konklusjonsvaliditet.

Cook & Campbell (1979) konkluderer med at det ikke finnes noe substitutt for multiple operasjonaliseringer av eksperimentvariablene, "dersom dette er mulig". Når vi til tross for diskusjonen ovenfor opererer med mono-operasjonaliseringer av objekt- og subjektforankrete instruksjoner og -vurderingsskalaer har dette flere årsaker. For det første har vi valgt å benytte det vi velger å kalle ytterpunktene i objektforankrete og subjektforankrete instruksjoner og -vurderingsskalaer. Vi har definert objektforankrete instruksjoner og -vurderingsskalaer til å reflektere situasjoner der respondenter blir bedt om å være "dommer" og vurdere et objekt eller ulike egenskaper ved et objekt. I en slik sammenheng vil "bra-dårlig" instruksjon og benevnelse på skalaen representere et ytterpunkt. Subjektforankrete instruksjoner og -vurderingsskalaer har vi definert til å reflektere de indre følelsesmessige tilstander et objekt kan gi opphav til. "Tilfreds-utilfreds" instruksjon og benevnelse på måleskalaen vil her representere ytterpunkter. Ytterpunktene er valgt for å maksimere forskjellene i instruksjon som benyttes i de ulike eksperimentgruppene. Alternative operasjonaliseringer av objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer bør imidlertid benyttes i senere studier for å øke generaliserbarheten av funnene i studien vi har gjennomført i forbindelse med avhandlingen.

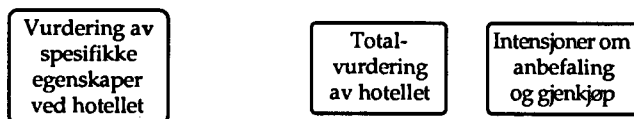
Et annet argument for å benytte kun en operasjonalisering av de eksperimentelle årsaksvariablene er begrensninger i mulighetene for å gjennomføre studien innenfor de tids- og kostnadsrammer vi har til rådighet. Økning av antall operasjonaliseringer av objektforankrete- og subjektforankrete instruksjoner ville medført behov for en dobling, eller tredobling, av nødvendig antall eksperimentgrupper. I den empiriske settingen vi har valgt å gjennomføre studien, ville dette ha vært vanskelig å få til.

Et tredje argument, som henger sammen med det foregående, for å beholde mono-operasjonaliseringer er at antall respondenter i eksperimentgruppene ville ha blitt lavt dersom vi skulle ha økt antall eksperimentgrupper i det faktorielle designet.

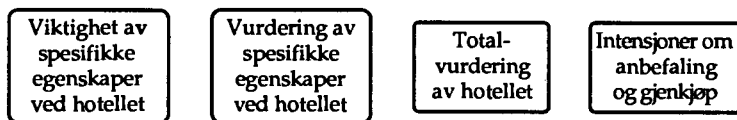
Selv om vi er enig med Cook & Campbell (1979) i at det ikke finnes “gode nok” unnskyldninger for å gjennomføre studier med mono operasjonalisering av uavhengige- og/eller avhengige variabler, har vi ved gjennomføring av denne studien foretatt en avveining av det som ideelt sett er ønskelig med det som er praktisk gjennomførbart. Til tross for svakhetene mono-operasjonaliseringer kan medføre, hevder vi at de valg vi har foretatt med hensyn til operasjonalisering av eksperimentvariablene representerer en svak trussel mot begrepsvaliditet.

FIGUR 6.2
Manipulering av rekkefølge på vurderingsspørsmål og viktighetsspørsmål i spørreskjemaet

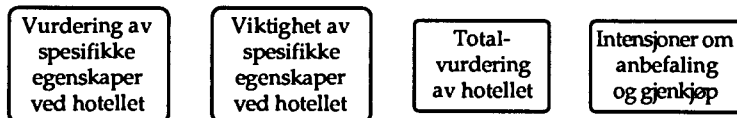
Spørreskjema uten viktighetsspørsmål



Spørreskjema med viktighetsspørsmål før spesifikke vurderingsspørsmål



Spørreskjema med viktighetsspørsmål etter spesifikke vurderingsspørsmål



I tillegg til instruksjon og vurderingsskala, manipuleres inkludering og plassering av viktighetsspørsmål. I figur 6.2 illustrerer vi hvordan dette ble gjennomført i spørreskjemaene. Figuren viser at det kun er tilstedeværelse og plassering av viktighetsspørsmål som manipuleres i spørreskjemaet. Totalvurderingss spørsmål og spørsmål som fanger opp intensjoner om anbefalinger og gjenkjøp har fast plassering, og kommer alltid etter vurderinger av spesifikke egenskaper og eventuelle viktighetsspørsmål.

Hvordan viktighetsspørsmålene stilles i spørreskjemaet beskrives nærmere i avsnitt 6.2.2. Her gjør vi rede for hvordan tillegging av viktighet ved ulike egenskaper ved hotellproduktet måles og benyttes som avhengig variabel.

6.2 Avhengige variabler

Avhandlingens avhengige variabler som finnes på høyre side i modellen i figur 5.1, er: vurderinger av spesifikke egenskaper ved hotellproduktet, tillegging av viktighet til de spesifikke egenskapene, totalvurderinger og atferdsintensjoner, sammenheng mellom vurdering av spesifikke egenskaper ved hotellproduktet og rapportering av viktighet ved de samme egenskapene, sammenheng mellom vurdering av spesifikke egenskaper ved hotellproduktet og totalvurdering og atferdsintensjoner, sikkerhet i vurderinger og relativ betydning av immaterielle produkt egenskaper. Nedenfor vil vi gjøre nærmere rede for hvordan de avhengige variablene blir operasjonalisert i studien.

6.2.1 Vurderinger av spesifikke egenskaper

I dette avsnittet skal vi beskrive hvordan gjestenes vurderinger av de spesifikke egenskapene et hotellprodukt kan bestå av, blir operasjonalisert og målt i spørreskjemaet.

Basert på tidligere undersøkelser⁷⁵ ble gjestene bedt om å vurdere 27 ulike aspekter ved hotellproduktet. For å bestemme om manipuleringene, som ble foretatt ved å endre instruksjon og vurderingsskala og plasseringen av viktighetsspørsmålene, påvirker subjektene vurderinger blir følgende sammenligninger foretatt:

⁷⁵Gjennom undersøkelser foretatt ved Stiftelsen for samfunns- og næringslivsforskning er spørreskjema for måling av gjesters vurderinger av hotellprodukter utviklet og testet (Henjesand 1991a, Henjesand 1991b, Henjesand og Troye 1991)

1. Sammenligning av gjennomsnittsvurderinger for individuelle egenskaper mellom de ulike eksperimentgruppene i det faktorielle undersøkelsesopplegget. Disse sammenligningene gjør det mulig å bestemme om vurderinger foretatt av subjekter i en eksperimentgruppe er konsistent høyere enn vurderinger foretatt av subjekter tilhørende en annen eksperimentgruppe.
2. Sammenligning av gjennomsnittet av samtlige 27 vurderte egenskaper⁷⁶. Gjennom disse sammenligningene vil eventuelle problemer knyttet til tilstedeværelse av tilfeldige fluktuasjoner i vurderinger av individuelle egenskaper forventes å bli eliminert. Reliabilitet ved gjennomsnittet av samtlige 27 egenskaper forventes å være høyere enn reliabilitet for de individuelle indikatorene (Lord og Novick, 1968). Eventuelle "sanne" forskjeller mellom eksperimentgruppene vil derfor lettere kunne avdekkes.
3. Sammenligning av spesifikke egenskaper gruppert sammen i summerte gjennomsnittsskaler. Disse sammenligningene setter oss i stand til å fokusere på forskjeller relatert til vurderinger av egenskaper knyttet til f.eks. service (immaterielle egenskaper) og det fysiske produktet som leveres (materielle egenskaper).

Gjennom tidligere studier utført i forbindelse med reiselivsprogrammet (SiR) ved SNF (Henjesand, 1991a, Henjesand, 1991b, Henjesand og Troye, 1991) har vi forsøkt å operasjonalisere og måle hotellgjestenes vurderinger av de fire produktelementene, samhandlings-, backstage-, struktur- og kundeavhengige elementer (Troye, 1990). Disse fire produktelementene representerer en avbildning av et produkt bestående av ulike egenskaper slik det kan konseptualiseres fra et tilbyderperspektiv. De fire produktelementene fremkommer som et resultat av kombinasjoner av produksjonsfaktorene som inngår i produksjonen av produktet, der kunden, i tillegg til de tradisjonelle produksjonsfaktorene, inkluderes som en produksjonsfaktor ("medprodusent")

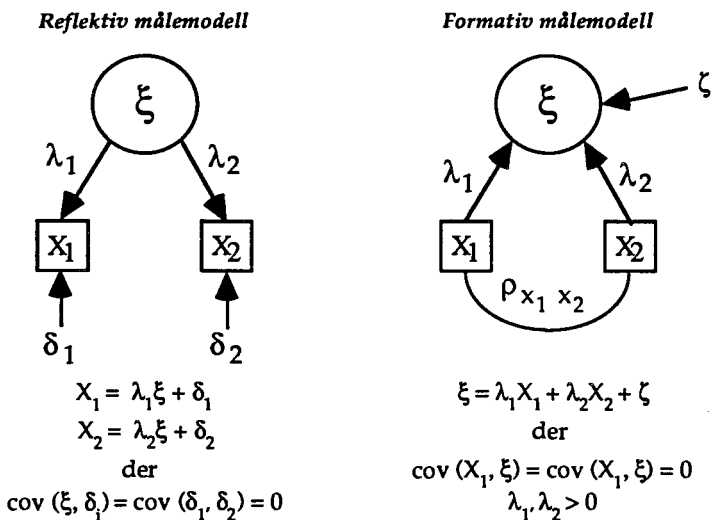
Som vi gjorde rede for i kapittel 3, representerer samhandlingselementene immaterielle egenskaper, mens strukturelementene representerer materielle egenskaper ved det samme produktet. De to andre produktelementene, backstage og kundeavhengige elementer, kan bestå av både materielle- og immaterielle egenskaper (Troye, 1990). Vi

⁷⁶Grand mean

velger å bygge videre på disse inndelingene i denne studien og grupperer de 27 spørsmålene i spørreskjemaet i de fire produktelementene; samhandlingselementer, strukturelementer, backstageelementer og kundeavhengige produktelementer. I avhandlingen fokuserer vi på hvordan materielle og immaterielle egenskaper vurderes forskjellig avhengig av om det benyttes subjektforankrete- eller objektforankrete instruksjoner og -vurderingsskalaer. Det er derfor bare serviceelementene og strukturelementene i produktmodellen som benyttes i analyser knyttet til hypotesene som omhandler vurderinger av immaterielle og materielle egenskaper.

Før vi diskuterer validitet og reliabilitet ved målinger av respondentenes vurderinger gruppert i de fire produktelementene, vil vi drøfte målemodellene som vi har valgt å benytte. Valg av målemodell får store konsekvenser for hvilke kriterier som benyttes for å vurdere målingenes validitet og reliabilitet. Det blir i litteraturen gjort et skille mellom såkalte *reflektive*- og *formative* målemodeller (se figur 6.3). Dette skillet er basert på sammenhengene mellom de latente teoretiske begrepene man ønsker å måle og de observerte variablene, eller indikatorene⁷⁷, man benytter for å operasjonalisere de latente teoretiske begrepene. Hovedforskjellen mellom målemodellene ligger i retningen på disse sammenhengene. Reflektive- og formative målemodeller er illustrert i figur 6.3 nedenfor. Her har vi også tatt med hvordan modellene kan uttrykkes på ligningsform.

FIGUR 6.3
Formative og reflektive målemodeller



⁷⁷I avhandlingene benyttes begrepene indikatorer, variabler, item og spørsmål omhverandre

Dersom vi opererer med en teoretisk latent variabel som påvirker svargivningen på de observerte variablene kalles indikatorene som benyttes *effektindikatorer* eller *reflektive indikatorer*, og man taler gjerne om en *reflektiv målemodell* (Bollen og Lennox, 1991; Bollen, 1984; Bollen, 1989). Reflektive målemodeller finner vi f.eks. ved måling av personlighetstrekk som pesimisme, der vi antar at forandringer i et individs grad av pesimisme medfører endringer i svargivning på spørsmål som er valgt som indikatorer på grad av pesimisme, og ikke vice versa.

Dersom vi i motsatt fall opererer med en teoretisk latent variabel, som kan oppfattes som en effekt av indikatorene som benyttes for å operasjonlisere og måle den latente variabelen, kalles indikatorene *årsaksindikatorer*, eller *formative indikatorer*, og målemodellen som benyttes kalles ofte en *formativ målemodell* (Blalock, 1964; Bollen og Lennox, 1991; Bollen, 1984; Bollen, 1989). Bruk av formative indikatorer og -målemodeller finner vi f.eks. ved måling av sosioøkonomisk status, der indikatorer som utdanning, inntekt og yrkestittel eller stillingsbetegnelse antas å påvirke et individs sosioøkonomiske status. Endring i utdanning og/eller inntekt og/eller yrkestittel vil medføre endringer i sosioøkonomisk status, og ikke vice versa.

Faktoranalyse og klassisk test-teori har hatt en sterk påvirkning på valg av perspektiver for måling, og påfølgende validitets og reliabilitetsdrøftelser i samfunnsvitenskapelige studier (Churchill, 1979). Tradisjonelt har fem konvensjonelle retningslinjer blitt benyttet. Nedenfor gir vi en kort innføring i disse, og drøfter anvendbarheten av dem for de to målemodellene som er presentert ovenfor.

Indikatorer for samme begrepet bør være internt konsistente

Det mest aksepterte premiss i klassisk måleteori er at indikatorer som er positivt assosiert med samme begrep, også bør være positivt korrelert med hverandre (Bollen og Lennox, 1991). Basert på dette er det en vanlig praksis å inspisere korrelasjonsmatriser for indikatorer som "klynger" seg sammen, og utelukke indikatorer som har nær null, eller negative korrelasjoner med andre mål på det samme begrepet. I en situasjon med effektindikatorer og reflektive målemodeller vil et kriterium om intern konsistens være legitimt. Dette er ikke nødvendigvis tilfelle i en situasjon med formative målemodeller og årsaksindikatorer. Årsaksindikatorer for det samme begrepet kan ha positiv-, negativ- eller ingen korrelasjon. Bruk av faktoranalyse og/eller inspeksjon av korrelasjonsmatriser for å velge indikatorer kan derfor føre til at man overser valide mål på et begrep dersom indikatorene bestemmer eller påvirker den latente teoretiske variabelen.

Det bør være "optimale" korrelasjoner mellom indikatorer for det samme begrepet

En vanlig anbefaling ved vurdering av mål på teoretiske begreper er at man bør søke å oppnå et optimalt nivå på korrelasjon mellom indikatorene. Hva dette *optimale* nivået er hersker det en viss uenighet om i litteraturen. Sellitz, Wrightsman, og Cook (1976) argumenterer f.eks. for at alle variabler som måler det samme begrepet bør være høyt korrelert. Cattell (1965) påstår det motsatte, nemlig at korrelasjoner mellom indikatorer for det samme begrepet ikke bør være for høye. Til slutt finner vi forskere som hevder at optimalt nivå på homogenitet i målingene oppnås når korrelasjoner mellom indikatorene ligger i området 0,2 til 0,4 (Briggs og Cheek, 1986). Når det benyttes effektindikatorer og reflektive målemodeller anbefaler Bollen og Lennox (1991) høye korrelasjoner mellom indikatorene for det samme begrepet, fremfor moderate eller lave korrelasjoner. I en situasjon der et begrep måles med årsaksindikatorer, og målemodellen er formativ, er ikke korrelasjoner mellom indikatorene et resultat av målemodellen. Høye korrelasjoner mellom indikatorene i målemodellen kan gjøre det vanskelig å skille betydningen av hver enkelt indikator for det latente teoretiske begrepet. I den grad en eller flere av indikatorene er en perfekt lineær kombinasjon av de andre, skaper dette et multikollinearitetsproblem. Moderate eller lave korrelasjoner kan redusere disse problemene.

Bollen og Lennox (1991) anbefaler høye korrelasjoner mellom indikatorer i situasjoner med effektindikatorer. Siden korrelasjoner mellom årsaksindikatorer er forklart av forhold utenfor målemodellen er det til sammenligning vanskelig å gi anbefalinger for optimalt nivå på korrelasjon for formative målemodeller.

Det bør velges indikatorer som reflekterer alle fasetter ved et begrep

For å unngå at man underminerer begrepsvaliditet anbefales det at man forsøker å velge indikatorer som representerer alle fasetter ved begrepet. Konsekvensene av å fjerne indikatorer er forskjellige for målemodeller med effektindikatorer og målemodeller med årsaksindikatorer (Bollen og Lennox, 1991). For reflektive målemodeller vil effektindikatorer for et endimensjonalt begrep være utskiftbare, da alle indikatorene som benyttes ved målingen påvirkes av det samme latente teoretiske begrepet. Imidlertid skal man merke seg at dette kan få betydning for reliabilitet ved begrepet. Dersom et begrep har flere dimensjoner, bør hver dimensjon behandles separat med sitt eget sett av effektindikatorer. I situasjoner med formative målemodeller og årsaksindikatorer vil ekskludering av en indikator endre sammensetningen av den latente teoretiske variabelen. Dette medfører at alle indikatorer som påvirker den latente variabelen bør inkluderes i målemodellen.

Korrelasjoner innenfor et begrep bør være høyere enn korrelasjoner mellom begreper

Når man skal velge indikatorer for å måle et begrep er det vanlig å lete etter indikatorer som “henger sammen”. Mer spesifikt anbefales det at korrelasjoner mellom indikatorer for det samme begrepet bør være høyere enn korrelasjoner mellom indikatorer for forskjellige begreper (Campbell og Fiske, 1959). For reflektive målemodeller med effektindikatorer kan høye korrelasjoner mellom indikatorer for forskjellige latente teoretiske begreper være et resultat av at begrepene i seg selv er korrelerte. I slike tilfeller vil ikke regelen om høyere korrelasjoner innenfor et bestemt begrep enn mellom begreper være gjeldende. Når et latent teoretisk begrep er målt med årsaksindikatorer, og en formativ målemodell benyttes, vil regelen ovenfor ha enda mindre gyldighet. Ved bruk av årsaksindikatorer settes det ingen restriksjoner på størrelsen på korrelasjonene mellom indikatorene. Det er derfor vanskelig å komme med eksakte anbefalinger om forholdet mellom korrelasjoner innenfor og mellom begreper for formative målemodeller (Bollen og Lennox, 1991) .

I utgangspunktet operasjonaliseres hotellgjesters vurderinger av et hotell i en multiattributtmodell bestående av 27 produktegenskaper. Produktegenskapene representerer indikatorer på de fire produktelementene som fremkommer gjennom å kombinere ulike produksjonsfaktorer. Vi har altså brutt en multiattributtmodell bestående av 27 egenskaper ned i fire mindre multiattributtmodeller som hver representerer ulike aspekter ved hotellproduktet.

Redegjørelsen vi foretar nedenfor vedrørende gruppering av de individuelle vurderings-spørsmålene i de fire produktelementene: service, struktur, backstage og kunde-avhengige produktelementer avviker på mange måter fra en tradisjonell validitets- og reliabilitetsdrøfting. Det kan derfor av enkelte fremsettes påstander om at den valgte fremgangsmåten er et resultat av “dårlige mål med svak begrepsvaliditet og reliabilitet”. Dette mener vi ikke er tilfellet i vår studie. Redegjørelsen er heller et forsøk på å ta konsekvensene av hvilke målemodeller vi har med å gjøre.

6.2.1.1 Vurdering av samhandlingsegenskaper (immaterielle egenskaper)

Gjestenes vurderinger av serviceegenskaper ved hotellproduktet omfatter variabler som har med deres interaksjon med personalet på hotellet å gjøre (Troye, 1990). Det kan diskuteres hvorvidt spørsmål som fanger opp ulike aspekt ved service, eller interaksjon mellom gjester og personale skal oppfattes som årsaks- eller effektindikatorer. På den ene siden kan det argumenteres for at endringer i oppfatninger av service påvirkes av

endringer i forhold som personalets imøtekommenhet, høflighet, serviceinnstilling og evne til å løse problemer. Dersom sammenhengene mellom disse indikatorene og det latente begrepet "service" oppfattes på denne måten vil vi ha en formativ målemodell med årsaksindikatorer. På den andre siden kan det argumenteres for at svargivning på spørsmål om vurdering av personalets imøtekommenhet, høflighet, serviceinnstilling og evne til å løse problemer reflekterer og påvirkes av gjestenes oppfatning av service og interaksjonen med personalet. Dersom indikatorene på service betraktes på denne måten vil vi stå overfor en situasjon med en reflektiv målemodell med effektindikatorer. Det kan argumenteres for at vurderinger av de ulike indikatorene på service har samme utgangspunkt, nemlig hotellets personale. Selv om disse kan være forskjellige og det derfor ikke nødvendigvis trenger å være høye korrelasjoner mellom indikatorer som måler ulike aspekter ved service, vil vi betrakte servicevurderinger som reflektive effektindikatorer i denne avhandlingen. Hovedargumentet for dette er som nevnt ovenfor at servicevurderinger baseres på samme utgangspunkt. I tillegg finnes det grunnlag for å anta at måling av ulike aspekter ved service kan være basert på halo-effekter⁷⁸, som også kan danne en felles kilde for vurderingene.

TABELL 6.1

Samhandlingsskalaen (immaterielle egenskaper) med spørsmål og statistiske egenskaper

| SPØRSMÅL* | α DERSOM SPM. EKSKL. | GJ.SN. KORR. MED ØVRIGE SPM. INNENF. SKALAEN | GJ.SN. KORR. MED SPM. UTENFOR SKALAEN |
|---|---|---|--|
| VU1 Personalets imøtekommenhet ved innsjekking på hotellet | 0,7945 | 0,46 | 0,29 |
| VU5 Personalets høflighet | 0,7921 | 0,50 | 0,35 |
| VU9 Personalets serviceinnstilling | 0,7797 | 0,52 | 0,35 |
| VU13 Personalets evne til å løse problemer | 0,7799 | 0,45 | 0,33 |
| VU17 Personaltes evne til å informere om aktivitetstilbud på stedet | 0,7998 | 0,42 | 0,35 |
| VU21 Personalets evne til å yte rask service | 0,7707 | 0,48 | 0,35 |
| Gjennomsnittskorrelasjon | | <u>0,47</u> | <u>0,34</u> |
| Chronbach's α: 0,8154 | | | |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendiks B

Med utgangspunkt i at vi betrakter indikatorene for service som effektindikatorer kan tradisjonelle kriterier for validitet- og reliabilitet benyttes. I tabell 6.1 rapporteres reliabilitet representert ved Chronbach's α , og serviceindikatorenes gjennomsnittskorrelasjoner innenfor skalaen, sammenlignet med serviceindikatorenes korrelasjoner

⁷⁸Den mest benyttede og aksepterte teori om halo-effekter impliserer at globale inntrykk påvirker vurderinger av spesifikke aspekter ved et objekt (Cooper 1981, Murphy og Anhalt 1992, Murphy og Reynolds 1988)

med indikatorer utenfor serviceskalaen som grunnlag for drøfting av konvergent og diskriminant validitet.

I følge Nunnally (1978) er en α på 0,7 eller høyere akseptabel. Dette betyr at en α på 0,8154, som observert for serviceskalaen, må sies å være svært akseptabel. Det er heller ingen spørsmål som er slik at α økes dersom de ekskluderes fra skalaen.

Konvergent og diskriminant validitet for serviceskalaen ble målt ved å kalkulere gjennomsnittskorrelasjoner. Som vi diskuterte ovenfor er et mye benyttet mål på konvergent validitet at indikatorer innenfor samme skala er høyt korrelert, mens et kriterium for diskriminant validitet impliserer⁷⁹ at indikatorer innenfor skalaen ikke korrelerer i nevneverdig grad med indikatorer utenfor skalaen (Campbell og Fiske, 1959). En nødvendig betingelse for konvergent og diskriminant validitet blir derfor at hver variabel har en høyere gjennomsnittskorrelasjon med andre variabler innenfor serviceskalaen enn med variabler utenfor serviceskalaen. Gjennomsnittskorrelasjoner beregnes med utgangspunkt i korrelasjonsmatrisen som er gjengitt i appendiks C⁷⁹. Legg merke til at alle variablene innenfor serviceskalaen har relativt høye gjennomsnittskorrelasjoner med andre variabler innenfor skalaen, og til sammenligning lavere gjennomsnittskorrelasjoner med variabler utenfor skalaen. Dette indikerer at serviceskalaen har konvergent og diskriminant validitet.

6.2.1.2 Vurdering av strukturegenskaper (materielle egenskaper)

Gjestenes vurderinger av de fysiske, materielle egenskapene ved hotellproduktet kalles vurdering av strukturelementer (Troye, 1990). Eksempler på strukturelementer er forhold som hotellets beliggenhet, badetrommet, hotellrommet, ventilasjon og parkeringsmuligheter. Vi velger å betrakte indikatorer på de fysiske, materielle strukturelementene som årsaksindikatorer. Vi antar at gjestenes vurderinger av f.eks. hotellets beliggenhet, badetrommet, hotellrommet, ventilasjon og parkeringsmuligheter påvirker deres oppfatning av det fysiske produktet, og ikke vice versa. Det finnes lite grunnlag for å anta at det i målemodellsammenheng eksisterer en generell oppfatning av det fysiske produktet som gir seg utslag i korrelerte vurderinger av uensartete forhold som f.eks. hotellrommet og parkeringsmulighetene.

⁷⁹Ved bruk av korrelasjonskoeffisientene i korrelasjonsmatrisen i appendiks 2 kalkuleres gjennomsnittskorrelasjonen mellom f.eks spørsmål 1 og de andre spørsmålene i serviceskalaen på følgende måte: $(0,60+0,59+0,37+0,35+0,37)/6 = 0,46$, mens gjennomsnittskorrelasjon mellom det samme spørsmålet og spørsmål utenfor serviceskalaen beregnes som følger: $(0,34+0,35+0,32+\dots+0,26+0,27)/21 = 0,29$.

Årsaksindikatorer for det samme begrepet kan altså ha positiv-, negativ- eller ingen korrelasjon. Dette forventer vi vil være tilfelle for indikatorer som benyttes for å måle ulike aspekter ved strukturelementer ved hotellproduktet. Bruk av faktoranalyse eller inspeksjon av korrelasjonsmatriser for å velge hvilke indikatorer som skal inkluderes i målet kan derfor føre til at man overser valide mål når indikatorene bestemmer, eller påvirker den latente teoretiske variabelen. Siden korelasjoner mellom årsaksindikatorer er forklart av forhold utenfor målemodellen er det også vanskelig å gi anbefalinger for optimalt nivå på korrelasjon for indikatorene som benyttes for å måle struktur-egenskapene ved produktet. Det er også vanskelig å komme med eksakte anbefalinger om forholdet mellom korrelasjoner innenfor strukturbegrepet og mellom struktur-begrepet og andre begreper.

TABELL 6.2

Struktur skalaen (materielle egenskaper) med spørsmål og statistiske egenskaper

| SPØRSMÅL* | GJ.SN. KORR. MED ØVRIGE SPM. INNENF. SKALAEN | GJ.SN. KORR. MED SPM. UTENFOR SKALAEN |
|--|---|--|
| VU3 Hotelllets beliggenhet | 0,29 | 0,31 |
| VU7 Hotellrommenes utseende/innendørsarkitektur | 0,38 | 0,31 |
| VU11 Baderommet | 0,34 | 0,29 |
| VU15 Sengens komfort | 0,31 | 0,29 |
| VU19 Rengjøring av rommet | 0,33 | 0,36 |
| VU22 Ventilasjon på rommet | 0,40 | 0,33 |
| VU23 Hotelllets utseende/arkitektur | 0,37 | 0,33 |
| VU25 Hotelllets utendørsanlegg (f.eks. hotellhage) | 0,27 | 0,29 |
| VU26 Parkeringsmulighetene | 0,29 | 0,23 |
| VU27 Mulighet for rolig og fredelig opphold | 0,33 | 0,29 |
| Gjennomsnittskorrelasjon | <u>0,33</u> | <u>0,30</u> |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendiks B

For å illustrere momentene i drøftingen ovenfor presenterer vi i tabell 6.2 gjennomsnittskorrelasjoner for indikatorer benyttet for å måle vurderinger av struktur-egenskapene ved hotellproduktet. Gjennomsnittskorrelasjonene mellom indikatorer som er benyttet for å måle ulike aspekter ved strukturelementene er gjennomgående lave. Vi observerer også at forskjellene mellom korrelasjoner innenfor strukturskalaen og korrelasjoner utenfor skalaen er små. I flere tilfeller korrelerer strukturindikatorer høyere med indikatorer utenfor skalaen (VU3, VU19 og VU25).

Selv om resultatene, i tradisjonell forstand, samlet indikerer svak konvergent og divergent validitet, er det vanskelig å si om dette skyldes forhold som ligger utenfor målemodellen, eller om det skyldes intern konsistens. Reliabilitet for årsaksindikatorer

kan ikke vurderes på bakgrunn av intern konsistens (Bollen, 1989). Det er derfor lite hensiktsmessig å rapportere slike mål på reliabilitet for struktur indikatorene. Gjennomsnittskorrelasjonene i tabell 6.2 er likevel tatt med for å illustrere hvordan indikatorene i strukturskalaen, eller de materielle egenskapene ved hotellproduktet, oppfører seg i forhold til indikatorer på egenskaper som ligger utenfor

6.2.1.3 Vurdering av backstage- og kundeavhengige egenskaper

Troye (1991) definerer backstageelementer som den delen av produksjonen som finner sted uten kundens tilstedeværelse. Disse produktelementene er i regelen materielle, selv om de ofte krever at personalet er til stede. Kunden er mer en passiv observatør som observerer denne delen av produktet. Eksempler på backstage elementer ved et hotellprodukt er tilberedelse av mat, informasjon, åpningstider, og rutiner for betaling. Målemodellen for backstageelementer ved hotellproduktet velger vi å betrakte som en formativ målemodell med årsaksindikatorer. Det er ingen grunn til å forvente at en latent teoretisk variabel som kalles backstage, og som har med "teknologien" hotellet benytter for å produsere det produktet gjesten opplever, vil påvirke gjestenes vurdering av forhold så forskjellige som informasjon om aktivitetstilbud, maten på hotellet, åpningstider osv. Det synes mer logisk å forestille seg at endringer i vurderinger av de individuelle indikatorene som benyttes for å måle backstage elementene fører til endringer i vurdering av backstage elementene. Basert på diskusjonen om hvilke kriterier som bør benyttes for å vurdere reliabilitet og validitet for formative målemodeller konkluderer vi med at de tradisjonelle og konvensjonelle kriteriene er lite egnet for validitets- og reliabilitetsvurderinger av backstageelementer.

Gjestenes egen aktivitet gjennom bruk av det fysiske produktet kalles kundeavhengige produktelementer (Troye, 1990). Denne delen av hotellproduktet omfatter egenskaper som f.eks. aktivitetstilbudet på hotellet. Målemodellen som benyttes for å måle de kundeavhengige produktelementene er på samme måte som for struktur- og backstageelementene definert som formativ. De individuelle indikatorene som benyttes for å måle vurderinger av kundeavhengige produktelementer er dermed årsaksindikatorer. Dette impliserer at endringer i vurdering av aktivitetstilbud, kultur og underholdningstilbud, turmuligheter og natteliv og underholdning om kvelden oppfattes å føre til endringer i den latente teoretiske variabelen, "vurdering av kundeavhengige produktelementer", og ikke vice versa.

Implikasjonene for validitets- og reliabilitetsdiskusjoner av å betrakte indikatorene for vurdering av de kundeavhengige elementene som årsaksindikatorer er de samme som for strukturelementene og backstageelementene. Gjennomsnittskorrelasjoner for måling av backstage- og kundeavhengige produktelementer finnes i appendiks C.

6.2.2 Viktighet av egenskapene som vurderes

Ved operasjonalisering av multiattributtmodeller for måling av kunders produktvurderinger er det vanlig å inkludere ett sett viktighetsspørsmål i tillegg til vurderingene i spørreskjemaet. Et problem med bruk av slike viktighetsspørsmål er at det kan være vanskelig å oppnå varians i svarene fordi respondenten vil oppfatte alle egenskaper som viktige. Tidligere i dette kapitlet beskrev vi hvordan manipulering av rekkefølgen på viktighetsspørsmål vs. vurderingss spørsmål i spørreskjemaet benyttes som en uavhengig eksperimentvariabel i det faktorielle undersøkelsesopplegget som er valgt for studien.

FIGUR 6.4
Måling av viktighet ved egenskapene som vurderes

**Hvor viktig mener De følgende faktorer er for et vellykket hotelopphold?
Sett en sirkel rundt tallet som uttrykker Deres oppfatning.**

| | <i>Helt uvesentlig</i> | | | | | <i>Svært vesentlig</i> | |
|---|----------------------------|---|---|---|---|----------------------------|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1. Personalets imøtekommenhet ved innsjekking på hotellet | | | | | | | |

Tilsvarende oppsett benyttes for de øvrige egenskapene.

Basert på litteraturstudier og erfaringer fra tidligere undersøkelser⁸⁰, operasjonaliseres og måles gjestenes tillegging av viktighet til de ulike egenskapene som de blir bedt om å vurdere som illustrert i figur 6.4.

Denne måten å operasjonalisere viktighetsspørsmål på har tidligere vist seg å gi gode resultater med hensyn til varians på viktighetsspørsmålene (Henjesand, 1991a). For å bestemme om manipuleringene, som ble foretatt ved å endre instruksjon og rekkefølge på spørsmålene, påvirker subjektene vektlegging av de 27 produktegenskapene som

⁸⁰Gjennom undersøkelser foretatt ved Stiftelsen for samfunns- og næringslivsforskning er spørreskjema for måling av gjesters vurderinger av hotellprodukter utviklet og testet (Henjesand 1991a, Henjesand 1991b, Henjesand og Troye 1991)

ble inkludert i spørreskjemaet, blir følgende operasjonaliseringer av viktighet benyttet som grunnlag for sammenligninger:

1. Gjennomsnittsvektlegging for individuelle egenskaper mellom de ulike eksperimentgruppene i det faktorielle undersøkelsesopplegget. Disse operasjonaliseringene gjør det mulig å bestemme om vektlegging foretatt av subjekter i en eksperimentgruppe er konsistent forskjellig fra vektlegging foretatt av subjekter tilhørende en annen eksperimentgruppe.
2. Vektlegging av spesifikke egenskaper gruppert sammen i summerte gjennomsnittsskaler. Her benyttes de samme grupperingene som ble benyttet for de spesifikke vurderingene. Disse sammenligningene vil gjøre oss i stand til å fokusere på forskjeller knyttet til vektlegging av egenskaper knyttet til f.eks. service (immaterielle egenskaper) og det fysiske produktet som leveres (materielle egenskaper).

6.2.3 Totalvurderinger og atferdsintensjoner

I avhandlingen har vi fremsatt hypoteser vedrørende de spesifikke vurderingenes evne til å predikere totalvurdering av produktet og atferdsintensjoner overfor produktet. Disse hypotesene er i hovedsak basert på perspektiver fra holdningsteori (Fishbein, 1963; Fishbein, 1972; Fishbein og Ajzen, 1975; Lutz, 1991) der sammenhenger mellom kognitive affektive og konative komponenter er viet stor oppmerksomhet. Både totalvurderinger og atferdsintensjoner måles med multiple mål. Nedenfor viser vi først hvordan totalvurderinger av hotellet blir operasjonalisert i spørreskjemaet. Deretter beskriver vi hvordan atferdsintensjoner blir målt.

Som illustrert i figur 6.5 blir gjestenes totalvurdering målt med tre spørsmål. I de to første spørsmålene ber vi gjestene vurdere hvor bra eller dårlig hotellet vurderes. Med parallell til instruksjonene som blir gitt for de spesifikke vurderingene, reflekterer disse totalvurderingsspørsmålene objektforankrede totalvurderinger. I det tredje totalvurderingsspørsmålet ber vi subjektene gi uttrykk for sin tilfredshet med den totalopplevelsen hoteloppholdet har gitt opphav til. Med referanse til instruksjonen for de spesifikke vurderingene reflekterer dette totalvurderingsspørsmålet en subjektforankret totalvurdering.

TABELL 6.3
Totalvurderingsskalaen med spørsmål og statistiske egenskaper

| SPØRSMÅL* | α DERSOM SPM. EKSKL. | GJ.SN. KORR. MED ØVRIGE SPM. INNENF. SKALAEN | GJ.SN. KORR. MED SPM. UTENFOR SKALAEN |
|---|---|---|--|
| S22 Sammenlignet med andre hotell i Norge i samme prisklasse var dette hotellet (bra-dårlig) | 0,8624 | 0,79 | 0,65 |
| S23 I forhold til den prisen De har betalt hvordan vil de vurdere dette hotellet (bra-dårlig) | 0,8317 | 0,81 | 0,64 |
| S24 Basert på Deres totalopplevelse, hvor tilfreds eller utilfreds er De med dette hotellet | 0,9147 | 0,76 | 0,68 |
| Gjennomsnittskorrelasjon | | <u>0,79</u> | <u>0,66</u> |
| Chronbach's α: 0,9122 | | | |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendiks B

Reliabilitetstallene i tabellen gir videre indikasjoner på at spørsmål 24, kan ekskluderes fra skalaen. Basert på argumentasjonen for å ta dette spørsmålet med velger vi imidlertid å beholde spørsmålet i skalaen da dette måler en subjektforankret totalvurdering, mens de to første spørsmålene reflekterer objektforankrete totalvurderinger. En økning i Chronbach's α fra 0,9122 til 0,9147 indikerer også at det er lite å hente på å ekskludere spørsmålet fra skalaen. For analysene der vi tester forskjeller mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer vil vi behandle objektforankrete- og subjektforankrete totalvurderinger separat.

Gjestenes atferdsintensjoner etter at de har benyttet og vurdert hotellproduktet er representert ved sannsynligheten for at de kommer til å besøke hotellet igjen ved en senere anledning, og ved sannsynligheten for at de kommer til å anbefale hotellet til en god venn. Atferdsintensjoner kan dermed betraktes som nøytrale totalvurderings-spørsmål som reflekterer både de stimulusforankrete og de subjektforankrete totalvurderingene. Disse sannsynlighetene måles ved å be respondentene sette en sirkel på en 11-punktsskala der 1 er "svært usannsynlig" og 11 er "svært sannsynlig". Det ble vurdert å benytte en 10 punktsskala for disse spørsmålene for å hjelpe respondentene til å tenke i sannsynligheter og prosent. 11 punktsskalaen ble imidlertid valgt for å beholde samme skalalengde som på vurderingss spørsmålene.

På samme måte som for totalvurderingss spørsmålene argumenterer vi for at indikatorene som benyttes for å måle atferdsintensjoner kan betraktes som effektindikatorer. Da vi

har en reflektiv målemodell velger vi å rapportere konvensjonelle mål på validitet og reliabilitet.

TABELL 6.4
Atferdsintensjonsskalaen med spørsmål og statistiske egenskaper

| SPØRSMÅL* | | α DERSOM SPM. EKSKL. | GJ.SN. KORR. MED ØVRIGE SPM. INNENF. SKALAEN | GJ.SN. KORR. MED SPM. UTENFOR SKALAEN |
|-------------------------------|--|--------------------------------------|---|--|
| S25 | Hvis De skulle på ferie til dette området ved en senere anledning, hvor sannsynlig er det at De vil velge det samme hotellet | - | 0,90 | 0,63 |
| S26 | Hvis en god venn av Dem skulle på ferie til dette området, hvor sannsynlig er det da at De ville anbefale nettopp dette hotellet | - | 0,90 | 0,67 |
| Gjennomsnittskorrelasjon | | | <u>0,90</u> | <u>0,65</u> |
| Chronbach's α : 0,9455 | | - | | |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendiks B

Tallene i tabell 6.4 viser at atferdintensjonsskalaen har både konvergent og diskriminant validitet. Reliabilitet representert ved Chronbach's α er også svært akseptabel.

6.2.4 Sikkerhet i vurderinger

I kapittel 3 ble sikkerhet i vurderinger (*confidence*) beskrevet som en modererende variabel ved måling av holdninger. Sikkerheten i vurderinger av holdningsspørsmål kan blant annet være avhengig av i hvilken grad et individ har utviklet holdninger overfor et objekt (Sample og Warland, 1973). Sikkerhet i vurderinger er tidligere operasjonalisert på forskjellige måter. Sample & Warland (1973) ber respondentene om å indikere hvor sikre de er i hver av responsene på holdningsspørsmålene som benyttes i spørreskjemaet. Det benyttes en skala fra 1 (svært usikker) til 5 (svært sikker). Fishbein & Ajzen (1975) og Marks og Kamins (1988) måler *belief confidence* for hver egenskap ved å be respondentene gi uttrykk for sin sikkerhet i at vurdering av *beliefs* er korrekt. Det ble benyttet en sju-punkt-skala der benevnelsene er 1 (svært usikker) og 7 (svært sikker). Antil (1983) instruerer respondentene til å indikere hvor sikre de er i hver av responsene på holdningsspørsmålene som benyttes i spørreskjemaet. Her benyttes en skala fra 1 (svært usikker) til 5 (svært sikker).

Som et alternativ til sikkerhetsmålinger for hver vurdert egenskap, foreslår Antil (1983) at man benytter ett, eller et mindre antall spørsmål for å måle sikkerhet i vurderinger. Prosedyren som foreslås er å benytte et spørsmål der man ber respondenten gi uttrykk for hvor sikker han/hun er i svarene på vurderingsspørsmålene sett under ett. På denne måten unngår man problemer forbundet med å koble sikkerhetsspørsmål til hver vurdering.

FIGUR 6.6
Måling av sikkerhet i vurderinger for de individuelle egenskapene som vurderes

Hvor sikker er De i Deres vurdering?
 Sett sirkel rundt tallet som uttrykker Deres svar

| | | | | | |
|----------------|---|---|---|---|---------------|
| <i>svært</i> | | | | | <i>svært</i> |
| <i>usikker</i> | | | | | <i>sikker</i> |
| 1 | 2 | 3 | 4 | 5 | |

(Dette spørsmålet blir gjentatt for hver egenskapsvurdering)

I avhandlingen er vi i tillegg til å analysere forskjeller i gjennomsnittlig sikkerhet for samtlige av de 27 egenskapene som vurderes, også interessert i å studere sikkerhet i vurderinger knyttet til de individuelle vurderingene og til vurderinger gruppert i summerte skalaer slik disse er presentert tidligere i dette kapitlet. Vi har derfor valgt å operasjonalisere sikkerhet i vurderinger som illustrert i figur 6.6.

6.2.5 Relativ betydning av immaterielle- og materielle egenskaper

Vi har i avhandlingen gjort et skille mellom vurdering av materielle og immaterielle egenskaper ved et produkt. Med relativ betydning vil vi forstå hvordan disse egenskapene inngår i kundens totalvurdering av produktet. Basert på litteraturgjennomgangen og teoridrøftelsene i kapittel 3, argumenterte vi for at materielle (struktur) egenskaper ved produktet vil ha sterkere prediksjonsevne for kundens totalvurdering og atferdsintensjoner enn immaterielle (service) egenskaper. I hypotese 5a, 5b, 8a og 8b fremmet vi, med dette som utgangspunkt, påstander om at betydningen av materielle og immaterielle egenskaper for kundens totalvurdering og atferdsintensjoner vil være avhengig av ordlyd i spørsmålsstilling, benevnelse på svarskala og produktkunnskap.

Med bakgrunn i denne drøftingen velger vi å beskrive relativ betydning av materielle- og immaterielle produktegenskaper på følgende måte:

$$\frac{\text{Totalvurdering/}}{\text{Atferdsintensjoner}} = f(\text{vurdering av materielle egenskaper,} \\ \text{vurdering av immaterielle egenskaper})$$

Kundens totalvurdering av produktet blir definert som en funksjon av materielle og immaterielle egenskaper. I regresjonstermer kan dette beskrives på følgende måte:

$$\frac{\text{Totalvurdering/}}{\text{Atferdsintensjoner}} = \text{konstantledd} + \\ \text{B}_1 \text{ vurdering av materielle egenskaper} + \\ \text{B}_2 \text{ vurdering av immaterielle egenskaper}$$

Relativ betydning vil med utgangspunkt i denne regresjonsligningen bli målt ved betydningen av materielle og immaterielle egenskaper, reflektert ved størrelsen og fortegnet på B₁ og B₂. En alternativ måte å operasjonalisere struktur i vurderingene, er korrelasjon mellom totalvurdering og vurdering av henholdsvis materielle og immaterielle egenskaper. Ved operasjonalisering av materielle- og immaterielle egenskaper tar vi utgangspunkt i de fire produktelementene som vi benyttet for å beskrive et produkt i kapittel 3⁸¹. Vi er her spesielt interessert i å fokusere på betydningen av materielle og immaterielle egenskaper for totalvurderingene, og vil derfor kun benytte samhandlingsegenskapene og strukturegenskapene for å operasjonalisere henholdsvis immaterielle og materielle produktegenskaper. Samhandlingsegenskaper representerer i denne sammenheng "rene" immaterielle egenskaper, mens strukturegenskaper representerer "rene" materielle egenskaper ved hotellproduktet. De øvrige produkt-egenskapene som respondentene ble bedt om å vurdere er vanskelig å klassifisere som enten materielle-, eller immaterielle egenskaper. Ved å fokusere på samhandlings-egenskaper og strukturegenskaper har vi på denne måten forsøkt å "rendyrke" operasjonaliseringen av immaterielle og materielle egenskaper for analysene.

6.3 Modererende variabler

I avhandlingen opererer vi med to modererende variabler; respondentenes produkt-kunnskap og kvalitetsnivået på produktet som vurderes. I det følgende vil vi gjøre rede for hvordan disse variablene er operasjonalisert i spørreskjemaet som benyttes for innsamling av data.

⁸¹Samhandlings-, struktur-, backstage-, og kundeavhengige produktelementer

6.3.1 Respondentenes produktkunnskap

Alba (1987) betrakter erfaring som en komponent i et multidimensjonalt produktkunnskapsbegrep, der kunnskap har to komponenter: *kjennskap* og *ekspertise* (Alba og Hutchinson, 1987; Jacoby et al., 1983). Kjennskap defineres som “antall produktrelaterte erfaringer som har blitt akkumulert av konsumenten”. Ekspertise er definert som “evnen til å utføre produktrelaterte oppgaver på en suksessfull måte”. Produktrelaterte erfaringer omfatter forhold som eksponering for reklame, informasjonssøking, interaksjon med selgere, valg og beslutningstaking, kjøp og bruk av produkter i ulike situasjoner. Ekspertise inkluderer både konsumentens kognitive strukturer, som f.eks. antakelser om egenskaper ved produktet, og kognitive prosesser, som f.eks. beslutningsregler for atferd på bakgrunn av antagelsene. Økt produktkjennskap fører til økt ekspertise (Alba og Hutchinson, 1987). Med utgangspunkt i oversikten som blir gitt av Alba og Hutchinson (1987) ser vi at kjennskap til produktet og produktrelatert erfaring betraktes som årsaksfaktorer som fører til produktkunnskap og ekspertise.

Ved måling av produktkunnskap som et multidimensjonalt begrep er det i hovedsak tre forskjellige innfallsvinkler som er benyttet (Brucks, 1985; Brucks, 1986; Cole, Gaeth, og Singh, 1986):

1) Objektive mål på kunnskap. Formålet med objektive mål er å finne ut hvor mye en person vet om et bestemt domene basert på objektive kriterier. Måling av kunnskap ved bruk av objektive mål innebærer f.eks. bruk av en “riktig/galt”- test for å finne ut hvor mye respondentene vet om et bestemt domene. Alternativt benyttes et åpent spørsmål hvor man ber respondentene rapportere hvilke egenskaper de vil benytte for å vurdere produktet. Antall egenskaper, i tillegg til hvilke egenskaper som fremkommer vil her representere subjektens kognitive kompleksitet ved vurdering av produktet .

2) Subjektive selvrapporterte mål på kunnskap. Her blir respondentene bedt om å vurdere sin egen kunnskap om et domene. Måling av kunnskap ved bruk av selvrapporterte mål innebærer at vi lar respondentene rapportere hvor mye de vet om et bestemt domene ved bruk av spørsmål som f.eks. “jeg vet mye om hvordan jeg skal vurdere et hotell” eller “når jeg skal vurdere et hotell regner jeg meg som ekspert”.

3) Frekvensmål på tidligere kjøps- og/eller bruksatferd. Antagelsen bak frekvensmålene er at hvis man har brukt et produkt, kjøpt et produkt eller søkt informasjon om et produkt, vil dette føre til økt produktkunnskap (Bettman og Park, 1980). Måling av kunnskap ved bruk av frekvensmål innebærer bruk av åpne spørsmål som f.eks. “hvor mange ganger har De bodd på hotell i løpet av de siste tre årene?”. Selnes og Grønhaug

(1986) påpeker to konseptuelle problemer forbundet med bruk av frekvensbaserte erfaringsmål på kunnskap.

For det første kan produktkunnskap utvikles gjennom informasjonssøking og bruk av informasjon. Kunnskap er på denne måten ikke nødvendigvis relatert til direkte erfaring. For det andre kan erfaring i form av økt bruk av et produkt forekomme, uten at kunnskapen om produktet øker. Dette er forhold som vi må være oppmerksom på ved operasjonalisering av kunnskapsbegrepet gjennom bruksfrekvens.

Ytterligere et forhold som gjør frekvensmålinger problematiske er de kognitive oppgavene slike mål pålegger respondentene (Schwarz, 1990). Ideelt sett bør det benyttes åpne spørsmål der respondentene gis mulighet til å rapportere det eksakte svaret. Dette vil bidra til en kontinuerlig skala med optimale skalaegenskaper. Som vi drøftet i kapittel 4 må respondentene utføre forskjellige oppgaver ved svar på spørsmål i et spørreskjema.

Først må spørsmålene tolkes og det må klargjøres hvilke svar som forventes. Bruk av lukkede svaralternativer der man gir respondentene kategorier å velge mellom har vist seg å forenkle denne fasen (Schwarz, 1990). Neste oppgave for respondenten blir å forsøke å huske relevante tilfeller. Bruk av lukkede svaralternativer har også vist seg å forenkle denne fasen. Å tidfeste de tilfellene som er fremkallt i minnet til den relevante tidsperioden blir neste oppgave. Hvor fremtredende episodene som huskes er, blir ofte brukt som indikator for tidfesting (Schwarz, 1990). Til slutt må svarene uttrykkes i spørreskjemaet. Et problem kan da være at svarkategoriene som er gitt i skjemaet ikke passer med respondentenes egne kategorier (Schwarz, 1990). For å øke gyldigheten av frekvensrapporteringer anbefales det å benytte lukkede svarkategorier og begrense rapporteringen til et bestemt tidsrom (Schwarz, 1990).

I tidligere forskning, der man har benyttet en multidimensjonal tilnærming til måling av kunnskap, har man vist at effekter av kunnskap for konsumenters atferd er en funksjon av hvordan kunnskapen er målt (Alba og Marmorstein, 1986). Basert på (Alba og Hutchinson, 1987) beskrivelse av produktrelatert erfaring som en årsaksfaktor som fører til produktkunnskap og ekspertise velger vi å måle kunnskap ved bruk av frekvensmål som sier noe om respondentens erfaring med bruk av produktet, i tillegg til subjektive og objektive mål på produktkunnskap.

Målene på som benyttes for å måle erfaring, og subjektiv- og objektiv produktkunnskap blir på bakgrunn av drøftingen ovenfor som illustrert i figur 6.7:

FIGUR 6.7
Måling av erfaring og produktkunnskap

Fekvenssmål:

Hvor mange ganger har De overnattet på hotell i løpet av de siste tre årene?

- | | | | |
|----|-----------------------|----|------------------------------|
| 1. | 0 - 5 overnattinger | 4. | 16 - 25 overnattinger |
| 2. | 6 - 10 overnattinger | 5. | 26 - 35 overnattinger |
| 3. | 11 - 15 overnattinger | 6. | 36 overnattinger eller flere |

Subjektive kunnskapsmål:

I hvilken grad er De enig i følgende utsagn: "Når jeg skal vurdere et hotell, vet jeg hvilke egenskaper som er viktige, og som må være tilstede for at jeg skal bli fornøyd med oppholdet"

| | | |
|--|--------------------|-------------------|
| Sett en sirkel rundt tallet som uttrykker Deres svar | <i>Svært uenig</i> | <i>Svært enig</i> |
| | 1 2 3 4 5 | 6 7 |

Hvordan tror De Deres venner vurderer Deres kunnskaper når De skal vurdere et hotell?

| | | |
|--|---------------------------------|------------------------------|
| Sett en sirkel rundt tallet som uttrykker Deres svar | <i>Svært dårlige kunnskaper</i> | <i>Svært gode kunnskaper</i> |
| | 1 2 3 4 5 | 6 7 |

Objektive kunnskapsmål

La oss anta at De skal velge hotell som overnattingssted for neste feriereise. Hvilke egenskaper ville De ha benyttet for å skille mellom de ulike alternativene basert på Deres nåværende kunnskaper?

- | | |
|---------|----------|
| 1. | 10. |
| 2. | 11. |
| 3. | 12. |
| 4. | 13. |
| 5. | 14. |
| 6. | 15. |
| 7. | 16. |
| 8. | 17. |
| 9. | 18. |

For å forenkle tolking av analysene der produktkunnskap inngår som en modererende variabel, har vi valgt å benytte de tre målene på produktkunnskap, som er presentert i tabellen ovenfor, som grunnlag for å dele respondentene inn i tre grupper: En gruppe med lav produktkunnskap, en gruppe med middels produktkunnskap og en gruppe med høy produktkunnskap.

For identifisering av de tre gruppene ble det benyttet clusteranalyse. Clusteranalysen gjør oss i stand til å identifisere respondenter som er homogene innenfor gruppen de tilhører, og som samtidig er maksimalt forskjellige fra respondenter i de øvrige gruppene⁸². Før clusteranalysen ble gjennomført ble det subjektive kunnskapsmålet

⁸²For en nærmere beskrivelse av clusteranalyse, se: Hair jr., Andreson, og Tatham (1987).

konstruert gjennom å lage en summert skala bestående av de to spørsmålene som måler subjektiv kunnskap. I tillegg ble det objektive kunnskapsmålet redusert til seks kategorier. Resultatet av clusteranalysen gav følgende gruppestørrelser for de tre produktkunnskapsnivåene; Lav produktkunnskap: 195 respondenter, Middels produktkunnskap: 154 respondenter og Høy produktkunnskap: 252 respondenter. Disse tre gruppene benyttes som grunnlag for å analysere forskjeller mellom respondenter med ulikt nivå av produktkunnskap.

6.3.2 Kvalitetsnivå på produktet som vurderes

Som grunnlag for operasjonalisering av kvalitetsnivå på produktene som vurderes kan det skilles mellom to alternative fremgangsmåter. Man kan enten velge å benytte såkalte ekspertvurderinger, eller basere operasjonaliseringen av kvalitetsnivå på gjennomsnittlige totalvurderinger for hotellene som deltok i undersøkelsen. Felles for begge fremgangsmåtene er en gruppering av hotellene i henhold til en inndeling i lavkvalitetshotell, hotell med gjennomsnittskvalitet og høykvalitetshotell.

Vi har valgt å basere operasjonaliseringen av kvalitetsnivå på gjennomsnittlige totalvurderinger for hvert hotell, slik disse fremkommer gjennom respondentenes svargivning på totalvurderingsspørsmålene i spørreskjemaet. Først ble gjennomsnittlig totalvurdering beregnet for hvert hotell. Deretter ble hotellene gruppert ved hjelp av variansanalyse. På denne måten kategoriseres hotellene i tre grupper basert på signifikantesting av forskjeller i respondentenes totalvurderinger. Som følge av variansanalyse plasseres tre hotell i gruppen *lavkvalitetshotell*, fem hotell plasseres i gruppen *gjennomsnittshotell* og tre hotell plasseres i gruppen *høykvalitetshotell*.

Som nevnt ovenfor er bruk av ekspertvurderinger en alternativ fremgangsmåte for kategorisering av hotellene som lavkvalitetshotell, hotell med gjennomsnittskvalitet og høykvalitetshotell. Her ville det ha vært nødvendig å beregne konsistens i ekspertenes vurderinger av hotellene som et mål på godheten av kategoriseringen⁸³. Når vi har valgt å benytte respondentenes totalvurderinger, slik disse kommer til uttrykk i spørreskjemaet, som mål på kvalitetsnivå på produktene som vurderes, kan respondentene oppfattes som eksperter. Bruk av variansanalyse som grunnlag for å kategorisere hotellene basert på signifikante forskjeller i totalvurderinger kan derfor på mange måter sies å tilsvare ekspertvurderinger med et større antall eksperter.

⁸³*Inter rater* reliabilitet.

DEL
IV
BESKRIVELSE AV
DATA-
MATERIALET
OG
HYPOTESETESTING

Denne delen av avhandlingen består av 2 kapitler. Først, i kapittel 7, underlegger vi det innsamlete datamaterialet en grundig inspeksjon med hensyn på datakvalitet. I denne forbindelse diskuterer vi også forutsetninger for valg av statistiske analysemetoder. I kapittel 8 foretar vi deretter en grundig gjennomgang av resultatene fra hypotesetestingen.

7

Beskrivelse av data og valg av statistiske analyseteknikker

I dette kapittelet presenteres resultater fra innledende dataanalyser som ble foretatt for å vurdere kvaliteten på de innsamlede dataene, som senere vil bli benyttet for å teste hypotesene som ble fremsatt i kapittel 4. Først presenteres en oversikt som viser respondentenes manglende evne, eller vilje til å svare på de spesifikke vurderings-spørsmålene. Deretter presenteres beskrivende statistikk for de avhengige variablene det fokuseres på i studien. Til slutt i kapittelet drøftes valg av statistiske analyseteknikker. I denne forbindelse presenteres resultater fra tester som ble foretatt for å undersøke om de innsamlede dataene tilfredsstillende forutsetningene som ligger til grunn for analyseteknikkene som er valgt for hypotesetestingen.

7.1 Manglende svargivning⁸⁴

Når vi ber hotellgjester vurdere en rekke spesifikke egenskaper ved hotellproduktet forventer vi ikke at alle respondenter har grunnlag for å foreta vurderinger av samtlige produktegenskaper. Den mest fremtredende årsaken til manglende svargivning på enkeltspørsmål i studien som er gjennomført i forbindelse med denne avhandlingen, vil være at respondentene ikke har erfart de sidene ved hotellproduktet som de blir bedt om å vurdere. I en slik sammenheng vil det være riktigere å la være å svare på spørsmålet, enn å forsøke å gjette seg frem til et svar.

⁸⁴*Item nonresponse* er en feilkilde som oppstår når respondentene mangler evne eller vilje til å svare på et bestemt spørsmål i et spørreskjema (Churchill 1991).

Ved gjennomføring av dataanalyser kan manglende svargivning på enkeltspørsmål behandles på forskjellige måter. En mulighet er å erstatte en respondents manglende svargivning på et spørsmål med gjennomsnittet for de som har svart. En alternativ mulighet er å ekskludere respondenter som har unnlatt å svare på ett eller flere spørsmål fra analysene⁸⁵. En tredje mulighet er å inkludere respondenter som har unnlatt å svare på ett eller flere spørsmål, og dermed beholde informasjonen som ligger i de spørsmålene de har svart på for analysene⁸⁶. Til slutt kan manglende svargivning på enkeltspørsmål erstattes med estimerte verdier, basert på sannsynlighetsbetraktninger.

TABELL 7.1
Manglende svargivning på enkeltspørsmål

| SPØRSMÅL* | ANTALL MANGLENDE SVAR | % MANGLENDE SVAR |
|--|----------------------------------|-----------------------------|
| <i><u>Samhandling /servicevurderinger</u></i> | | |
| VU1 Personalets imøtekomenhet ved innsjekking | 2 | 0,3 |
| VU5 Personalets høflighet | 14 | 2,3 |
| VU9 Personalets serviceinnstilling | 18 | 3,0 |
| VU13 Personalets evne til å løse problemer | 96 | 15,8 |
| VU17 Personaltes evne til å informere om aktivitetstilbud | 77 | 12,7 |
| VU21 Personalets evne til å yte rask service | 71 | 11,7 |
| <i><u>Struktur vurderinger</u></i> | | |
| VU3 Hotelllets beliggenhet | 13 | 2,1 |
| VU7 Hotellrommenes utseende/innendørsarkitektur | 2 | 0,3 |
| VU11 Baderommet | 30 | 4,9 |
| VU15 Sengens komfort | 17 | 2,8 |
| VU19 Rengjøring av rommet | 32 | 5,3 |
| VU22 Ventilasjon på rommet | 20 | 3,3 |
| VU23 Hotelllets utseende/arkitektur | 23 | 3,8 |
| VU25 Hotelllets utendørsanlegg (f.eks. hotellhage) | 50 | 8,2 |
| VU26 Parkeringsmulighetene | 21 | 3,5 |
| VU27 Mulighet for rolig og fredelig opphold | 19 | 3,1 |
| <i><u>Backstage vurderinger</u></i> | | |
| VU2 Informasjon om aktivitetstilbudet på stedet via brosjyrer, oppslagstavler o.l. på hotellet | 24 | 3,9 |
| VU6 Maten på hotellet | 34 | 5,6 |
| VU10 Åpningstidene i bar og restaurant | 60 | 9,9 |
| VU14 Menyens variasjon | 61 | 10,0 |
| VU18 Rutine prosedyre for betaling | 61 | 10,0 |
| VU24 Informasjon om sikkerhetsrutiner | 34 | 5,6 |
| <i><u>Kundeavhengig vurderinger</u></i> | | |
| VU4 Aktivitetstilbudet på hotellet | 55 | 9,0 |
| VU8 Kultur og underholdning på stedet | 78 | 12,8 |
| VU12 Turmulighetene på stedet | 55 | 9,0 |
| VU16 Natteliv og underholdningstilbudet om kvelden | 74 | 12,2 |
| VU20 Shoppingtilbud på stedet | 57 | 9,4 |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendiks B

⁸⁵Listwise deletion of missing variables

⁸⁶Pairwise deletion of missing variables

I tabell 7.1 gir vi en oversikt over manglende svargivning for hver egenskap som gjestene blir bedt om å vurdere. Basert på argumentasjonen om at vi ikke forventer at alle respondenter har grunnlag for å svare på samtlige vurderingsspørsmål, konkluderer vi med at tallene som rapporteres for manglende svargivning på vurderingsspørsmålene ligger innenfor akseptable grenser. Vi velger derfor å inkludere alle respondenter i analysene, selv om de ikke har svart på ett eller flere vurderingsspørsmål. I analyser der vi benytter de individuelle vurderingene vil dette føre til ulike utvalgsstørrelser for hver vurdering. Når vi benytter summerte gjennomsnittsskalaer i analysene, vil manglende svargivning på enkeltspørsmål bety at grunnlaget for beregning av gjennomsnittene bli forskjellig.

Manglende svargivning på enkeltspørsmål kan medføre skjevheter i datamaterialet dersom forekomsten av manglende svar er systematisk forskjellig i eksperimentgruppene. Dersom dette er tilfellet kan det stilles spørsmål om manipuleringene, som er foretatt ved utforming av de ulike spørreskjematypene, påvirker respondentenes evne eller vilje til å svare på vurderingsspørsmålene. Som grunnlag for å kontrollere for denne uønskete kilden til variasjon, ble det gjennomført en variansanalyse, der antall manglende svar ble betraktet som avhengig variabel, og der manipuleringene i spørreskjemaet betraktes som uavhengige variabler. Variansanalysen er gjengitt i tabell 7.2 nedenfor.

TABELL 7.2
ANOVA, summert antall manglende svar på enkeltspørsmål mellom eksperimentgruppene

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|----------------------------------|----------------|-----|-------------|-------|------------|
| Within+residual | 8303,946 | 602 | 13,794 | | |
| Eksperimentgruppene ^a | 27,153 | 5 | 5,431 | 0,394 | 0,853 |
| Explained | 27,153 | 5 | 5,431 | 0,394 | 0,853 |
| Total | 8331,099 | 607 | 13,725 | | |

^a Gruppegjennomsnitt:

| | |
|--|------------------|
| Objektforankret instruksjon - uten viktighetsspørsmål | : 1,69 (n = 113) |
| Objektforankret instruksjon - viktighetsspørsmål før vurderingsspørsmål | : 2,14 (n = 103) |
| Objektforankret instruksjon - viktighetsspørsmål etter vurderingsspørsmål | : 1,76 (n = 100) |
| Subjektforankret instruksjon - uten viktighetsspørsmål | : 1,48 (n = 101) |
| Subjektforankret instruksjon - viktighetsspørsmål før vurderingsspørsmål | : 2,00 (n = 84) |
| Subjektforankret instruksjon - viktighetsspørsmål etter vurderingsspørsmål | : 1,81 (n = 107) |

Variansanalysene viser at antall manglende svar er relativt jevnt fordelt i de seks eksperimentgruppene. Manglende svar på vurderingsspørsmål oppfattes derfor ikke som noen trussel mot studiens validitet.

7.2 Beskrivende statistikk

Før vi går videre til en mer detaljert analyse av de innsamlete dataene med hensyn på grunnlaget for hypotesetesting, vil vi studere enkle tabeller som viser beskrivende statistikk. Hensikten er her å foreta en kontroll av dataene, blant annet med hensyn på feil som kan ha oppstått ved punching av dataene eller ved konstruksjon av summerte skalaer. Beskrivende statistikk for hele utvalget som ble oppnådd i studien er vist i tabell 7.3.

TABELL 7.3
Beskrivende statistikk for hele utvalget

| VARIABEL | GJ.SN. | ST.AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|--|--------|----------|-------|-------|-------|-------|-----|
| Summert gjennomsnittsvurdering | 3,15 | 1,16 | 1,76 | -0,98 | -3,69 | 4,96 | 608 |
| Samhandling vurdering (immaterielle egenskaper) | 3,54 | 1,24 | 0,80 | -0,98 | -2,17 | 5,00 | 607 |
| Struktur vurdering (materielle egenskaper) | 3,21 | 1,30 | 2,32 | -1,27 | -4,10 | 5,00 | 608 |
| Backstage vurdering | 3,17 | 1,37 | 1,93 | -1,12 | -4,50 | 5,00 | 604 |
| Kundeavhengig vurdering | 2,44 | 1,58 | 0,66 | -0,65 | -5,00 | 5,00 | 598 |
| Samhandling viktighet (immaterielle egenskaper) | 5,93 | 0,78 | 1,05 | -0,83 | 2,67 | 7,00 | 381 |
| Struktur viktighet (materielle egenskaper) | 5,70 | 0,68 | 1,71 | -0,79 | 2,30 | 7,00 | 380 |
| Backstage viktighet | 5,29 | 0,87 | 0,67 | -0,59 | 2,00 | 7,00 | 380 |
| Kundeavhengig viktighet | 4,20 | 1,09 | 0,34 | -0,28 | 1,00 | 7,00 | 378 |
| Totalvurdering | 3,35 | 1,54 | 3,52 | -1,59 | -4,33 | 5,00 | 598 |
| Objektforankret | 3,23 | 1,72 | 3,36 | -1,62 | -4,50 | 5,00 | 595 |
| Subjektforankret | 3,59 | 1,42 | 5,92 | -1,87 | -5,00 | 5,00 | 590 |
| Atferdsintensjoner | 9,08 | 2,24 | 2,00 | -1,52 | 1,00 | 11,00 | 597 |
| Summert gjennomsnittssikkerhet | 4,16 | 0,68 | 1,04 | -0,88 | 1,00 | 5,00 | 582 |
| Samhandlings sikkerhet (immaterielle egenskaper) | 4,13 | 0,81 | 0,41 | -0,89 | 1,00 | 5,00 | 582 |
| Struktur sikkerhet (materielle egenskaper) | 4,42 | 0,57 | 2,46 | -1,25 | 1,00 | 5,00 | 573 |
| Backstage sikkerhet | 4,10 | 0,79 | 0,40 | -0,85 | 1,00 | 5,00 | 570 |
| Kundeavhengig sikkerhet | 3,74 | 1,04 | -0,39 | -0,66 | 1,00 | 5,00 | 564 |

I tabellen presenteres gjennomsnitt, standardavvik, skewness, kurtosis, minimumsverdier og maksimumsverdier for de summerte gjennomsnittsskalaene (samtlige 27 vurderte egenskaper), og for de grupperte gjennomsnittsskalaene (samhandlings-, struktur-, backstage- og kundeavhengige egenskaper). Beskrivende statistikk for de individuelle indikatorene er presentert i tabell D.1 i appendix D.

Ved inspeksjon av tabell 7.3 ser vi først på minimums- og maksimumsverdiene for hver av de avhengige variablene. Tabellen viser at ingen av variablene har minimums- eller maksimumsverdier som kan indikere feilpunching eller kalkulasjonsfeil ved konstruksjon av de summerte skalaene.

Rapporterte minimumsverdier samsvarer ikke med minimumsverdiene som benyttes i måleskalaen i spørreskjemaet. Det er altså ingen respondenter som systematisk har markert de laveste verdiene på spørsmålene i spørreskjemaet. Det er imidlertid verd å merke seg at det for samtlige variabler, med unntak av summert gjennomsnittsvurdering, rapporteres maksimumsverdier som samsvarer med endepunktene på skalaen. Dette betyr at det finnes respondenter i utvalget som har markert de høyeste verdiene på alle spørsmålene som inngår i de summerte skalaene. Årsaken til maksimumsverdien på 4,96 for summert gjennomsnittsvurdering er at respondenter som systematisk har markert maksimumsverdiene (+5) på samtlige vurderingsspørsmål ble ekskludert fra utvalget som *careless respondents*.

Vi observerer videre at gjennomsnittsscorene for samtlige variabler er positive og relativt høye. Dette indikerer f.eks. at få respondenter har benyttet den negative delen av skalaen når de har fortetatt sine vurderinger. Dette resultatet er som forventet, og i samsvar med studier som har vist at svargivning i forbindelse med tilfredshetsmålinger og produktvurderinger tenderer mot den positive delen av skalaen (Peterson og Wilson, 1992).

De høye gjennomsnittsverdiene kan ha flere årsaker. For det første kan høye gjennomsnittsvurderinger være et tegn på at hotellene som er med i studien har relativt høy kvalitet, og at vurderingene fra gjest til gjest derfor ikke svinger i nevneverdig grad. Dersom dette er tilfellet er det kun forandringer av måleskalene som kan bidra til lavere gjennomsnittsvurderinger.

For det andre kan de høye gjennomsnittsvurderingene være et resultat av at det er et flertall av gjester som er fornøyde som svarer på spørreskjemaet. Misfornøyde gjester kan ha unnlatt å fylle ut spørreskjemaet. Dersom dette er tilfellet er ikke resultatene en "sann" beskrivelse av populasjonen "hotellgjester", og vurderingene kan ikke

generaliseres til denne populasjonen. Formålet med studien er imidlertid ikke å generalisere vurderinger av et hotell foretatt av et utvalg av gjester til å gjelde for samtlige gjester som har bodd på hotellet i tidsperioden undersøkelsen ble gjennomført. Formålet er å finne ut om manipulering av instruksjon og vurderingskala for spesifikke vurderingsspmå, og inkludering og alternativ plassering av viktighets-spmå i et spørreskjema fører til systematiske forskjeller i respondentenes svar-givning. Det kritiske forholdet ved utdeling av spørreskjemaene er derfor ikke å ha respondenter som er representative for populasjonen av hotellgjester som sådan, men å sørge for at det er tilfeldig hvilke gjester som får en bestemt type spørreskjema. Basert på drøftingen av undersøkelsesopplegget i kapittel 5, har vi valgt å konkludere vi med at disse kravene med stor sannsynlighet ser ut til å være oppfylt.

Når vi studerer gjennomsnitt og standardavvik for vurderinger av de ulike produkt-elementene kan vi observere et mønster som har vist seg å være konsistent med en rekke studier der et produkt er konseptualisert på denne måten (Troye, 1990). Gjennomsnitt for vurdering av samhandlingsegenskapene (immaterielle egenskaper) er høyest, mens gjennomsnittsvurderingene er lavest for de kundeavhengige produktegenskapene. Samtidig ser vi at standardavviket er størst for de kundeavhengige egenskapene, mens dette er minst for samhandlingssegenskapene. Disse resultatene indikerer en viss grad av reliabilitet/stabilitet i målinger av de samme variablene på tvers av undersøkelser.

Inspeksjon av standardavvik for samtlige variabler viser at disse er store nok til at vi konkluderer med at det er tilstrekkelig varians til å differensiere mellom respondentene i studien.

Som grunnlag for å vurdere distribusjonsegenskapene i de innsamlete dataene ble det kalkulert verdier for *skewness* og *kurtosis*. *Skewness* er en indikator for grad av manglende symmetri i dataene, mens *kurtosis* er et mål for hvor flate eller spisse fordelingene er (Bollen, 1989; Ferguson og Takane, 1989; Meek et al., 1987; Nunnally og Bernstein, 1994).

De negativt skjeve dataene indikerer fordelinger med en lang hale til venstre. For perfekt normalfordelte data vil *skewness* være lik null (Bollen, 1989; Ferguson og Takane, 1989; Meek et al., 1987; Nunnally og Bernstein, 1994). Verdier mellom -1 og 1 antas ikke å være kritiske, mens verdier som er større eller mindre enn disse indikerer skjeve fordelinger. Fra tabellen ser vi at variablene: struktur vurderinger, backstage vurderinger, totalvurdering, atferdsintensjoner og struktur sikkerhet viser høyest skjevhet.

Verdien for kurtosis for normalfordelte data er 0 (Bollen, 1989; Ferguson og Takane, 1989; Meek et al., 1987; Nunnally og Bernstein, 1994). Fordelinger med kurtosisverdier som er høyere enn 0 har tynnere haler og er spissere enn normalfordelingskurven (leptokurtisk), mens fordelinger med kurtosisverdier som er lavere enn 0 har fetere haler og er flatere enn normalfordelingskurven (platykurtisk). Verdiene for kurtosis som rapporteres i tabell 7.2 indikerer at dataene, med unntak av totalvurderingene, har fordelinger som er flatere og har fetere haler enn normalfordelingskurven. Totalvurderingene ser ut til å ha tynnere haler, og er samtidig spissere enn normalfordelingskurven. Nedenfor vil fordelingsegenskapene ved de innsamlete dataene danne et viktig utgangspunkt for diskusjon i forbindelse med valg av statistiske analyseteknikker.

7.3 Valg av statistiske analyseteknikker

Hypotesene som er formulert i denne avhandlingen krever analyser som gjør oss i stand til å undersøke forskjeller i svargivning som resultat av at respondenter har fått utdelt spørreskjema der manipulasjoner av ordlyd i spørsmålsstilling og rekkefølge på spørsmål har blitt foretatt. Samtlige hypoteser er derfor formulert med hensyn på å teste forskjeller mellom eksperimentgrupper basert på de avhengige variablene, og et *between subjects* design er benyttet i det faktorielle undersøkelsesopplegget.

Hovedprinsippet for valg av statistiske analyseteknikker har vært å velge den enkleste statistiske analyseteknikk som gir en tilfredsstillende og valid test av hypotesene. Det ble på bakgrunn av dette besluttet å benytte variansanalyser (ANOVA) for å teste hypotesene. Variansanalyser vil nettopp gjøre oss i stand til å teste forskjeller i svargivning mellom eksperimentgrupper med ulik instruksjon og vurderingsskala for de spesifikke vurderingsspørsmålene, og for forskjeller i svargivning mellom eksperimentgrupper der tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet ble manipulert⁸⁷.

For valg av analysemetode der man opererer med mer enn én avhengig variabel står valget mellom univariat ANOVA, der forskjeller mellom grupper kan undersøkes for hver avhengig variabel separat, og multivariat ANOVA (eller MANOVA), der de samlede effekter på samtlige avhengige variabler undersøkes på tvers av eksperiment-

⁸⁷For identifisering og analyse av modererende variabler er et alternativ til bruk av variansanalyse å benytte moderert regresjonsanalyse (Sharma, Durand, og Gur-arie 1981) der eksperimentgruppene kodes som dummyvariabler. Resultatene ved bruk av de to analysemetodene vil bli identiske.

grupper. I litteraturen anbefales ofte MANOVA i situasjoner med multiple avhengige variabler med høye inter-variabel korrelasjoner (Bray og Maxwell, 1985). I lys av dette vil MANOVA være den mest egnete analyseteknikken basert på de relativt høye korrelasjonene vi observerer mellom de avhengige variablene i studien. (Se tabell C.1. i Appendiks C).

Forskningsspørsmålene og de korresponderende hypotesene er imidlertid formulert med hensyn på univariate forskjeller mellom eksperimentgrupper. Det vil derfor være mest riktig å teste effekter av gruppedlemskap for hver enkelt avhengig variabel. På denne måten kan det gjennomføres en direkte test av hypotesene. Med dette utgangspunktet vil det være mest hensiktsmessig å velge univariat ANOVA fremfor multivariat MANOVA for å teste hypotesene⁸⁸.

I tillegg til rapportering av resultater fra variansanalyser, vil hypotesene knyttet til forskjeller i de spesifikke produktvurderingenes evne til å forklare totalvurderinger og atferdsintensjoner bli utdypet gjennom presentasjon av parameterestimater. Parameterestimaterne tilsvarer ustandardiserte (rå) regresjonskoeffisienter i en regresjonsanalyse og gjør oss i stand til å si noe om retning på de observerte effektene.

For ytterligere å illustrere forskjeller mellom eksperimentgruppene presenteres standardiserte beta-koeffisienter og forklart varians for hver eksperimentgruppe. Beta-koeffisientene reflekterer relativ betydning av en endring i ett standardavvik i den uavhengige variabelen på den avhengige variabelen. Ved bruk av beta-koeffisienter for å illustrere forskjeller mellom eksperimentgrupper må man ta hensyn til et viktig forhold, nemlig det at størrelsen på beta-koeffisientene henger nøye sammen med variansen i gruppene (høyere varians medfører høyere beta koeffisient). Det blir derfor frarådet å benytte standardiserte beta-koeffisienter som grunnlag for å sammenligne parameterestimater mellom eksperimentgrupper (Bollen, 1989; Hair jr., Andreson, og Tatham, 1987; Lewis-Beck, 1980). I hypotese 2 har vi fremsatt påstand om høyere konsistens i vurderinger i eksperimentgruppen som har benyttet spørreskjema med

⁸⁸Når vi til tross for dette har valgt å benytte MANOVA for hypotesetestingen har dette sin årsak i at vi i hypotesene har fremsatt påstander om forskjeller mellom eksperimentgruppene når det gjelder de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner. I hypotesene vi har formulert med hensyn på instruksjon, vurderingsskala og plassering av viktighetsspørsmål i spørreskjemaet, betraktes disse forholdene i hovedsak som variabler som modererer de spesifikke vurderingss spørsmålenes evne til å forklare totalvurdering og atferdsintensjoner. Dette innebærer at vi i tillegg til de direkte effektene av manipulering av instruksjon og vurderingsskala, og tilstedeværelse og plassering av viktighetsspørsmål for nivået på avhengige variabler som totalvurderinger, atferdsintensjoner og sikkerhet i vurderinger, også må ta hensyn til interaksjon mellom disse og de spesifikke vurderingene. Interaksjon mellom kontinuerlige variabler, som f.eks gjennomsnittet av de 27 spesifikke vurderingene, og gruppevariabler, som f.eks hvilken instruksjonstype og vurderingsskala som er benyttet i spørreskjemaet, er mulig å analysere ved å benytte subrutinen "analysis" i MANOVA prosedyren i programpakken SPSS-X for Windows (versjon 6.1).

objektforankret instruksjon og -vurderingsskala enn i eksperimentgruppen som har benyttet subjektforankret instruksjon og -vurderingsskala. Vi benytter derfor de standardiserte beta-koeffisientene kun for illustrasjon⁸⁹.

Neste trinn vil bli å foreta analyser for å bestemme om forutsetningene som ligger til grunn for å benytte ANOVA er tilfredsstillt i de innsamlete dataene.

7.3.1 Test av forutsetninger for å gjennomføre variansanalyser

Det er spesielt tre forutsetninger som de innsamlete dataene må oppfylle for at en klassisk ANOVA modell (og tilhørende variansanalyse) skal kunne benyttes for å teste hypotesene som er fremsatt i denne avhandlingen. Disse forutsetningene er (Keppel, 1982):

1. Forutsetning om at populasjonene som utsettes for ulik manipulasjon/ behandling er normalfordelte
2. Forutsetning om at populasjonene som utsettes for ulik manipulasjon/ behandling har lik varians (homogenitet i feilvariens)
3. Forutsetning om ukorrelerte feilledd.

I praksis vil forutsetningen om ukorrelerte feilledd være oppfylt dersom tilordning av subjekter til eksperimentgruppene er randomisert (Keppel, 1982). Vi har tidligere drøftet hvorvidt dette er tilfelle, og konkluderte med at det er stor sannsynlighet for at det er tilfeldig hvilke gjester som har fått utdelt en bestemt type spørreskjema. Forutsetningen om ukorrelerte feilledd antar vi derfor er oppfylt.

Forutsetningen om ukorrelerte feilledd er direkte relatert til utforming og gjennomføring av det faktorielle undersøkelsesopplegget. Hvorvidt de to andre forutsetningene for å benytte variansanalyse (forutsetning om at populasjonene som utsettes for ulik manipulasjon/behandling er normalfordelte og at de har lik varians) er oppfylt er et empirisk spørsmål, som må besvares på bakgrunn av de innsamlete dataene. Disse forutsetningene impliserer at det for hver hypotese må testes for normalitet *innenfor* hver eksperimentgruppe og for homogenitet i varians *på tvers* av eksperimentgrupper.

⁸⁹En rekke forhold bør nevnes i forbindelse med rapportering av standardiserte beta-koeffisienter (Bollen 1989, Hair jr., Andreson, og Tatham 1987, Lewis-Beck 1980): Beta-koeffisientene viser kun den relative betydningen av individuelle uavhengige variabler når kolinearitets er minimal. Beta-koeffisienter kan kun tolkes i forhold til andre variabler i regresjonsligningen. Nivået på variablene påvirker beta-koeffisientene, og tolking må derfor knyttes til området observasjonene er avgrenset til.

Testing av hypotesene: H1 - H9b innebærer sammenligning av de avhengige variablene for respondenter som har fått utdelt spørreskjema med henholdsvis objektforankrete- og subjektforankrete instruksjoner og vurderingsskalaer. Testing av hypotesene: H10a - H14b innebærer sammenligninger mellom respondenter som har fått utdelt spørreskjema uten viktighetsspørsmål, respondenter som har fått utdelt spørreskjema med viktighetsspørsmål før vurderingsspørsmålene og respondenter som har fått utdelt spørreskjema med viktighetsspørsmål etter vurderingsspørsmålene. I de neste avsnittene diskuteres resultater fra tester for normalitet og homogenitet i varians for de to gruppene av sammenligninger.

7.3.1.1 Forutsetning om normalitet i eksperimentpopulasjonene

Relevant statistikk for å teste om de innsamlete dataene tilfredsstiller forutsetningen om normalfordelte eksperimentpopulasjoner, er rapportering av verdier for skewness og kurtosis. I denne studien er disse verdiene kalkulert både for instruksjonssammenligningene og for rekkefølgesammenligningene. I tabellene D.2 til D.6 i appendiks D presenters beskrivende statistikk som inkluderer indikatorer for skewness og kurtosis. Som rapportert i disse tabellene er verdiene for skewness og kurtosis innenfor akseptable grenser for de fleste avhengige variablene. Det er imidlertid verdt å merke seg høye kurtosisverdier for totalvurderingene og for sikkerhet i vurdering av strukturegenskaper (materielle egenskaper) for respondenter som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala. Vi ser også at for respondenter som har benyttet spørreskjema med subjektforankrete instruksjon og -vurderingsskala rapporteres realt høye og positive kurtosisverdier for vurderingsskalaene. Verdiene for skewness er også gjennomgående mer negative for denne eksperimentgruppen.

For eksperimentgruppene der det ble foretatt manipulasjoner med tilstedeværelse og plassering av viktighetsspørsmål observerer vi akseptable verdier for skewness og kurtosis for eksperimentgruppen som benyttet spørreskjema uten viktighetsspørsmål. Med unntak av sikkerhet i vurdering av strukturegenskaper, som har kurtosis = 4,19 er alle variablene innenfor akseptable grenser. For de to andre eksperimentgruppene observerer vi mer problematiske verdier på skewness og kurtosis.

De rapporterte verdiene for skewness og kurtosis for samtlige eksperimentgrupper indikerer avvik fra normalitet for en rekke av de avhengige variablene i studien. Dette skulle isolert sett føre til at en forutsetning for å gjennomføre variansanalyser ikke er til stede. Imidlertid har man gjennom en rekke eksperimentelle studier vist at F-verdiene

som rapporteres i klassiske ANOVA design er forholdsvis robuste mot avvik fra forutsetningen om normalfordelte eksperimentpopulasjoner⁹⁰. Keppel (1982) hevder f.eks. at:

“violations of the normality assumption do not constitute a serious problem, except if the violations are severe. Under these circumstances, we need only worry about F’s that fall close to the critical value of F defining the start of the rejection region” (s. 86).

I lys av bevisene på robusthet for F-testen for moderate avvik fra normalitet, konkluderer vi med at avvikene fra normalitet som rapporteres i denne studien ikke representerer alvorlige trusler mot statistisk konklusjonsvaliditet i forbindelse med de slutninger som foretas på bakgrunn av F-tester i ANOVA.

7.3.1.2 Forutsetning om homogenitet i varians

I det foregående siterte vi litteratur der det argumenteres for at F-tester i klassiske ANOVA design er relativt robuste mot brudd på forutsetningene om normalitet i eksperimentpopulasjonene. Det samme kan ikke sies om forutsetningen om homogenitet i varians mellom eksperimentpopulasjonene. Selv om det finnes litteratur der det argumenteres for at F-tester også er relativt robuste mot denne forutsetningen er dette et forhold det hersker større uenighet om. F.eks. hevder Winer, Brown, og Michels (1991)⁹¹ at:

“even though..... the experimental design literature is prone to emphasize that the analysis of variance F-test is robust with regard to this assumption the situation regarding violations of homogeneity of variance is much more complex” (s. 101 - 103)

I situasjoner der forutsetningen om homogenitet i varians er brutt foreslår Winer, Brown og Michels (1991) at man bør “sørge for å oppnå store og like utvalgsstørrelser i de ulike eksperimentgruppene som en alternativ fremgangsmåte for å sikre nøyaktigheten i det nominelle signifikansnivået”. Alternativt anbefales det at man vurderer alternativer til variansanalyse og bruk av F-tester (s.110). Nedenfor diskuteres resultater fra tester for homogenitet i varians mellom de ulike eksperimentgruppene i studien som er gjennomført.

⁹⁰For en oversikt se f.eks.(Glass, Peckham og Sanders 1972)

⁹¹Som siterer(Wilcox 1987)

En rekke tester for å kontrollere forutsetningen om homogenitet i varians er foreslått i litteraturen (Winer, Brown, og Michels, 1991). I denne avhandlingen baserer vi drøfting av denne forutsetningen på rapportering av Cochrans C og Bartlett-Box F⁹². Resultatene fra disse testene er presentert i tabell D7 og D8 i appendiks D. I tabell D.7 rapporteres Cochrans C og Bartlett-Box F for testing av homogenitet i varians for de avhengige variablene for respondenter som har fått utdelt spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer, og for respondenter som har fått utdelt spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer. Testing av hypotese H10a - H14b innebærer sammenligninger mellom respondenter som har fått utdelt spørreskjema uten viktighetsspørsmål, respondenter som har fått utdelt spørreskjema med viktighetsspørsmål før spesifikke vurderingsspørsmål og respondenter som har fått utdelt spørreskjema med viktighetsspørsmål etter de spesifikke vurderingsspørsmålene. I tabell D.8 i appendiks D rapporteres Cochrans C og Bartlett-Box F for de avhengige variablene for disse eksperimentgruppene. Både Cochran testen og Bartlett-Box testen er basert på lik varians i eksperimentpopulasjonene som nullhypotese. Dette innebærer at høye verdier på testverdiene, og tilhørende lave p-verdier, medfører forkasting av forutsetningen om homogenitet i varians.

Ved inspeksjon av tabell D.7 i appendiks D ser vi at forutsetningen om homogenitet i varians ikke er tilfredsstillt for testing av forskjeller mellom respondenter som er tilordnet spørreskjema med objektforankret instruksjon og -vurderingsskala, og respondenter som er tilordnet spørreskjema med subjektforankret instruksjon og -vurderingsskala når det gjelder summert gjennomsnittsvurdering, samhandling-, struktur- og backstagevurderinger, viktighet ved samhandling-, struktur- og backstageegenskaper. Både Cochrans C og Bartlett-Box F testene rapporterer høye testverdier og lave p-verdier for disse variablene. For de øvrige variablene tilfredsstiller dataene forutsetningen om homogenitet i varians mellom de to eksperimentgruppene.

For testing av forutsetningen om homogenitet i varians mellom eksperimentgrupper som er tilordnet spørreskjema der tilstedeværelse og plassering av viktighetsspørsmål er manipulert, rapporteres Cochrans C og Bartlett-Box F for de avhengige variablene i tabell D.8 i appendiks D. Inspeksjon av tabellen indikerer at forutsetningen om homogenitet i varians i større grad er tilfredsstillt her. Med unntak av sikkerhet i vurdering av struktur- og samhandlingsegenskaper, viktighet av backstage- og struktur-

⁹²Beslutningen om å benytte disse testene for homogenitet i varians er basert på testenes popularitet i litteraturen. Siden testene i enkelte tilfeller kan lede til forskjellige konklusjoner (Winer, Brown, og Michels 1991) blir begge rapportert.

egenskaper rapporteres lave testverdier og høye p-verdier for både Cochrans C og Bartlett-Box F for samtlige variabler.

Implikasjonen av testene for homogenitet i varians mellom eksperimentgruppene i denne studien er at vi må se nærmere på Winer, Brown og Michel (1991) anbefalinger i situasjoner der denne forutsetningen ikke er tilfredsstillt. For å benytte en klassisk ANOVA model for å teste hypotesene om forskjeller mellom de ulike eksperimentgruppene i forbindelse med de fleste avhengige variablene, må vi sørge for at utvalgsstørrelsene i eksperimentgruppene er store, og så like som mulig. I tabell 5.2. rapporterte vi oppnådd utvalg i de ulike eksperimentgruppene. Det relativt høye antall respondenter, i tillegg til den jevne fordelingen av respondenter i hver gruppe, gjør at vi til tross for brudd på forutsetningen om homogenitet i varians for en rekke avhengige variabler, kan forsvare å benytte variansanalyser for å teste de fremsatte hypotesene.

8

Resultater av hypotesetesting

I kapittel 4 ble det formulert to sett hypoteser som grunnlag for den empiriske studien. Ett sett i tilknytning til forskjeller mellom bruk av subjektforankret- og objektforankret instruksjon og -vurderingsskala, og ett sett i forbindelse med inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet. Resultatene fra hypotesetestingen er i dette kapittelet organisert rundt disse to settene av hypoteser.

Som grunnlag for å teste hypotesene har vi gjennomført en kvasi-eksperimentell studie, der vi i et faktorielt undersøkelsesopplegg opererer med seks eksperimentgrupper. I avhandlingen har vi ikke formulert hypoteser for effekter av interaksjon mellom de to eksperimentvariablene vi manipulerer i spørreskjemaet. For testing av hypotesene i tilknytning til forskjeller mellom subjektforankret- og objektforankret instruksjon og -vurderingsskala, slår vi derfor sammen eksperimentgruppene der tilstedeværelse og plassering av viktighetsspørsmål manipuleres i spørreskjemaet. På denne måten har vi redusert seks eksperimentgrupper til to, basert på forskjeller i instruksjon og vurderingsskala for de spesifikke vurderingsspørsmålene. For testing av hypotesene i tilknytning til tilstedeværelse og alternativ plassering av viktighetsspørsmål, blir det på tilsvarende måte ikke foretatt noe skille mellom eksperimentgrupper med ulik instruksjon og vurderingsskala for vurderingsspørsmålene. For å teste disse hypotesene opererer vi derfor med tre eksperimentgrupper som reflekterer tilstedeværelse og ulik plassering av viktighetsspørsmål i spørreskjemaet.

Resultater fra analyser for testing av forskjeller mellom objektforankrete- og subjektforankrete vurderinger av spesifikke produktegenskaper, rapporteres i avsnitt 8.1. Deretter, i avsnitt 8.2, undersøker vi betydningen av å inkludere viktighetsspørsmål i spørreskjemaet.

8.1 Objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer

I dette avsnittet presenteres først resultater fra analyser der det fokuseres på forskjeller i nivå og konsistens for de spesifikke vurderingsspørsmålene, mellom respondenter som er tilordnet spørreskjema med henholdsvis objektforankret- og subjektforankret instruksjon og -vurderingsskala. Deretter introduseres kvalitetsnivået på produktene som vurderes som modererende variabel, og vi ser nærmere på forskjeller mellom eksperimentgruppene for nivået på de spesifikke vurderingsspørsmålene. De spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner betraktes som en viktig indikator på vurderingenes prediktive validitet. Analyser av prediktiv validitet foretas først uten å inkludere respondentenes produktkunnskap som modererende variabel. Etter at vi har rapportert resultater fra analyser av forskjeller mellom eksperimentgruppene for respondentenes sikkerhet i vurderinger, inkluderes produktkunnskap som modererende variabel, både for prediktiv validitet og sikkerhet i vurderinger.

8.1.1 Forskjeller i nivå på spesifikke vurderingsspørsmål

Subjektforankrete (“tilfreds-utilfreds”) og objektforankrete (“bra-dårlig”) instruksjoner og -vurderingsskalaer blir benyttet om hverandre, både ved gjennomføring av praktiske markedsundersøkelser og i akademisk forskning. Det kan derfor tyde på at man ikke forventer forskjeller i nivå på vurderinger foretatt ved bruk av de to instruksjonene og vurderingsskalaene. Litteraturgjennomgangen som er foretatt i forbindelse med denne avhandlingen gir heller ikke støtte for at slike forskjeller skal eksistere. Basert på disse “ikke-teoretiske” betraktningene formulerte vi følgende hypotese i kapittel 4.

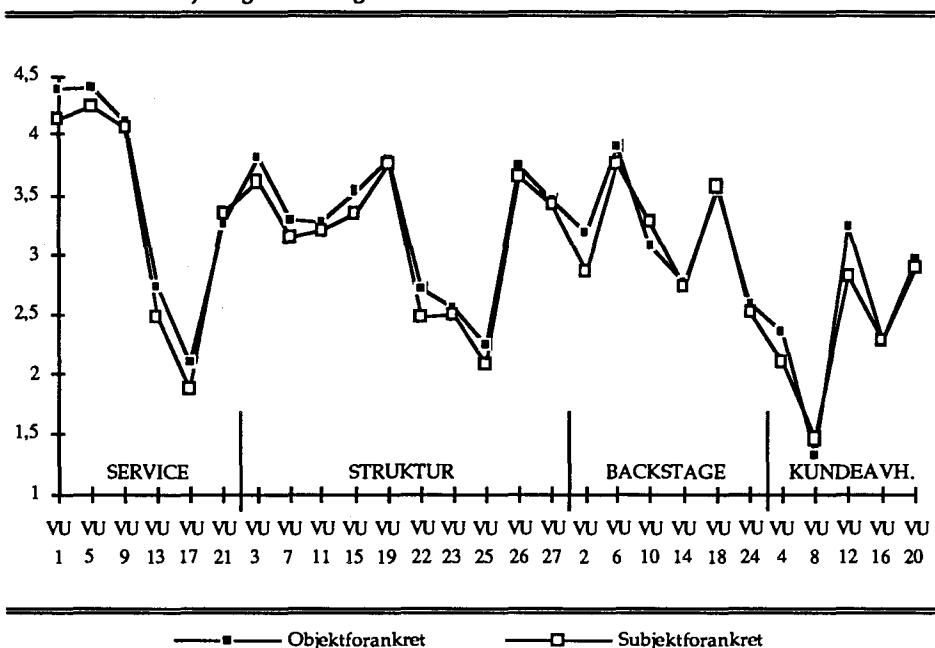
H1: Nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer vil ikke avvike fra nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer⁹³

⁹³Hypotese 1 kan betraktes som en null-hypotese i den forstand at vi ikke forventer forskjeller mellom eksperimentgruppene. Dette får konsekvenser for testlogikken vi benytter ved hypotesetestingen. Et signifikansnivå $\alpha < 0,05$ impliserer at rapportering av signifikante forskjeller mellom eksperimentgrupper med liten sannsynlighet kan antas å skyldes tilfeldigheter. For testing av hypotese 1 vil eventuelle forskjeller mellom eksperimentgruppene derfor betraktes som tilfeldige dersom rapporterte p-verdier er høyere enn valgt signifikansnivå.

En måte å teste hypotese 1, er å sammenligne gruppegjennomsnitt for hver av de individuelle vurderingsvariablene mellom de to eksperimentgruppene. Som grunnlag for disse sammenligningene illustreres gruppegjennomsnitt for samtlige 27 spesifikke vurderinger av hotellproduktet, foretatt ved bruk av de to forskjellige instruksjonene- og vurderingsskalaene, i figur 8.1.

På X-aksen i figuren finner vi de 27 forskjellige aspektene ved hotellproduktet som respondentene ble bedt om å vurdere. Rekkefølgen disse er presentert i samsvarer med gruppering av egenskapene i de fire produktelementene samhandling-, struktur-, backstage- og kundeavhengige produktelementer⁹⁴.

FIGUR 8.1
Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala



Y-aksen i figuren representerer nivået på vurderingene. Måleskalaen som ble benyttet i spørreskjemaet har verdier fra -5 til 5. I figur 8.1 har vi avgrenset måleskalaen til å gjelde området vurderingene varierer innenfor. På denne måten fremkommer eventuelle forskjeller mellom eksperimentgruppene på en klarere måte.

⁹⁴Se spørreskjema i appendiks B

Figur 8.1 indikerer små eller ingen forskjeller mellom eksperimentgruppene når det gjelder nivået på de individuelle vurderingene. Basert på en slik første overflateinspeksjon, ledes vi til å konkludere med at hypotesen om ingen forskjeller i nivå støttes av de innsamlede dataene.

En nærmere inspeksjon av figuren indikerer imidlertid at de spesifikke vurderingene gjennomgående er noe høyere i eksperimentgruppen som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala, enn i eksperimentgruppen som har benyttet spørreskjema med subjektforankret instruksjon og -vurderingsskala. Av 27 spesifikke vurderinger er 17 vurderinger høyest for respondenter som har benyttet objektforankret instruksjon og -vurderingsskala, mens bare to vurderinger (VU8 og VU10) er høyest for respondenter som har benyttet subjektforankret instruksjon og -vurderingsskala. For de resterende åtte spesifikke vurderingene er gruppegjennomsnittene tilnærmet identiske for de to eksperimentgruppene. Selv om vi observerer et mønster som beskrevet ovenfor, er det imidlertid viktig å merke seg at kun to av de observerte forskjellene er statistisk signifikant med $\alpha < 0,05$ (VU1 og VU12), og at kun to forskjeller er statistisk signifikant med $\alpha < 0,10$ (VU2 og VU5)⁹⁵.

En alternativ måte å teste hypotesen om forskjeller i nivå på de 27 individuelle vurderingene, er å betrakte vurderingene som en tilnærmet binomisk serie. Med andre ord kan vi stille spørsmål om sannsynligheten for at minst 17 av 27 vurderinger er høyest i eksperimentgruppen med objektforankret instruksjon og -vurderingsskala av ren tilfeldighet. Resultatene fra en ikke-parametrisk binomisk test gir en to-halet sannsynlighet på 0,09 for at det observerte mønsteret er tilfeldig dersom de to eksperimentgruppene er trukket fra samme populasjon⁹⁶. Basert på disse analysene kan det konkluderes med marginal retningssøtte for høyere nivå på objektforankrete vurderinger enn på subjektforankrete vurderinger.

Et problem ved bruk av individuelle variabler som grunnlag for sammenligning mellom eksperimentgruppene, er tilstedeværelse av tilfeldig feilvarians i hver individuell variabel. Dette problemet kan reduseres ved å kalkulere gjennomsnitt på tvers av alle 27

⁹⁵Det kan argumenteres for at responsskalaene som er benyttet for måling av vurderinger i spørreskjemaet kun tilsvarer et ordinalt målenivå, og at de derfor ikke tilfredsstiller forutsetningene for å gjennomføre parametriske tester, som f.eks ANOVA eller T-tester. Det ble derfor besluttet å supplere de parametriske testene med ikke-parametriske Mann-Witney tester. Resultatene fra disse testene gir de samme konklusjonene som de parametriske testene.

⁹⁶Et problem forbundet med å benytte en ikke-parametrisk binomisk test på denne måten, er at mange av vurderingsvariablene er korrelert. Dette er et klart brudd på forutsetningen om uavhengighet i en binomisk serie

vurderinger for de to eksperimentgruppene, og sammenligne disse⁹⁷. Resultat fra en T-test der disse forskjellene er testet er presentert i tabell 8.1.

TABELL 8.1

T-test, samtlige 27 vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| VARIABLE | | NUMBER OF CASES | MEAN | STANDARD DEVIATION | SE OF MEAN | | |
|---|-------------|----------------------------|--------------------|----------------------------|------------|--------------------|-------------|
| Gjennomsnitt av samtlige 27 vurderinger | | | | | | | |
| Objektforankret | | 316 | 3,2128 | 1,070 | 0,060 | | |
| Subjektforankret | | 292 | 3,0863 | 1,253 | 0,073 | | |
| | | * POOLED VARIANCE ESTIMATE | * | SEPARATE VARIANCE ESTIMATE | * | | |
| F VALUE | 2-TAIL PROB | * T VALUE | DEGREES OF FREEDOM | 2-TAIL PROB | * T VALUE | DEGREES OF FREEDOM | 2-TAIL PROB |
| 4,755 | 0,030 | * 1,34 | 606 | 0,180 | * 1,33 | 574,51 | 0,183 |

T-testen som presenteres i tabell 8.1 viser at de observerte forskjellene i gjennomsnitt mellom eksperimentgruppene heller ikke her kan betegnes som statistisk signifikante. Resultatene fra analysene ovenfor gir dermed klar støtte for hypotese 1.

En tredje måte å teste hypotese 1, er å gruppere de 27 vurderingene ved konstruksjon av summerte skalaer. I kapittel 6 operasjonaliserte vi fire slike summerte skalaer der produktens egenskaper grupperes i samhandlings-, struktur-, backstage- og kundeavhengige produktelementer. Som grunnlag for å bestemme hvorvidt gjennomsnittsvurderinger for de fire summerte skalaene er forskjellige i de to eksperimentgruppene, ble det gjennomført multivariat variansanalyse, MANOVA, der de fire produktelementene behandles som korrelerte avhengige variabler.

TABELL 8.2

MANOVA, nivå på vurderinger av produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala. Rapportering av Pillais-T

| SOURCE OF VARIATION | VALUE | EXACT F | HYP. DF | ERROR F | SIGN. OF F |
|--|---------|---------|---------|---------|------------|
| Instruksjon/vurderingsskala Pillais-T | 0,00472 | 0,70096 | 4,00 | 591.00 | 0,592 |

⁹⁷Grand mean

I tabell 8.2 rapporteres Pillais-T som grunnlag for testing av forskjeller mellom eksperimentgruppene for samtlige avhengige variabler. Som det fremgår av tabellen finnes det heller ingen signifikante forskjeller i nivå på vurderinger mellom eksperimentgruppene når de summerte skalaene benyttes som avhengige variabler. Hypotese 1 støttes dermed også av disse analysene.

Som oppsummering av resultatene for testing av forskjeller mellom eksperimentgruppene, med hensyn på nivå på spesifikke produktvurderinger, har vi valgt å konkludere med støtte for hypotese 1. Denne konklusjonen gjelder til tross for at nivået på vurderingene, basert på en binomisk test, gjennomgående synes å være marginalt høyere i eksperimentgruppen som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala, enn i eksperimentgruppen som har benyttet spørreskjema med subjektforankret instruksjon og -vurderingsskala.

8.1.2 Forskjeller i konsistens for spesifikke vurderinger

I kapittel 3 argumenterte vi for at det kan eksistere situasjoner der objektive eller intersubjektive standarder danner grunnlag for vurderinger. Slike standarder kan f.eks. skapes av tilbyderen gjennom informasjon til markedet som kan føre til at det etableres en felles norm for vurderinger. Valg av objekter for gjennomføring av denne studien er blant annet motivert med utgangspunkt i dette. Det kan argumenteres for at en hotellkjede, gjennom sine markedsaktiviteter og sin profilering skaper en objektiv eller intersubjektiv norm i markedet som alle hotellene i kjeden vil bli vurdert etter.

Objektforankret (“bra-dårlig”) instruksjon og -vurderingsskala inviterer respondentene til å basere sine vurderinger på disse objektive eller intersubjektive standardene, i den grad de eksisterer. Ved bruk av subjektforankret (“tilfreds-utilfreds”) instruksjon og -vurderingsskala knyttes vurderingene nærmere respondentens indre følelser overfor sin egen erfaring med produktet, og vurderingene blir dermed mer subjektive. Basert på disse teoretiske betraktningene formulerte vi i kapittel 4 følgende hypotese:

H2: *Bruk av objektforankrete instruksjoner og -vurderingsskalaer vil føre til høyere konsistens i respondentenes vurdering av spesifikke produkttegenskaper enn bruk av subjektforankrete instruksjoner og -vurderingsskalaer*

En måte å teste hypotese 2 er å inspisere standardavvikene⁹⁸ for samtlige 27 individuelle vurderinger. Standardavvikene gir uttrykk for variasjoner i svargivning mellom respondentene for hvert enkelt vurderingsspørsmål. Det knytter seg imidlertid visse problemer til å benytte standardavviket som et mål på konsistens i vurderinger.

For det første er skalaen som benyttes for måling av vurderinger i spørreskjemaet en såkalt *censored scale*. I dette ligger at skalaen har en absolutt topp- og bunnverdi. Dette gjør at gjennomsnittsverdier som ligger opp mot ekstremverdiene, på grunn av *ceiling* effekter, er uforenlige med høy varians. Eksempelvis vil en gjennomsnittsverdi nær -5 eller +5 nødvendigvis innebære at variansen blir nær 0. Variansen vil dermed tendere til å øke når vi beveger oss fra skalaens ekstremverdier mot skalaens midtpunkt.

For det andre er respondentene fullstendig nestet innenfor hvert av de 11 hotellene som deltok i undersøkelsen. For å ta hensyn til at hotellene er forskjellige, fant vi det derfor mest hensiktsmessig å beregne standardavvik for hver enkelt variabel innenfor hvert hotell. Gjennomsnittlig standardavvik i hver eksperimentgruppe ble deretter beregnet ved å veie standardavviket for hvert hotell med antall respondenter⁹⁹. Standardavvikene for hver variabel kan deretter grupperes, enten for samtlige 27 vurderinger, eller i henhold til inndelingen av produktet i samhandlings-, struktur-, backstage- og kunde-avhengige produktegenskaper.

I figur 8.2 presenteres veiet standardavvik for hver variabel i de to eksperimentgruppene. På X-aksen i figuren finner vi de 27 spesifikke egenskapene ved hotellproduktet som respondentene ble bedt om å vurdere. Y-aksen i figuren representerer størrelsen på standardavvikene.

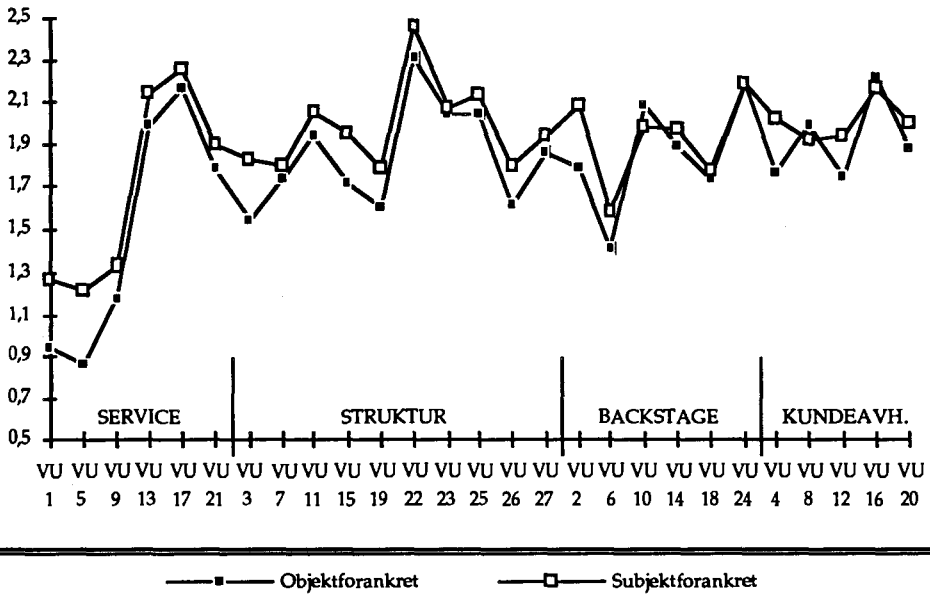
Vi ser av figur 8.2 at det er små forskjeller i standardavvik mellom eksperimentgruppene. Imidlertid kan det observeres en tendens til at standardavvikene gjennomgående er noe høyere i eksperimentgruppen som er tilordnet spørreskjema med subjektforankret instruksjon og -vurderingsskala, enn i eksperimentgruppen er tilordnet spørreskjema med objektforankret instruksjon og -vurderingsskala. Av totalt 27 individuelle vurderinger er standardavvikene høyest i eksperimentgruppen som har benyttet spørreskjema med subjektforankret instruksjon og -vurderingsskala for 21 individuelle vurderinger.

⁹⁸Et alternativt mål for konsistens i vurderinger er varians, vi har imidlertid her valgt å benytte standardavvik for testing av hypotese 2.

⁹⁹Selv om antall respondenter innenfor hver eksperimentgruppe er balansert for hvert hotell har vi valgt å beregne veiet standardavvik for å ta hensyn til ulikt antall respondenter i hotellene. Veiet standardavvik avviker imidlertid i liten grad fra uveiet standardavvik.

FIGUR 8.2

Standardavvik for spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala



I og med at standardavvikene som er rapportert i figur 8.2 er veiet i forhold til antall respondenter for hvert hotell, er det vanskelig å foreta en test av signifikans for de observerte forskjellene. Det ble derfor gjennomført F-tester for forskjellene separat for hvert hotell. Resultatet av disse er rapportert i tabell E.1 i appendiks E. Resultatene fra F-testene viser forskjeller i standardavvik som ikke er tilstrekkelige til at vi velger å konkludere med støtte for hypotese 2.

Hypotesen om høyest standardavvik for de spesifikke vurderingsspørsmålene i eksperimentgruppen som er tilordnet spørreskjema med subjektforankret instruksjon og -vurderingsskala kan også testes gjennom å benytte en ikke-parametrisk binomisk test, dersom vi antar at vurderingene er uavhengige. Resultatene fra en slik test gir en enhalet sannsynlighet på 0,0007 for at standardavvikene, av ren tilfeldighet, er høyest i eksperimentgruppen som har benyttet subjektforankret instruksjon og -vurderingsskala i tilknytning til de spesifikke vurderingsspørsmålene. Dette resultatet indikerer sterk retningsstøtte for hypotese 2.

En alternativ fremgangsmåte for å teste hypotesen om høyest standardavvik for vurderinger foretatt i eksperimentgruppen som er tilordnet spørreskjema med subjekt-

forankret instruksjon og -vurderingsskala, er å sammenligne standardavvikene for samtlige 27 vurderingsvariabler mellom de to eksperimentgruppene. Veiet gjennomsnittlig standardavvik i eksperimentgruppen som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala rapporteres til 1,78, mens standardavviket i eksperimentgruppen som har benyttet spørreskjema med subjektforankret instruksjon og -vurderingsskala rapporteres til 1,91. Den observerte forskjellen i standardavvik er, i henhold til hypotese 2, i riktig retning, men statistisk signifikant for bare ett av hotellene ($\alpha < 0,05$).

En tredje måte å teste hypotese 2, er å ta utgangspunkt i standardavvikene for vurdering av produktegenskapene gruppert i henhold til samhandlings-, struktur-, backstage- og kundeavhengige produktelementer. I tabell 8.3 rapporteres veiet standardavvik for de summerte skalaene for de to eksperimentgruppene.

TABELL 8.3
Standardavvik for spesifikke vurderinger gruppert i produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SKALA | GRUPPESTANDARDAVVIK | |
|---------------------|---------------------|------------------|
| | OBJEKTFORANKRET | SUBJEKTFORANKRET |
| Samhandling/service | 1,49 | 1,69 |
| Struktur | 1,84 | 1,98 |
| Backstage | 1,85 | 1,93 |
| Kundeavhengig | 1,92 | 2,01 |

Resultatene fra tabellen ovenfor indikerer høyest standardavvik i eksperimentgruppen med subjektforankret instruksjon og -vurderingsskala. Det er imidlertid kun for ett hotell at forskjellene mellom eksperimentgruppene er statistisk signifikante med $\alpha < 0,05$ (en-hale test). Dette gjelder i tillegg kun for samhandlings- og strukturegenskaper. For backstage- og kundeavhengige egenskaper, rapporteres ikke statistisk signifikante forskjeller i standardavvik mellom eksperimentgruppene. Vi kan legge merke til at de observerte forskjellene gjennomgående er relativt små, også der de er statistisk signifikante, både når vi sammenligner standardavvik for samtlige 27 vurderinger, og for de summerte skalaene. Dette kan skyldes reduksjon i betydningen av tilfeldig feilvarians som kan tilskrives hver individuell variabel.

Et alternativt mål på konsistens i vurderinger er grad av enighet mellom respondentene som har vurdert hvert hotell. En test for slik enighet er den ikke-parametriske *Kendall's*

*coefficient of concordance (W)*¹⁰⁰. Kendall's coefficient of concordance (W) tester hvorvidt k relaterte utvalg er fra den samme populasjonen, og er et mål på enighet mellom subjekter der hvert case representerer et subjekts vurdering av flere variabler. Verdier for de 27 spesifikke produktvurderingene rangeres fra 1 til 27 for hvert case, og gjennomsnittlig rangering kalkuleres for hver vurdering på tvers av alle case. Verdier for W ligger mellom 0 (ingen enighet) og 1 (fullstendig enighet). I tabell 8.4 er veiet Kendall's W rapportert for de to eksperimentgruppene¹⁰¹.

TABELL 8.4

Kendall's coefficient of concordance (W) for vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| INSTRUKSJON OG VURDERINGSSKALA | KENDALL'S W |
|--------------------------------|-------------------|
| Objektforankret | 0,32 ^a |
| Subjektforankret | 0,30 ^a |

^a Signifikant $\alpha < ,01$

Rapportert Kendall's W viser en noe høyere verdi for eksperimentgruppen med objektforankret instruksjon og -vurderingsskala enn for den andre eksperimentgruppen. Dette kan tolkes som noe større enighet mellom respondentene i denne gruppen. Forskjellen mellom eksperimentgruppene er imidlertid så marginal at den ikke kan vektlegges som støtte for hypotese 2.

Selv om dataene, basert på signifikanstester, ikke gir støtte for hypotese 2, gir den ikke-parametriske testen en klar indikasjon på en tendens til høyest konsistens for objektforankrete vurderinger. Vi velger derfor å konkludere med retningsstøtte for hypotese 2.

8.1.3 Forskjeller i nivå på spesifikke vurderingsspørsmål - kvalitetsnivået på produktene som vurderes som modererende variabel

Tidligere i dette kapittelet fikk vi støtte for hypotesen om at det ikke er forskjell i nivå på vurderinger av spesifikke produktegenskaper foretatt av respondenter som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala, og

¹⁰⁰Kendall's W

¹⁰¹Rapportert Kendall's W er veiet med samme begrunnelse som for standardavvikene. Tallene som fremkommer i tabellen er derfor beregnet individuelt for hvert hotell. Deretter er veiet gjennomsnittlig W beregnet

respondenter som har benyttet spørreskjema med subjektforankret instruksjon og -vurderingsskala. I kapittel 3 argumenterte vi imidlertid for at det kan finnes situasjoner der gjennomsnittsvurderingene kan være forskjellige.

Når produktene som vurderes enten overstiger, eller oppleves som dårligere enn forventningene respondentene har, er det grunn til å forvente at gjennomsnittsvurderinger ved bruk av de to instruksjonene og vurderingsskalaene vil være forskjellige. Både høy kvalitet, og lav kvalitet kan føre til kontrasteffekter siden høykvalitetsprodukter vil tendere til å overstige forventningene, mens lavkvalitetsprodukter vil tendere til ikke å innfri forventningene.

Subjektforankrete vurderinger kan oppfattes som mer egoinvolverende enn objektforankrete vurderinger. Vi kan derfor forvente at produktvurderinger foretatt i spørreskjema med subjektforankret instruksjon og -vurderingsskala, medfører høyere nivå på vurderinger for høykvalitetsprodukter, enn vurderinger basert på objektforankret instruksjon og -vurderingsskala for de samme produktene.

Med samme argumentasjon forventer vi at vurderinger foretatt med subjektforankret instruksjon og -vurderingsskala medfører lavere nivå på vurderinger for lavkvalitetsprodukter, enn vurderinger foretatt med objektforankret instruksjon og -vurderingsskala for de samme produktene. For produkter med gjennomsnittskvalitet forventer vi ingen forskjell i nivå på vurderinger mellom de to typene instruksjon og -vurderingsskala. Basert på denne drøftingen formulerte vi i kapittel 4 følgende hypoteser.

H3a: *For lavkvalitetsprodukter vil nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer overstige nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer*

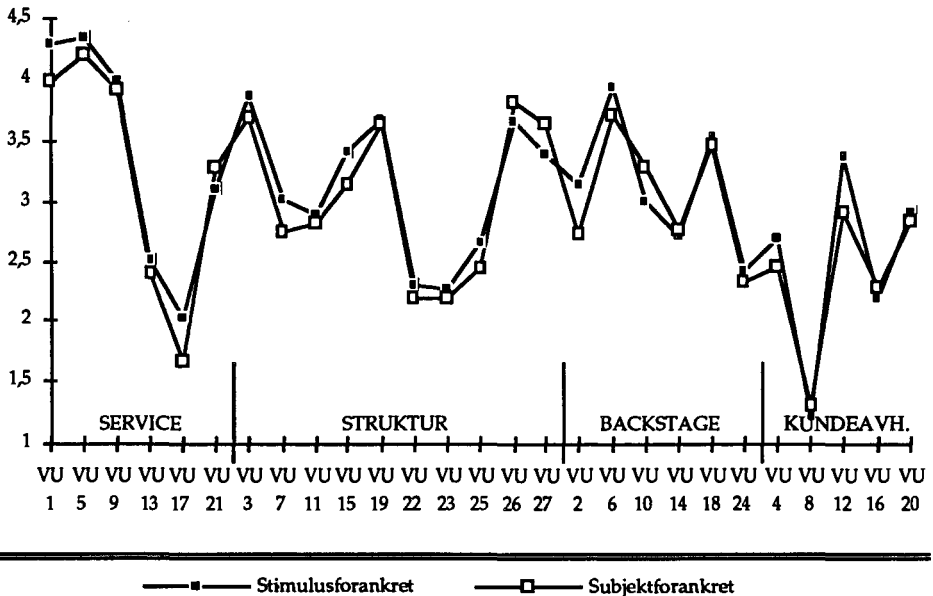
H3b: *For produkter med kvalitet lik gjennomsnittet, vil nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer ikke avvike fra nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer*

H3c: For høykvalitetsprodukter vil nivå på vurderinger foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer overstige nivå på vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer

Som grunnlag for å teste disse hypotesene ble de 11 hotellene som deltok i studien delt inn i tre grupper basert på respondentenes totalvurderinger¹⁰². Tre hotell plasseres på denne måten i gruppen *lavkvalitetshotell*, fem hotell plasseres i gruppen *gjennomsnittshotell* og tre hotell plasseres i gruppen *høykvalitetshotell*. En måte å teste hypotesene er å inspisere forskjeller i gjennomsnitt for de individuelle vurderingsvariablene. I figur 8.3a, 8.3b og 8.3c nedenfor oppsummeres disse analysene i diagrammer der gruppegjennomsnitt sammenlignes for hotellene representert ved de tre kvalitetsnivåene.

FIGUR 8.3a

Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - lavkvalitetshotell

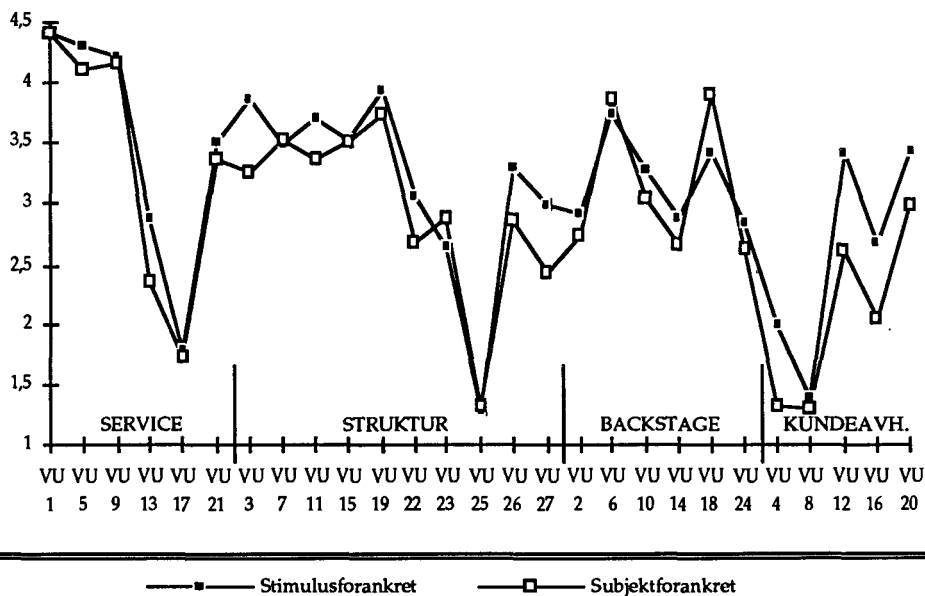


¹⁰²Se kapittel 6 for operasjonalisering av totalvurdering av hotellene, og kvalitetsnivå som moderende variabel. En alternativ fremgangsmåte vil være å gruppere hotellene i ulike kvalitets kategorier basert på "ekspertvurderinger".

I figur 8.3a rapporteres forskjeller i gruppegjennomsnitt for lavkvalitetshotellene. 18 av 27 vurderinger er i henhold til hypotese 3a i riktig retning. Imidlertid er bare tre av disse forskjellene statistisk signifikante med $\alpha < 0,05$ (en-hale test) (VU1, VU2 og VU12), mens tre av forskjellene er statistisk signifikante med $\alpha < 0,10$ (en-hale test) (VU17, VU15 og VU6). Selv om T-tester indikerer at gruppegjennomsnittene sett under ett, ikke kan betraktes som statistisk signifikant høyere i eksperimentgruppen som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala, enn i eksperimentgruppen som har benyttet subjektforankret instruksjon og -vurderingsskala, kan vi observere en tendens til at objektforankrete vurderinger ligger over subjektforankrete vurderinger. Dersom vi er villig til å akseptere brudd på forutsetningen om uavhengighet mellom vurderingene, kan vi betrakte de 27 vurderingene som en binomisk serie. Den en-halete sannsynligheten for at 18 av 27 vurderinger er høyest i eksperimentgruppen med objektforankret instruksjon og -vurderingsskala av ren tilfeldighet er 0,0094. På bakgrunn av disse analysene velger vi å konkludere med at hypotese 3a ikke støttes, men at det observeres en tendens til forskjeller i gruppegjennomsnitt i riktig retning. Disse funnene samsvarer også med inspeksjonen av de individuelle variablene vi gjennomførte under drøftingen av hypotese 1.

FIGUR 8.3b

Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - hotell med gjennomsnittskvalitet



For hypotese 3b og 3c ble det gjennomført en liknende eksersis som for hypotese 3a. Inspeksjon av figur 8.3b i viser at kun 6 av 27 gruppegjennomsnitt overlapper hverandre fullstendig i de to eksperimentgruppene. Imidlertid viser to-haete T-tester at kun ett av gruppegjennomsnittene er statistisk signifikant forskjellige for eksperimentgruppene med $\alpha < 0,05$ (VU 12), og at kun to gruppegjennomsnitt er statistisk signifikant forskjellige med $\alpha < 0,10$ (VU3 og VU 16).

FIGUR 8.3c

Nivå på spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - høykvalitetshotell



I figur 8.3c observerer vi at 10 av 27 vurderinger i følge hypotese 3c er forskjellige i riktig retning. Av disse er imidlertid kun en gruppeforskjell statistisk signifikant (en-halet test) med $\alpha < 0,10$ (VU13). Hypotese 3c støttes dermed heller ikke av dataene. Resultatene som rapporteres i diagrammene ovenfor leder oss, isolert sett, til å konkludere med støtte for hypotese 3b. Sett i relasjon til forskjellene som ble rapportert for hypotese 3a og 3c velger vi imidlertid, så langt, å konkludere med at hypotesene 3a, 3b og 3c som henger sammen, som helhet ikke støttes av de innsamlede dataene.

En alternativ måte å teste hypotesene ovenfor, er å rapportere gruppegjennomsnitt for samtlige 27 vurderingsvariabler, og for vurderingene gruppert i henhold til samhandlings-, struktur-, backstage- og kundeavhengige produkttegenskaper. I henhold til gruppegjennomsnittene som rapporteres i tabell 8.5, er det kun for hypotese 3b vi isolert

sett kan hevde at de innsamlede dataene gir støtte. Her antok vi at det ikke ville være forskjell i gruppegjennomsnittene for hotell med gjennomsnittskvalitet.

Selv om det for de kundeavhengige vurderingene rapporteres statistisk signifikante forskjeller, velger vi å akseptere at dataene sett under ett gir støtte for hypotese 3b. Imidlertid ser vi også av tabellen at forskjellene i gruppegjennomsnitt er i samme retning for samtlige variabler uavhengig av kvalitetsnivået på hotellene, med unntak av vurderinger av kundeavhengige produktegenskaper for hotell med kvalitet over gjennomsnittet.

TABELL 8.5

Gruppegjennomsnitt for spesifikke produktvurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala - kvalitetsnivå på hotellene som vurderes som modererende variabel

| SKALA | INSTRUKSJON OG VURDERINGSSKALA | KVALITET | | |
|-------------------------|--------------------------------|--------------|-------------------|-------------|
| | | UNDER GJ.SN. | GJSN. | OVER GJ.SN. |
| Samtlige 27 vurderinger | Objektforankret | 3,12 | 3,22 | 3,53 |
| | Subjektforankret | 3,00 | 3,00 | 3,45 |
| Samhandling/service | Objektforankret | 3,48 | 3,64 | 4,04 |
| | Subjektforankret | 3,34 | 3,49 | 3,81 |
| Struktur | Objektforankret | 3,14 | 3,21 | 3,82 |
| | Subjektforankret | 3,05 | 3,00 | 3,64 |
| Backstage | Objektforankret | 3,15 | 3,18 | 3,42 |
| | Subjektforankret | 3,06 | 3,14 | 3,37 |
| Kundeavhengig | Objektforankret | 2,52 | 2,72 | 2,25 |
| | Subjektforankret | 2,40 | 2,09 ^a | 2,54 |

^a Forskjellen er statistisk signifikant med $\alpha < 0,05$

Ingen av forskjellene er statistisk signifikante, og siden hypotese 3a, 3b og 3c henger sammen resonnementmessig, velger vi å også her å konkludere med at settet av hypoteser som helhet er falsifisert¹⁰³.

¹⁰³En forklaring på manglende støtte for hypotesene kan være at vi i studien bare har med 11 hotell fra Inter Nor kjeden. Gjennomsnittene som er rapportert i tabell 8.4 indikerer at disse hotellene, i henhold til hypotesene, kun er representative for hotell med kvalitet rundt, og under gjennomsnittet. I en tidligere undersøkelse (Troye, Henjesand, og Breivik, 1992) der 31 Inter Nor Hotell deltok, ble de samme hypotesene testet, og det ble funnet støtte for hypotese 3a og 3c. Dette kan tyde på at vi i denne undersøkelsen klarte å fange opp nivåforskjeller kvalitet mellom de 31 hotellene som var store nok til at hele utfallsrommet for vurderinger var representert. Funnene fra denne studien vil bli drøftet nærmere i relasjon til resultatene fra testingen av hypotese 3a, 3b og 3c under diskusjonen i kapittel 9.

8.1.4 Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner

Vi har i teoridrøftelsen hevdet at bruk av objektforankret- vs. subjektforankret instruksjon og -vurderingsskalaer vil pålegge respondentene kognitive oppgaver som er forskjellige. Bruk av objektforankret ("bra-dårlig") instruksjon og -vurderingsskala innebærer at respondenten opptrer som "dommer" og foretar en bedømmelse av vurderingsobjektet (produktet), mens subjektforankrete ("tilfreds-tilfreds") vurderinger reflekterer respondentens egne følelser, eller affekt, overfor objektet (produktet).

I kapittel 3 benyttet vi sentrale teoribidrag fra informasjonsintegrasjonsteori, holdningsteori og teori om kundetilfredshet for å fremsette en modell for hvordan kunder kan tenkes å foreta produktvurderinger etter kjøp og bruk av produktet. Et sentralt problemområde innenfor alle disse teoriretningene har vært modellenes prediktive validitet. I denne forbindelse ønsker vi i avhandlingen å undersøke hvor godt vurderinger av spesifikke egenskaper ved et produkt, foretatt med subjektforankret- vs. objektforankret instruksjon og -vurderingsskala, predikerer respondentenes totalvurderinger og intensjoner om atferd, som f.eks. gjenkjøp og *word of mouth*.

Basert på teoridrøftelsen i kapittel 3 argumenterte vi for følgende hypotese knyttet til de spesifikke produktvurderingenes evne til å forklare totalvurderinger av produktet:

H4a: Vurderinger av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer vil ha større evne til å forklare totalvurderinger enn vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer

Som beskrevet i kapittel 6 er totalvurderingene i spørreskjemaet operasjonalisert med tre spørsmål. De to første spørsmålene ble definert som objektforankrete totalvurderingss spørsmål der hotellgjestene blir bedt om å vurdere hvor bra eller dårlig hotellet er. I det tredje totalvurderingss spørsmålet ber vi hotellgjestene gi uttrykk for sin tilfredshet med den totalopplevelsen hotelloppholdet har gitt opphav til. Med referanse til de spesifikke vurderingene reflekterer dette siste totalvurderingss spørsmålet en subjektforankret totalvurdering. I analysene som følger behandles subjektforankrete- og objektforankrete totalvurderinger derfor som separate avhengige variabler.

I figur 3.5 i kapittel 3 presenterte vi en modell der totalvurdering av et produkt, og intensjoner om gjenkjøp og *word of mouth*, kan betraktes som en transformasjon av kundens vektlegging og vurdering av oppfattede produktegenskaper. Når vi nå skal undersøke prediktiv validitet knyttet til de spesifikke produktegenskapene, avhengig av hvilken instruksjon og vurderingsskala som er benyttet, har vi valgt å se bort fra viktighetskomponenten i modellen. Vurderingsmodellen vi benytter som grunnlag for testing av hypotesene nedenfor kan dermed uttrykkes matematisk på følgende måte:

$$TV_p = \sum_{i=1}^n E_i$$

der TV_p er totalvurderingen av produktet, E_i er vurdering av oppfattet egenskap i ved produkt p og n er antall fremtredende egenskaper. For analysene har vi i tillegg benyttet gjennomsnittsscoren for vurderingene som uavhengig variabel. De samme gjennomsnittsscorene benyttes som grunnlag for testing av hypotesene der atferdsintensjoner betraktes som avhengig variabel.

I kapittel 7 konkluderte vi med at forutsetningen om homogenitet i varians mellom de to eksperimentgruppene er oppfylt for begge typene totalvurderinger ($p > 0,10$ for både Cochran's C og Bartlett-Box F). Hypotese 4a kan derfor testes gjennom bruk av univariat variansanalyse.

Nedenfor rapporteres resultater fra analyser der betydningen av ulike instruksjon og vurderingsskala på de spesifikke vurderings spørsmålene analyseres separat for begge typene totalvurderings spørsmål. Først presenteres resultater fra analyser for testing av hypotese 4a der objektforankret totalvurdering betraktes som avhengig variabel.

Fra variansanalysen i tabell 8.6a ser vi først hvordan respondentenes svargivning i forbindelse med de 27 spesifikke vurderings spørsmålene bidrar til å forklare den *objektforankrede totalvurderingen* av hotellene. Den rapporterte p -verdien, på mindre enn 0,001, indikerer at de 27 spesifikke spørsmålene har prediktiv validitet. I henhold til hypotese 4a er vi interessert i å analysere i hvilken grad forskjellig instruksjon og vurderingsskala i forbindelse med de 27 spesifikke vurderings spørsmålene i spørreskjemaet forklarer forskjeller i nivået på respondentenes totalvurdering. Hvilken instruksjon og vurderingsskala som benyttes i spørreskjemaet kan på denne måten ha en direkte effekt for nivået på totalvurderingen. I tillegg kan de 27 spesifikke produkt egenskapenes evne til å forklare nivået på totalvurderingen modereres av instruksjonen og vurderingsskalaen som benyttes i spørreskjemaet.

TABELL 8.6a

ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 1089,75 | 591 | 1,84 | | |
| Konstant | 9,62 | 1 | 9,62 | 5,22 | 0,023 |
| Samtlige 27 vurderinger | 657,82 | 1 | 657,82 | 356,75 | < 0,001 |
| Instruksjon/vurderingsskala ^a | 11,98 | 1 | 11,98 | 6,50 | 0,011 ^b |
| Samtlige 27 vurderinger x instruksjon/vurderingsskala | 7,68 | 1 | 7,68 | 4,17 | 0,042 ^c |
| (Model) | 659,72 | 3 | 219,91 | 119,26 | < 0,001 |
| (Total) | 1749,47 | 594 | 2,95 | | |
| R-squared = | 0,377 | | | | |
| Adjusted R-squared = | 0,374 | | | | |

^a Gruppegjennomsnitt objektforankret totalvurdering:
 Objektforankret instruksjon og -vurderingsskala : 3,18 (n = 307)
 Subjektforankret instruksjon og -vurderingsskala : 3,29 (n = 288)

^b p = 0,005 for en-hale test

^c p = 0,021 for en-hale test

| <u>Parameterestimer:</u> | <u>Konstant</u> | <u>Samtlige 27 vurderinger</u> | <u>Instruksjon/vurderingsskala</u> | <u>Samtlige 27 vurderinger x instruksjon/vurderingsskala</u> |
|--------------------------|-----------------|--------------------------------|------------------------------------|--|
| Objektforankret: | 0,369 | 0,913 | -0,412 | 0,099 |
| Subjektforankret: | 0,369 | 0,913 | 0,412 | -0,099 |

Fra fotnote a i tabellen ser vi at nivået på den objektforankrete totalvurderingen er 3,18 for eksperimentgruppen som har benyttet subjektforankret instruksjon og -vurderingsskala for de spesifikke vurderingsspørsmålene, mens nivået på den samme totalvurderingen i eksperimentgruppen med objektforankret instruksjon og -vurderingsskala er 3,29. Under nullhypotesen om at det ikke finnes forskjell i nivå på totalvurderingen mellom de to eksperimentgruppene, viser ANOVA tabellen en F-verdi på 6,50 med en korresponderende to-halet sannsynlighet på 0,011. Denne p-verdien indikerer at de observerte forskjellene i nivå på totalvurdering i de to eksperimentgruppene med liten sannsynlighet kan påstås å skyldes tilfeldigheter.

Interaksjonen mellom hvilken instruksjon og vurderingsskala som benyttes i spørreskjemaet, og svargivningen på de 27 spesifikke vurderingsspørsmålene er rapportert som det tredje leddet i variansanalysen i tabell 8.6a. Fra tabellen ser vi at det her rapporteres en to-halet sannsynlighet på 0,042 for nullhypotesen om at denne interaksjonseffekten ikke har betydning for nivået på den objektforankrete totalvurderingen. Med en hypotese som spesifiserer at bruk av subjektforankret instruksjon og -vurderingsskala i forbindelse med spesifikke vurderingsspørsmål vil øke de 27

spesifikke egenskapenes evne til å forklare totalvurderingen, har vi en retningsbestemt hypotese. Den relevante en-alete sannsynligheten blir dermed 0,021 dersom funnene er i riktig retning.

Ved inspeksjon av parameterestimaterne for de to eksperimentgruppene, fremgår det at den direkte effekten av instruksjon og vurderingsskala er positiv, mens interaksjonen mellom samtlige 27 vurderinger og instruksjon og vurderingsskala er liten og negativ for eksperimentgruppen som er tilordnet spørreskjema med subjektforankret instruksjon og -vurderingsskala. Dette innebærer at bruk av subjektforankret instruksjon og -vurderingsskala i forbindelse med spesifikke vurderinger direkte bidrar til høyere objektforankret totalvurdering. Samtidig reduseres de spesifikke produktvurderingenes evne til å forklare objektforankret totalvurdering. Parameterestimaterne for de signifikante effektene i variansanalysen indikerer dermed at objektforankret vurdering av spesifikke produktegenskaper forklarer objektforankret totalvurdering bedre enn subjektforankrete vurderinger. Disse funnene er i motsatt retning av hypotesen vi har formulert. Basert på parameterestimaterne må vi derfor konkludere med at det ikke finnes støtte for hypotese 4a i dataene når objektforankrete totalvurderinger benyttes som avhengig variabel.

For å undersøke betydningen av de observerte forskjellene ble det gjennomført en styrketest. Med denne testen kan effektstørrelsen (et mål på forskjellen mellom nullhypotesen og en spesifikk alternativ hypotese), og styrken ved den statistiske testen (dvs. testens evne til å forkaste en uriktig nullhypotese) beregnes. Ved å spesifisere signifikansnivå $\alpha = 0,05^{104}$ rapporteres verdiene for styrke ved de statistiske testene til 0,718 for den direkte effekten av instruksjon og -vurderingsskala, og 0,529 for interaksjonen mellom instruksjon og vurderingsskala og de spesifikke vurderingene. *Eta-square* som et mål på effektstørrelse rapporteres til 0,011 og 0,007. Som vi ser er de rapporterte effektstørrelsene små, mens verdiene for styrke indikerer over middels statistisk styrke for F-testene¹⁰⁵.

I tabell 8.6b rapporteres resultater fra variansanalyse der vi ser nærmere på hvordan de spesifikke vurderingsspørsmålene bidrar til å forklare den *subjektforankrete totalvurderingen* av hotellene.

¹⁰⁴Det tradisjonelle signifikansnivået som benyttes i de fleste studier.

¹⁰⁵*Eta squared* kan tolkes som andelen av varians som kan tilskrives gruppedlemskap (Sawyer og Ball 1981, Winer, Brown, og Michels 1991) presenterer en tabell der verdier for eta squared på 0,01, 0,059 og 0,138 for en F-test henholdsvis kan oppfattes som små, middels og store effektstørrelser (Table 1, s. 276). De observerte verdiene for styrke på 0,718 og 0,529 ligger begge over gjennomsnittet for middels styrke rapportert i *Journal of Marketing Research* og en rekke andre journaler (Sawyer og Ball, 1981; Table 3, s.283).

TABELL 8.6b

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 720,34 | 586 | 1,23 | | |
| Konstant | 103,33 | 1 | 103,33 | 84,06 | < 0,001 |
| Samtlige 27 vurderinger | 455,15 | 1 | 455,15 | 370,27 | < 0,001 |
| Instruksjon/vurderingsskala ^a | 1,07 | 1 | 1,07 | 0,87 | 0,350 ^b |
| Samtlige 27 vurderinger x instruksjon/vurderingsskala | 0,99 | 1 | 0,99 | 0,80 | 0,371 ^c |
| (Model) | 458,58 | 3 | 152,86 | 124,35 | < 0,001 |
| (Total) | 1178,92 | 589 | 2,00 | | |
| R-squared = | 0,389 | | | | |
| Adjusted R-squared = | 0,386 | | | | |

^a Gruppegjennomsnitt: subjektforankret totalvurdering:

Objektforankret instruksjon og -vurderingsskala : 3,26 (n = 305)

Subjektforankret instruksjon og -vurderingsskala : 3,55 (n = 285)

^b p = 0,175 for en-hale test^c p = 0,186 for en-hale test

| <u>Parameterestimer:</u> | <u>Konstant</u> | <u>Samtlige 27 vurderinger</u> | <u>Instruksjon/vurderingsskala</u> | <u>Samtlige 27 vurderinger x instruksjon/vurderingsskala</u> |
|--------------------------|-----------------|--------------------------------|------------------------------------|--|
| Objektforankret: | 1,210 | 0,759 | - 0,123(ns) | 0,035(ns) |
| Subjektforankret: | 1,210 | 0,759 | 0,123(ns) | - 0,035(ns) |

Den rapporterte p-verdien, på mindre enn 0,001 for de 27 spesifikke vurderingsspørsmålene indikerer også her at disse har prediktiv validitet. Fra fotnote a i tabellen ser vi videre at nivået på den subjektforankrete totalvurderingen er 3,26 for eksperimentgruppen som har benyttet subjektforankret instruksjon og -vurderingsskala i forbindelse med de spesifikke vurderingsspørsmålene, mens gjennomsnittet for subjektforankret totalvurdering i eksperimentgruppen som har benyttet objektforankret instruksjon og -vurderingsskala er 3,55. ANOVA tabellen viser en F-verdi på 0,87 med en korresponderende to-halet sannsynlighet på 0,350 for at det ikke finnes forskjell i nivået på totalvurdering mellom de to eksperimentgruppene. Denne p-verdien indikerer at de observerte forskjellene i nivå på subjektforankret totalvurdering mellom de to eksperimentgruppene ikke er signifikant.

Fra tabellen ser vi videre at det rapporteres en to-halet sannsynlighet på 0,371 for at interaksjonen mellom samtlige 27 vurderinger og instruksjon og vurderingsskala ikke har betydning for nivået på totalvurderingen. Selv med en retningsbestemt hypotese, og en en-halet sannsynlighet på 0,186, kan vi ikke påstå at denne interaksjonseffekten er signifikant.

Inspeksjon av tabell 8.6a og tabell 8.6b har vist at de innsamlete dataene ikke støtter hypotesen om at bruk av subjektforankret instruksjon og -vurderingsskala for spesifikke vurderingsspmårl forklarer totalvurderinger av produktet bedre enn spesifikke vurderinger foretatt med objektforankret instruksjon og -vurderingsskala. På bakgrunn av analysene kan vi så langt konkludere med at vurderinger av spesifikke egenskaper ved produktet foretatt med objektforankret instruksjon og -vurderingsskala forklarer objektforankrete totalvurderinger marginalt bedre enn spesifikke vurderinger foretatt med subjektforankret instruksjon og -vurderingsskala. For subjektforankret totalvurdering rapporteres det ingen forskjell i forklaringssevne mellom vurderinger av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankret- og objektforankret instruksjon og -vurderingsskala.

I tillegg til variansanalysene og de korresponderende parameterestimatene som er rapportert ovenfor, har vi valgt å rapportere standardiserte beta-koeffisienter og forklart varians for ytterligere å illustrere de forskjellige bruk av objektforankret- vs. subjektforankret instruksjon og -vurderingsskala i forbindelse med spesifikke vurderingsspmårl kan gi opphav til. Standardiserte beta-koeffisienter og forklart varians er presentert i tabell 8.7 nedenfor.

TABELL 8.7

Standardiserte beta-koeffisienter og forklart varians for objektforankrete- og subjektforankrete totalvurderinger

| TOTALVURDERING/ INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA | (ADJ. R SQ.) |
|---|-----------------------|--------------|
| Objektforankret totalvurdering | | |
| Objektforankret instruksjon og -vurderingsskala | 0,60 | (0,36) |
| Subjektforankret instruksjon og -vurderingsskala | 0,63 | (0,39) |
| Subjektforankret totalvurdering | | |
| Objektforankret instruksjon og -vurderingsskala | 0,61 | (0,37) |
| Subjektforankret instruksjon og -vurderingsskala | 0,64 | (0,40) |

Ved inspeksjon av tabell 8.7 ovenfor, ser vi for det første at forskjellene mellom eksperimentgruppene er små. For det andre leder tabellen oss til å konkludere med retningsstøtte for hypotese 4a, en konklusjon som ikke er i overensstemmelse med våre tidligere analyser presentert i tabell 8.6a og 8.6b. Årsaken til forskjellene mellom parameterestimatene fra variansanalysene og de standardiserte beta-koeffisientene kan være at variansen i de spesifikke vurderingene er størst i eksperimentgruppen med subjektforankret instruksjon og -vurderingsskala. Dette vil isolert sett bidra til høyere standardiserte beta-koeffisienter for denne eksperimentgruppen. Med bakgrunn i dette

anbefales det ikke å benytte standardiserte løsninger for sammenligning mellom eksperimentgrupper (Bollen, 1989; Hair jr., Andreson, og Tatham, 1987; Lewis-Beck, 1980). Vi har derfor valgt å konkludere med at det ikke finnes støtte for hypotese 4a i dataene.

I tillegg til de spesifikke vurderingenes evne til å forklare totalvurderinger har vi i avhandlingen fokusert på atferdsintensjoner som avhengig variabel. I motsetning til totalvurderinger, som enten vil være objektforankrete- eller subjektforankrete, kan atferdsintensjoner oppfattes som en nøytral avhengig variabel som reflekterer respondentenes totalvurdering av produktet. I kapittel 4 formulerte vi følgende hypotese for de spesifikke vurderingenes evne til å forklare respondentenes atferdsintensjoner.

H4b: Vurderinger av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer vil ha større evne til å forklare atferdsintensjoner enn vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer

I kapittel 7 viste vi at forutsetningen om homogenitet i varians mellom de to eksperimentgruppene er oppfylt også for respondentenes atferdsintensjoner. Vi kan dermed teste hypotese 4b ved bruk av univariat variansanalyse.

Fra variansanalysen i tabell 8.8 ser vi først hvordan hotellgjestenes vurderinger av de 27 spesifikke produktegenskapene forklarer deres atferdsintensjoner. Den rapporterte p-verdien, på mindre enn 0,001, indikerer at de 27 spesifikke spørsmålene har prediktiv validitet for gjestenes ønske om å besøke hotellet ved en senere anledning, og sannsynligheten for at de vil anbefale hotellet til andre.

Fra fotnote a i tabellen ser vi at nivået på atferdsintensjoner rapporteres til 9,12 for spesifikke produktvurderinger foretatt med subjektforankret instruksjon og -vurderingsskala, mens nivået på atferdsintensjonene i eksperimentgruppen med objektforankret instruksjon og -vurderingsskala er 9,04. Under nullhypotesen om at det ikke finnes forskjell i nivået på totalvurdering mellom de to eksperimentgruppene viser ANOVA tabellen en F-verdi på 2,03 med en korresponderende to-halet sannsynlighet på 0,155 for at denne er sann. De observerte forskjellene i atferdsintensjoner mellom de to eksperimentgruppene er dermed ikke signifikante.

TABELL 8.8

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 2147,42 | 593 | 3,62 | | |
| Konstant | 2444,64 | 1 | 2444,64 | 675,08 | < 0,001 |
| Samtlige 27 vurderinger | 833,94 | 1 | 833,94 | 230,29 | < 0,001 |
| Instruksjon/vurderingsskala ^a | 7,33 | 1 | 7,33 | 2,03 | 0,155 ^b |
| Samtlige 27 vurderinger x instruksjon/vurderingsskala | 7,24 | 1 | 7,24 | 2,00 | 0,158 ^c |
| (Model) | 835,56 | 3 | 278,52 | 76,91 | < 0,001 |
| (Total) | 2982,98 | 596 | 5,00 | | |
| R-squared = | 0,280 | | | | |
| Adjusted R-squared = | 0,276 | | | | |

^a Gruppegjennomsnitt atferdsintensjoner:

Objektforankret instruksjon og -vurderingsskala : 9,12 (n = 309)
 Subjektforankret instruksjon og -vurderingsskala : 9,04 (n = 288)

^b p = 0,076 for en-hale test

^c p = 0,079 for en-hale test

| Parameterestimer: | Konstant | Samtlige 27 vurderinger | Instruksjon/vurderingsskala | Samtlige 27 vurderinger x instruksjon/vurderingsskala |
|-------------------|----------|-------------------------|-----------------------------|---|
| Objektforankret: | 5,868 | 1,023 | - 0,321(ns) | 0,095(ns) |
| Subjektforankret: | 5,868 | 1,023 | 0,321(ns) | - 0,095(ns) |

Fra tabellen ser vi videre at det rapporteres en to-halet sannsynlighet på 0,158 for at interaksjonen mellom samtlige 27 vurderinger og instruksjon og vurderingsskala ikke har betydning for atferdsintensjonene. Med en retningsbestemt hypotese blir den relevante en-halete sannsynligheten dermed 0,079 som er marginalt signifikant, dersom vi velger $\alpha < 0,10$ og funnene er i riktig retning.

Inspeksjon av parameterstimatene for de to eksperimentgruppene viser imidlertid at interaksjonen mellom samtlige 27 vurderinger og instruksjon og vurderingsskala er liten og negativ, når det benyttes subjektforankret instruksjon og -vurderingsskala. Disse funnene er i motsatt retning av hva vi har spesifisert i hypotesene. Dette innebærer at det ikke blir riktig å benytte en-halet p-verdi. Resultatene fra hypotesetestingen indikerer dermed at bruk av objektforankret-, vs. subjektforankret instruksjon og vurderingsskala i forbindelse med spesifikke vurderingsspmåll ikke får betydning for respondentenes atferdsintensjoner.

TABELL 8.9
Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner

| TOTALVURDERING/ INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA | (ADJ. R SQ.) |
|---|-----------------------|--------------|
| Objektforankret instruksjon og -vurderingsskala | 0,54 | (0,29) |
| Subjektforankret instruksjon og -vurderingsskala | 0,51 | (0,27) |

I tabell 8.9 rapporteres standardiserte beta-koeffisienter og forklart varians for ytterligere å illustrere forskjeller mellom eksperimentgruppene. Forskjellene i forklart varians og beta-koeffisienter er svært små. I motsetning til resultatene for totalvurderingene ser vi imidlertid her at konklusjonene for atferdsintensjoner blir de samme enten vi betrakter parameterestimaterne fra variansanalysen eller de standardiserte beta-koeffisientene. Basert på resultatene fra analysene ovenfor konkluderer vi med det ikke finnes støtte for hypotese 4b i dataene.

I kapittel 3 gjorde vi et skille mellom materielle og immaterielle egenskaper ved produktet. Basert på teoridrøftelsene argumenterte vi for at materielle egenskaper vil være viktigere, og bety mer for respondentenes totalvurderinger og atferdsintensjoner enn immaterielle egenskaper. Ved bruk av subjektforankret instruksjon og -vurderingsskala hevdet vi imidlertid at forskjellen i betydning av materielle-, og immaterielle egenskaper vil være mindre. I hypoteseform ble dette, i kapittel 4, uttrykt på følgende måte:

H5a: Når det benyttes subjektforankrete instruksjoner og -vurderingsskalaer i tilknytning til spesifikke vurderings-spørsmål i et spørreskjema, vil immaterielle egenskaper ved produktet ha sterkere relativ betydning for prediksjon av totalvurderinger enn ved bruk av objektforankrete instruksjoner og -vurderingsskalaer

For testing av hypotese 5a behandler vi som tidligere objektforankrete- og subjektforankrete totalvurderinger i separate analyser. Ved operasjonalisering av materielle- og immaterielle produktegenskaper for disse analysene tar vi utgangspunkt i de fire produktelementene som vi benyttet for å beskrive et produkt i kapittel 3¹⁰⁶, og som har dannet utgangspunkt for utvikling av spørreskjemaet som ble benyttet i studien. Tidligere har vi rapportert resultater for samtlige fire produktelementer. Vi er nå spesielt

¹⁰⁶Samhandlings-, struktur-, backstage-, og kundeavhengige produktelementer

interessert i å fokusere på betydningen av materielle og immaterielle egenskaper for totalvurderingene. For disse analysene vil vi derfor kun benytte samhandlingsegenskapene og strukturegenskapene. Samhandlingsegenskaper representerer i denne sammenheng "rene" immaterielle egenskaper, mens strukturegenskaper representerer "rene" materielle egenskaper ved hotellproduktet. De øvrige produkttegenskapene som respondentene ble bedt om å vurdere er vanskelig å klassifisere som enten materielle-, eller immaterielle egenskaper. Ved å fokusere på samhandlingsegenskaper og strukturegenskaper, har vi på denne måten forsøkt å rendyrke operasjonaliseringen av immaterielle- og materielle produkttegenskaper for analysene.

TABELL 8.10a

ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 976,79 | 588 | 1,66 | | |
| Konstant | 0,63 | 1 | 0,63 | 0,38 | 0,539 |
| Immaterielle vurderinger | 34,54 | 1 | 34,54 | 20,79 | < 0,001 |
| Materielle vurderinger | 253,16 | 1 | 253,16 | 152,39 | < 0,001 |
| Instruksjon/vurderingsskala ^a | 12,69 | 1 | 12,69 | 7,64 | 0,006 ^b |
| Immaterielle vurderinger x instruksjon/ vurderingsskala | 0,02 | 1 | 0,02 | 0,01 | 0,914 ^c |
| Materielle vurderinger x instruksjon/vurderingsskala | 7,08 | 1 | 7,08 | 4,26 | 0,039 ^d |
| (Model) | 772,09 | 5 | 154,42 | 92,96 | < 0,001 |
| (Total) | 1748,88 | 593 | 2,95 | | |
| R-squared = | 0,441 | | | | |
| Adjusted R-squared = | 0,437 | | | | |

^a Gruppegjennomsnitt objektforankret totalvurdering:
 Objektforankret instruksjon og -vurderingsskala : 3,18 (n = 307)
 Subjektforankret instruksjon og -vurderingsskala : 3,28 (n = 287)

^b p = 0,003 for en-hale test

^c p = 0,457 for en-hale test

^d p = 0,020 for en-hale test

| Parameterestimer: | Konstant | Immaterielle vurderinger | Materielle vurderinger | Instruksjon/vurderingsskala | Immaterielle vurd. x instruksjon/vurderingsskala | Materielle vurd. x instruksjon/vurderingsskala |
|-------------------|----------|--------------------------|------------------------|-----------------------------|--|--|
| Objektforankret: | 0,102 | 0,267 | 0,686 | - 0,456 | - 0,006(ns) | 0,115 |
| Subjektforankret: | 0,102 | 0,267 | 0,686 | 0,456 | 0,006(ns) | - 0,115 |

Variansanalysen for testing av hypotese 5a i tabell 8.10a, viser først hvordan respondentenes vurderinger av de immaterielle-, og materielle produkttegenskapene,

uavhengig av hvilken instruksjon og vurderingsskala som benyttes, bidrar til å forklare den objektforankrete totalvurderingen av hotellene. De rapporterte p-verdiene, på mindre enn 0,001, indikerer at både materielle-, og immaterielle produktegenskaper har høy prediktiv validitet. Det kan imidlertid merkes at de materielle egenskapene forklarer en vesentlig større del av variansen i objektforankrete totalvurderinger enn de immaterielle egenskapene.

For testing av hypotese 5a fokuserer vi spesielt på de to interaksjonseffektene i tabellen. Disse sier noe om hvorvidt betydningen av immaterielle-, eller materielle egenskaper er forskjellig ved bruk av spørreskjema med henholdsvis objektforankret- og subjektforankret instruksjon og -vurderingsskala for de spesifikke vurderingsspørsmålene. Hypotesene er formulert med hensyn på den relative betydningen av immaterielle egenskaper for totalvurderingene. Dette innebærer at betydningen av immaterielle egenskaper må øke, eller at betydningen av materielle egenskaper må reduseres i eksperimentgruppen med subjektforankret instruksjon og -vurderingsskala for at vi skal konkludere med støtte for hypotese 5a basert på de innsamlede dataene. Fra tabellen ser vi videre at det rapporteres en to-halet sannsynlighet på 0,039 for null-hypotesen om at interaksjonen mellom instruksjon og vurderingsskala og vurdering av materielle egenskaper ikke har betydning for nivået på totalvurderingen. Med en retningsbestemt hypotese, blir den relevante en-halete p-verdien dermed 0,020 dersom funnene er i riktig retning. Interaksjonseffekten for immaterielle vurderinger og instruksjon og vurderingsskala er ikke signifikant.

Inspeksjon av parameterestimatene for de to eksperimentgruppene viser at bruk av subjektforankret instruksjon og -vurderingsskala i forbindelse med de spesifikke vurderingsspørsmålene, reduserer betydningen av materielle egenskaper for den objektforankrete totalvurderingen. Dette innebærer at den relative betydningen av immaterielle egenskaper øker. Vi har dermed valgt å konkludere med at det finnes støtte for hypotese 5a i dataene når det gjelder de objektforankrete totalvurderingene.

Ved å spesifisere signifikansnivå $\alpha = 0,05$ rapporteres verdien for styrke til 0,538 for interaksjonen mellom materielle vurderinger og instruksjon og -vurderingsskala. Den korresponderende *eta-square* rapporteres til 0,007.

Vi har også gjennomført de samme analysene for subjektforankrete totalvurderinger for testing av hypotese 5b. Variansanalysen i tabell 8.10b viser i denne forbindelse først hvordan respondentenes vurderinger av de immaterielle og materielle produktegenskapene, uavhengig av instruksjon og vurderingsskala, bidrar til å forklare den objektforankrete totalvurderingen av hotellene. De rapporterte p-verdiene, på mindre

enn 0,001, indikerer at både materielle og immaterielle produkttegenskaper har prediktiv validitet. De materielle produkttegenskapene forklarer også her respondentenes totalvurdering best.

TABELL 8.10b

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 639,82 | 584 | 1,10 | | |
| Konstant | 62,52 | 1 | 62,52 | 57,06 | < 0,001 |
| Immaterielle vurderinger | 20,10 | 1 | 20,10 | 18,35 | < 0,001 |
| Materielle vurderinger | 181,06 | 1 | 181,06 | 165,27 | < 0,001 |
| Instruksjon/vurderingsskala ^a | 0,79 | 1 | 0,79 | 0,72 | 0,397 ^b |
| Immaterielle vurderinger x instruksjon/vurderingsskala | 3,42 | 1 | 3,42 | 3,12 | 0,078 ^c |
| Materielle vurderinger x instruksjon/vurderingsskala | 8,05 | 1 | 8,05 | 7,34 | 0,007 ^d |
| (Model) | 539,10 | 5 | 107,82 | 98,41 | < 0,001 |
| (Total) | 1178,92 | 589 | 2,00 | | |
| R-squared = | 0,457 | | | | |
| Adjusted R-squared = | 0,453 | | | | |

^a Gruppergjennomsnitt subjektforankret totalvurdering:
 Objektforankret instruksjon og -vurderingsskala : 3,62 (n = 305)
 Subjektforankret instruksjon og -vurderingsskala : 3,55 (n = 285)

^b p = 0,199 for en-hale test

^c p = 0,039 for en-hale test

^d p = 0,003 for en-hale test

| Parameterestimater: | Konstant | Immaterielle vurderinger | Materielle vurderinger | Instruksjon/vurderingsskala | Immaterielle vurd. x instruksjon/vurderingsskala | Materielle vurd. x instruksjon/vurderingsskala |
|---------------------|----------|--------------------------|------------------------|-----------------------------|--|--|
| Objektforankret: | 1,014 | 0,204 | 0,582 | -0,114(ns) | -0,084 | 0,123 |
| Subjektforankret: | 1,014 | 0,204 | 0,582 | 0,114(ns) | 0,084 | -0,123 |

Interaksjonen mellom hvilken instruksjon og vurderingsskala som benyttes i spørreskjemaet, og vurderinger av henholdsvis immaterielle-, og materielle egenskaper er rapportert som det tredje og fjerde leddet i variansanalysen i tabell 8.9b. Fra tabellen ser vi at det rapporteres en to-halet sannsynlighet på 0,078 for nullhypotesen om at interaksjonen mellom instruksjon og vurderingsskala og immaterielle vurderinger ikke har betydning for nivået på den subjektforankrete totalvurderingen. Den relevante en-halete sannsynligheten blir dermed 0,039 dersom funnene er i riktig retning. Fra tabellen ser vi videre at det rapporteres en to-halet sannsynlighet på 0,007 for null-

hypotesen om at interaksjonen mellom instruksjon og vurderingsskala og materielle vurderinger ikke har betydning for nivået på totalvurderingen. Med en retningsbestemt hypotese, blir den en-haite sannsynligheten 0,003 dersom funnene er i riktig retning.

Inspeksjon av korresponderende parameterstimater for de to eksperimentgruppene viser at bruk av subjektforankret instruksjon og -vurderingsskala reduserer betydningen av materielle egenskaper for den subjektforankrete totalvurderingen, mens betydningen av immaterielle produktegenskaper øker. Dette innebærer at den relative betydningen av immaterielle egenskaper øker, og vi kan konkludere med at det finnes støtte for hypotese 5a i dataene også når subjektforankrete totalvurderinger benyttes som avhengig variabel.

Ved å spesifisere signifikansnivå $\alpha = 0,05$ rapporteres verdiene for styrke til 0,422 for interaksjonseffekten mellom immaterielle vurderinger og instruksjon og vurderingsskala, og 0,769 for interaksjonseffekten mellom materielle vurderinger og instruksjon og vurderingsskala. *Eta-square*, som mål på effektstørrelse rapporteres henholdsvis lik 0,005 og 0,012. Effektstørrelsene må også her karakteriseres som små, mens verdiene for styrke indikerer henholdsvis middels, og over middels, statistisk styrke for F-testene.

TABELL 8.11

Standardiserte beta-koeffisienter og forklart varians for objektforankret- og subjektforankret totalvurdering, betydning av materielle- og immaterielle produktegenskaper

| TOTALVURDERING/ INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA | (ADJ. R SQ.) |
|---|-----------------------|--------------|
| Objektforankret totalvurdering | | |
| Objektforankret instruksjon og -vurderingsskala | | |
| Immaterielle vurderinger | 0,17 | (0,44) |
| Materielle vurderinger | 0,55 | |
| Subjektforankret instruksjon og -vurderingsskala | | |
| Immaterielle vurderinger | 0,23 | (0,44) |
| Materielle vurderinger | 0,49 | |
| Subjektforankret totalvurdering | | |
| Objektforankret instruksjon og -vurderingsskala | | |
| Immaterielle vurderinger | 0,10 | (0,47) |
| Materielle vurderinger | 0,62 | |
| Subjektforankret instruksjon og -vurderingsskala | | |
| Immaterielle vurderinger | 0,27 | (0,45) |
| Materielle vurderinger | 0,44 | |

For ytterligere å illustrere hvordan valg av instruksjon og vurderingsskala i forbindelse med spesifikke vurderings spørsmål i et spørreskjema får innvirkning på den relative betydningen av materielle og immaterielle produkt egenskaper, når det gjelder disse egenskapenes relative evne til å forklare totalvurderinger av produktet, rapporteres forklart varians og standardiserte beta-koeffisienter i tabell 8.11.

Tabell 8.11 indikerer at de rapporterte beta-koeffisientene, både for objektforankrete og subjektforankrete totalvurderinger, viser et mønster som støtter hypotese 5a. Forklart varians endrer seg ikke nevneverdig mellom eksperimentgruppene.

Multikollinearitet kan være et problem i forbindelse med analyser der vurderinger av immaterielle-, og materielle produkt egenskaper inngår som uavhengige variabler. For å kontrollere for dette ble først korrelasjonen mellom disse variablene inspisert. Korrelasjonen mellom immaterielle-, og materielle vurderinger rapporteres til 0,65 i eksperimentgruppen med objektforankret instruksjon og -vurderingsskala, og 0,71 i eksperimentgruppen med subjektforankret instruksjon og -vurderingsskala. I henhold til et konvensjonelt kriterium som benyttes litteraturen (se f.eks.: Berry og Feldman (1985) og Lewis-Beck (1980)), der korrelasjoner mellom uavhengige variabler høyere enn 0,8 betraktes som problematiske, konkluderer vi med at multikollinearitet ikke utgjør noe stort problem for analysene¹⁰⁷.

I tillegg til å betrakte objektforankrete- og subjektforankrete totalvurderinger som avhengige variabler, ble de samme implikasjonene av valg mellom objektforankret- og subjektforankret instruksjon og -vurderingsskala, i kapittel 3 også diskutert i forbindelse med respondentenes atferdsintensjoner. I kapittel 4 ble følgende hypotese formulert.

H5b: Når det benyttes subjektforankrete instruksjoner og -vurderingsskalaer i tilknytning til spesifikke vurderings spørsmål i et spørreskjema, vil immaterielle egenskaper ved produktet ha sterkere relativ betydning for prediksjon av atferdsintensjoner enn ved bruk av objektforankrete instruksjoner og -vurderingsskalaer

¹⁰⁷ En alternativ måte å teste for betydningen av multikollinearitet er å betrakte den ene uavhengige variabelen som avhengig av den andre (Berry og Feldman 1985, Lewis-Beck 1980). Dersom forklart varians ikke nærmer seg 1, betraktes ikke korrelasjon mellom de uavhengige variablene som noe vesentlig problem for analysene. For eksperimentgruppen med objektforankret instruksjon og -vurderingsskala rapporteres forklart varians lik 0,41, mens vi i eksperimentgruppen med subjektforankret instruksjon og -vurderingsskala observerer forklart varians lik 0,51. Ingen av testene for multikollinearitet indikerer at dette vil være et problem for analysene.

De samme analysene som ble gjennomført for totalvurderingene, ble også foretatt for respondentenes atferdsintensjoner. Variansanalysen som er rapportert i tabell 8.12 viser for det første hvordan hotellgjestenes vurdering av de immaterielle-, og materielle produkttegenskapene, uavhengig av instruksjon og vurderingsskala, bidrar til å forklare atferdsintensjonene. De rapporterte p-verdiene, på mindre enn 0,001, indikerer at både materielle og immaterielle produkttegenskaper har prediktiv validitet for gjestenes ønske om å besøke hotellet igjen, og for sannsynligheten for at de vil anbefale hotellet til andre. Også her forklarer de materielle egenskapene en vesentlig større del av variansen i den avhengige variabelen enn de immaterielle egenskapene.

TABELL 8.12

ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 2067,83 | 590 | 3,50 | | |
| Konstant | 1921,85 | 1 | 1921,85 | 548,35 | < 0,001 |
| Immaterielle vurderinger | 69,29 | 1 | 69,29 | 19,77 | < 0,001 |
| Materielle vurderinger | 242,76 | 1 | 242,76 | 69,26 | < 0,001 |
| Instruksjon/vurderingsskala ^a | 10,07 | 1 | 10,07 | 2,87 | 0,091 ^b |
| Immaterielle vurderinger x instruksjon/vurderingsskala | 0,02 | 1 | 0,02 | 0,01 | 0,944 ^c |
| Materielle vurderinger x instruksjon/vurderingsskala | 8,11 | 1 | 8,11 | 2,31 | 0,129 ^c |
| (Model) | 913,97 | 5 | 182,79 | 52,16 | < 0,001 |
| (Total) | 2981,80 | 595 | 5,01 | | |
| R-squared = | 0,307 | | | | |
| Adjusted R-squared = | 0,301 | | | | |

^a Gruppegjennomsnitt atferdsintensjoner:

Objektforankret instruksjon og -vurderingsskala : 9,12 (n = 309)

Subjektforankret instruksjon og -vurderingsskala : 9,05 (n = 287)

^b p = 0,045 for en-hale test

^c p = 0,472 for en-hale test

^d p = 0,065 for en-hale test

| Parameterestimer: | Konstant | Immaterielle vurderinger | Materielle vurderinger | Instruksjon/vurderingsskala | Immaterielle vurd. x instruksjon/vurderingsskala | Materielle vurd. x instruksjon/vurderingsskala |
|-------------------|----------|--------------------------|------------------------|-----------------------------|--|--|
| Objektforankret: | 5,605 | 0,378 | 0,672 | - 0,406 | - 0,006(ns) | 0,123 |
| Subjektforankret: | 5,605 | 0,378 | 0,672 | 0,406 | 0,006(ns) | - 0,123 |

Fra tabellen ser vi videre at det rapporteres en to-halet sannsynlighet på 0,129 for nullhypotesen om at interaksjonen mellom instruksjon og vurderingsskala og materielle

vurderinger ikke har betydning for atferdsintensjonene. Den relevante en-halete sannsynligheten blir dermed 0,065 som er statistisk signifikant med $\alpha < 0,10$ dersom funnene er i riktig retning. Interaksjonen mellom immaterielle vurderinger og instruksjon og vurderingsskala er ikke signifikant.

Parameterstimatene for de to eksperimentgruppene viser at bruk av subjektforankret instruksjon og -vurderingsskala i forbindelse med de spesifikke produktvurderingene reduserer betydningen av materielle egenskaper for respondentenes atferdsintensjoner. Dette kan tolkes som at den relative betydningen av immaterielle egenskaper økes. Vi velger derfor å konkludere med at det finnes svak støtte for hypotese 5b i dataene når respondentenes atferdsintensjoner benyttes som avhengig variabel.

Ved å spesifisere signifikansnivå $\alpha = 0,05$ rapporteres verdien for styrke til 0,329 for interaksjonen mellom materielle vurderinger og instruksjon og -vurderingsskala. Den korresponderende *eta-square* rapporteres til 0,004 Både den rapporterte effektstørrelsen og verdien for styrke må karakteriseres som svært liten.

Også her suppleres resultatene fra variansanalysen med rapportering av forklart varians og standardiserte beta-koeffisienter for ytterligere å illustrere hvordan valg av instruksjon og vurderingsskala i forbindelse med spesifikke vurderingsspmåål i et spørreskjema kan få innvirkning på den relative betydningen av materielle-, og immaterielle produktenskaper.

TABELL 8.13

Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner, betydning av materielle- og immaterielle produktenskaper

| INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT | |
|--|---------------|--------------|
| | BETA | (ADJ. R SQ.) |
| Objektforankret instruksjon og -vurderingsskala | | |
| Immaterielle vurderinger | 0,19 | (0,33) |
| Materielle vurderinger | 0,44 | |
| Subjektforankret instruksjon og -vurderingsskala | | |
| Immaterielle vurderinger | 0,23 | (0,27) |
| Materielle vurderinger | 0,34 | |

Tabell 8.13 indikerer at de rapporterte beta-koeffisientene viser et mønster som støtter hypotese 5b. Betydningen av immaterielle produktenskaper øker, mens betydningen av materielle produktenskaper reduseres for atferdsintensjonene når det benyttes subjektforankret instruksjon og -vurderingsskala for de spesifikke vurderingene.

Sammenlignet med variansanalysen i tabell 8.12 er det imidlertid kun effekten for materielle egenskaper som er signifikant. Forklart varians er i liten grad forskjellig mellom eksperimentgruppene.

8.1.5 Forskjeller i sikkerhet ved individuelle vurderinger

I kapittel 3 ble det argumentert for at subjektforankrete vurderinger vil være mer tilgjengelige for respondentene enn objektforankrete vurderinger. Tilgjengelighet til holdninger og tilhørende vurderinger vil f.eks. være avhengig av at holdninger og vurderinger er utviklet. I hvilken grad et individ har utviklet holdninger overfor et objekt vil også få betydning for sikkerhet i vurdering av objektet. "Tilfreds-tilfreds" vurderinger vil i mange tilfeller representere en lettere oppgave for respondentene enn "bra-dårlig" vurderinger.

I kapittel 4 ble følgende hypotese formulert i forbindelse med respondentenes sikkerhet i vurderinger foretatt i spørreskjema med henholdsvis subjektforankret- og objektforankret instruksjon og -vurderingsskala .

H6: Respondenter vil være sikrere på vurdering av spesifikke produktegenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer enn vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer

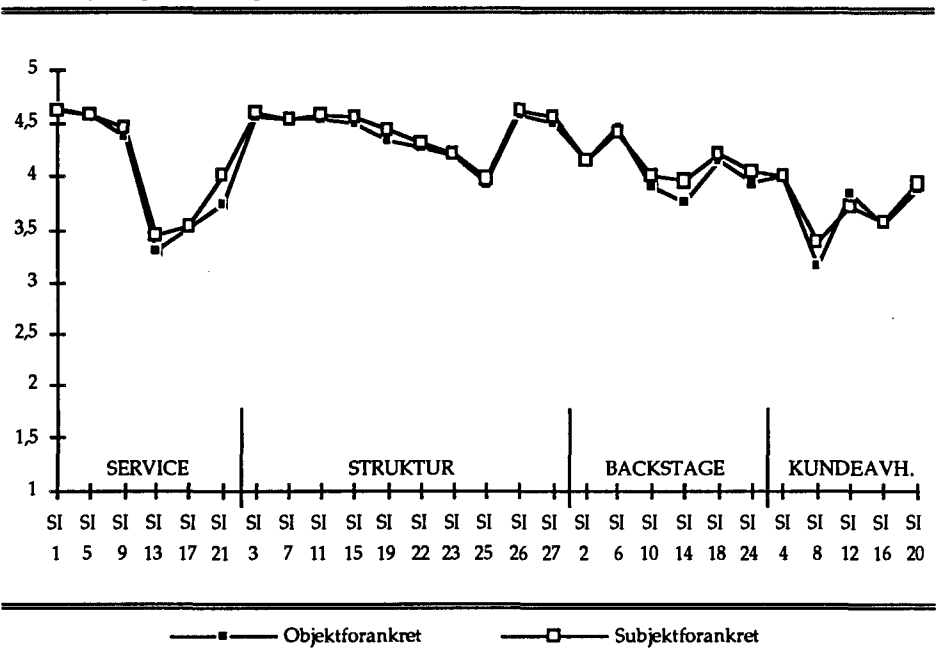
En måte å teste hypotese 6 er å sammenligne gruppegjennomsnitt for sikkerhet i vurderinger for hver av de individuelle vurderingsvariablene i de to eksperimentgruppene.

På X-aksen i figur 8.4 finner vi igjen de 27 forskjellige aspektene ved hotellproduktet som respondentene ble bedt om å vurdere. Rekkefølgen disse er presentert i samsvarer med grupperingen av egenskapene i henhold til de fire produktelementene samhandlings-, struktur-, backstage- og kundeavhengige produktelementer. Y-aksen i figuren representerer nivå på sikkerhet der skalaen i spørreskjemaet går fra 1 ("svært usikker") til 5 ("svært sikker")¹⁰⁸.

¹⁰⁸Se spørreskjema i appendix B

FIGUR 8.4

Sikkerhet i vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala



Visuell inspeksjon av figur 8.4 indikerer små, eller ingen forskjeller mellom eksperimentgruppene når det gjelder gjennomsnittsverdier for respondentenes sikkerhet i forbindelse med de spesifikke vurderingene. En nærmere inspeksjon av figuren indikerer at respondentene, kun for fem av de individuelle vurderingene, ser ut til å være sikrere i sine vurderinger når det benyttes subjektforankret instruksjon og -vurderingsskala. Imidlertid er ingen av disse forskjellene statistisk signifikante når vi velger $\alpha < 0,05$.

En alternativ måte å teste hypotesen om forskjeller i gruppegjennomsnitt for sikkerhet i de spesifikke produktvurderingene, er å betrakte rapportering av sikkerhet som en tilnærmet binomisk serie. Det kan altså stilles spørsmål om sannsynligheten for at det rapporteres høyest sikkerhet i vurderinger i eksperimentgruppen med subjektforankret instruksjon og -vurderingsskala for minst 5 av 27 vurderinger av ren tilfeldighet. Resultatene fra en ikke-parametrisk binomisk test gir en en-halet sannsynlighet nær 1 for at det observerte mønsteret er tilfeldig dersom de to eksperimentgruppene er trukket fra samme populasjon.

En tredje måte å teste hypotese 6 er å kalkulere gjennomsnittlig sikkerhet på tvers av alle 27 vurderinger for de to eksperimentgruppene. Resultater fra en T-test der forskjellene mellom eksperimentgruppene er analysert er presentert i tabell 8.14.

TABELL 8.14

T-test, sikkerhet i vurderinger av de 27 spesifikke produkttegenskapene foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala.

| VARIABLE | | NUMBER OF CASES | MEAN | STANDARD DEVIATION | SE OF MEAN | | |
|---------------------------------------|-------------|-----------------|------------------------------|--------------------|------------|--------------------|-------------|
| Sikkerhet ved samtlige 27 vurderinger | | | | | | | |
| Objektforankret | | 305 | 4,1431 | 0,667 | 0,038 | | |
| Subjektforankret | | 277 | 4,1792 | 0,689 | 0,041 | | |
| * POOLED VARIANCE ESTIMATE | | | * SEPARATE VARIANCE ESTIMATE | | | | |
| F VALUE | 2-TAIL PROB | T VALUE | DEGREES OF FREEDOM | 2-TAIL PROB | T VALUE | DEGREES OF FREEDOM | 2-TAIL PROB |
| 0,338 | 0,561 | * -0,64 | 580 | 0,522 | * -0,64 | 570,71 | 0,522 |

Ved å kalkulere gjennomsnittlig sikkerhet for samtlige 27 vurderinger, reduseres tilfeldige fluktuasjoner som inngår ved rapportering av sikkerhet for de individuelle vurderingene. Resultatene som rapporteres i tabell 8.14 viser at forskjellene i sikkerhet mellom eksperimentgruppene, selv med en en-halet test med p-verdi lik 0,261, ikke kan betegnes som statistisk signifikante.

En fjerde måte å teste hypotese 6 er å gruppere sikkerhet i de 27 vurderingene ved konstruksjon av summerte skalaer. For å bestemme hvorvidt gjennomsnittene for de fire summerte skalaene er forskjellige for de to eksperimentgruppene ble det gjennomført multivariat variansanalyse, MANOVA, der de fire produkttelementene samhandlings-, struktur-, backstage- og kundeavhengige produkttelementer behandles som korrelerte avhengige variabler. Resultatene fra denne analysen presenteres i tabell 8.15.

TABELL 8.15

MANOVA, sikkerhet i vurderinger av produkttelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala. Rapportering av Pillais-T

| SOURCE OF VARIATION | VALUE | EXACT F | HYP. DF | ERROR F | SIGN. OF F |
|-----------------------------|---------|---------|---------|---------|------------|
| Instruksjon/vurderingsskala | | | | | |
| Pillais-T | 0,00383 | 0,53342 | 4,00 | 555,00 | 0,711 |

I tabell 8.15 rapporteres Pillais-T¹⁰⁹ for test av forskjeller mellom eksperimentgruppene for samtlige avhengige variabler. Som det fremgår av tabellen finnes det ingen signifikante forskjeller i sikkerhet mellom eksperimentgruppene når de summerte skalaene benyttes som avhengige variabler. Resultatene fra analysene som er foretatt for å undersøke forskjeller i sikkerhet i spesifikke produktvurderinger mellom eksperimentgruppene leder oss til å konkludere med at det ikke finnes støtte for hypotese 6 i de innsamlede dataene.

8.1.6 Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner - respondentenes produktkunnskap som modererende variabel

I kapittel 3 drøftet vi hvordan respondentenes brukserfaring og produktkunnskap kan påvirke både produktvurderingene og respons i spørreskjemaet. Økt brukserfaring og produktkunnskap kan bidra til å øke respondentenes tilgjengelighet både til holdninger og produktvurderinger. Produktkunnskap kan på denne måten bidra til å redusere forskjellene som vi fremmet i hypotese 4a og 4b, angående de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner mellom eksperimentgruppene som er tilordnet spørreskjema med henholdsvis subjektforankret- og objektforankret instruksjon og -vurderingsskala. Nedenfor undersøker vi først betydningen av produktkunnskap for respondentenes totalvurderinger. I kapittel 4 ble det i forbindelse med totalvurderingene formulert følgende hypotese:

H7a: Forskjellene mellom spesifikke vurderinger foretatt i spørreskjema med subjektforankrete - vs. objektforankrete instruksjoner og -vurderingsskalaer, når det gjelder de spesifikke vurderingenes evne til å forklare totalvurderinger, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

Hovedformålet med hypotese 7a er å finne ut hvordan valg av instruksjon og vurderingsskala for spesifikke vurderinger i et spørreskjema, får betydning for disse vurderingenes evne til å forklare respondentenes totalvurderinger når vi tar hensyn til respondentenes produktkunnskap. For testing av hypotese 7a har vi derfor valgt å

¹⁰⁹Pillais-T er mest robust for brudd på forutsetningene for bruk av MANOVA

gjennomføre variansanalyser innenfor ulike nivå av produktkunnskap. Variansanalysene kan da enten gjennomføres separat for hvert nivå av produktkunnskap, eller gjennom å inkludere produktkunnskap som en faktor i variansanalysene¹¹⁰.

I dette avsnittet rapporteres resultater fra variansanalyser det respondentenes produktkunnskap inkluderes. I tabell E.2 - tabell E.7 i appendiks E har vi i tillegg rapportert variansanalyser gjennomført separat innenfor hvert av de tre produktkunnskapsnivåene som ble operasjonalisert i kapittel 6. Parameterestimatene i tabell 8.16a og 8.16b er hentet fra disse analysene. Dette forenkler tolkningen av interaksjonen mellom de spesifikke vurderingene og instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap. I tillegg til rapportering av resultater fra variansanalyser illustreres forskjellene mellom eksperimentgruppene, når vi tar hensyn til respondentenes produktkunnskap, også her ved hjelp av standardiserte beta-koeffisienter og forklart varians.

I henhold til hypotese 7a, er vi som tidligere nevnt interessert i å analysere i hvilken grad respondentenes produktkunnskap moderer effektene valg av objektforankret- vs. subjektforankret instruksjon og -vurderingsskala i forbindelse med de 27 spesifikke vurderingsspørsmålene i spørreskjemaet, får for disse spørsmålenes evne til å forklare respondentenes totalvurderinger. Produktkunnskap inngår som en faktor i variansanalysen, og interaksjonen mellom samtlige 27 vurderinger og instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap er det relevante leddet som rapporteres for testing av hypotese 7a.

Først vil vi se nærmere på de øvrige funnene fra variansanalysen i tabell 8.16a. Tabellen viser at valg av instruksjon og vurderingsskala har en direkte betydning for nivået på den objektforankrete totalvurderingen, mens det samme ikke er tilfellet for produktkunnskap. Videre ser vi at interaksjonen mellom instruksjon og vurderingsskala og respondentenes produktkunnskap er svært signifikant med to-halet p-verdi lik 0,012. Dette betyr at produktkunnskap direkte modererer effekten av instruksjon og vurderingsskala for nivået på den objektforankrete totalvurderingen.

Under nullhypotesen om at det ikke finnes forskjell i de 27 produkttegenskapenes evne til å forklare objektforankrete totalvurderinger mellom de to eksperimentgruppene når vi tar hensyn til respondentenes produktkunnskap, viser ANOVA tabellen en F-verdi på 3,64 med en korresponderende to-halet sannsynlighet på 0,013. Med en retningsbestemt hypotese blir den relevante en-halete sannsynligheten 0,006 dersom funnene er i riktig

¹¹⁰Dersom produktkunnskap inkluderes som en faktor kan hypotesen testes i et "nested design" gjennom å benytte subrutinen *within* for produktkunnskapsleddet i MANOVA i programpakken SPSSX.

retning. Denne p-verdien indikerer at produktkunnskap har betydning for effektene av valg av instruksjon og vurderingsskala for de spesifikke vurderingenes evne til å predikere objektforankrete totalvurderinger i spørreskjemaet.

TABELL 8.16a

ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 1067,46 | 581 | 1,84 | | |
| Konstant | 9,59 | 1 | 9,59 | 5,22 | 0,023 |
| Samtlige 27 vurderinger | 650,35 | 1 | 650,35 | 353,97 | < 0,001 |
| Instruksjon/vurderingsskala | 9,65 | 1 | 9,65 | 5,25 | 0,022 ^b |
| Produktkunnskap | 0,04 | 2 | 0,02 | 0,01 | 0,988 ^c |
| Instruksjon/vurderingsskala x produktkunnskap ^a | 16,30 | 2 | 8,15 | 4,44 | 0,012 ^d |
| Samtlige 27 vurderinger x instruksjon/vurderingsskala within produktkunnskap | 20,05 | 3 | 6,68 | 3,64 | 0,013 ^e |
| (Model) | 679,13 | 9 | 75,46 | 41,07 | < 0,001 |
| (Total) | 1746,59 | 590 | 2,96 | | |
| R-squared = | 0,389 | | | | |
| Adjusted R-squared = | 0,379 | | | | |

^a Gruppegjennomsnitt objektforankret totalvurdering:
 Objektforankret instruksjon og -vurderingsskala
 Subjektforankret instruksjon og -vurderingsskala

Lav produktkunnskap: 3,26 (n = 90)
 Middels produktkunnskap: 3,03 (n = 88)
 Høy produktkunnskap: 3,20 (n = 125)
 Lav produktkunnskap: 3,17 (n = 102)
 Middels produktkunnskap: 3,50 (n = 63)
 Høy produktkunnskap: 3,27 (n = 123)

^b p = 0,011 for en-hale test

^c p = 0,494 for en-hale test

^d p = 0,006 for en-hale test

^e p = 0,006 for en-hale test

| Parameterestimer: | Konstant | Samtlige 27 vurderinger | Instruksjon / vurderingsskala | Samtlige 27 vurderinger x instruksjon/vurderingsskala |
|-------------------------|----------|-------------------------|-------------------------------|---|
| Lav produktkunnskap | | | | |
| Objektforankret: | 0,385 | 0,906 | 0,271(ns) | -0,071(ns) |
| Subjektforankret:: | 0,385 | 0,906 | -0,271(ns) | 0,071(ns) |
| Middels produktkunnskap | | | | |
| Objektforankret: | 0,389 | 0,911 | -0,608 | 0,108(ns) |
| Subjektforankret:: | 0,389 | 0,911 | 0,608 | -0,108(ns) |
| Høy produktkunnskap | | | | |
| Objektforankret: | 0,350 | 0,919 | -0,809 | 0,218 |
| Subjektforankret:: | 0,350 | 0,919 | 0,809 | -0,218 |

Ved inspeksjon av parameterstimatene for de to eksperimentgruppene innenfor de tre nivåene av produktkunnskap fremgår det at interaksjonen mellom de spesifikke

vurderingene og instruksjon og vurderingsskala for ulike nivå av produktkunnskap kun er signifikant for respondenter med høy produktkunnskap. For disse respondentene forklarer spesifikke produktvurderinger, foretatt med objektforankret instruksjon og -vurderingsskala, objektforankret totalvurdering bedre enn spesifikke vurderinger foretatt med subjektforankret instruksjon og -vurderingsskala.

Ved å spesifisere signifikansnivå $\alpha = 0,05$ rapporteres verdiene for styrke til 0,762 for to-veis-interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap, og 0,797 for interaksjonen mellom de 27 vurderingsspørsmålene og instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap. De korresponderende verdiene for *eta-square* rapporteres lik 0,015 og 0,018. Effektstørrelsene er små, mens verdiene for styrke indikerer over middels statistisk styrke for F-testene.

En tilsvarende variansanalyse ble gjennomført med hensyn på subjektforankret totalvurdering. Resultatene fra denne rapporteres i tabell 8.16b. I motsetning til resultatene som ble rapportert for objektforankrete totalvurderinger, er ikke den direkte effekten av instruksjon og vurderingsskala signifikant for nivået på subjektforankrete totalvurderinger. Det rapporteres heller ikke signifikante effekter av produktkunnskap (tohalet p-verdi på 0,121). Tabellen viser imidlertid signifikant interaksjon mellom instruksjon og vurderingsskala og produktkunnskap, og signifikant interaksjon mellom de 27 spesifikke vurderingsspørsmålene og instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap. Begge disse er signifikante med $\alpha < 0,01$ dersom funnene er i riktig retning.

Gjennom inspeksjon av parameterstimatene for de to eksperimentgruppene fremkommer det et interessant mønster når vi tar hensyn til respondentenes produktkunnskap. For respondenter med lav produktkunnskap forklarer de spesifikke vurderingsspørsmålene subjektforankret totalvurdering bedre når det benyttes subjektforankret instruksjon og -vurderingsskala, enn når det benyttes objektforankret instruksjon og -vurderingsskala. Det motsatte er tilfellet for respondenter med høy produktkunnskap. Her forklarer spesifikke spørsmål foretatt med bakgrunn i objektforankret instruksjon og -vurderingsskala den subjektforankrete totalvurderingen best. For respondenter med middels produktkunnskap rapporteres det ingen signifikante forskjeller i de spesifikke vurderingenes forklaringsvegne mellom eksperimentgruppene.

Ved signifikansnivå $\alpha = 0,05$ rapporteres verdiene for styrke ved de statistiske testene til 0,900 for to-veis-interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap, og 0,787 for to-veis-interaksjonen mellom de 27 vurderingsspørsmålene og instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap. *Eta-square*,

som et mål på effektstørrelse rapporteres lik 0,022 og 0,018. Effektstørrelsene er fortsatt relativt små, mens verdiene for styrke indikerer høy statistisk styrke for F-testene.

TABELL 8.16b

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 699,00 | 576 | 1,21 | | |
| Konstant | 106,01 | 1 | 106,01 | 87,36 | < 0,001 |
| Samtlige 27 vurderinger | 446,72 | 1 | 446,72 | 368,11 | < 0,001 |
| Instruksjon/vurderingsskala | 0,78 | 1 | 0,78 | 0,64 | 0,423 ^b |
| Produktkunnskap | 5,14 | 2 | 2,57 | 2,12 | 0,121 ^c |
| Instruksjon/vurderingsskala x produktkunnskap ^a | 15,44 | 2 | 7,72 | 6,36 | 0,002 ^d |
| Samtlige 27 vurderinger x instruksjon/vurderingsskala within produktkunnskap | 12,93 | 3 | 4,31 | 3,55 | 0,014 ^e |
| (Model) | 479,23 | 9 | 53,25 | 43,88 | < 0,001 |
| (Total) | 1178,23 | 585 | 2,01 | | |
| R-squared = | 0,407 | | | | |
| Adjusted R-squared = | 0,397 | | | | |

^a Gruppegjennomsnitt objektforankret totalvurdering:
Objektforankret instruksjon og -vurderingsskala

Lav produktkunnskap: 3,67 (n = 87)
Middels produktkunnskap: 3,69 (n = 89)
Høy produktkunnskap: 3,54 (n = 125)
Lav produktkunnskap: 3,46 (n = 101)
Middels produktkunnskap: 3,84 (n = 62)
Høy produktkunnskap: 3,48 (n = 122)

Subjektforankret instruksjon og -vurderingsskala

^b p = 0,212 for en-hale test

^c p = 0,061 for en-hale test

^d p = 0,001 for en-hale test

^e p = 0,007 for en-hale test

| Parameterestimater: | Konstant | Samtlige 27 vurderinger | Instruksjon / vurderingsskala | Samtlige 27 vurderinger x instruksjon / vurderingsskala |
|-------------------------|----------|-------------------------|-------------------------------|---|
| Lav produktkunnskap | | | | |
| Objektforankret: | 1,292 | 0,731 | 0,552 | - 0,142 |
| Subjektforankret: | 1,292 | 0,731 | - 0,552 | 0,142 |
| Middels produktkunnskap | | | | |
| Objektforankret: | 1,421 | 0,740 | - 0,365 | 0,079(ns) |
| Subjektforankret: | 1,421 | 0,740 | 0,365 | - 0,079(ns) |
| Høy produktkunnskap | | | | |
| Objektforankret: | 1,043 | 0,785 | - 0,515 | 0,145 |
| Subjektforankret: | 1,043 | 0,785 | 0,515 | - 0,145 |

For ytterligere å illustrere forskjellene mellom eksperimentgruppene, rapporteres standardiserte beta-koeffisienter og forklart varians i tabell 8.17.

TABELL 8.17

Standardiserte beta-koeffisienter og forklart varians for objektforankret- og subjektforankret totalvurdering, innenfor ulike nivåer av produktkunnskap

| TOTALVURDERING/ INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA / (ADJ. R SQ.) | | |
|---|--------------------------------------|---------------|-------------------|
| | L.P. | M.P. | H.P. ^a |
| Objektforankret totalvurdering/ | | | |
| Objektforankret instruksjon og -vurderingsskala | 0,60 / (0,35) | 0,55 / (0,29) | 0,65 / (0,42) |
| Subjektforankret instruksjon og -vurderingsskala | 0,75 / (0,56) | 0,70 / (0,48) | 0,52 / (0,26) |
| Subjektforankret totalvurdering/ | | | |
| Objektforankret instruksjon og -vurderingsskala | 0,54 / (0,28) | 0,54 / (0,29) | 0,71 / (0,51) |
| Subjektforankret instruksjon og -vurderingsskala | 0,72 / (0,52) | 0,73 / (0,53) | 0,54 / (0,29) |

^a L.P. = Lav produktkunnskap M.P. = Middels produktkunnskap H.P. = Høy produktkunnskap

Forskjellene i forklart varians og beta-koeffisienter er relativt store. Det samme mønsteret som fremkommer i variansanalysene finnes igjen i tabell 8.17, med unntak av tallene som rapporteres for respondenter med middels produktkunnskap.

Basert på resultatene fra variansanalysene ovenfor, de rapporterte parameterestimaterne, beta-koeffisientene og forskjellene i forklart varians kan vi konkludere med at det ikke finnes støtte for hypotese 7a i dataene når det gjelder objektforankrete totalvurderinger. Analysene indikerer at forskjellene mellom eksperimentgruppene, i motsetning til hva vi hadde forventet, er større for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap. Analysene indikerer derimot at det til en viss grad finnes støtte for hypotese 7a i dataene for subjektforankrete totalvurderinger. Forskjellene mellom eksperimentgruppene er i henhold til hypotesen i riktig retning for respondenter med lav produktkunnskap. For respondenter med middels produktkunnskap blir forskjellene mellom eksperimentgruppene mindre og ikke signifikante. Dette er også i samsvar med hypotese 7a. Et noe overraskende funn er det imidlertid at objektforankrete vurderinger av spesifikke produkttegenskaper forklarer subjektforankrete totalvurderinger bedre enn spesifikke vurderinger foretatt med subjektforankret instruksjon og -vurderingsskala.

I forbindelse med teoridrøftingen ble atferdsintensjoner betraktet som en nøytral avhengig variabel. Respondentenes ønske om å besøke hotellet igjen ved en senere anledning, og sannsynligheten for at de vil anbefale hotellet til andre kan oppfattes som å reflektere totalvurderinger av hotellet. I kapittel 4 ble følgende hypotese formulert i

forbindelse med produktkunnskap og de spesifikke vurderingenes evne til å forklare atferdsintensjoner:

H7b: Forskjellene mellom spesifikke vurderinger foretatt i spørreskjema med subjektforankrete- vs. objektforankrete instruksjoner og -vurderingsskalaer, når det gjelder de spesifikke vurderingenes evne til å forklare atferdsintensjoner, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

For å teste hypotese 7b ble tilsvarende analyser, som de vi gjennomførte for objektforankrete- og subjektforankrete totalvurderinger, også gjennomført for respondentenes atferdsintensjoner. Variansanalyse der produktkunnskap inkluderes som en faktor i analysen er rapportert i tabell 8.18. I tabell E.8 - tabell E.10 i appendiks E presenteres i tillegg separate variansanalyser for hvert nivå av produktkunnskap.

På samme måte som for de subjektforankrete totalvurderingene, er det også her kun toveis-interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap og toveis-interaksjonen mellom de spesifikke vurderingsspørsmålene og instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap som er signifikant med $\alpha < 0,05$.

ANOVA tabellen viser en F-verdi på 7,79 med en to-halet sannsynlighet mindre enn 0,001 for toveis-interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap, og en F-verdi på 7,05 med to-halet p-verdi på 0,001 for interaksjonen mellom de spesifikke vurderingsspørsmålene og instruksjon og vurderingsskala når vi tar hensyn til respondentenes produktkunnskap. Med retningsbestemte hypoteser blir de relevante en-halete p-verdiene, begge mindre enn 0,001 dersom funnene er i riktig retning. P-verdiene indikerer at valg av instruksjon og vurderingsskala i spørreskjemaet får betydning for de spesifikke produkttegenskapens evne til å forklare atferdsintensjoner når vi tar hensyn til respondentenes produktkunnskap. For respondenter med lav produktkunnskap forklarer de spesifikke vurderingene atferdsintensjoner bedre når det benyttes subjektforankret instruksjon og -vurderingsskala, enn når det benyttes objektforankret instruksjon og -vurderingsskala. Det motsatte er tilfellet for respondenter med høy produktkunnskap. For disse forklarer spesifikke produktvurderinger foretatt på bakgrunn av objektforankret instruksjon og -vurderingsskala

atferdsintensjoner best. For respondenter med middels produktkunnskap rapporteres det ikke signifikante forskjeller i forklaringsevne mellom eksperimentgruppene.

TABELL 8.18

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivåer av produktkunnskap.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 2076,07 | 583 | 3,56 | | |
| Konstant | 2376,97 | 1 | 2376,97 | 667,50 | < 0,001 |
| Samtlige 27 vurderinger | 825,46 | 1 | 825,46 | 231,80 | < 0,001 |
| Instruksjon/vurderingsskala | 2,96 | 1 | 2,96 | 0,83 | 0,362 ^b |
| Produktkunnskap | 3,86 | 2 | 1,93 | 0,54 | 0,582 ^c |
| Instruksjon/vurderingsskala x produktkunnskap ^a | 58,72 | 2 | 29,36 | 8,25 | < 0,001 |
| Samtlige 27 vurderinger x instruksjon/vurderingsskala within produktkunnskap | 60,04 | 3 | 20,01 | 5,62 | 0,001 ^d |
| (Model) | 898,68 | 9 | 99,85 | 28,04 | < 0,001 |
| (Total) | 2974,76 | 592 | 5,02 | | |
| R-squared = | 0,302 | | | | |
| Adjusted R-squared = | 0,291 | | | | |

^a Gruppergjennomsnitt atferdsintensjoner:

| | |
|--|--|
| Objektforankret instruksjon og -vurderingsskala | Lav produktkunnskap: 3,14 (n = 90) |
| | Middels produktkunnskap: 3,99 (n = 88) |
| | Høy produktkunnskap: 3,17 (n = 127) |
| Subjektforankret instruksjon og -vurderingsskala | Lav produktkunnskap: 3,93 (n = 102) |
| | Middels produktkunnskap: 3,96 (n = 63) |
| | Høy produktkunnskap: 3,18 (n = 123) |

^b p = 0,181 for en-hale test

^c p = 0,291 for en-hale test

^d p < 0,001 for en-hale test

| Parameterestimer: | Konstant | Samtlige 27 vurderinger | Instruksjon / vurderingsskala | Samtlige 27 vurderinger x instruksjon/vurderingsskala |
|-------------------------|----------|-------------------------|-------------------------------|---|
| Lav produktkunnskap | | | | |
| Objektforankret: | 5,865 | 1,015 | 0,902 | - 0,254 |
| Subjektforankret: | 5,865 | 1,015 | - 0,902 | 0,254 |
| Middels produktkunnskap | | | | |
| Objektforankret: | 5,418 | 1,125 | - 0,331(ns) | 0,094(ns) |
| Subjektforankret: | 5,418 | 1,125 | 0,331(ns) | - 0,094(ns) |
| Høy produktkunnskap | | | | |
| Objektforankret: | 6,083 | 0,978 | - 1,160 | 0,334 |
| Subjektforankret: | 6,083 | 0,978 | 1,160 | - 0,334 |

For interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap, og for interaksjonen mellom de 27 vurderingsskjemaene, instruksjon og vurderingsskala innenfor ulike nivå av produktkunnskap, rapporteres verdiene for styrke ved de

statistiske testene til 0,961 og 0,945 ved spesifisert signifikansnivå $\alpha = 0,05$. *Eta-square* rapporteres henholdsvis lik 0,028 og 0,028. De rapporterte effektstørrelsene er også her relativt små, mens verdiene for styrke indikerer høy statistisk styrke for F-testene.

For ytterligere å illustrere forskjellene mellom eksperimentgruppene når det gjelder atferdsintensjoner, når vi tar hensyn til respondentenes produktkunnskap, rapporteres standardiserte beta-koeffisienter og forklart varians i tabell 8.19.

TABELL 8.19

Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner, innenfor ulike nivåer av produktkunnskap

| INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA / (ADJ. R SQ.) | | |
|--|--------------------------------------|---------------|-------------------|
| | L.P. | M.P. | H.P. ^a |
| Objektforankret instruksjon og -vurderingsskala | 0,41 / (0,16) | 0,53 / (0,27) | 0,63 / (0,40) |
| Subjektforankret instruksjon og -vurderingsskala | 0,73 / (0,53) | 0,53 / (0,27) | 0,36 / (0,12) |

^a L.P. = Lav produktkunnskap M.P. = Middels produktkunnskap H.P. = Høy produktkunnskap

Forskjellene i beta-koeffisienter og forklart varians mellom eksperimentgruppene er relativt store når vi tar hensyn til respondentenes produktkunnskap. Tallene i tabell 8.19 illustrerer funnene fra variansanalysen i tabell 8.18 på en svært god måte. For respondenter med lav produktkunnskap ser vi at spesifikke produktvurderinger foretatt med utgangspunkt i subjektforankret instruksjon og -vurderingsskala forklarer en vesentlig større del av variansen i atferdsintensjoner enn spesifikke vurderinger foretatt med utgangspunkt i objektforankret instruksjon og -vurderingsskala. For respondenter med høy produktkunnskap viser tabellen et motsatt resultat med tilsvarende store forskjeller mellom eksperimentgruppene. For respondenter med middels produktkunnskap er forklart varians og standardiserte beta-koeffisienter identisk i de to eksperimentgruppene.

Variansanalysen og parameterestimatene som ble rapportert i tabell 8.18, samt beta-koeffisientene og forklart varians rapportert i tabell 8.19 leder oss til å konkludere med at det til en viss grad finnes støtte for hypotese 7b i de innsamlete dataene. De observerte forskjellene mellom eksperimentgruppene er i riktig retning for respondenter med lav produktkunnskap. For respondenter med middels produktkunnskap blir forskjellene mellom eksperimentgruppene borte. Dette er også i samsvar med hypotese 7b. Et overraskende funn er det imidlertid at vurderinger av spesifikke produkt-egenskaper foretatt med objektforankret instruksjon og -vurderingsskala, forklarer

atferdsintensjoner bedre enn vurderinger foretatt med subjektforankret instruksjon og -vurderingsskala.

Gjennom testing av hypotese 5a og 5b fant vi støtte for at immaterielle egenskaper ved produktet har større relativ betydning ved kundenes vurderinger, og større evne til å forklare respondentenes totalvurdering og atferdsintensjoner når det benyttes subjektforankret instruksjon og -vurderingsskala. Produktkunnskap kan bidra til å redusere disse effektene. Følgende hypotese ble derfor fremmet i kapittel 4.

H8a: Forskjellen mellom spesifikke produktvurderinger foretatt i spørreskjema med objektforankrete- vs. subjektforankrete instruksjoner og -vurderingsskalaer, når det gjelder den relative betydningen av materielle- og immaterielle egenskaper ved produktet for forklaring av totalvurderinger, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

For testing av hypotese 8a ble det først gjennomført variansanalyser der produktkunnskap ble inkludert som en faktor. Disse er rapportert i tabell 8.20a og tabell 8.20b. I tabell E.11 - tabell E.16 i appendiks E rapporteres i tillegg separate variansanalyser for ulike nivå av produktkunnskap. Parameterestimatene som presenteres i tabell 8.20a og 8.20b er hentet fra disse. På samme måte som tidligere illustreres også her forskjeller mellom eksperimentgruppene ved hjelp av standardiserte beta-koeffisienter og forklart varians.

For testing av hypotese 8a fokuserer vi spesielt på de to interaksjonseffektene i tabell 8.20a. Disse sier noe om hvorvidt betydningen av immaterielle-, og materielle egenskaper er forskjellig ved bruk av spørreskjema der det er benyttet henholdsvis objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap.

Fra tabellen ser vi at det rapporteres en to-halet sannsynlighet på 0,186 for nullhypotesen om at interaksjon mellom instruksjon og vurderingsskala og materielle vurderinger ikke har betydning for nivået på totalvurderingen når vi tar hensyn til respondentenes produktkunnskap. Med en retningsbestemt hypotese, blir den relevante en-halete sannsynligheten 0,093 dersom funnene er i riktig retning. Interaksjonen mellom immaterielle vurderinger og instruksjon og vurderingsskala er ikke signifikant når vi tar hensyn til respondentens produktkunnskap.

TABELL 8.20a

ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegnenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 952,55 | 576 | 1,65 | | |
| Konstant | 0,53 | 1 | 53 | 0,32 | 0,573 |
| Immaterielle vurderinger | 33,93 | 1 | 33,93 | 20,51 | < 0,001 |
| Materielle vurderinger | 250,03 | 1 | 250,03 | 151,19 | < 0,001 |
| Instruksjon/vurderingsskala | 9,49 | 1 | 9,49 | 5,74 | 0,017 ^b |
| Produktkunnskap | 0,10 | 2 | 0,05 | 0,03 | 0,970 ^c |
| Instruksjon/vurderingsskala x produktkunnskap ^a | 16,69 | 2 | 8,34 | 5,05 | 0,007 ^d |
| Immaterielle vurderinger x instruksjon/vurderingsskala within produktkunnskap | 3,00 | 3 | 1,00 | 0,60 | 0,612 ^e |
| Materielle vurderinger x instruksjon/vurderingsskala within produktkunnskap | 7,99 | 3 | 2,66 | 1,61 | 0,186 ^f |
| (Model) | 793,45 | 13 | 61,03 | 36,91 | < 0,001 |
| (Total) | 1745,99 | 589 | 2,96 | | |
| R-squared = | 0,454 | | | | |
| Adjusted R-squared = | 0,442 | | | | |

^a Gruppergjennomsnitt objektforankret totalvurdering:
 Objektforankret instruksjon og -vurderingsskala
 Subjektforankret instruksjon og -vurderingsskala

Lav produktkunnskap: 3,26 (n = 90)
 Middels produktkunnskap: 3,03 (n = 88)
 Høy produktkunnskap: 3,26 (n = 125)
 Lav produktkunnskap: 3,16 (n = 101)
 Middels produktkunnskap: 3,50 (n = 63)
 Høy produktkunnskap: 3,27 (n = 123)

^b p = 0,009 for en-hale test
^c p = 0,485 for en-hale test
^d p = 0,004 for en-hale test
^e p = 0,306 for en-hale test
^f p = 0,093 for en-hale test

| Parameterestimer: | Konstant | Immaterielle vurderinger | Materielle vurderinger | Instruksjon/vurderingsskala | Immaterielle vurd. x instruksjon/vurderingsskala | Materielle vurd. x instruksjon/vurderingsskala |
|-------------------------|----------|--------------------------|------------------------|-----------------------------|--|--|
| Lav produktkunnskap | | | | | | |
| Objektforankret: | 0,119 | 0,280 | 0,672 | 0,274(ns) | -0,091(ns) | 0,030(ns) |
| Subjektforankret: | 0,119 | 0,280 | 0,672 | -0,274(ns) | 0,091(ns) | -0,030(ns) |
| Middels produktkunnskap | | | | | | |
| Objektforankret: | 0,408 | 0,075 | 0,803 | -0,770 | 0,006(ns) | 0,144(ns) |
| Subjektforankret: | 0,408 | 0,075 | 0,803 | 0,770 | -0,006(ns) | -0,144(ns) |
| Høy produktkunnskap | | | | | | |
| Objektforankret: | 0,001 | 0,336 | 0,645 | -0,900 | 0,088(ns) | 0,144 |
| Subjektforankret: | 0,001 | 0,336 | 0,645 | 0,900 | -0,088(ns) | -0,144 |

Parameterestimaten for de to eksperimentgruppene viser at bruk av subjektforankret instruksjon og -vurderingsskala reduserer betydningen av materielle egenskaper for den objektforankrede totalvurderingen for respondenter med høy produktkunnskap. Den relative betydningen av immaterielle egenskaper økes altså for disse respondentene. Det observerte resultatet avviker fra påstanden vi formulerte i hypotese 8a, og indikerer dermed at det ikke finnes støtte for denne hypotesen i dataene når det gjelder objektforankrede totalvurderinger.

Ved å spesifisere signifikansnivå $\alpha = 0,05$ rapporteres verdien for styrke til 0,817 for interaksjonen mellom instruksjon og -vurderingsskala og respondentenes produktkunnskap. Den korresponderende verdien for *eta-square*, som mål på effektstørrelse rapporteres lik 0,017. For den marginalt signifikante to-veis-interaksjonen mellom vurdering av materielle produktegenskaper og instruksjon og -vurderingsskala, når vi kontrollerer for produktkunnskap rapporteres styrke på 0,424, med korresponderende *eta square* lik 0,008. De rapporterte effektstørrelsene må karakteriseres som små, mens verdiene for styrke indikerer over middels statistisk styrke for F-testen for interaksjonen mellom produktkunnskap og instruksjon og -vurderingsskala, og svak styrke for interaksjonen mellom vurdering av materielle produktegenskaper og instruksjon og vurderingsskala, kontrollert for respondentenes produktkunnskap.

Tilsvarende variansanalyse ble gjennomført med hensyn på subjektforankrede totalvurderinger. Fra tabell 8.20b ser vi først at det rapporteres en to-halet sannsynlighet på 0,001 for null-hypotesen om at interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap ikke har betydning for nivået på respondentenes subjektforankrede totalvurderinger.

Videre ser vi at både interaksjonen mellom instruksjon og vurderingsskala og vurdering av immaterielle egenskaper, og interaksjonen mellom instruksjon og vurderingsskala og vurdering av materielle produktegenskaper er signifikant med en-halet p-verdi $< 0,05$ når vi tar hensyn til respondentenes produktkunnskap og funnene er i riktig retning.

Parameterstimaten for de to eksperimentgruppene viser at bruk av subjektforankret instruksjon og -vurderingsskala øker betydningen av immaterielle produktegenskaper for subjektforankrede totalvurderinger for respondenter med lav produktkunnskap. For respondenter med middels- og høy produktkunnskap reduseres betydningen av materielle egenskaper. Den relative betydningen av immaterielle egenskaper økes altså for samtlige respondenter når det benyttes subjektforankret instruksjon og -vurderingsskala i forbindelse med spesifikke vurderingsskåler.

TABELL 8.20b

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 616,88 | 572 | 1,08 | | |
| Konstant | 64,12 | 1 | 64,12 | 59,46 | < 0,001 |
| Immaterielle vurderinger | 17,45 | 1 | 17,45 | 16,18 | 0,001 |
| Materielle vurderinger | 183,79 | 1 | 183,79 | 170,42 | < 0,001 |
| Instruksjon/vurderingsskala | 0,45 | 1 | 0,45 | 0,42 | 0,512 ^b |
| Produktkunnskap | 4,17 | 2 | 2,09 | 1,93 | 0,146 ^c |
| Instruksjon/vurderingsskala x produktkunnskap ^a | 16,46 | 2 | 8,23 | 7,63 | 0,001 ^d |
| Immaterielle vurderinger x instruksjon/vurderingsskala within produktkunnskap | 7,38 | 3 | 2,46 | 2,28 | 0,078 ^e |
| Materielle vurderinger x instruksjon/vurderingsskala within produktkunnskap | 8,73 | 3 | 2,91 | 2,70 | 0,045 ^f |
| (Model) | 561,35 | 13 | 43,18 | 40,04 | < 0,001 |
| (Total) | 1178,23 | 585 | 2,01 | | |
| R-squared = | 0,476 | | | | |
| Adjusted R-squared = | 0,465 | | | | |

^a Gruppegjennomsnitt objektforankret totalvurdering:
Objektforankret instruksjon og -vurderingsskala

| | |
|--------------------------|----------------|
| Lav produktkunnskap: | 3,67 (n = 87) |
| Middels produktkunnskap: | 3,69 (n = 89) |
| Høy produktkunnskap: | 3,54 (n = 125) |
| Lav produktkunnskap: | 3,46 (n = 101) |
| Middels produktkunnskap: | 3,84 (n = 62) |
| Høy produktkunnskap: | 3,48 (n = 122) |

^b p = 0,256 for en-hale test

^c p = 0,073 for en-hale test

^d p < 0,001 for en-hale test

^e p = 0,004 for en-hale test

^f p = 0,023 for en-hale test

| Parameterestimer: | Konstant | Immaterielle vurderinger | Materielle vurderinger | Instruksjon/vurderingsskala | Immaterielle vurd. x instruksjon/vurderingsskala | Materielle vurd. x instruksjon/vurderingsskala |
|-------------------------|----------|--------------------------|------------------------|-----------------------------|--|--|
| Lav produktkunnskap | | | | | | |
| Objektforankret: | 1,004 | 0,333 | 0,445 | 0,597 | -0,184 | 0,051(ns) |
| Subjektforankret: | 1,004 | 0,333 | 0,445 | -0,597 | 0,184 | -0,051(ns) |
| Middels produktkunnskap | | | | | | |
| Objektforankret: | 1,389 | 0,018 | 0,715 | -0,473 | -0,072(ns) | 0,187 |
| Subjektforankret: | 1,389 | 0,018 | 0,715 | 0,473 | 0,072(ns) | -0,187 |
| Høy produktkunnskap | | | | | | |
| Objektforankret: | 0,928 | 0,150 | 0,641 | -0,505 | 0,017(ns) | 0,120 |
| Subjektforankret: | 0,928 | 0,150 | 0,641 | 0,505 | -0,071(ns) | -0,120 |

Ved signifikansnivå $\alpha = 0,05$ rapporteres verdien for styrke til 0,946 for interaksjonen mellom instruksjon og -vurderingsskala og respondentenes produktkunnskap. Den korresponderende *eta-square*, som et mål på effektstørrelse, rapporteres lik 0,026. For to-veis-interaksjonene mellom vurdering av immaterielle og materielle produkt-egenskaper og instruksjon og vurderingsskala når vi kontrollerer for produktkunnskap rapporteres styrke henholdsvis lik 0,575 og 0,657, med korresponderende *eta square* 0,012 og 0,014 .

For ytterligere å illustrere forskjellene mellom eksperimentgruppene rapporteres standardiserte beta-koeffisienter og forklart varians i tabell 8.21. De standardiserte beta-koeffisientene viser det samme mønsteret som ble rapportert i parameterestimaten fra variansanalysene.

TABELL 8.21

Standardiserte beta-koeffisienter og forklart varians for objektforankret- og subjektforankret totalvurdering, betydning av materielle- og immaterielle produkttegenskaper, innenfor ulike nivå av produktkunnskap

| TOTALVURDERING/ INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA / (ADJ. R SQ.) | | |
|---|--------------------------------------|---------------|-------------------|
| | L.P. | M.P. | H.P. ^a |
| Objektforankret totalvurdering/ | | | |
| Objektforankret instruksjon og -vurderingsskala | | | |
| Immaterielle vurderinger | 0,15 | 0,04 | 0,26 |
| Materielle vurderinger | 0,57 / (0,44) | 0,56 / (0,33) | 0,54 / (0,53) |
| Subjektforankret instruksjon og -vurderingsskala | | | |
| Immaterielle vurderinger | 0,32 | 0,05 | 0,20 |
| Materielle vurderinger | 0,52 / (0,60) | 0,69 / (0,51) | 0,40 / (0,31) |
| Subjektforankret totalvurdering/ | | | |
| Objektforankret instruksjon og -vurderingsskala | | | |
| Immaterielle vurderinger | 0,15 | -0,03 | 0,14 |
| Materielle vurderinger | 0,51 / (0,36) | 0,65 / (0,38) | 0,69 / (0,62) |
| Subjektforankret instruksjon og -vurderingsskala | | | |
| Immaterielle vurderinger | 0,48 | 0,09 | 0,12 |
| Materielle vurderinger | 0,35 / (0,58) | 0,71 / (0,58) | 0,49 / (0,33) |

^a L.P. = Lav produktkunnskap M.P. = Middels produktkunnskap H.P. = Høy produktkunnskap

I hypotese 8a har vi fremsatt påstand om at betydningen av instruksjon og vurderingsskala for immaterielle-, og materielle produkttegenskapers evne til å forklare

totalvurderinger reduseres for respondenter med høy produktkunnskap. Basert på resultatene fra hypotesetestingen har vi derfor valgt å konkludere med at hypotese 8a ikke støttes av de innsamlete dataene

Som tidligere gjennomføres tilsvarende analyser for ulike nivå av produktkunnskap der respondentenes atferdsintensjoner betraktes som avhengig variabel. Disse analysene baseres på følgende hypotese:

H8b: *Forskjellen mellom spesifikke produktvurderinger foretatt i spørreskjema med objektforankrete- vs. subjektforankrete instruksjoner og -vurderingsskalaer, når det gjelder den relative betydningen av materielle- og immaterielle egenskaper ved produktet for forklaring av atferdsintensjoner, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap*

I tabell 8.22 rapporteres resultater fra ANOVA der produktkunnskap inkluderes som faktor i analysen. Separate analyser presenteres i tabell E.17 - tabell E.19 i appendiks E.

En to-halet sannsynlighet på 0,001 viser at interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap har betydning for respondentenes atferdsintensjoner. Videre viser tabellen at både interaksjonen mellom instruksjon og vurderingsskala og vurdering av immaterielle egenskaper, og interaksjonen mellom instruksjon og vurderingsskala og vurdering av materielle produkttegenskaper er marginalt signifikant med en-halete p-verdier lik 0,116 når vi tar hensyn til respondentenes produktkunnskap og funnene er i riktig retning.

Ved gjennomføring av separate variansanalyser med atferdsintensjoner som avhengig variabel innenfor ulike nivå av produktkunnskap (appendiks E) rapporteres interaksjon mellom instruksjon og vurderingsskala og immaterielle og materielle produkttegenskaper som mer signifikante enn de som rapporteres i tabell 8.22. Årsaken til dette kan være at analysen der produktkunnskap innkluderes som en forklaringsvariabel gir en høyere kritisk verdi med mindre sannsynlighet for å avsløre signifikante effekter i de innsamlete dataene.

Parameterstimatene for de to eksperimentgruppene viser at bruk av subjektforankret instruksjon og -vurderingsskala øker betydningen av immaterielle produkttegenskaper for atferdsintensjonene for respondenter med lav produktkunnskap.

TABELL 8.22

ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, innenfor ulike nivå av produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 1993,28 | 578 | 3,45 | | |
| Konstant | 1835,21 | 1 | 1835,21 | 532,17 | < 0,001 |
| Immaterielle vurderinger | 72,64 | 1 | 72,64 | 21,06 | < 0,001 |
| Materielle vurderinger | 230,38 | 1 | 230,38 | 66,80 | < 0,001 |
| Instruksjon/vurderingsskala | 5,50 | 1 | 5,50 | 1,60 | 0,207 ^b |
| Produktkunnskap | 5,85 | 2 | 2,92 | 0,85 | 0,429 ^c |
| Instruksjon/vurderingsskala x produktkunnskap ^a | 54,72 | 2 | 27,36 | 7,93 | 0,001 ^d |
| Immaterielle vurderinger x instruksjon/vurderingsskala within produktkunnskap | 14,88 | 3 | 4,96 | 1,44 | 0,231 ^e |
| Materielle vurderinger x instruksjon/vurderingsskala within produktkunnskap | 14,84 | 3 | 4,95 | 1,43 | 0,232 ^f |
| (Model) | 980,32 | 13 | 75,41 | 21,87 | < 0,001 |
| (Total) | 2973,60 | 591 | 5,03 | | |
| R-squared = | 0,330 | | | | |
| Adjusted R-squared = | 0,315 | | | | |

^a Gruppegjennomsnitt objektforankret totalvurdering:
 Objektforankret instruksjon og -vurderingsskala
 Subjektforankret instruksjon og -vurderingsskala

Lav produktkunnskap: 9,14 (n = 90)
 Middels produktkunnskap: 8,99 (n = 88)
 Høy produktkunnskap: 9,17 (n = 127)
 Lav produktkunnskap: 8,94 (n = 101)
 Middels produktkunnskap: 8,96 (n = 63)
 Høy produktkunnskap: 9,18 (n = 123)

^b p = 0,104 for en-hale test
^c p = 0,215 for en-hale test
^d p < 0,001 for en-hale test
^e p = 0,116 for en-hale test
^f p = 0,116 for en-hale test

| Parameterestimater: | Konstant | Immaterielle vurderinger | Materielle vurderinger | Instruksjon/vurderingsskala | Immaterielle vurd. x instruksjon/vurderingsskala | Materielle vurd. x instruksjon/vurderingsskala |
|-------------------------|----------|--------------------------|------------------------|-----------------------------|--|--|
| Lav produktkunnskap | | | | | | |
| Objektforankret: | 5,487 | 0,449 | 0,630 | 0,794 | -0,235 | 0,042(ns) |
| Subjektforankret: | 5,487 | 0,449 | 0,630 | -0,794 | 0,235 | -0,042(ns) |
| Middels produktkunnskap | | | | | | |
| Objektforankret: | 5,090 | 0,505 | 0,636 | -0,436(ns) | 0,129(ns) | -0,025(ns) |
| Subjektforankret: | 5,090 | 0,505 | 0,636 | 0,436(ns) | -0,129(ns) | 0,025(ns) |
| Høy produktkunnskap | | | | | | |
| Objektforankret: | 5,875 | 0,308 | 0,687 | -0,251 | 0,099(ns) | 0,247 |
| Subjektforankret: | 5,875 | 0,308 | 0,687 | 0,251 | -0,099(ns) | -0,247 |

For respondenter med høy produktkunnskap reduseres betydningen av materielle egenskaper, mens forskjellene mellom eksperimentgruppene ikke er signifikante for respondenter med middels produktkunnskap. Den relative betydningen av immaterielle egenskaper for forklaring av respondentenes atferdsintensjoner økes altså både for respondenter med lav produktkunnskap og for respondenter med høy produktkunnskap. Ved signifikansnivå $\alpha = 0,05$ rapporteres verdien for styrke til 0,954 for interaksjonen mellom instruksjon og vurderingsskala og respondentenes produktkunnskap. Den korresponderende *eta-square* rapporteres lik 0,027. Effektstørrelsen er her relativt liten, mens verdien for styrke er høy.

For to-veis-interaksjonene mellom vurdering av immaterielle- og materielle produkt-egenskaper og instruksjon og vurderingsskala, når vi kontrollerer for produktkunnskap rapporteres styrke på 0,381 og 0,380, med tilhørende *eta square* lik 0,007 for begge interaksjonseffektene. De rapporterte effektstørrelsene må karakteriseres som små, og verdiene for styrke indikerer svak statistisk styrke for F-testene når det kontrolleres for respondentenes produktkunnskap.

For oppsummering av resultatene fra variansanalysen i tabell 8.22, rapporteres standardiserte beta-koeffisienter og forklart varians for respondentenes atferdintensjoner innenfor ulike nivå av produktkunnskap i tabell 8.23.

TABELL 8.23

Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner, betydning av materielle- og immaterielle produkttegenskaper innenfor ulike nivå av produktkunnskap

| INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA / (ADJ. R SQ.) | | |
|--|--------------------------------------|---------------|-------------------|
| | L.P. | M.P. | H.P. ^a |
| Objektforankret instruksjon og -vurderingsskala | | | |
| Immaterielle vurderinger | 0,13 | 0,27 | 0,21 |
| Materielle vurderinger | 0,41 / (0,29) | 0,29 / (0,24) | 0,53 / (0,47) |
| Subjektforankret instruksjon og -vurderingsskala | | | |
| Immaterielle vurderinger | 0,44 | 0,17 | 0,13 |
| Materielle vurderinger | 0,35 / (0,54) | 0,41 / (0,27) | 0,27 / (0,13) |

^a L.P. = Lav produktkunnskap M.P. = Middels produktkunnskap H.P. = Høy produktkunnskap

De standardiserte beta-koeffisientene viser det samme mønsteret som ble illustrert ved hjelp av parameterestimaterne i variansanalysen i tabell 8.22. For respondenter med lav produktkunnskap økes betydningen av immaterielle vurderinger for atferdintensjonene.

For respondenter med høy produktkunnskap reduseres betydningen av materielle produktegenskaper. Tabellen viser i tillegg at forklart varians dobles når det benyttes subjektforankret instruksjon og -vurderingsskala i forbindelse med de spesifikke vurderingsspørsmålene for respondenter med lav produktkunnskap, mens forklart varians reduseres til en tredjedel når subjektforankret instruksjon og -vurderingsskala benyttes for respondenter med høy produktkunnskap.

I henhold til hypotese 8b forventer vi at forskjellene mellom de immaterielle-, og materielle egenskapenes evne til å forklare respondentenes atferdsintensjoner reduseres når produktkunnskapen øker. Resultatene som rapporteres i tabell 8.22 og 8.23 gir til en viss grad støtte for denne hypotesen. Forskjellene mellom eksperimentgruppene er i riktig retning for respondenter med lav produktkunnskap, og reduseres for respondenter med middels produktkunnskap. Dette er i samsvar med hypotesen. For respondenter med høy produktkunnskap økes imidlertid forskjellene gjennom den signifikante reduksjonen i betydning av materielle produktegenskaper. Dette kan tolkes som økt relativ betydning av de immaterielle egenskapene, og vi konkluderer med at hypotese 8b sett under ett ikke støttes av de innsamlete dataene.

Gjennom testing av hypotese 6 fant vi ingen forskjeller i respondentenes sikkerhet ved vurdering av spesifikke produktegenskaper basert på hvilken instruksjon og vurderingsskala som ble benyttet i spørreskjemaet. I kapittel 3 argumenterte vi for at sikkerhet ved vurderinger av et objekt vil være avhengig av i hvilken grad et individ har utviklet holdninger og tilhørende vurderinger overfor objektet.

Økt brukserfaring og produktkunnskap kan være forhold som kan bidra til å øke tilgjengeligheten til holdninger og vurderinger. Økt produktkunnskap kan dermed være bestemmende for hvor sikre respondentene vil være i sine vurderinger. Respondenter med høy produktkunnskap antas å være sikrere ved vurderinger av spesifikke egenskaper ved produktet, enn respondenter med lav produktkunnskap. Dette kan moderere effektene av instruksjon og vurderingsskala. Med utgangspunkt i den teoretiske drøftingen i kapittel 3 ble følgende hypoteser formulert i kapittel 4.

H9a: Respondenter med lav produktkunnskap vil være sikrere på sine vurderinger av spesifikke produktegenskaper fore-tatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer enn vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer

H9b: Respondenter med høy produktkunnskap vil være tilnærmet like sikre på sine vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer som vurderinger foretatt i spørreskjema med objektforankrete instruksjoner og -vurderingsskalaer

For testing av hypotese 9a og 9b ble det først foretatt variansanalyse der produktkunnskap inkluderes som en faktor. Resultatene fra denne analysen presenteres i tabell 8.24.

TABELL 8.24

ANOVA, sikkerhet i vurderinger av de 27 spesifikke produktegenskapene foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, avhengig av respondentenes produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|-------|--------------------|
| Within + residual | 264,714 | 570 | 0,464 | | |
| Instruksjon/vurderingsskala | 0,223 | 1 | 0,223 | 0,479 | 0,489 ^b |
| Produktkunnskap | 1,195 | 2 | 0,598 | 1,287 | 0,277 ^c |
| Instruksjon/vurderingsskala x Produktkunnskap ^a | 0,006 | 2 | 0,003 | 0,007 | 0,993 ^d |
| (Model) | 1,424 | 5 | 0,285 | 0,613 | 0,690 |
| (Total) | 266,138 | 575 | 0,463 | | |
| Adjusted R-squared = | 0,295 | | | | |

^a Gruppegjennomsnitt sikkerhet i vurderinger:

Objektforankret instruksjon og -vurderingsskala

Lav produktkunnskap: 4,13 (n = 88)

Middels produktkunnskap: 4,08 (n = 88)

Høy produktkunnskap: 4,19 (n = 124)

Subjektforankret instruksjon og -vurderingsskala

Lav produktkunnskap: 4,16 (n = 95)

Middels produktkunnskap: 4,12 (n = 62)

Høy produktkunnskap: 4,23 (n = 119)

^b p = 0,245 for en-hale test

^c p = 0,139 for en-hale test

^d p = 0,497 for en-hale test

Produktkunnskap kan ha både en direkte effekt på sikkerhet i vurderinger, og en effekt via interaksjonen med instruksjon og vurderingsskala. Testing av hypotese 9a og 9b innebærer at vi fokuserer spesielt på interaksjonen mellom produktkunnskap og instruksjon og vurderingsskala.

Resultatene fra variansanalysen i tabell 8.24 viser at hverken instruksjon og vurderingsskala, produktkunnskap eller interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap har betydning for respondentenes rapportering av sikkerhet i vurderinger når vi betrakter samtlige vurderinger under ett. Gruppegjennomsnittene som rapporteres i fornote a til tabellen gir heller ikke indikasjoner på at produktkunnskap skal ha noen betydning for respondentenes sikkerhet i vurderinger avhengig av hvilken instruksjon og vurderingsskala som er benyttet for å måle vurderingene.

For ytterligere å undersøke mulige effekter av produktkunnskap som spesifisert i hypotese 9a og 9b ble det gjennomført variansanalyser der respondentenes sikkerhet i vurderinger ble gruppert i henhold til samhandlings-, struktur-, backstage- og kundeavhengige produktelementer. I tabell 8.25 rapporteres resultater fra MANOVA for forskjeller mellom eksperimentgruppene.

Tabell 8.25 viser at produktkunnskap direkte bidrar til forskjeller for respondentenes rapportering av sikkerhet ved vurderinger. Den to-halete p-verdien på 0,009 indikerer at forskjellene er svært signifikante. For den direkte effekten av instruksjon og vurderingsskala, og for interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap rapporteres ikke signifikante forskjeller mellom eksperimentgruppene for noen av de avhengige variablene. Som tidligere nevnt er det kun interaksjonen mellom instruksjon og vurderingsskala og respondentenes produktkunnskap som er relevant for testing av hypotese 9a og 9b. Interaksjonseffekten er ikke signifikant selv når det benyttes en-hale test, som gir en p-verdi på 0,420.

TABELL 8.25

MANOVA, sikkerhet i vurdering av produktelementer foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, avhengig av respondentenes produktkunnskap. Rapportering av Pillais-T

| SOURCE OF VARIATION | VALUE | EXACT F | HYP. DF | ERROR F | SIGN. OF F |
|---|---------|------------|------------|------------|---------------|
| Instruksjon/vurderingsskala Pillais-T | 0,00221 | 0,301 | 4,00 | 545,00 | 0,877 |
| Produktkunnskap Pillais-T | 0,03794 | 2,547 | 8,00 | 1092,00 | 0,009 |
| Instruksjon/vurderingsskala x produktkunnskap Pillais-T | 0,00762 | 0,522 | 8,00 | 1092,00 | 0,840 |

Basert på resultatene som rapporteres i tabell 8.25 kan vi konkludere med at hypotese 9a og 9b heller ikke finner støtte i de innsamlede datene når vi grupperer de 27 produkt-

egenskapene i henhold til de fire produktelementene samhandlings-, struktur-, back-stage- og kundeavhengige produktelementer.

Testing av hypotesene der vi predikerer forskjeller mellom eksperimentgruppene når det gjelder repondentenes sikkerhet i vurderinger (H6, H9a og H9b), leder oss til å konkludere med at hverken instruksjon og vurderingsskala, eller interaksjonen mellom instruksjon og vurderingsskala og produktkunnskap fører til forskjeller i respondentenes sikkerhet i vurdering av forskjellige egenskaper ved hotellproduktet.

8.2 Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet

I dette avsnittet presenteres resultater fra analyser av forskjeller mellom eksperimentgrupper som henholdsvis er tilordnet spørreskjema uten viktighetsspørsmål, spørreskjema med viktighetsspørsmål før spesifikke vurderingss spørsmål og spørreskjema med viktighetsspørsmål etter spesifikke vurderingss spørsmål. Først ser vi nærmere på hvilke forskjeller tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet gir opphav til for respondentenes rapportering av viktighet, og for nivået på de spesifikke vurderingene. I tillegg fokuserer vi på sammenhenger mellom viktighetsspørsmålene og vurderingss spørsmålene. Deretter undersøker vi forskjeller knyttet til de spesifikke vurderingenes evne til å forklare respondentenes totalvurderinger og atferdsintensjoner avhengig av om viktighetsspørsmål er inkludert i spørreskjemaet, og i tilfelle hvor disse er plassert i forhold til de spesifikke vurderingss spørsmålene.

8.2.1 Forskjeller i nivå på spesifikke vurderingss spørsmål

I kapittel 3 diskuterte vi hvordan *salience* effekter kan sees i sammenheng med assimilasjons- og kontrasteffekter. Assimilasjonseffekter oppstår når svargivning på etterfølgende spørsmål bringes nærmere svargivning på forutgående spørsmål enn de ville gjort dersom de forutgående spørsmålene hadde vært utelatt fra spørreskjemaet. Assimilasjonseffekter er på denne måten et resultat av at respondenten ønsker å være konsistent i sin svargivning. I motsetning til assimilasjonseffekter kan kontrasteffekter føre til større forskjeller i svargivning som resultat av spørsmålsrekkefølgen. Når viktighetsspørsmål plasseres før de spesifikke vurderingene i spørreskjemaet, kan det tenkes

at viktighetsspørsmålene som et resultat av kontrast- og assimilasjonseffekter vil påvirke vurderingene. Det motsatte kan være tilfellet dersom vurdering av spesifikke produktegenskaper foretas før man tillegger egenskapene vekt gjennom svar på viktighetsspørsmålene. Assimilasjons- og kontrasteffekter kan i første rekke påvirke nivået både på viktighetsspørsmål og vurderingss spørsmål.

Basert på litteraturgjennomgangen og teordiskusjonen i kapittel 3 var vi ikke i stand til å formulere generelle hypoteser i tilknytning til nivå på vurderingss spørsmål og viktighetsspørsmål, avhengig av om viktighetsspørsmål blir inkludert i spørreskjemaet, og i tilfelle hvor viktighetsspørsmålene plasseres i forhold til vurderingss spørsmålene. Kun ved å skille mellom ulike kvalitetsnivåer på produktet kunne vi fremsette hypoteser i tilknytning til nivået på vurderingss spørsmålene.

Et alternativ til svake hypoteser, der vi hevder at det vil være forskjeller mellom eksperimentgrupper, er å formulere eksplorative forskningsspørsmål, som belyses med de innsamlete datene. Selv om vi ikke har formulert generelle hypoteser, vil vi derfor i det følgende rapportere resultater fra analyser der nivå på vurderingss spørsmål og viktighetsspørsmål sammenlignes mellom eksperimentgruppene. Resultatene fra disse analysene, som på mange måter kan karakteriseres som eksplorative, vil danne grunnlag for videre diskusjon i kapittel 9.

I figur 8.5 illustreres først nivå på de spesifikke vurderingene av hotellproduktet, foretatt ved bruk av spørreskjema henholdsvis uten viktighetsspørsmål, med viktighetsspørsmål før vurderingss spørsmål og med viktighetsspørsmål plassert etter vurderingss spørsmålene.

På X-aksen i figur 8.5 finner vi de 27 forskjellige aspektene ved hotellproduktet som respondentene ble bedt om å vurdere. Rekkefølgen disse er presentert i samsvarer med gruppering av egenskapene i de fire produktelementene samhandlings-, struktur-, backstage- og kundeavhengige produktelementer.

Y-aksen i figuren viser nivået på vurderingene. Måleskalaen som ble benyttet i spørreskjemaet har verdier fra -5 til 5. I figuren har vi som tidligere avgrenset måleskalaen til å gjelde området vurderingene varierer innenfor, slik at eventuelle forskjeller mellom eksperimentgruppene fremkommer på en klarere måte.

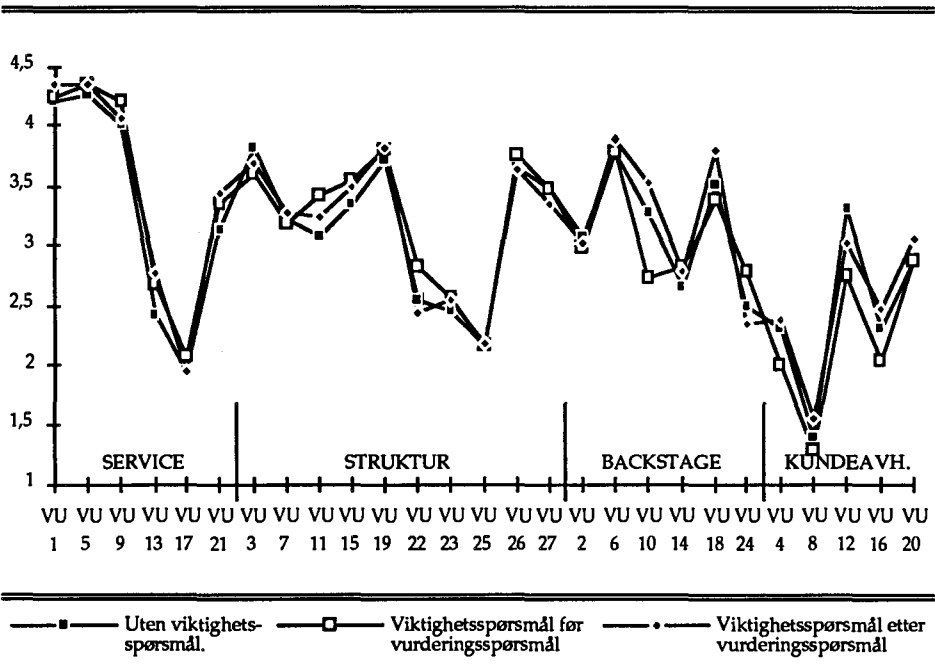
Visuell inspeksjon av figur 8.5 indikerer små, eller ingen systematiske forskjeller mellom eksperimentgruppene når det gjelder gjennomsnittsverdier for de individuelle vurderingene. Basert på en slik første overflate inspeksjon ledes vi derfor til å

konkludere med at nivå på vurderinger ikke er forskjellig avhengig av inkludering og alternativ plassering av viktighets spørsmål i spørreskjemaet.

En nærmere inspeksjon av figuren indikerer heller ikke tilstedeværelse av noe spesielt mønster når det gjelder forskjeller mellom eksperimentgruppene. Av 27 spesifikke vurderinger er kun én vurdering signifikant forskjellig mellom eksperimentgruppene med $\alpha < 0,05$ (VU10). For de resterende vurderingene er gruppegjennomsnittene tilnærmet identiske for de to eksperimentgruppene.

FIGUR 8.5

Nivå på vurderinger foretatt i spørreskjema uten viktighets spørsmål, med viktighets spørsmål før spesifikke vurderingsspørsmål og med viktighets spørsmål etter spesifikke vurderingsspørsmål



Et problem ved bruk av individuelle variabler som grunnlag for sammenligning mellom eksperimentgruppene, er som tidligere nevnt, tilstedeværelse av tilfeldig feilvarians i hver individuell variabel. Dette problemet kan reduseres ved å kalkulere gjennomsnitt på tvers av alle 27 vurderinger for de to eksperimentgruppene, og sammenligne disse.

Resultat fra variansanalyse der disse sammenligningene er foretatt, er presentert i tabell 8.26. Tabellen viser at forskjellene mellom eksperimentgruppene heller ikke her kan betegnes som statistisk signifikante.

TABELL 8.26

ANOVA, nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|-------|------------|
| Within + residual | 819,801 | 605 | 1,355 | | |
| Tilstedeværelse og plassering av viktighetsspørsmål ^a | 0,363 | 2 | 0,181 | 0,134 | 0,875 |
| (Model) | 0,363 | 2 | 0,181 | 0,134 | 0,875 |
| (Total) | 820,164 | 607 | 1,351 | | |

^a Gruppegjennomsnitt nivå på vurderinger:

| | | |
|---|---|----------------|
| Uten viktighetsspørsmål | : | 4,13 (n = 220) |
| Viktighetsspørsmål før vurderingsspørsmål | : | 3,15 (n = 204) |
| Viktighetsspørsmål etter vurderingsspørsmål | : | 3,19 (n = 184) |

En tredje måte å sammenligne nivå på vurderinger mellom de tre eksperimentgruppene er å gruppere de 27 vurderingene i samhandlings-, struktur-, backstage- og kunde-avhengige produktenskaper.

Som grunnlag for å bestemme hvorvidt gjennomsnittsvurderingene er forskjellige mellom eksperimentgruppene, ble det gjennomført MANOVA der vurderinger gruppert i henhold til de fire produktelementene ble betraktet som avhengige variabler. Resultater fra denne variansanalysen er vist i tabell 8.27a

TABELL 8.27a

MANOVA, nivå på vurderinger gruppert i produktelementer, foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. Rapportering av Pillais-T

| SOURCE OF VARIATION | VALUE | EXACT F | HYP. DF | ERROR F | SIGN. OF F |
|---|---------|---------|---------|---------|------------|
| Tilstedeværelse og plassering av viktighetsspørsmål | | | | | |
| Pillais-T | 0,02923 | 2,192 | 8,00 | 1182,00 | 0,026 |

Tabell 8.27a viser at det finnes forskjell i nivå på vurderinger når disse grupperes i henhold til de fire produktelementene. Tabellen viser ikke for hvilke produktelementer og hvordan eksperimentgruppene er forskjellige. For å få svar på dette presenteres univariate forskjeller mellom eksperimentgruppene i tabell 8.27b, som viser at forskjellene mellom eksperimentgruppene er signifikante for vurderinger av kunde-avhengige produktenskaper.

TABELL 8.27b

MANOVA, nivå på vurderinger gruppert i produktelementer foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. Univariate F-tester

| SOURCE OF VARIANCE/ VARIABLE | HYP. SS | ERROR SS | HYP. MS | ERROR MS | F | SIGN. OF F |
|---|------------|-------------|------------|-------------|-------|---------------|
| Tilstedeværelse og plassering av viktighetsspørsmål | | | | | | |
| Samhandling/service vurd. | 2,106 | 917,058 | 1,053 | 1,546 | 0,681 | 0,506 |
| Struktur vurd. | 1,128 | 1013,815 | 0,564 | 1,710 | 0,330 | 0,719 |
| Backstage vurd. | 1,152 | 118,335 | 0,576 | 1,886 | 0,305 | 0,737 |
| Kundeavhengig vurd. ^a | 14,735 | 1470,185 | 7,367 | 2,479 | 2,972 | 0,052 |

^a Gruppergjennomsnitt kundeavhengige produktegenskaper:
 Uten viktighetsspørsmål : 2,54 (n = 220)
 Viktighetsspørsmål før vurderingsspørsmål : 2,21 (n = 204)
 Viktighetsspørsmål etter vurderingsspørsmål : 2,55 (n = 184)

Fra fotnote a til tabellen ser vi at nivået på vurderinger av de kundeavhengige egenkapene er lavest når viktighetsspørsmål inkluderes før vurderingsspørsmålene i spørreskjemaet. Nivået er identisk for vurderinger foretatt i spørreskjema uten viktighetsspørsmål og i spørreskjema med viktighetsspørsmål etter vurderingsspørsmål. For de øvrige produktelementene finner vi ikke forskjeller mellom eksperimentgruppene.

For å oppsummere resultatene for analysene av nivå på spesifikke produktvurderinger velger vi å konkludere med at dataene sett under ett, til tross for funnene som rapporteres i tabell 8.27a og 8.27b for de kundeavhengige produktegenskapene, ikke i tilstrekkelig grad indikerer forskjeller mellom eksperimentgruppene basert på tilstedeværelse og ulik plassering av viktighetsspørsmål i spørreskjemaet.

8.2.2 Forskjeller i nivå på spesifikke vurderingsspørsmål - kvalitetsnivået på produktene som vurderes som modererende variabel

I forrige avsnitt fant vi ingen systematiske forskjeller i nivå på spesifikke vurderingsspørsmål mellom eksperimentgruppene som henholdsvis har benyttet spørreskjema uten viktighetsspørsmål, spørreskjema med viktighetsspørsmål før de spesifikke vurderingsspørsmålene og spørreskjema med viktighetsspørsmål etter de spesifikke vurderingsspørsmålene.

I kapittel 3 argumenterte vi imidlertid for at det kan finnes situasjoner der gjennomsnittsvurderingene kan være forskjellige. På samme måte som for hypotesene i

tilknytning til nivå på vurderinger avhengig av hvilken instruksjon og vurderingsskala som benyttes, skilte vi også mellom ulike kvalitetsnivåer på produktene ved drøfting av assimilasjons- og kontrasteffekter for nivået på vurderinger avhengig av om viktighets-spørsmål inkluderes, og i tilfelle hvor disse plasseres i forhold til de spesifikke vurderingene. Med dette som utgangspunkt kan formulerte vi følgende hypoteser:

H10a: For lavkvalitetsprodukter vil nivå på vurderinger av spesifikke produktegenskaper, foretatt i spørreskjema der viktighetsspørsmål er utelatt, overstige nivå på vurderinger av spesifikke produktegenskaper foretatt i spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene (kontrasteffekt)

H10b: For produkter med kvalitet lik gjennomsnittet, vil nivå på vurderinger av spesifikke produktegenskaper, foretatt i spørreskjema der viktighetsspørsmål er utelatt, ikke avvike fra nivå på vurderinger av spesifikke produktegenskaper foretatt i spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene

H10c: For høykvalitetsprodukter vil nivå på vurderinger av spesifikke produktegenskaper, foretatt i spørreskjema der viktighetsspørsmål er inkludert og plassert før de spesifikke vurderingsspørsmålene, overstige eller være tilnærmet lik nivå på vurderinger av spesifikke produktegenskaper foretatt i spørreskjema der viktighetsspørsmål er utelatt (assimilasjonseffekt)

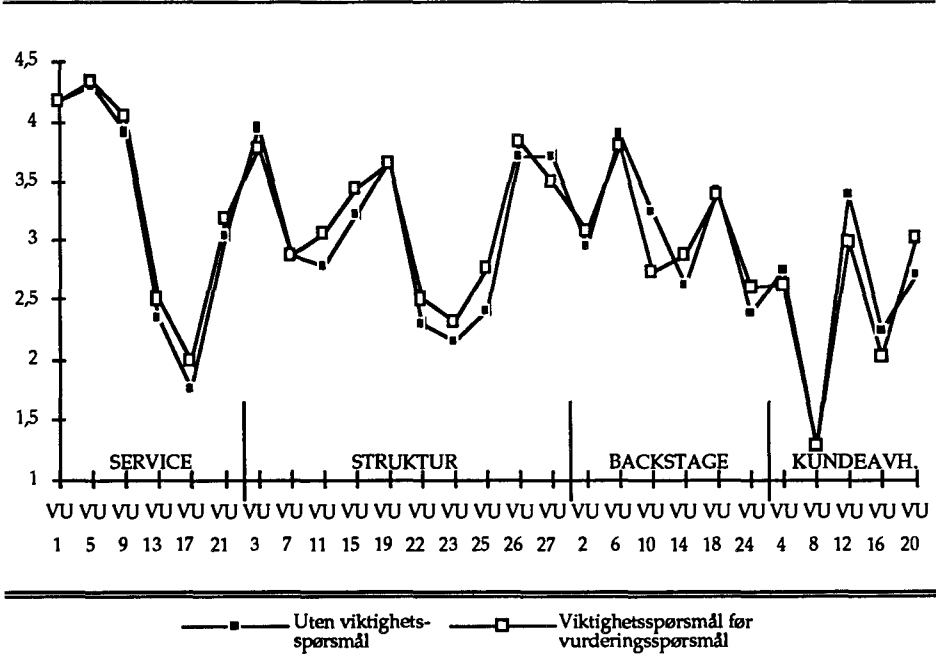
Som grunnlag for å teste disse hypotesene ble de 11 hotellene som deltar i studien delt inn i tre grupper basert på respondentenes totalvurderinger¹¹¹. Tre hotell plasseres på denne måten i gruppen *lavkvalitetshotell*, fem hotell plasseres i gruppen *gjennomsnittshotell* og tre hotell plasseres i gruppen *høykvalitetshotell*. En måte å teste hypotesene er å inspisere forskjeller i gjennomsnitt for de individuelle vurderingsspørsmålene. I figur 8.6a, 8.6b og 8.6c nedenfor oppsummeres resultater for disse

¹¹¹Se kapittel 6 for operasjonalisering av totalvurdering av hotellene. En alternativ måte å gruppere hotellene i ulike kvalitetskategorier ville ha vært å benytte "ekspertvurderinger". Se også analysene i forbindelse med hypotese 3a, 3b og 3c.

analysene i diagrammer der gruppegjennomsnitt sammenlignes for hotellene gruppert i henhold til de tre kvalitetsnivåene. I disse figurene fokuserer vi i henhold til hypoteser 10a, 10b og 10c, kun på forskjeller i svargivning mellom respondenter som har benyttet spørreskjema uten viktighetsspørsmål og respondenter som har benyttet spørreskjema med viktighetsspørsmål plassert før spesifikke vurderingsspørsmål.

Dersom respondentene foretar en sekvensiell utfylling av spørreskjemaet, antar vi at svargivning på spesifikke vurderingsspørsmål når disse er plassert før viktighets-spørsmålene, ikke vil avvike fra svargivning på spesifikke vurderingsspørsmål i et spørreskjema der viktighetsspørsmålene er utelatt. Det finnes imidlertid studier der det argumenteres for at utfylling av et spørreskjema i en situasjon der respondenten er overlatt til seg selv ikke nødvendigvis foregår sekvensielt. Dersom dette er tilfelle kan det tenkes at viktighetsspørsmålene kan påvirke svargivning på de spesifikke vurderingsspørsmålene selv om viktighetsspørsmålene er plassert bak etter disse i spørreskjemaet. For å kontrollere for dette har vi i appendiks E presentert figurer der eksperimentgruppen som har benyttet spørreskjema der viktighetsspørsmålene er plassert etter de spesifikke vurderingsspørsmålene er inkludert.

FIGUR 8.6a
Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål og med viktighets-spørsmål før spesifikke vurderingsspørsmål - lavkvalitetshotell



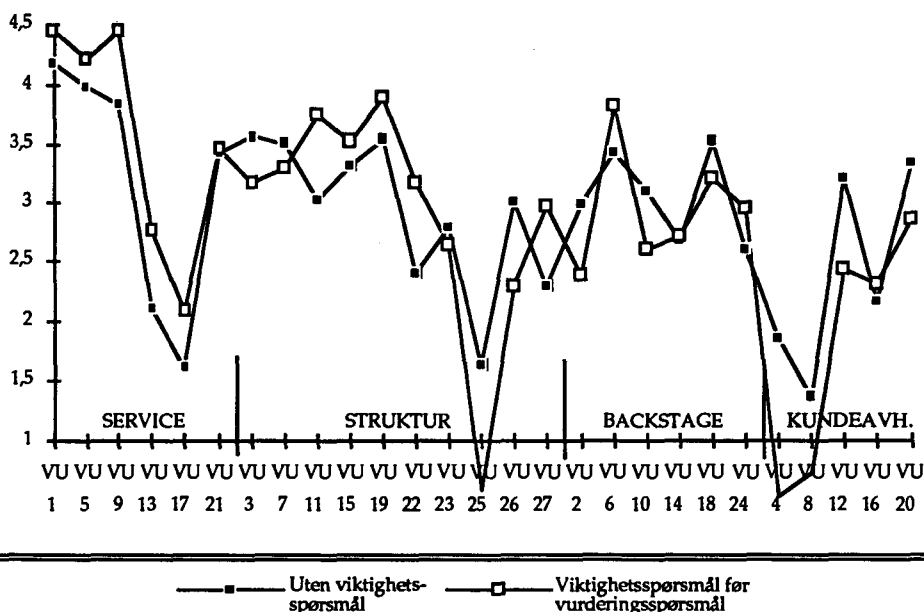
I figur 8.6a rapporteres forskjeller i gruppegjennomsnitt for lavkvalitetshotellene. På X-aksen i figuren finner vi de 27 forskjellige aspektene ved hotellproduktet som respondentene ble bedt om å vurdere. Rekkefølgen disse er presentert i samsvarer med gruppering av egenskapene i de fire produktelementene samhandlings-, struktur-, backstage- og kundeavhengige produkttegniskaper¹¹². Y-aksen i figur 8.6a representerer nivå på vurderingene.

Ved inspeksjon av figur 8.6a ser vi at kun 2 av 27 vurderinger, i henhold til hypotese 10a, er i riktig retning (VU8 og VU10). I tillegg er bare en av disse forskjellene (VU10) statistisk signifikant med $\alpha < 0,05$ (en-hale T-test). På bakgrunn av denne figuren og analysene velger vi å konkludere med at hypotese 10a ikke støttes av de innsamlete dataene.

For hypotese 10b og 10c ble det gjennomført en liknende øksersis som for hypotese 10a. Resultatene for hypotese 10b presenteres i figur 8.6b.

FIGUR 8.6b

Nivå på vurderinger foretatt i spørreskjema uten viktighetspørsmål og med viktighetspørsmål for spesifikke vurderingspørsmål - hotell med gjennomsnittskvalitet



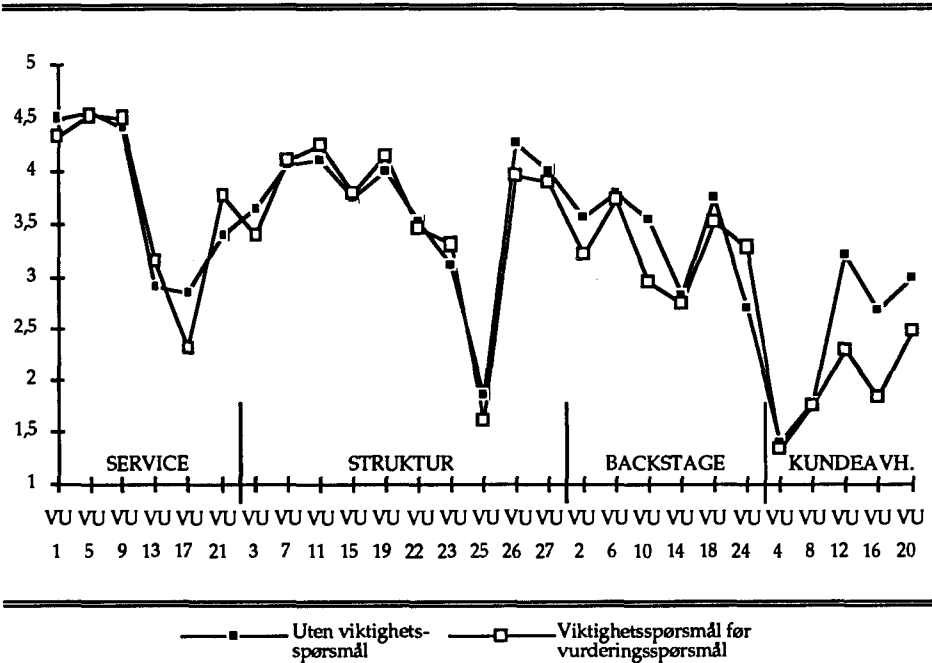
¹¹²Se spørreskjema i appendix B

Inspeksjon av figur 8.6b viser at det for kun 2 av 27 gruppegjennomsnitt er fullstendig overlapp mellom de to eksperimentgruppene (VU14 og VU21) når vi fokuserer på hotell med gjennomsnittskvalitet. Imidlertid viser T-tester at det kun er ett gruppegjennomsnitt som er signifikant forskjellig (to-halet T-test og $\alpha < 0,05$) mellom eksperimentgruppene (VU4). Til tross for at de observerte forskjellene for hotell med gjennomsnittskvalitet er relativt store, velger vi på bakgrunn av signifikanstestene å konkludere med støtte for hypotese 10b.

Figur 8.6c viser forskjeller i gruppegjennomsnitt for høykvalitetshotellene. Her har vi i hypotese 10c, hevdet at det skal være høyest gjennomsnittsvurderinger i eksperimentgruppen som har benyttet spørreskjema der viktighets spørsmål er plassert før de spesifikke vurderingsspørsmålene. I figuren observerer vi at kun 7 av 27 vurderinger i følge hypotesen er forskjellige i riktig retning (VU7, VU9, VU13, VU11, VU21, VU23 og VU24). Imidlertid kan det merkes at ingen av disse forskjellene er statistisk signifikante. Hypotese 10c støttes dermed heller ikke av de innsamlede dataene.

Med bakgrunn i analysene ovenfor er det kun hypotese 10b som finner støtte i de innsamlede dataene. Dette til tross for at det er her vi finner de største forskjellene.

FIGUR 8.6c
 Nivå på vurderinger foretatt i spørreskjema uten viktighets spørsmål og med viktighets spørsmål før spesifikke vurderingsspørsmål - høykvalitetshotell



Vi har tidligere argumentert for at hypotese 10a, 10b og 10c henger sammen resonnementmessig. Dette innebærer at det vil være mest riktig å betrakte dem samlet når vi diskuterer resultatene. Sett i relasjon til resultatene som ble rapportert for hypotese 10a, 10b og 10c, sett under ett, velger vi derfor, så langt, å konkludere med at hypotesene som helhet ikke støttes av de innsamlede dataene.

En alternativ måte å teste hypotesene ovenfor er å rapportere gruppegjennomsnitt for samtlige 27 vurderingsvariabler, og for vurderinger gruppert i samhandlings-, struktur-, backstage- og kundeavhengige produktegenskaper. Forskjeller i gruppegjennomsnitt mellom eksperimentgruppene for de ulike kvalitetskategoriene er presentert i tabell 8.28.

TABELL 8.28

Gruppegjennomsnitt for vurderinger foretatt i spørreskjema uten viktighetsspørsmål og med viktighetsspørsmål før spesifikke vurderingsspørsmål, for hotellene i studien gruppert etter kvalitetsnivå

| SKALA | TILSTEDEVÆRELSE OG PLASSERING AV VIKTIGHETSSPØRSMÅL | KVALITET | | |
|-------------------------|---|-----------------|----------------------|----------------------|
| | | UNDER GJ.SN. | GJSN. | OVER GJ.SN. |
| Samtlige 27 vurderinger | Uten viktighetsspørsmål | 3,08 | 2,98 | 3,45 |
| | Viktighetsspørsmål før vurderingsspørsmål | 3,12 | 3,00 | 3,36 |
| Samhandling/service | Uten viktighetsspørsmål | 3,39 | 3,27 | 3,83 |
| | Viktighetsspørsmål før vurderingsspørsmål | 3,50 | ^b 3,71 | 3,85 |
| Struktur | Uten viktighetsspørsmål | 3,10 | 2,95 | 3,68 |
| | Viktighetsspørsmål før vurderingsspørsmål | 3,22 | 3,08 | 3,64 |
| Backstage | Uten viktighetsspørsmål | 3,12 | 3,03 | 3,41 |
| | Viktighetsspørsmål før vurderingsspørsmål | 3,11 | 3,01 | 3,33 |
| Kundeavhengig | Uten viktighetsspørsmål | 2,55 | 2,55 | 2,49 |
| | Viktighetsspørsmål før vurderingsspørsmål | 2,42 | ^a 1,76 | ^b 1,98 |

^a Forskjellen er statistisk signifikant med $\alpha < 0,05$ ^b Forskjellen er statistisk signifikant med $\alpha < 0,10$

I følge resultatene i tabell 8.28, er det kun for hypotese 10b vi isolert sett til en viss grad kan hevde at de innsamlede dataene gir støtte. Her antok vi at det ikke ville være forskjell i gruppegjennomsnittene for hotell med gjennomsnittskvalitet. Selv om det for

samhandlings- og kundeavhengige produkttegenskaper blir rapportert statistisk signifikante forskjeller, velger vi å akseptere at dataene sett under ett gir støtte for hypotese 10b.

Siden hypotese 10a, 10b og 10c henger sammen resonnementmessig, velger vi også her å konkludere med at settet av hypoteser som helhet er falsifisert. Det finnes altså ikke grunnlag for å hevde at de forventete forskjeller i gjennomsnitt for spesifikke vurderinger, foretatt henholdsvis i spørreskjema uten viktighetsspørsmål og i spørreskjema der viktighetsspørsmål er plassert før de spesifikke vurderingsspørsmålene, for produkter med ulikt kvalitetsnivå, finner støtte i de innsamlete datene.

I appendiks E har vi i presentert figurer der også eksperimentgruppen som har benyttet spørreskjema der viktighetsspørsmålene er plassert etter de spesifikke vurderingsspørsmålene er inkludert. Dersom det er slik at respondenter leser og svarer på spørsmålene i et spørreskjema i den rekkefølgen de er presentert i spørreskjemaet, antar vi at svargivning på de spesifikke vurderingsspørsmålene i vår studie vil være tilnærmet identisk for respondenter som har benyttet spørreskjema der viktighetsspørsmålene er utelatt, og for respondenter som har benyttet spørreskjema der viktighetsspørsmålene er plassert etter de spesifikke vurderingsspørsmålene.

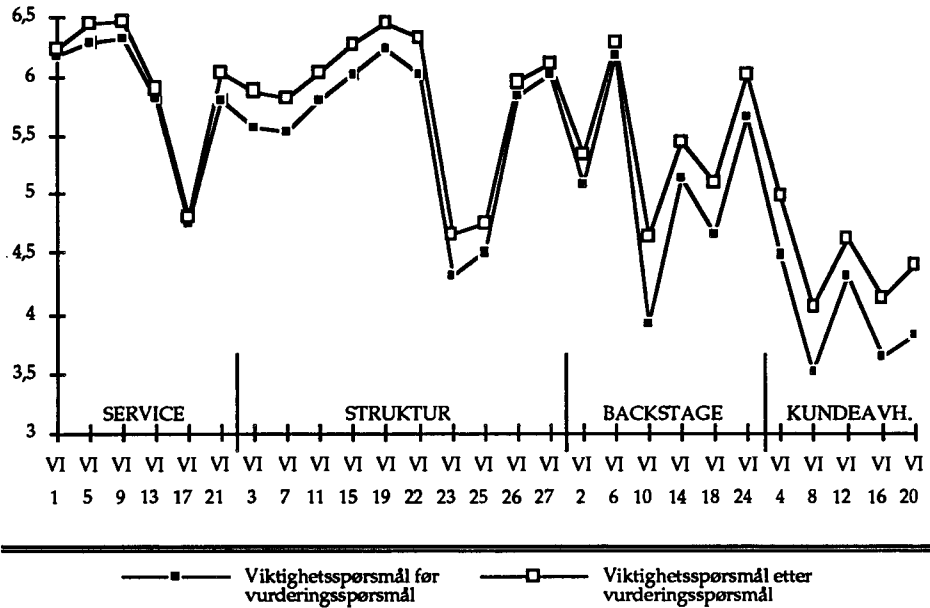
Figurene som presenteres i appendiks E viser at dette ikke er tilfellet. For samtlige kvalitetskategorier er de observerte forskjellene i nivå på de spesifikke vurderingsspørsmålene relativt store mellom respondenter som har benyttet spørreskjema der viktighetsspørsmålene henholdsvis er utelatt, eller plassert etter de spesifikke vurderingsspørsmålene. Disse resultatene kan betraktes som støtte til tidligere studier der det hevdes at betydningen av rekkefølgen på spørsmålene som stilles i en intervjuundersøkelse, med sikkerhet kun kan studeres når man har kontroll over rekkefølgen spørsmålene stilles og besvares. Dette betinger at intervjuet gjennomføres som et personlig intervju eller som et telefonintervju.

8.2.3 Forskjeller i nivå på viktighetsspørsmål

I figur 8.7 illustreres respondentenes rapportering av viktighet for de 27 spesifikke vurderingene av hotellene, foretatt ved bruk av spørreskjema, henholdsvis uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter de spesifikke vurderingsspørsmålene.

FIGUR 8.7

Nivå på viktighetsspørsmål plassert henholdsvis før-, og etter spesifikke vurderings-spørsmål i spørreskjemaet



Inspeksjon av figur 8.7 indikerer at rapportert viktighet for de individuelle vurderingene gjennomgående ser ut til å være høyest i eksperimentgruppen med viktighetsspørsmål før spesifikke vurderingsspørsmål i spørreskjemaet. Av 27 spesifikke egenskaper er rapportert viktighet høyest i denne eksperimentgruppen for 26 egenskaper. For den siste vurderingen er gruppegjennomsnittene tilnærmet identiske for de to eksperimentgruppene. De observerte forskjellene er statistisk signifikante med $\alpha < 0,05$ for 17 egenskaper, mens tre av forskjellene er statistisk signifikante med $\alpha < 0,10$.

En alternativ måte å analysere forskjeller i viktighet for de individuelle vurderingene er å betrakte vurderingene som en tilnærmet binomisk serie. Med andre ord kan vi stille spørsmål om sannsynligheten for at minst 26 av 27 vurderinger er høyest i eksperimentgruppen med viktighetsspørsmål før spesifikke vurderingsspørsmål av ren tilfeldighet. Resultatene fra en ikke-parametrisk binomisk test gir en to-halet sannsynlighet på mindre enn 0,001 for at det observerte mønsteret er tilfeldig dersom de to eksperimentgruppene er trukket fra samme populasjon¹¹³. Disse resultatene indikerer sterk retnings-

¹¹³Et problem med å benytte en ikke-parametrisk binomisk test på denne måten, er som tidligere at mange av vurderingsvariablene er korrelert. Dette er et klart brudd på forutsetningen om uavhengighet i en binomisk serie.

støtte for høyere nivå på viktighetsspørsmål når disse plasseres før de spesifikke vurderingsspørsmålene.

En tredje måte å analysere forskjeller ved rapportering av viktighet mellom de to eksperimentgruppene, er å gruppere respondentenes rapportering av viktighet for de 27 vurderingene ved konstruksjon av summerte skalaer.

TABELL 8.29a

MANOVA, nivå på viktighet gruppert i produktelementer for vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. Rapportering av Pillais-T

| SOURCE OF VARIATION | VALUE | EXACT F | HYP. DF | ERROR F | SIGN. OF F |
|---|---------|------------|------------|------------|---------------|
| Tilstedeværelse og plassering av viktighetsspørsmål | | | | | |
| Pillais-T | 0,06633 | 6,60671 | 4,00 | 372,00 | > 0,001 |

Tabell 8.29a viser at det finnes forskjell i nivå på viktighet når produktegenskapene grupperes i henhold til de fire produktelementene. Tabellen viser ikke for hvilke produktelementer, og hvordan eksperimentgruppene er forskjellige. For å få svar på dette presenteres univariate tester for forskjeller mellom eksperimentgruppene i tabell 8.29b

TABELL 8.29b

MANOVA, nivå på viktighet gruppert i produktelementer for vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. Univariate F-tester

| SOURCE OF VARIANCE/ VARIABLE | HYP. SS | ERROR SS | HYP. MS | ERROR MS | F | SIGN. OF F |
|--|------------|-------------|------------|-------------|--------|---------------|
| Tilstedeværelse og plassering av viktighetsspørsmål ^a | | | | | | |
| Samhandling | 1,281 | 224,109 | 1,281 | 0,598 | 2,144 | 0,144 |
| Struktur | 4,933 | 168,362 | 4,933 | 0,449 | 10,988 | 0,001 |
| Backstage | 11,764 | 268,342 | 11,764 | 0,716 | 16,441 | > 0,001 |
| Kundeavhengig | 21,211 | 428,871 | 21,211 | 1,144 | 18,634 | > 0,001 |

^a Gruppegjennomsnitt viktighet :

| | | |
|---|---|----------------|
| Samhandling | : | 5,86 (n = 203) |
| Viktighetsspørsmål før vurderingsspørsmål | : | 5,98 (n = 178) |
| Viktighetsspørsmål etter vurderingsspørsmål | : | |
| Struktur | : | 5,59 (n = 203) |
| Viktighetsspørsmål før vurderingsspørsmål | : | 5,82 (n = 177) |
| Viktighetsspørsmål etter vurderingsspørsmål | : | |
| Backstage | : | 5,11 (n = 203) |
| Viktighetsspørsmål før vurderingsspørsmål | : | 5,47 (n = 177) |
| Viktighetsspørsmål etter vurderingsspørsmål | : | |
| Kundeavhengig | : | 3,98 (n = 203) |
| Viktighetsspørsmål før vurderingsspørsmål | : | 4,45 (n = 175) |
| Viktighetsspørsmål etter vurderingsspørsmål | : | |

Alle de summerte skalaene viser høyest nivå for viktighet i eksperimentgruppen med viktighetsspørsmål etter vurderingsspørsmålene. For samhandlingsegenskapene er forskjellene imidlertid ikke statistisk signifikante med $\alpha < 0,10$. For de øvrige produkt-elementene er forskjellene statistisk signifikant med $\alpha < 0,001$. Basert på analysene ovenfor har vi derfor valgt å konkludere med at det finnes systematiske forskjeller mellom eksperimentgruppene når det gjelder nivå på viktighetsspørsmålene.

8.2.4 Sammenheng mellom svargivning på viktighetsspørsmål og vurderingsspørsmål

I analysene vi har rapportert ovenfor, har vi så langt kun fokusert på nivå på viktighets-spørsmål og vurderingsspørsmål. Dersom vi sammenholder resultatene som illustreres i figur 8.5 for nivået på vurderingsspørsmålene med nivået på viktighetsspørsmålene slik dette er illustrert i figur 8.7, fremkommer et interessant mønster. I gjennomsnitt blir de produktsegenskapene som vurderes mest positivt, også rapportert til å være de viktigste for respondentene når disse skal vurdere et hotell. Dette ser ut til å gjelde uavhengig av plasseringen av viktighetsspørsmålene og vurderingsspørsmålene. Disse analysene representerer imidlertid aggregerte størrelser, og gjør oss ikke i stand til å si noe om hvordan rekkefølgen på spørsmålene kan påvirke vurderingsprosessene på individnivå.

I kapittel 3 diskuterte vi hvordan en respondents svargivning på et spørsmål i et spørreskjema kan virke retningsgivende for svargivning på et etterfølgende spørsmål. Her ble korrelasjon mellom spørsmålene fremhevet som indikator på diagnose-verdien¹¹⁴ ved spørsmålet som stilles først. Dersom spørsmålene har symmetrisk diagnoseverdi vil ikke korrelasjonen mellom spørsmålene påvirkes av rekkefølgen. Dersom spørsmålene har asymmetrisk diagnoseverdi vil korrelasjonen mellom spørsmålene påvirkes av rekkefølgen i spørreskjemaet. Høyere korrelasjon mellom spørsmålene indikerer at spørsmålet som er stillt først har høyere diagnoseverdi. Basert på denne teoridrøftelsen formulerte vi i kapittel 4 følgende hypotese.

H11: Sammenhengen mellom viktighet og vurderinger vil være sterkere når viktighetsspørsmål er plassert før spesifikke vurderingsspørsmål i spørreskjemaet, enn når viktighets-spørsmål er plassert etter de spesifikke vurderings-spørsmålene

¹¹⁴Diagnosticity

For å teste denne hypotesen ble korrelasjoner mellom viktighetsspørsmål og vurderings- spørsmål først beregnet separat for hver respondent. På denne måten finner vi et uttrykk for sammenhengen mellom viktighetsspørsmål og vurderingsspørsmål på individnivå. Respondentene ble deretter gruppert i henhold til eksperimentgruppene som ble til- ordnet spørreskjema der viktighetsspørsmål er inkludert og plassert henholdsvis før-, og etter de spesifikke vurderingsspørsmålene.

For sammenligning mellom eksperimentgruppene ble først de individuelle korrelasjons- koeffisientene Z-transformert¹¹⁵. Deretter ble gjennomsnittlig korrelasjon mellom viktighet og vurderinger beregnet. Prosedyren for testing av hypotese 11, blir med utgangspunkt i dette å teste om gjennomsnittlig korrelasjon er større i den ene eksperimentgruppen enn den andre. Dette innebærer at vi kan benytte vanlige T-tester for å teste om de antatte forskjellene er signifikante. I tabell 8.30 rapporteres resultater fra T-test for forskjeller i gjennomsnittlige Z-transformerte korrelasjoner mellom viktighetsspørsmål og vurderingsspørsmål avhengig av rekkefølgen på spørsmålene.

TABELL 8.30

T-test, gjennomsnittskorrelasjon mellom viktighetsspørsmål og vurderingsspørsmål foretatt i spørreskjema med viktighetsspørsmål før vs. etter spesifikke vurderingsspørsmål.

| VARIABLE | | NUMBER OF | MEAN | STANDARD | SE OF | | |
|-------------------------------------|--------|----------------------------|------------|------------------------------|---------|------------|--------|
| | | CASES | | DEVIATION | MEAN | | |
| Gjennomsnittskorrelasjon | | | | | | | |
| Viktighetsspm. før vurderingsspm. | | 123 | 0,1635 | 0,135 | 0,012 | | |
| Viktighetsspm. etter vurderingsspm. | | 95 | 0,1303 | 0,136 | 0,014 | | |
| | | * POOLED VARIANCE ESTIMATE | | * SEPARATE VARIANCE ESTIMATE | | | |
| F | 2-TAIL | * T | DEGREES OF | 2-TAIL | * T | DEGREES OF | 2-TAIL |
| VALUE | PROB | * VALUE | FREEDOM | PROB | * VALUE | FREEDOM | PROB |
| 0,020 | 0,887 | * 1,79 | 216 | 0,074 | * 1,79 | 201,24 | 0,075 |

Resultatene fra T-testen i tabell 8.30 viser at Z-transformert gjennomsnittskorrelasjon mellom viktighetsspørsmål og vurderingsspørsmål er størst i eksperimentgruppen som er tilordnet spørreskjema der viktighetsspørsmålene er plassert før vurderings- spørsmålene. I henhold til hypotese 11 er forskjellen mellom eksperimentgruppene i riktig retning.

¹¹⁵ $Z = 0,5 \log (1+r/1-r)$, der r = korrelasjonskoeffisienten

Standardavvikene i de to eksperimentgruppene er tilnærmet identiske. Signifikans-testing kan dermed gjennomføres ved å benytte "pooled variance estimate". For forskjellene mellom eksperimentgruppene rapporteres det her en tosidig p-verdi lik 0,074. Hypotese 11 er imidlertid retningsbestemt. Den relevante p-verdien blir dermed 0,037 som indikerer at forskjellen mellom eksperimentgruppene er signifikant. Basert på disse analysene kan det konkluderes med støtte for hypotese 11.

8.2.5 Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner

Vi har tidligere argumentert for at en respondents tilgjengelighet til vurderinger av ulike produktegenskaper vil øke etterhvert som egenskapene blir mer fremtredende. Forutgående spørsmål regnes ofte som den viktigste kilde til aktivering av informasjon som er relevant for å svare på etterfølgende spørsmål. Den psykologiske prosessen som oppstår i slike tilfeller kalles *priming* eller fremkalling av informasjon. Fremkalling kan bidra til aktivering av ulike typer informasjon som kan være relevant for respondentens svargivning. Når kundens tilgjengelighet til vurderingene øker, f.eks. gjennom svar på tidligere spørsmål i spørreskjemaet argumenterte vi for at sammenhengene mellom vurdering av egenskaper, totalvurdering og atferdsintensjoner vil forsterkes. Effektene av spørsmålsrekkefølge for de spesifikke vurderingsspørsmålenes prediksjonsevne ble i kapittel 4 presisert i følgende hypotese:

H12a: Målinger der viktighetsspørsmål er plassert før spesifikke vurderingsspørsmål i spørreskjemaet vil gi bedre forklaring på totalvurderinger enn målinger der viktighetsspørsmålene er plassert etter de spesifikke vurderingsspørsmålene

Som beskrevet i kapittel 6 er totalvurderingene i spørreskjemaet operasjonalisert med to objektforankrete totalvurderingsspørsmål, og ett subjektforankret totalvurderingsspørsmål. For testing av hypotesene i tilknytning til effekter av valg av instruksjon og vurderingsskala i spørreskjemaet ble de objektforankrete- og subjektforankrete totalvurderingsspørsmålene behandlet som separate avhengige variabler. Når vi nå skal se nærmere på de spesifikke vurderingenes evne til å forklare totalvurderingene avhengig av tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet, skiller vi ikke mellom hvilken instruksjon og vurderingsskala som er benyttet i de ulike eksperiment-

gruppene. Totalvurdering vil derfor i analysene som følger operasjonaliseres som gjennomsnitt av både objektforankrete- og subjektforankrete totalvurderinger¹¹⁶.

I kapittel 3 presenterte vi i figur 3.5, en modell der totalvurdering av et produkt, og intensjoner om gjenkjøp og *word of mouth*, kan betraktes som en transformasjon av kundens vektlegging og vurdering av oppfattede produktegenskaper. I forbindelse med analysene av prediktiv validitet knyttet til de spesifikke produktegenskapene, avhengig av hvilken instruksjon og vurderingsskala som er benyttet, valgte vi å se bort fra viktighetskomponenten i modellen. Når vi nå skal undersøke prediktiv validitet for vurderinger av spesifikke produktegenskaper foretatt i spørreskjema uten viktighets-spørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighets-spørsmål etter de spesifikke vurderingsspørsmålene, inkluderes viktighetsspørsmålene i den uavhengige variabelen. Dette innebærer at vi beregner veiet gjennomsnitt av samtlige vurderingsspørsmål for respondentene som har benyttet spørreskjema der viktighetsspørsmål er inkludert. Vurderingsmodellen som benyttes som grunnlag for testing av hypotesene kan for disse respondentene dermed uttrykkes matematisk på følgende måte:

$$TV_p = \sum_{i=1}^n E_i V_i$$

der TV_p er totalvurderingen av produktet; E_i er vurdering av oppfattet egenskap i ved produkt p ; V_i er viktigheten av egenskap i ; og n er antall fremtredende egenskaper. For analysene har vi altså benyttet veiet gjennomsnitt for vurderingene som uavhengig variabel. For respondenter som har benyttet spørreskjema der viktighetsspørsmål ikke er inkludert benyttes gjennomsnittsverdier for vurderinger av samtlige produktegenskaper som uavhengig variabel. De samme gjennomsnittsverdiene benyttes som grunnlag for testing av hypotesene der atferdsintensjoner betraktes som avhengig variabel.

I kapittel 7 slo vi fast at forutsetningen om homogenitet i varians mellom de to eksperimentgruppene er oppfylt for respondentenes totalvurdering, også når denne operasjonaliseres som beskrevet ovenfor ($p > 0,10$ for både Cochrans C og Bartlett-Box F). Hypotese 12a kan derfor testes gjennom bruk av univariat variansanalyse.

Fra tabell 8.31 ser vi først hvordan respondentenes svargivning i forbindelse med de spesifikke vurderingsspørsmålene bidrar til å forklare totalvurdering av hotellene. Den

¹¹⁶Analysene ble også gjennomført separat for stimulusforankrete- og subjektforankrete totalvurderinger. Resultatene fra disse analysene fraviker ikke fra analysene der gjennomsnittet av disse benyttes som avhengig variabel.

rapporterte p-verdien, på mindre enn 0,001, indikerer at de 27 spesifikke spørsmålene har prediktiv validitet.

TABELL 8.31

ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 841,25 | 592 | 1,42 | | |
| Konstant | 40,29 | 1 | 40,29 | 28,35 | < 0,001 |
| Samtlige 27 vurderinger | 557,57 | 1 | 557,57 | 392,37 | < 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål ^a | 14,62 | 2 | 7,31 | 5,14 | 0,006 ^b |
| Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetsspørsmål | 8,09 | 2 | 4,05 | 2,85 | 0,059 ^c |
| (Model) | 577,51 | 5 | 115,50 | 81,28 | < 0,001 |
| (Total) | 1418,77 | 597 | 2,38 | | |
| R-squared = | 0,407 | | | | |
| Adjusted R-squared = | 0,402 | | | | |

^a Gruppergjennomsnitt totalvurdering:
 Uten viktighetsspørsmål : 3,23 (n = 220)
 Viktighetsspørsmål før vurderingsspørsmål : 3,36 (n = 199)
 Viktighetsspørsmål etter vurderingsspørsmål : 3,50 (n = 179)

^b p = ,003 for en-hale test

^c p = ,030 for en-hale test

| Parameterestimer: | Konstant | Samtlige 27 vurderinger | Tilstedeværelse og plassering av viktighetsspørsmål | Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetsspørsmål |
|---------------------------|----------|-------------------------|---|---|
| Uten viktighetsspørsmål: | 0,749 | 0,846 | - 0,513 | 0,113 |
| Viktighetsspørsmål før: | 0,749 | 0,846 | - 0,055 | 0,017 |
| Viktighetsspørsmål etter: | 0,749 | 0,846 | 0,568 | - 0,130 |

I henhold til hypotese 12a ønsker vi å studere forskjeller i svargivning mellom respondenter som har benyttet spørreskjema med viktighetsspørsmål før spesifikke vurderingsspørsmål, og med viktighetsspørsmål etter spesifikke vurderingsspørsmål. I variansanalysene har vi i tillegg inkludert eksperimentgruppen som er tilordnet spørreskjema uten viktighetsspørsmål. Vi opererer altså med en faktor som vi har kalt tilstedeværelse og plassering av viktighetsspørsmål som består av tre eksperimentgrupper. Tilstedeværelse og plassering av viktighetsspørsmål kan på denne måten ha en direkte effekt for nivået på totalvurderingen. I tillegg kan de spesifikke produktvurderingenes evne til å forklare nivå på totalvurderingen modereres av tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet.

For nullhypotesen om at det ikke finnes forskjell i nivået på totalvurdering mellom de tre eksperimentgruppene viser ANOVA tabellen en F-verdi på 5,14 med en korresponderende to-halet sannsynlighet på 0,006. Forskjellene i nivå på totalvurdering i de to eksperimentgruppene kan dermed med liten sannsynlighet påstås å skyldes tilfeldigheter.

I fotnote a i tabellen ser vi hvordan nivået på respondentenes totalvurderingen øker fra 3,23 til 3,36 når viktighetsspørsmål inkluderes, og plasseres før spesifikke vurderingss spørsmål i spørreskjemaet. For spørreskjema der viktighetsspørsmålene flyttes bak de spesifikke vurderingss spørsmålene, økes totalvurderingen ytterligere til 3,50. Det å inkludere viktighetsspørsmål i et spørreskjema fører altså direkte til at nivået på respondentenes totalvurdering øker jo nærmere totalvurderingen viktighetsspørsmålene plasseres.

Interaksjonen mellom inkludering og plassering av viktighetsspørsmål i spørreskjemaet og svargivningen på de spesifikke vurderingss spørsmålene er rapportert som det tredje leddet i variansanalysen i tabell 8.30. Fra tabellen ser vi at det rapporteres en to-halet sannsynlighet på 0,059 for nullhypotesen om at denne interaksjonen ikke har betydning for totalvurderingen. Med en hypotese som spesifiserer at plassering av viktighetsspørsmål før spesifikke vurderingss spørsmål i spørreskjemaet vil øke de 27 produktegenskapenes evne til å forklare totalvurderingen har vi en retningsbestemt hypotese. Den relevante en-halets sannsynligheten blir dermed 0,030 dersom funnene er i riktig retning.

Ved inspeksjon av parameterestimatene fremgår det at den direkte effekten av å inkludere viktighetsspørsmål i spørreskjemaet, som vist ved gruppegjennomsnittene i fotnote a, er positiv, mens interaksjonseffekten mellom samtlige 27 vurderinger og inkludering og plassering av viktighetsspørsmål er liten og negativ når viktighets spørsmålene plasseres før de spesifikke vurderingss spørsmålene. Når viktighetsspørsmålene plasseres etter de spesifikke vurderingss spørsmålene reduseres de 27 produkt egenskapenes evne til å forklare totalvurderingen ytterligere. Dette er i samsvar med hypotese 12a.

For å undersøke betydningen av de observerte forskjellene ble det også for disse analysene gjennomført en styrketest. Ved signifikansnivå $\alpha = 0,05$ rapporteres verdiene for styrke ved de statistiske testene rapporteres til 0,824 for den direkte effekten av inkludering og alternativ og plassering av viktighetsspørsmål, og 0,558 for interaksjons effekten mellom de spesifikke vurderingene og tilstedeværelse og plassering av viktighetsspørsmål. De korresponderende *eta-square*, som et mål på effektstørrelse,

rapporteres lik 0,017 og 0,010. Som vi ser er de rapporterte effektstørrelsene små, mens verdiene for styrke indikerer middels statistisk styrke for F-testene.

For ytterligere å illustrere resultatene fra variansanalysen rapporteres standardiserte beta-koeffisienter og forklart varians for respondentenes totalvurdering i tabell 8.32 nedenfor.

TABELL 8.32

Standardiserte beta-koeffisienter og forklart varians for totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål for spesifikke vurderingss spørsmål og med viktighetsspørsmål etter spesifikke vurderingss spørsmål.

| TILSTEDEVÆRELSE OG Plassering av Viktighetsspørsmål | Standardisert Beta (Adj. R SQ.) | |
|--|---------------------------------|--------|
| Uten viktighetsspørsmål | 0,68 | (0,46) |
| Viktighetsspørsmål før vurderingss spørsmål | 0,65 | (0,42) |
| Viktighetsspørsmål etter vurderingss spørsmål | 0,55 | (0,30) |

Tabell 8.32 viser det samme mønsteret som fremkommer i variansanalysen i tabell 8.29. Spesifikke produktvurderinger foretatt i spørreskjema uten viktighetsspørsmål forklarer mest respondentenes totalvurdering best. Dersom viktighetsspørsmål inkluderes i spørreskjemaet reduseres forklart varians og standardisert beta. Reduksjonen i forklart varians er imidlertid størst dersom viktighetsspørsmålene plasseres etter de spesifikke vurderingss spørsmålene i spørreskjemaet. Basert på resultatene fra analysene ovenfor har vi valgt å konkludere med støtte for hypotese 12a i de innsamlede dataene.

Betydningen av å inkludere og variere plassering av viktighetsspørsmål i spørreskjemaet for respondentenes atferdsintensjoner ble i kapittel 4 uttrykt i hypoteseform på følgende måte.

H12b: Målinger der viktighetsspørsmål er plassert før spesifikke vurderingss spørsmål i spørreskjemaet vil gi bedre forklaring på atferdsintensjoner enn målinger der viktighets spørsmålene er plassert etter de spesifikke vurderingss spørsmålene

I kapittel 7 konkluderte vi med at forutsetningen om homogenitet i varians mellom de tre eksperimentgruppene er oppfylt også for respondentenes atferdsintensjoner ($p > 0,10$)

for både Cochrans C og Bartlett-Box F). Hypotese 12b kan derfor testes gjennom bruk av univariat variansanalyse.

For nullhypotesen om at det ikke finnes forskjell i nivå på atferdsintensjoner mellom de tre eksperimentgruppene, viser variansanalysen i tabell 8.33 en F-verdi på 1,99 med en korresponderende to-halet sannsynlighet på 0,138. Forskjellene i nivå på totalvurdering mellom eksperimentgruppene er dermed så små at vi ikke kan se bort fra at de kan skyldes tilfeldigheter.

TABELL 8.33

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 2188,02 | 591 | 3,70 | | |
| Konstant | 2565,54 | 1 | 2565,54 | 692,97 | < 0,001 |
| Samtlige 27 vurderinger | 780,71 | 1 | 780,71 | 210,87 | < 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål ^a | 14,72 | 2 | 7,36 | 1,99 | 0,138 ^b |
| Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetsspørsmål | 10,41 | 2 | 5,21 | 1,41 | 0,246 ^c |
| (Model) | 794,95 | 5 | 158,99 | 42,94 | < 0,001 |
| (Total) | 2982,98 | 596 | 5,00 | | |
| R-squared = | 0,266 | | | | |
| Adjusted R-squared = | 0,260 | | | | |

^a Gruppegjennomsnitt atferdsintensjoner:

| | | |
|---|---|----------------|
| Uten viktighetsspørsmål | : | 9,05 (n = 220) |
| Viktighetsspørsmål før vurderingsspørsmål | : | 9,03 (n = 199) |
| Viktighetsspørsmål etter vurderingsspørsmål | : | 9,20 (n = 178) |

^b p = 0,069 for en-hale test

^c p = 0,123 for en-hale test

| Parameterestimer: | Konstant | Samtlige 27 vurderinger | Tilstedeværelse og plassering av viktighetssp. | Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetssp. |
|---------------------------|----------|-------------------------|--|--|
| Uten viktighetsspørsmål: | 5,989 | 1,002 | - 0,231 | 0,049(ns) |
| Viktighetsspørsmål før: | 5,989 | 1,002 | - 0,404 | 0,112(ns) |
| Viktighetsspørsmål etter: | 5,989 | 1,002 | 0,635 | - 0,163(ns) |

Interaksjonen mellom tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet og svargivningen på de 27 spesifikke vurderingsspørsmålene er heller ikke signifikant. Fra tabellen ser vi at det rapporteres en to-halet sannsynlighet på 0,246 for nullhypotesen om at denne interaksjonen ikke har betydning for totalvurderingen. Den

relevante en-halete sannsynligheten blir dermed 0,123 dersom funnene er i riktig retning. Resultatene fra variansanalysen som rapporteres i tabell 8.33 indikerer at forskjellene mellom eksperimentgruppene ikke er signifikante.

For ytterligere å illustrere resultatene fra variansanalysen ovenfor, rapporteres også her standardiserte beta-koeffisienter og forklart varians for respondentenes totalvurdering i tabell 8.34 nedenfor.

TABELL 8.34

Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål.

| TILSTEDEVÆRELSE OG Plassering av Viktighetsspørsmål | Standardisert beta (Adj. R SQ.) | |
|--|---------------------------------|--------|
| Uten viktighetsspørsmål | 0,51 | (0,26) |
| Viktighetsspørsmål før vurderingsspørsmål | 0,55 | (0,29) |
| Viktighetsspørsmål etter vurderingsspørsmål | 0,48 | (0,23) |

Selv om tabell 8.34 indikerer en viss økning i forklart varians ved inkludering av viktighetsspørsmål før de spesifikke vurderingene i spørreskjemaet, leder rapportering av de standardiserte beta-koeffisientene og forklart varians til de samme konklusjonene som variansanalysen i tabell 8.33. Hypotese 12b finner altså ikke støtte i de innsamlede dataene.

8.2.6 Sammenheng mellom svargivning på viktighetsspørsmål og vurderingsspørsmål - respondentenes produktkunnskap som modererende variabel

En respondents tilgjengelighet (Fazio, 1989) til vurderinger av spesifikke egenskaper antas å øke etterhvert som egenskapene blir mer fremtredende. Respondenter med høy produktkunnskap vil dermed ha vurderinger som er mer tilgjengelige enn det som er tilfellet for kunder med lav produktkunnskap. Påvirkningen fra tidligere spørsmål i et spørreskjema antas å være mindre for disse respondentene.

I kapittel 3 argumenterte vi for at diagnoseverdien av viktighetsspørsmålene for vurderingene, representert ved sammenhengen mellom viktighet og vurderinger vil være sterkere for respondenter med lav produktkunnskap enn for respondenter med høy

produktkunnskap. Med bakgrunn i dette fremsatte vi følgende hypotese vedrørende betydningen av respondentenes produktkunnskap når det gjelder plassering av viktighets spørsmål og spesifikke vurderingsspørsmål i et spørreskjema:

H13: *Forskjellen mellom målinger der viktighetsspørsmål er plassert før, vs. målinger der viktighetsspørsmål er plassert etter spesifikke vurderingsspørsmål i et spørreskjema, når det gjelder sammenhengen mellom viktighetsspørsmål og vurderingsspørsmål vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap*

I henhold til hypotese 13 er vi interessert i å sammenligne forskjeller mellom eksperimentgruppene når det gjelder sammenhengen mellom viktighet og vurderinger for respondenter med henholdsvis lav og høy produktkunnskap. For testing av hypotese 13 ble respondentene innenfor hver eksperimentgruppe derfor gruppert i henhold til høy og lav produktkunnskap. Deretter ble det foretatt T-tester for forskjeller mellom eksperimentgruppene. I tabell 8.35, presenteres først T-test for forskjeller i gjennomsnittlig Z-transformert korrelasjon mellom viktighetsspørsmål og vurderingsspørsmål for respondenter med lav produktkunnskap.

TABELL 8.35

T-test, gjennomsnittskorrelasjon mellom viktighetsspørsmål og spesifikke vurderingsspørsmål foretatt i spørreskjema med viktighetsspørsmål før vs. etter spesifikke vurderingsspørsmål. Respondenter med lav produktkunnskap.

| VARIABLE | | NUMBER OF CASES | MEAN | STANDARD DEVIATION | SE OF MEAN | | |
|---|-------------|-----------------|--------------------|--------------------|------------|--------------------|-------------|
| Gjennomsnittskorrelasjon | | | | | | | |
| Viktighetsspm. før vurderingsspm. | | 37 | 0,1737 | 0,124 | 0,020 | | |
| Viktighetsspm. etter vurderingsspm. | | 23 | 0,1020 | 0,126 | 0,026 | | |
| * POOLED VARIANCE ESTIMATE * SEPARATE VARIANCE ESTIMATE * | | | | | | | |
| F VALUE | 2-TAIL PROB | T VALUE | DEGREES OF FREEDOM | 2-TAIL PROB | T VALUE | DEGREES OF FREEDOM | 2-TAIL PROB |
| 0,292 | 0,591 | 2,16 | 58 | 0,035 | 2,15 | 46,16 | 0,037 |

Resultatene som presenteres i tabell 3.35 støtter hypotese 13 for respondenter med lav produktkunnskap. T-testen viser at Z-transformert gjennomsnittskorrelasjon mellom

respondenter med høy produktkunnskap, ikke påvirkes av plasseringen av viktighets-spørsmålene i spørreskjemaet.

8.2.7 Forskjeller i spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner - respondentenes produktkunnskap som modererende variabel

Gjennom testing av hypotese 12a og 12b tidligere i dette kapittelet fant vi at tilstedeværelse og plassering av viktighetsspørsmål har betydning for respondentenes totalvurderinger, men ikke for rapportering av atferdsintensjoner. Vi forklarte disse effektene med en økning i tilgjengelighet til vurderingene. Gjennom svar på tidligere spørsmål i spørreskjemaet, vil fremkallingseffekter kunne oppstå og sammenhengene mellom vurderinger av spesifikke produktegenskaper, totalvurderinger og atferdsintensjoner forsterkes. Respondentenes produktkunnskap kan opptre som modererende variabel i forbindelse med inkludering av-, og rekkefølge på viktighetsspørsmål. Respondenter med høy produktkunnskap antar vi vil ha vurderinger som er mer tilgjengelige enn det som er tilfellet for respondenter med lav produktkunnskap. Dette kan føre til at sammenhengen mellom vurdering av spesifikke egenskaper, totalvurdering og atferdsintensjoner vil være sterkere for kunder med høy produktkunnskap enn for kunder med lav produktkunnskap. Disse betraktningene ledet oss til å formulere følgende hypotese i kapittel 4:

H14a: Forskjellen mellom målinger der viktighetsspørsmål er plassert før, vs. målinger der viktighetsspørsmål er plassert etter spesifikke vurderingssørsmål i et spørreskjema, når det gjelder de spesifikke vurderingssørsmålenes evne til å forklare totalvurderinger, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

For testing av hypotese 14a ble det gjennomført variansanalyse der produktkunnskap ble inkludert som en faktor i analysen. Variansanalysen er rapportert i tabell 8.37 på neste side. I tabell E.20 - tabell E.22 i appendiks E rapporteres i tillegg separate variansanalyser for ulike nivå av produktkunnskap. Parameterestimaten i tabell 8.37 er hentet fra disse. På samme måte som tidligere illustreres også her forskjeller mellom eksperimentgruppene ved hjelp av standardiserte beta-koeffisienter og forklart varians.

TABELL 8.37

ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingss spørsmål og med viktighetsspørsmål etter spesifikke vurderingss spørsmål, innenfor ulike nivå av produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 826,59 | 578 | 1,43 | | |
| Konstant | 40,20 | 1 | 40,20 | 28,11 | < 0,001 |
| Samtlige 27 vurderinger | 540,17 | 1 | 540,17 | 377,71 | < 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 7,23 | 2 | 3,62 | 2,53 | 0,081 ^b |
| Produktkunnskap | 0,43 | 2 | 0,21 | 0,15 | 0,861 ^c |
| Tilstedeværelse og plassering av viktighetsspørsmål x produktkunnskap ^a | 12,95 | 4 | 3,24 | 2,26 | 0,061 ^d |
| Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetsspørsmål within produktkunnskap | 18,82 | 6 | 3,14 | 2,19 | 0,042 ^e |
| (Model) | 590,27 | 15 | 39,35 | 27,52 | < 0,001 |
| (Total) | 1416,86 | 593 | 2,39 | | |
| R-squared = | 0,417 | | | | |
| Adjusted R-squared = | 0,401 | | | | |

^a Gruppegjennomsnitt totalvurdering:
Uten viktighetsspørsmål

Viktighetsspørsmål før vurderingss spørsmål

Viktighetsspørsmål etter vurderingss spørsmål

| | |
|--------------------------|---------------|
| Lav produktkunnskap: | 3,12 (n = 74) |
| Middels produktkunnskap: | 3,44 (n = 48) |
| Høy produktkunnskap: | 3,18 (n = 94) |
| Lav produktkunnskap: | 3,44 (n = 61) |
| Middels produktkunnskap: | 3,45 (n = 56) |
| Høy produktkunnskap: | 3,23 (n = 82) |
| lav produktkunnskap: | 3,47 (n = 57) |
| Middels produktkunnskap: | 3,33 (n = 48) |
| Høy produktkunnskap: | 3,63 (n = 74) |

^b p = 0,041 for en-hale test

^c p = 0,431 for en-hale test

^d p = 0,031 for en-hale test

^e p = 0,021 for en-hale test

| Parameterestimer: | Konstant | Samtlige 27 vurderinger | Tilstedeværelse og plassering av viktighetsspørsmål | Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetsspørsmål |
|---------------------------|----------|-------------------------|---|---|
| Lav produktkunnskap | | | | |
| Uten viktighetsspørsmål: | 1,001 | 0,774 | -0,859 | 0,222 |
| Viktighetsspørsmål før: | 1,001 | 0,774 | 0,200 | -0,061 |
| Viktighetsspørsmål etter: | 1,001 | 0,774 | 0,659 | -0,161 |
| Middels produktkunnskap | | | | |
| Uten viktighetsspørsmål: | 0,560 | 0,914 | 0,037(ns) | -0,037(ns) |
| Viktighetsspørsmål før: | 0,560 | 0,914 | 0,353(ns) | -0,098(ns) |
| Viktighetsspørsmål etter: | 0,560 | 0,914 | -0,390(ns) | 0,135(ns) |
| Høy produktkunnskap | | | | |
| Uten viktighetsspørsmål: | 0,677 | 0,861 | -0,506 | 0,095 |
| Viktighetsspørsmål før: | 0,677 | 0,861 | -0,498 | 0,147 |
| Viktighetsspørsmål etter: | 0,677 | 0,861 | 1,015 | -0,242 |

Interaksjonen mellom de spesifikke vurderingene og tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet er det relevante leddet i variansanalysen for testing av hypotese 14a. Dette leddet sier noe om hvorvidt de spesifikke produktegenskapenes evne til å forklare respondentenes totalvurdering er forskjellig ved bruk av spørreskjema der tilstedeværelse og plassering av viktighetsspørsmål manipuleres.

Fra tabellen ser vi først at det rapporteres en to-halet sannsynlighet på 0,081 for null hypotesen om at tilstedeværelse og plassering av viktighetsspørsmål ikke har betydning for nivået på totalvurderingen. Disse resultatene indikerer at effekten av å inkludere og variere plasseringen av viktighetsspørsmål i spørreskjemaet er marginalt signifikant.

Videre rapporteres en to-halet sannsynlighet på 0,042 for nullhypotesen om at interaksjonen mellom tilstedeværelse og plassering av viktighetsspørsmål og spesifikke vurderinger ikke har betydning for nivået på totalvurderingen når vi tar hensyn til respondentenes produktkunnskap. Den relevante en-halete p-verdien er her 0,021 dersom funnene er i riktig retning.

Parameterstimatene for de tre eksperimentgruppene viser at inkludering av viktighets-spørsmål før de spesifikke vurderingsspørsmålene i spørreskjemaet reduserer de spesifikke vurderingenes evne til å forklare totalvurderingen for respondenter med lav produktkunnskap. Dersom viktighetsspørsmål plasseres etter spesifikke vurderings-spørsmål viser parameterestimaten at dette ytterligere reduserer de spesifikke vurderingenes evne til å forklare totalvurderingen. For respondenter med middels produktkunnskap rapporteres ingen signifikante effekter av tilstedeværelse og plassering av viktighetsspørsmål.

For respondenter med høy produktkunnskap viser parameterestimaten i tabell 8.37 at de spesifikke produktvurderingenes evne til å forklare totalvurderingen øker dersom viktighetsspørsmålene plasseres før de spesifikke vurderingsspørsmålene. For disse respondentene reduseres de spesifikke vurderingenes evne til å forklare respondentenes totalvurdering, dersom viktighetsspørsmålene plasseres etter de spesifikke vurderings-spørsmålene.

Ved signifikansnivå $\alpha = 0,05$ rapporteres verdien for styrke til 0,662 for interaksjonen mellom tilstedeværelse og plassering av viktighetsspørsmål og respondentenes produktkunnskap. Den korresponderende *eta-square*, som et mål på effektstørrelse rapporteres lik 0,015. For to-veis-interaksjonen mellom de spesifikke produktvurderingene og tilstedeværelse og plassering av viktighetsspørsmål innenfor ulike nivå av produktkunnskap rapporteres styrke på 0,778, med en korresponderende *eta square* lik 0,022.

De rapporterte effektstørrelsene må karakteriseres som små, mens verdiene for styrke indikerer over middels statistisk styrke for F-testene.

TABELL 8.38

Standardiserte beta-koeffisienter og forklart varians for totalvurdering avhengig av tilstedeværelse og plassering av viktighetsspørsmål innenfor ulike nivå av produktkunnskap

| INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA (ADJ. R SQ.) | | |
|---|---------------------------------|-------------|-------------------|
| | L.P. | M.P. | H.P. ^a |
| Uten viktighetsspørsmål | 0,72 (0,51) | 0,62 (0,37) | 0,67 (0,44) |
| Viktighetsspørsmål før vurderingsspørsmål | 0,74 (0,54) | 0,59 (0,34) | 0,65 (0,42) |
| Viktighetsspørsmål etter vurderingsspørsmål | 0,49 (0,23) | 0,65 (0,42) | 0,53 (0,28) |

^a L.P. = Lav produktkunnskap M.P. = Middels produktkunnskap H.P. = Høy produktkunnskap

I tabell 8.38 presenteres standardiserte beta-koeffisienter og forklart varians for ytterligere å illustrere resultatene fra variansanalysen. For respondenter med lav og høy produktkunnskap viser tabell 8.38 et resultat som ikke er konsistent med tallene som rapporteres i variansanalysen i tabell 8.37. Forskjeller i varians mellom eksperimentgruppene når det gjelder de spesifikke vurderingene kan være en forklaring på dette. For respondenter med middels kunnskap viser tabell 8.38 små forskjeller mellom eksperimentgruppene, et resultat som er konsistent med resultatene fra variansanalysen. Det er for denne hypotesen vanskelig å benytte standardiserte beta-koeffisienter og forklart varians for å illustrere forskjeller mellom eksperimentgruppene. Imidlertid kan vi med bakgrunn i analysene ovenfor konkludere med at hypotese 14a ikke støttes av de innsamlede dataene.

Produktkunnskap kan også få betydning for respondentenes atferdsintensjoner. Som tidligere gjennomføres derfor analyser der respondentenes atferdsintensjoner betraktes som avhengig variabel når vi tar hensyn til produktkunnskap. Disse analysene er basert på følgende hypotese.

H14b: Forskjellen mellom målinger der viktighetsspørsmål er plassert før, vs. målinger der viktighetsspørsmål er plassert etter spesifikke vurderingsspørsmål i et spørreskjema, når det gjelder de spesifikke vurderingsspørsmålenes evne til å forklare atferdsintensjoner, vil være mindre for respondenter med høy produktkunnskap enn for respondenter med lav produktkunnskap

TABELL 8.39

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, innenfor ulike nivå av produktkunnskap

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 2160,13 | 577 | 3,74 | | |
| Konstant | 2456,24 | 1 | 2456,24 | 656,10 | < 0,001 |
| Samtlige 27 vurderinger | 762,61 | 1 | 762,61 | 203,70 | < 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 7,62 | 2 | 3,81 | 1,02 | 0,362 ^b |
| Produktkunnskap | 5,33 | 2 | 2,67 | 0,71 | 0,491 ^c |
| Tilstedeværelse og plassering av viktighetsspørsmål x produktkunnskap | 12,24 | 4 | 3,06 | 0,82 | 0,515 ^d |
| Samtlige 27 vurderinger x tilstedeværelse og plassering av viktighetsspørsmål within produktkunnskap | 20,47 | 6 | 3,41 | 0,91 | 0,486 ^e |
| (Model) | 814,63 | 15 | 54,31 | 14,51 | < 0,001 |
| (Total) | 2974,76 | 592 | 5,02 | | |
| R-squared = | 0,274 | | | | |
| Adjusted R-squared = | 0,255 | | | | |

^a Gruppegjennomsnitt totalvurdering:
Uten viktighetsspørsmål

Viktighetsspørsmål før vurderingsspørsmål

Viktighetsspørsmål etter vurderingsspørsmål

| | |
|--------------------------|----------------|
| Lav produktkunnskap: | 8,845 (n = 74) |
| Middels produktkunnskap: | 9,083 (n = 48) |
| Høy produktkunnskap: | 9,133 (n = 94) |
| Lav produktkunnskap: | 9,098 (n = 61) |
| Middels produktkunnskap: | 9,027 (n = 56) |
| Høy produktkunnskap: | 8,970 (n = 82) |
| lav produktkunnskap: | 9,184 (n = 57) |
| Middels produktkunnskap: | 8,809 (n = 47) |
| Høy produktkunnskap: | 9,459 (n = 74) |

^b p = 0,181 for en-hale test

^c p = 0,246 for en-hale test

^d p = 0,258 for en-hale test

^e p = 0,243 for en-hale test

| Parameterestimater: | Konstant | Samtlige 27 vurderinger | Tilstedeværelse og plassering av viktighetsspørsmål | Tilstedeværelse og plassering av viktighetsspørsmål x instruksjon og vurderingsskala |
|---------------------------|----------|-------------------------|---|--|
| Lav produktkunnskap | | | | |
| Uten viktighetsspørsmål: | 6,054 | 0,986 | -0,320(ns) | 0,053(ns) |
| Viktighetsspørsmål før: | 6,054 | 0,986 | -0,171(ns) | 0,037(ns) |
| Viktighetsspørsmål etter: | 6,054 | 0,986 | 0,491(ns) | -0,090(ns) |
| Middels produktkunnskap | | | | |
| Uten viktighetsspørsmål: | 5,430 | 1,132 | 0,541(ns) | -0,174(ns) |
| Viktighetsspørsmål før: | 5,430 | 1,132 | -0,304(ns) | 0,122(ns) |
| Viktighetsspørsmål etter: | 5,430 | 1,132 | -0,137(ns) | 0,052(ns) |
| Høy produktkunnskap | | | | |
| Uten viktighetsspørsmål: | 6,212 | 0,959 | -0,556(ns) | 0,146(ns) |
| Viktighetsspørsmål før: | 6,212 | 0,959 | -0,566(ns) | 0,138(ns) |
| Viktighetsspørsmål etter: | 6,212 | 0,959 | 1,122(ns) | -0,284(ns) |

Separate variansanalyser for ulike nivå av produktkunnskap for testing av hypotese 14b er presentert i tabell E.23 - tabell E.25 i appendiks E. I tabell 8.39 rapporteres resultater fra ANOVA for testing av hypotese 14b. I denne tabellen er interaksjonen mellom de spesifikke vurderingene og tilstedeværelse og plassering av viktighetsspørsmål i spørreskjemaet innenfor ulike nivå av produktkunnskap, som for totalvurderingen, det relevante leddet i variansanalysen for testing av hypotesen. Dette leddet sier noe om hvorvidt de spesifikke produkttegenskapenes evne til å forklare respondentenes atferdintensjoner er forskjellig ved bruk av spørreskjema der tilstedeværelse og plassering av viktighetsspørsmål manipuleres, dersom vi tar hensyn til respondentenes produktkunnskap. Fra tabellen ser vi først at det rapporteres en to-halet sannsynlighet på 0,362 for en null-hypotese om at inkludering og alternativ plassering av viktighetsspørsmål ikke har betydning for nivået på totalvurderingen. Videre rapporteres en to-halet sannsynlighet på 0,486 for en nullhypotese om at interaksjonen mellom tilstedeværelse og plassering av viktighetsspørsmål og spesifikke vurderinger ikke har betydning for nivået på totalvurderingen når vi tar hensyn til respondentenes produktkunnskap. Den en-halete p-verdien er her 0,243 som indikerer at interaksjonseffekten ikke er signifikant. De innsamlete dataene gir ikke grunnlag for å påstå at respondentenes produktkunnskap modererer effektene av å inkludere og varierer plassering av viktighetsspørsmål i spørreskjemaet for atferdsintensjonene.

I tabell 8.40 har vi presentert standardiserte beta-koeffisienter og forklart varians for ytterligere å illustrere resultatene fra variansanalysen.

TABELL 8.40

Standardiserte beta-koeffisienter og forklart varians for atferdsintensjoner avhengig av tilstedeværelse og plassering av viktighetsspørsmål innenfor ulike nivå av produktkunnskap

| INSTRUKSJON OG VURDERINGSSKALA | STANDARDISERT BETA (ADJ. R SQ.) | | |
|---|---------------------------------|-------------|-------------------|
| | L.P. | M.P. | H.P. ^a |
| Uten viktighetsspørsmål | 0,55 (0,29) | 0,46 (0,19) | 0,51 (0,25) |
| Viktighetsspørsmål før vurderingsspørsmål | 0,61 (0,36) | 0,54 (0,28) | 0,51 (0,26) |
| Viktighetsspørsmål etter vurderingsspørsmål | 0,51 (0,24) | 0,51 (0,25) | 0,47 (0,21) |

^a L.P. = Lav produktkunnskap M.P. = Middels produktkunnskap H.P. = Høy produktkunnskap

Tabell 8.40 viser i hovedsak et mønster tilsvarende det som rapporteres i variansanalysen i tabell 8.39. Basert på analysene ovenfor har vi valgt å konkludere med at hypotese 14b ikke finner støtte i dataene.

8.3 Oppsummering av funn fra hypotesetestingen

I kapittel 4 oppsummerte vi hypotesene i tabellform. For å gi en bedre oversikt over resultatene, vil vi nå sammenholde hypotesene med funnene fra hypotesetestingen. Først oppsummeres hypoteser som omhandler valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer. Deretter oppsummeres hypoteser der vi tar for oss valg i forbindelse med inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet.

8.3.1 Objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer

I tabell 8.41 og 8.42 presenteres først en oversikt over hypoteser og funn i tilknytning til nivå på-, og konsistens i vurderinger, som følge av valg mellom subjektforankret- og objektforankret instruksjon og -vurderingsskala for spesifikke vurderingss spørsmål i spørreskjemaet.

TABELL 8.41

Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Nivå og konsistens i vurderinger

| | | UAVHENGIG VARIABEL: | | | | |
|----|--------------------------|--------------------------------|-----------------|------------------|-----------------|------------------------------|
| | | Instruksjon og vurderingsskala | | | | |
| | | HYPOTESE: | | FUNN: | | SIGN. NIVÅ |
| H: | AVHENGIG VARIABEL: | Subjektforankret | Objektforankret | Subjektforankret | Objektforankret | |
| H1 | Nivå på vurderinger | | = | | = | |
| H2 | Konsistens i vurderinger | | < | | = (<) | ($p < 0,001$) ^b |

> < : retning på hypoteser og funn, = : ingen forskjell mellom eksperimentgruppene, (> <) : retningsstøtte
a: to-hale test, b: en-hale test

Som det fremgår av oversikten i tabell 8.41, finner vi ikke støtte for å forkaste "nullhypotesen" om at det ikke vil være forskjeller i nivå på vurderinger av spesifikke produkttegenskaper, avhengig av om det benyttes subjektforankret- eller objektforankret instruksjon og -vurderingsskala for vurderingss spørsmålene. Til tross for at nivået på 17 av 27 vurderinger er noe høyere for respondenter som har foretatt subjektforankrete vurderinger, er forskjellene så små at vi har valgt å konkludere med at dataene ikke gir

grunnlag for å påstå forskjeller mellom eksperimentgruppene. For hypotesen om høyere konsistens i vurderinger for respondenter som har benyttet spørreskjema med objektforankret instruksjon og -vurderingsskala, finner vi kun retningsstøtte. Retningsstøtten fremkommer som resultat av at standardavvikene for 21 av 27 individuelle vurderinger, er høyest i eksperimentgruppen som har benyttet spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer¹¹⁷.

Når vi tar hensyn til kvalitetsnivået på produktene som vurderes, har vi argumentert for at hypotesene vi har fremsatt i forbindelse med forskjeller i nivå på subjektforankrete- og objektforankrete vurderinger, henger sammen. Oppsummering av resultater for testing av disse hypotesene presenteres i tabell 8.42.

TABELL 8.42

Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Nivå på vurderinger, kvalitetsnivået på produktene som vurderes som modererende variabel

| | | UAVHENGIG VARIABEL: | | | | |
|-----|--|--------------------------------|-----------------|------------------|-----------------|------------|
| | | Instruksjon og vurderingsskala | | | | |
| H: | AVHENGIG VARIABEL: | HYPOTESE: | | FUNN: | | SIGN. NIVÅ |
| | | Subjektforankret | Objektforankret | Subjektforankret | Objektforankret | |
| H3a | Nivå på vurderinger for lav kvalitetsprodukter | | < | | = | |
| H3b | Nivå på vurderinger for gjennomsnittsprodukter | | = | | = | |
| H3c | Nivå på vurderinger for høykvalitetsprodukter | | > | | = | |

> < : retning på hypoteser og funn, = : ingen forskjell mellom eksperimentgruppene, (> <) : retningsstøtte

a: to-hale test, b: en-hale test

Til tross for at vi ikke finner forskjeller i nivå på vurderinger mellom eksperimentgruppene for gjennomsnittsprodukter, har vi valgt å konkludere med at hypotesene som helhet ikke finner støtte i de innsamlede dataene.

For hypotesene vi har formulert i forbindelse med betydningen av valg mellom subjektforankrete- og objektforankrete instruksjoner og -vurderingsskalaer for sikkerhet i vurderinger og de spesifikke vurderingenes prediktive validitet, har vi introdusert respondentenes produktkunnskap som modererende variabel. Vi presenterer derfor først, i tabell 8.43, resultater og funn fra testing av hypoteser der produktkunnskap ikke

¹¹⁷Jfr. figur 8.2

er inkludert i hypotesene. Deretter oppsummeres resultater fra hypotesetesting der vi skiller mellom respondenter med høy og lav produktkunnskap, i tabell 8.44.

TABELL 8.43

Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Prediktiv validitet og sikkerhet i vurdering

| | | UAVHENGIG VARIABEL: | | | | |
|-----|--|--------------------------------|-----------------|------------------|-----------------|------------------------|
| | | Instruksjon og vurderingsskala | | | | |
| H: | AVHENGIG VARIABEL: | HYPOTESE: | | FUNN: | | SIGN. NIVÅ |
| | | Subjektforankret | Objektforankret | Subjektforankret | Objektforankret | |
| H4a | Spesifikke vurderingers evne til å forklare: | | | | | |
| | Objektforankret totalvurd. | | > | | < | p < 0,05 ^a |
| | Subjektforankret totalvurd. | | > | | ≈ | |
| H4b | Spesifikke vurderingers evne til å forklare atferdsintensjoner | | > | | ≈ | |
| H5a | Immaterielle egenskapers relative betydning for forklaring av: | | | | | |
| | Objektforankret totalvurd. | | > | | > | p < 0,025 ^b |
| | Subjektforankret totalvurd. | | > | | > | p < 0,005 ^b |
| H5b | Immaterielle egenskapers relative betydning for forklaring av atferdsintensjoner | | > | | > | p < 0,1 ^b |
| H6 | Sikkerhet i vurderinger | | > | | ≈ | |

> < : retning på hypoteser og funn, ≈ : ingen forskjell mellom eksperimentgruppene, (> <) : retningsstøtte

a: to-hale test, b: en-hale test

Resultatene fra hypotesetestingen gir ikke støtte for å hevde at subjektforankrete vurderinger av spesifikke produktegenskaper forklarer totalvurderinger og atferdsintensjoner bedre enn objektforankrete vurderinger. Når objektforankrete totalvurderinger benyttes som avhengig variabel indikerer resultatene funn som går i motsatt retning av det vi, med bakgrunn i teoridrøftningen, hadde forventet. Resultatene viser videre at immaterielle produktegenskaper får større relativ betydning, både for objektforankrete- og subjektforankrete totalvurderinger og for atferdsintensjoner, når det blir benyttet subjektforankret instruksjon og -vurderingsskala for de spesifikke vurderings-spørsmålene i spørreskjemaet. Til slutt i tabellen, ser vi at hypotesen om at respondenter vil være sikrere i sine vurderinger, når disse foretas i spørreskjema med subjektforankret instruksjon og -vurderingsskala, ikke støttes av de innsamlete dataene.

TABELL 8.44

Oversikt over hypoteser og funn i tilknytning til valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer - Prediktiv validitet og sikkerhet i vurdering, respondentenes produktkunnskap som modererende variabel

| H: AVHENGIG VARIABEL: | | UAVHENGIG VARIABEL: | | | | SIGN. NIVÅ |
|-----------------------|--|--------------------------------|----------------------|-----------------------|----------------------|------------------------|
| | | Instruksjon og vurderingsskala | | | | |
| | | HYPOTESE: | | FUNN: | | |
| | | Subjekt- forankret | Objekt- forankret | Subjekt- forankret | Objekt- forankret | |
| H7a | Spesifikke vurderingers evne til å forklare: | | | | | |
| | | <u>Lav produktkunnskap</u> | | | | |
| | Objektforankret totalvurd. | > | | ≈ | | |
| | Subjektforankret totalvurd. | > | | > | | p < 0,01 ^b |
| | | <u>Høy produktkunnskap</u> | | | | |
| | Objektforankret totalvurd. | ≈ | | < | | p < 0,025 ^a |
| | Subjektforankret totalvurd. | ≈ | | < | | p < 0,025 ^a |
| H7b | Spesifikke vurderingers evne til å forklare | | | | | |
| | | <u>Lav produktkunnskap</u> | | | | |
| | Atferdsintensjoner: | > | | > | | p < 0,001 ^b |
| | | <u>Høy produktkunnskap</u> | | | | |
| | Atferdsintensjoner | ≈ | | < | | p < 0,001 ^a |
| H8a | Immaterielle egenskapers relative betydning for forklaring av: | | | | | |
| | | <u>Lav produktkunnskap</u> | | | | |
| | Objektforankret totalvurd. | > | | ≈ | | |
| | Subjektforankret totalvurd. | > | | > | | p < 0,005 ^b |
| | | <u>Høy produktkunnskap</u> | | | | |
| | Objektforankret totalvurd. | ≈ | | > | | p < 0,1 ^a |
| | Subjektforankret totalvurd. | ≈ | | > | | p < 0,005 ^a |
| H8b | Immaterielle egenskapers relative betydning for forklaring av: | | | | | |
| | | <u>Lav produktkunnskap</u> | | | | |
| | Atferdsintensjoner ¹¹⁸ | > | | > | | p < 0,1 ^b |
| | | <u>Høy produktkunnskap</u> | | | | |
| | Atferdsintensjoner | ≈ | | > | | p < 0,1 ^a |
| H9a | Sikkerhet i vurderinger | | | | | |
| | | <u>Lav produktkunnskap</u> | | | | |
| H9b | | > | | ≈ | | |
| | | <u>Høy produktkunnskap</u> | | | | |
| | | ≈ | | ≈ | | |

> < : retning på hypoteser og funn, ≈ : ingen forskjell mellom eksperimentgruppene, (> <) : retningsstøtte

a: to-hale test, b: en-hale test

¹¹⁸Funnene som rapporteres for atferdsintensjoner er basert på separate analyser foretatt innenfor ulike nivå av produktkunnskap (se appendix I). P-verdiene som rapporteres i tabellen er hentet fra disse analysene.

Funnene som oppsummeres i tabell 8.44 viser resultater fra testing av hypoteser der vi tar hensyn til at respondentene kan være forskjellige når det gjelder kunnskapene de har som grunnlag for å foreta vurderinger av hotellproduktet. For det første ser vi at det for respondenter med lav produktkunnskap, finnes støtte for hypotesen om at subjektforankrete vurderinger av spesifikke produktegenskaper, forklarer subjektforankrete totalvurderinger bedre enn spesifikke vurderinger foretatt med objektforankrete instruksjoner og -vurderingsskalaer. For objektforankrete totalvurderinger finner vi ingen forskjeller mellom eksperimentgruppene når det gjelder de spesifikke produktvurderingenes prediksjonsevne. Når vi derimot foretar tilsvarende analyser for respondenter med høy produktkunnskap, finner vi at objektforankrete vurderinger av spesifikke produktegenskaper forklarer objektforankrete totalvurderinger bedre enn spesifikke vurderinger foretatt med subjektforankret instruksjon og -vurderingsskala. For respondenter med høy produktkunnskap forklares også subjektforankrete totalvurderinger best når de spesifikke vurderingene foretas med objektforankret instruksjon og -vurderingsskala. Sett under ett finner vi altså ikke støtte for hypotese 7a.

Respondentenes atferdsintensjoner er i avhandlingen betraktet som “nøytrale totalvurderinger”. For respondenter med lav produktkunnskap ser vi først at vurderinger av spesifikke produktegenskaper foretatt med subjektforankret instruksjon og -vurderingsskala forklarer atferdsintensjoner bedre enn spesifikke vurderinger foretatt med objektforankret instruksjon og -vurderingsskala. Dette resultatet er i samsvar med hypotesene vi har formulert, mens det motsatte er tilfellet for respondenter med høy produktkunnskap. For disse finner vi at objektforankrete vurderinger forklarer atferdsintensjoner bedre enn subjektforankrete vurderinger. Dette gjør at vi må konkludere med manglende støtte, også for hypotese 7b.

Oversikten i tabell 8.44 viser videre at våre tidligere funn, vedrørende den relative betydningen av immaterielle produktegenskaper for forklaring av totalvurderinger og atferdsintensjoner, ikke ser ut til å bli påvirket av respondentenes produktkunnskap. Med unntak av funnene for respondenter med lav produktkunnskap, der det ikke finnes forskjeller mellom eksperimentgruppene når objektforankrete totalvurderinger betraktes som avhengig variabel, fører subjektforankrete vurderinger til større relativ betydning av de immaterielle produktegenskapene. Hypotese 8a og 8b støttes dermed heller ikke av de innsamlete dataene

Respondenter med høy produktkunnskap antas å være sikrere på sine vurderinger av spesifikke egenskaper ved produktet, enn kunder med lav produktkunnskap. Dette kan bidra til å moderere effektene av instruksjon og vurderingsskala for respondentenes sikkerhet i sine vurderinger. I tabell 8.44 oppsummeres resultater fra testing av disse

antakelsene i hypotese 9a og 9b. Resultatene fra hypotestesting viser her at våre tidligere funn, der vi ikke fant støtte for forskjeller i sikkerhet mellom objektforankrete- og subjektforankrete vurderinger, ikke påvirkes av respondentenes produktkunnskap.

8.3.2 Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet

Før vi retter oppmerksomheten mot resultater fra testing av hypoteser i forbindelse med inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, vil vi oppsummere funn fra analyser der fokus har vært nivå på vurderingss spørsmål og viktighets spørsmål. Først vil vi rapportere resultater fra analyser som kan betraktes som eksplorative, i den forstand at de ble foretatt uten at vi på forhånd har vært i stand til å formulere teoribaserte hypoteser. Disse analysene gir imidlertid grunnlag for diskusjon av interessante forskjeller mellom eksperimentgruppene. Deretter oppsummeres resultater fra analyser der vi fokuserer på betydningen av kvalitetsnivået på produktene som vurderes, for forskjeller i nivå på vurderinger foretatt i spørreskjema der plassering av viktighetsspørsmålene manipuleres.

TABELL 8.45

Oversikt over funn i tilknytning til eksplorative analyser vedrørende inkludering og plassering av viktighetsspørsmål og spesifikke vurderingss spørsmål i et spørreskjema - Nivå på viktighetsspørsmål og vurderingss spørsmål

| UAVHENGIG VARIABEL | | | | |
|---|---------------------|-----------------------|----------------|------------------------|
| Inkludering og plassering av viktighetsspørsmål | | | | |
| AVHENGIG VARIABEL: | Før spes. vurd.spm. | Etter spes. vurd.spm. | Uten viktighet | SIGN. NIVÅ |
| Nivå på spesifikke vurderinger | = | = | | |
| Nivå på viktighetsspørsmål | < | | | p < 0,001 ^a |

> < : retning på funn, = : ingen forskjell mellom eksperimentgruppene, (> <) : retringsstøtte

a: to-hale test, b: en-hale test

De eksplorative analysene vi har foretatt, i forbindelse med forskjeller i nivå på viktighetsspørsmål og vurderingss spørsmål, oppsummeres i tabell 8.45. Tabellen viser først hvordan svargivning på vurderingss spørsmålene i spørreskjemaet i liten grad ser ut til å bli påvirket av viktighetsspørsmålene, og svarene som er gitt på disse. Nivået på de spesifikke vurderingss spørsmålene avviker i liten grad mellom eksperimentgruppene.

Analysene viser videre hvordan nivå på viktighetsspørsmålene blir forskjellig avhengig av hvor disse plasseres i forhold til de spesifikke vurderingsspørsmålene. Mer spesifikt rapporteres de spesifikke produktegenskapene til å være viktigere når respondentene har foretatt vurderinger før de tillegger egenskapene viktighet.

I tabellen har vi oppsummert funnene fra de eksplorative analysene sett under ett. Dersom vi går nærmere inn i analysene fremkommer et interessant mønster for nivået på viktighetsspørsmålene. Som illustrert i figur 8.5 er det kun for samhandlingsdelen av produktet at nivået på viktighetsspørsmål ikke påvirkes av plasseringen av disse i forhold til de spesifikke vurderingsspørsmålene i spørreskjemaet. For de øvrige produkt-elementene blir egenskapene i gjennomsnitt rapportert til å være viktigere når viktighetsspørsmålene plasseres etter de spesifikke vurderingsspørsmålene.

TABELL 8.46

Oversikt over hypoteser og funn i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Nivå på vurderingsspørsmål, kvalitetsnivået på produktene som vurderes som modererende variabel

| | | UAVHENGIG VARIABEL | | | | |
|------|--|---|----------------|---------------------|----------------|------------|
| | | Inkludering og plassering av viktighetsspørsmål | | | | |
| | | HYPOTESE: | | FUNN: | | SIGN. NIVÅ |
| H: | AVHENGIG VARIABEL | Før spes. vurd.spm. | Uten viktighet | Før spes. vurd.spm. | Uten viktighet | |
| H10a | Nivå på vurderinger for lavkvalitetsprodukter | | < | | ≈ | |
| H10b | Nivå på vurderinger for gjennomsnittsprodukter | | ≈ | | ≈ | |
| H10c | Nivå på vurderinger for høykvalitetsprodukter | | > | | ≈ | |

> < : retning på hypoteser og funn, ≈ : ingen forskjell mellom eksperimentgruppene, (> <) : retringsstøtte

a: to-hale test, b: en-hale test

I tabell 8.46 oppsummeres resultater fra analyser av forskjeller i nivå på vurderinger av spesifikke produktegenskaper, når vi tar hensyn til kvalitetsnivået på produktene som vurderes. Vi har argumentert for at hypotesene om forskjeller i nivå på vurderinger, avhengig av kvalitetsnivået på produktene som vurderes, henger sammen. Dette fører til at vi, til tross for at vi ikke fant støtte for forskjeller mellom eksperimentgruppene når det gjelder nivå på vurderinger for gjennomsnittsprodukter (H10b), har valgt å konkludere med at hypotesene som helhet ikke finner støtte i de innsamlede dataene.

For hypotesene vi har formulert i forbindelse med sammenhengen mellom viktighet og vurderinger, og de spesifikke vurderingenes prediktive validitet, har vi introdusert respondentenes produktkunnskap som modererende variabel. For å gi en oversikt over resultatene presenterer vi derfor først, i tabell 8.47, resultater og funn fra testing av hypoteser der produktkunnskap ikke er inkludert i hypotesene. Deretter inkluderes produktkunnskap som modererende variabel, og resultater fra hypotesetesting der vi skiller mellom respondenter med høy og lav produktkunnskap presenteres i tabell 8.48.

TABELL 8.47

Oversikt over hypoteser og funn i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Sammenheng mellom viktighet og vurdering og prediktiv validitet

| | | UAVHENGIG VARIABEL | | | | |
|------|--|----------------------------------|-----------------------|---------------------|-----------------------|-----------------------|
| | | Plassering av viktighetsspørsmål | | | | |
| H: | AVHENGIG VARIABEL | HYPOTESE: | | FUNN: | | SIGN. NIVÅ |
| | | Før spes. vurd.spm. | Etter spes. vurd.spm. | Før spes. vurd.spm. | Etter spes. vurd.spm. | |
| H11 | Sammenheng mellom viktighet og vurdering | | > | | > | p < 0,05 ^b |
| H12a | Spesifikke vurderingers evne til å forklare totalvurdering | | > | | > | p < 0,05 ^b |
| H12b | Spesifikke vurderingers evne til å forklare atferdsintensjoner | | > | | ≈ | |

> < : retning på hypoteser og funn, ≈ : ingen forskjell mellom eksperimentgruppene, (> <) : retningsstøtte

a: to-hale test, b: en-hale test

Resultatene som presenteres i tabell 8.47 viser først at våre antakelser om at det vil være sterkere sammenheng mellom viktighet og vurderinger når viktighetsspørsmål plasseres før spesifikke vurderingsspørsmål, enn når viktighetsspørsmålene plasseres etter de spesifikke vurderingsspørsmålene, støttes av de innsamlete dataene.

For hypotesene vi har fremsatt i forbindelse med de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner, finner vi støtte kun for hypotese 12a der totalvurderinger betraktes som avhengig variabel. De innsamlete dataene gir altså støtte for våre antakelser om at de spesifikke egenskapenes evne til å forklare respondentenes totalvurderinger, vil være høyere når viktighetsspørsmålene plasseres før de spesifikke vurderingsspørsmålene i spørreskjemaet, enn når viktighetsspørsmålene plasseres etter

de spesifikke vurderingsspørsmålene. Når vi betrakter atferdsintensjoner som avhengig variabel finner vi ingen forskjeller mellom eksperimentgruppene.

TABELL 8.48

Oversikt over hypoteser og funn i tilknytning til plassering av viktighetsspørsmål og spesifikke vurderingsspørsmål i et spørreskjema - Sammenheng mellom viktighet og vurdering og prediktiv validitet, respondentenes produktkunnskap som modererende variabel

| H: AVHENGIG VARIABEL | | UAVHENGIG VARIABEL | | | | SIGN. NIVÅ |
|----------------------|--|----------------------------------|-----------------------|---------------------|-----------------------|------------------------|
| | | Plassering av viktighetsspørsmål | | | | |
| | | HYPOTESE: | | FUNN: | | |
| | | Før spes. vurd.spm. | Etter spes. vurd.spm. | Før spes. vurd.spm. | Etter spes. vurd.spm. | |
| H13 | Sammenheng mellom viktighet og vurdering | <u>Lav produktkunnskap</u> | | | | p < 0,025 ^b |
| | | > | | | > | |
| | | <u>Høy produktkunnskap</u> | | | | |
| | | = | | | = | |
| H14a | Spesifikke vurderingers evne til å forklare: Totalvurdering | <u>Lav produktkunnskap</u> | | | | p < 0,05 ^b |
| | | > | | | > | |
| | | <u>Høy produktkunnskap</u> | | | | |
| | | = | | | > | p < 0,05 ^a |
| H14b | Spesifikke vurderingers evne til å forklare Atferdsintensjoner: | <u>Lav produktkunnskap</u> | | | | |
| | | > | | | = | |
| | | <u>Høy produktkunnskap</u> | | | | |
| | | = | | | = | |

> < : retning på hypoteser og funn, = : ingen forskjell mellom eksperimentgruppene, (> <) : retningsstøtte
a: to-hale test, b: en-hale test

I tabell 8.48 oppsummeres først resultater fra analyser der vi har foretatt sammenligninger av sammenhengen mellom viktighet og vurderinger for respondenter med lav og høy produktkunnskap. For respondenter med lav produktkunnskap viser tabellen at vi for hypotese 13 finner støtte for våre antakelser om at sammenhengen mellom viktighet og vurderinger vil være sterkere når viktighetsspørsmålene plasseres før de spesifikke vurderingsspørsmålene, enn når viktighetsspørsmålene plasseres etter de spesifikke vurderingsspørsmålene. For respondenter med høy produktkunnskap finner vi ikke forskjeller mellom eksperimentgruppene. Disse resultatene samsvarer også med hypotesen vi har formulert.

Resultatene fra hypotesetestingen viser videre at hypotese 14a og 14b ikke finner støtte i de innsamlede dataene. For respondenter med lav produktkunnskap finner vi støtte for hypotesen om at de spesifikke vurderingene vil ha større evne til å forklare totalvurderinger når viktighetsspørsmålene plasseres før de spesifikke vurderingsspmålene. For respondenter med høy produktkunnskap har vi argumentert for at det ikke skal være forskjeller mellom eksperimentgruppene. For disse respondentene finner vi imidlertid det samme resultatet som ble rapportert for respondenter med lav produktkunnskap. Dette fører til at vi sett under ett konkluderer med manglende støtte for hypotese 14a.

For hypotese 14b rapporteres ingen forskjell mellom de to eksperimentgruppene. Her har vi argumentert for at de spesifikke vurderingene vil ha større evne til å forklare respondentenes atferdsintensjoner når viktighetsspørsmålene plasseres før de spesifikke vurderingsspmålene i spørreskjemaet. For respondenter med høy produktkunnskap har vi argumentert for at det ikke vil være forskjeller mellom eksperimentgruppene. Selv om vi for respondenter med høy produktkunnskap, isolert kan si at funnene fra hypotesetestingen støtter disse antakelsene, har vi valgt å konkludere med at hypotese 14b som helhet heller ikke finner støtte i de innsamlede dataene.

I tillegg til resultatene som rapporteres i tabell 8.47 og tabell 8.48 for de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner, har vi ved testing av disse hypotesene betraktet respondenter som har benyttet spørreskjema der viktighetsspørsmål ikke er inkludert som en kontrollgruppe. Analysene vi har foretatt viser at vurderinger foretatt av disse respondentene forklarer både totalvurderinger og atferdsintensjoner bedre enn vurderinger foretatt av respondenter som har benyttet spørreskjema der viktighetsspørsmål er inkludert¹¹⁹. Inkludering av viktighetsspørsmål i spørreskjemaet reduserer altså vurderingenes prediktive validitet.

¹¹⁹Dette gjelder uavhengig av hvor viktighetsspørsmålene plasseres i spørreskjemaet.

DEL V DISKUSJON

I denne siste delen av avhandlingen diskuteres først funnene fra hypotesetestingen i kapittel 9. Her drøftes en rekke praktiske implikasjoner av resultatene for utvikling av spørreskjema. Til slutt, i kapittel 10, rettes oppmerksomheten mot begrensninger i den gjennomførte studien som grunnlag for drøfting av implikasjoner for videre forskning.

9

Diskusjon og praktiske implikasjoner

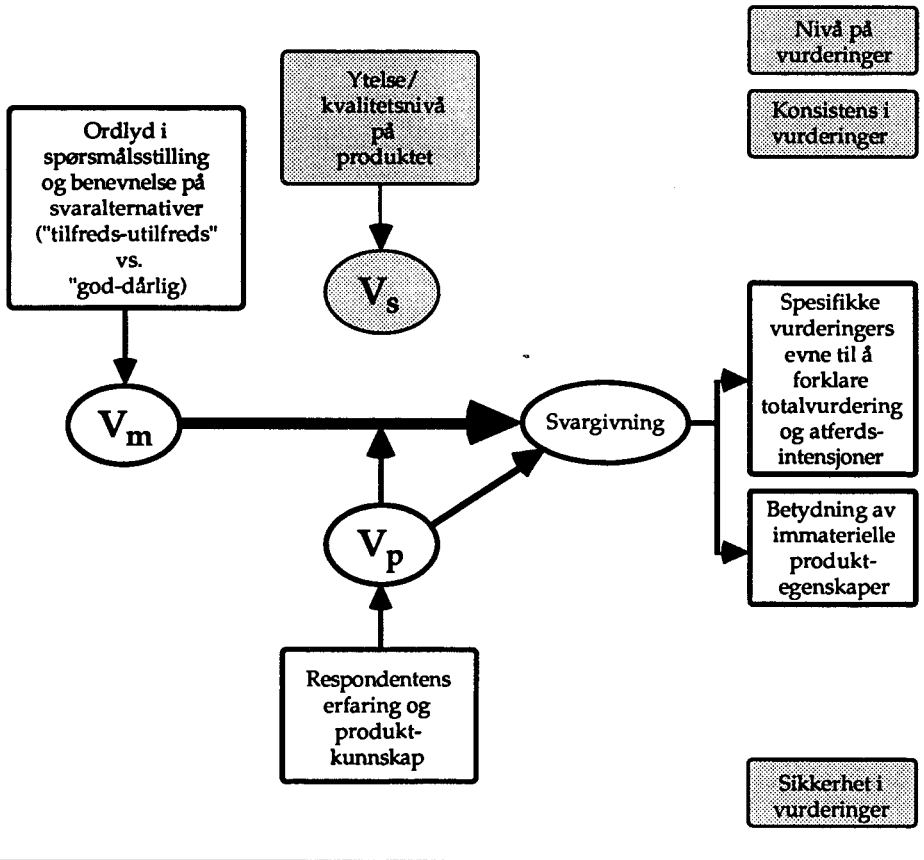
I dette kapitlet skal vi diskutere resultatene fra hypotesetestingen og ta for oss implikasjoner av funnene. Kapitlet består av to hovedblokker. I den første hovedblokken tar vi for oss problemstillingene vi har reist i tilknytning til valg mellom subjektforankrete- og objektforankrete instruksjoner og -vurderingsskalaer, til bruk ved måling av spesifikke produktvurderinger i et spørreskjema. I den andre hovedblokken retter vi oppmerksomheten mot problemstillinger i tilknytning til inkludering-, og alternativ plassering av viktighets spørsmål i spørreskjemaet ved operasjonalisering av multiattributtmodeller. Forsøkene på å anvende ulike teoretiske innfallsvinkler for å forklare og predikere responseffekter får konsekvenser, både for diskusjonen av funnene, og for hvilke implikasjoner vi vektlegger. Resultatene diskuteres således først i forhold til de teoretiske perspektivene vi har benyttet som grunnlag for å formulere hypotesene. Deretter vektlegger vi praktiske implikasjoner for utvikling av spørreskjema til bruk ved måling av kunders produktvurderinger.

9.1 Valg av subjektforankret- vs. objektforankret instruksjon og -vurderingsskala

Figur 9.1 illustrerer responseeffektene som fremkommer som følge av valg mellom å benytte objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer i et spørreskjema. I figuren reflekterer de delene av modellen som ikke er skravert funnene fra hypotesetestingen. En hovedkonklusjon basert på disse funnene er at respondentenes

produktkunnskap får stor betydning som modererende variabel for de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner.

FIGUR 9.1
Valg av subjektforankret- vs. objektforankret instruksjon og vurderingsskala - oversikt over funn fra hypotesetestingen



For respondenter med lav produktkunnskap finner vi støtte for at det bør benyttes subjektforankrete instruksjoner og -vurderingsskaler når man ønsker å måle vurderinger av spesifikke produktegenskaper i et spørreskjema. For disse respondentene er subjektforankrete vurderinger bedre egnet til å forklare totalvurderinger og atferdsintensjoner enn objektforankrete vurderinger. For respondenter med høy produktkunnskap finner vi at det motsatte er tilfellet. Når respondentenes produktkunnskap øker gir objektforankrete vurderinger best prediksjon, både på totalvurderinger og atferdsintensjoner.

Et annet viktig funn, som er i samsvar med hypotesene vi har fremsatt, er at det å benytte subjektforankrete instruksjoner og -vurderingsskalaer i forbindelse med spesifikke vurderingsspørsmål i spørreskjemaet fører til økt relativ betydning av immaterielle produktetegenskaper. Dette gjelder uavhengig av kunnskapen respondentene har om produktene vi ber dem vurdere. Dersom man i en undersøkelse skal måle kunders vurderinger av et produkt der service- eller tjenesteaspekter utgjør en vesentlig del, eller kan oppfattes som kjernen i produktet, bør disse funnene føre til at man velger å benytte subjektforankrete instruksjoner og -vurderingsskalaer i forbindelse med de spesifikke vurderingsspørsmålene i spørreskjemaet.

For hypotesen om høyere konsistens i vurderinger for respondenter som har benyttet objektforankret instruksjon og -vurderingsskala, finner vi kun retningsstøtte i de innsamlete dataene. Retningsstøtten fremkommer som resultat av at standardavvikene for 21 av 27 individuelle vurderinger, er høyest i eksperimentgruppen som har benyttet spørreskjema med subjektforankrete instruksjoner og -vurderingsskalaer¹²⁰. Selv om retningsstøtten i seg selv er relativt sterk, har vi valgt å konkludere med at vi ikke har tilstrekkelig grunnlag for å påstå at dette bør få konsekvenser når man skal velge instruksjon og vurderingsskala i forbindelse med vurderingsspørsmål i et spørreskjema.

Vi vil nå drøfte disse funnene mer i detalj, og utdype en rekke praktiske implikasjoner for utvikling av spørreskjema. I diskusjonen som følger vil vi også trekke frem mulige forklaringer på resultater der vi ikke finner støtte for våre teoribaserte hypoteser.

9.1.1 Valg av instruksjon og vurderingsskala, og spesifikke vurderingers evne til å forklare totalvurderinger og atferdsintensjoner

Som vi påpeker senere i dette kapittelet, kan det være vanskelig å ta hensyn til at respondenter har ulik kunnskap om produktene vi ber dem vurdere før vurderingene foretas i et standardisert spørreskjema. Vi innleder derfor drøftingen av valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer, med funn fra testing av hypoteser der respondentenes produktkunnskap ikke er inkludert.

Når vi ikke tar hensyn til forskjeller i respondentenes produktkunnskap, har vi konkludert med at objektforankrete vurderinger av spesifikke produktetegenskaper for-

¹²⁰Jfr. figur 8.2

klarer objektforankrete totalvurderinger bedre enn vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema med subjektforankret instruksjon og -vurderingsskala. Dersom det benyttes subjektforankret instruksjon og -vurderingsskala for totalvurderingsspørsmålene finner vi ingen forskjell mellom de to typene instruksjon og -vurderingsskala når det gjelder de spesifikke vurderingenes prediktive evne. Det samme gjelder når vi betrakter respondentenes atferdsintensjoner som avhengig variabel. Funnene som rapporteres for objektforankrete totalvurderinger er i motsatt retning av hva vi hadde forventet basert på de teoretiske perspektivene vi drøftet i kapittel 3. Her argumenterte vi for at bruk av subjektforankret instruksjon og -vurderingsskala for spesifikke vurderingsspørsmål, vil føre til bedre forklaring av både totalvurderinger og atferdsintensjoner.

En mulig årsak til at vi observerer funn som går i motsatt retning av det vi har argumentert for, basert på de teoretiske perspektivene vi har benyttet, kan være skjevheter i utvalget som danner grunnlag for hypotesetestingen. I kapittel 3 drøftet vi betydningen av produktkunnskap for valg mellom objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer for de spesifikke vurderingsspørsmålene i spørreskjemaet. Her argumenterte vi for at respondenter med høy produktkunnskap lettere vil kunne foreta objektforankrete vurderinger, enn respondenter med lav produktkunnskap. Dersom dataene som benyttes for å teste hypotesene har en overvekt av respondenter med høy produktkunnskap, vil et resultat som det observerte ikke være overraskende.

Med bakgrunn i funnene fra hypotesetestingen har vi konkludert med at respondentenes produktkunnskap bidrar til å nyansere konklusjonene fra våre tidligere analyser av de spesifikke vurderingenes evne til å forklare totalvurderinger og atferdsintensjoner. Funnene viser for det første at våre antakelser om at det ikke vil finnes forskjell mellom eksperimentgruppene for respondenter med høy produktkunnskap, ikke finner støtte i de innsamlede dataene. For disse respondentene er den entydige konklusjonen at objektforankrete vurderinger av spesifikke produkttegnegenskaper er best egnet for å forklare både objektforankrete- og subjektforankrete totalvurderinger og atferdsintensjoner. Med unntak av objektforankrete totalvurderinger, kan det for respondenter med lav produktkunnskap, i samsvar med våre hypoteser, konkluderes med at bruk av subjektforankret instruksjon og -vurderingsskala i forbindelse med spesifikke vurderingsspørsmål er best egnet for å forklare totalvurderinger og atferdsintensjoner.

Resultater fra analysene der vi tar hensyn til at respondenter har ulik produktkunnskap, bidrar på denne måten til å forklare hvorfor vi tidligere ikke fant støtte for hypotese 4a og 4b i de innsamlede dataene. Når vi foretar separate analyser for respondenter med

høy og lav produktkunnskap finner vi resultater som går i motsatt retning. Dersom vi ikke inkluderer produktkunnskap i analysene, vil effektene derfor oppheve hverandre dersom de er like sterke, eller gå i retning av den sterkeste effekten, slik tilfellet ser ut til å være for hypotese 4a og 4b.

For respondenter med lav produktkunnskap indikerer funnene fra hypotesetestingen at betydningen av konsistent instruksjon og skalabruk er størst når det benyttes subjektforankrete totalvurderingsspørsmål i spørreskjemaet. Når totalvurderinger derimot operasjonaliseres ved hjelp av objektforankret instruksjon og -vurderingsskala, kan resultatene fra hypotesetestingen tolkes dithen at man i større grad kan velge om man vil benytte objektforankret- eller subjektforankret instruksjon og -vurderingsskala for de spesifikke vurderingsspørsmålene, når disse stilles til respondenter med lav produktkunnskap.

Tidligere forskning har vist at økt produktkunnskap fører til sterkere sammenheng mellom vurderinger av spesifikke egenskaper ved et produkt, totalvurderinger av produktet og respondentenes atferdsintensjoner. Dette forklares blant annet med at økt produktkunnskap fører til at vurderingene blir mer tilgjengelige for respondentene (Fazio, 1981; Fazio og Zanna, 1981). I tillegg finnes det forskning som har vist at respondenter med høyere produktrelatert erfaring og produktkunnskap er mindre sensitive til forskningskontekst enn respondenter med lav erfaring og produktkunnskap (Lynch, Chakravatri og Mitra, 1991). Dette ledet oss til å argumentere for at vi for respondenter med høy produktkunnskap ikke skulle finne forskjeller mellom eksperimentgruppene, når det gjelder de spesifikke vurderingenes prediktive validitet.

Når vi likevel finner forskjeller, og forskjellene i tillegg er i motsatt retning av de vi fant for respondenter med lav produktkunnskap, kan det stilles spørsmål om respondenter med høy produktkunnskap lettere kan foreta objektforankrete vurderinger. Disse respondentene antar vi oftere har foretatt produktvurderinger, noe som vil reflekteres både i kognitive strukturer og kognitive prosesser. Blant annet kan produktkunnskap føre til at vurderingsprosessen blir automatisert. Dette kan i seg selv føre til bedre prediksjoner av totalvurderinger. Når vi i tillegg observerer at bruk av objektforankrete instruksjoner og -vurderingsskalaer forsterker denne effekten, kan dette tyde på at objektforankrete vurderinger for respondenter med høy produktkunnskap er mer grunnleggende for totalvurderingene.

Vi har valgt å betrakte respondentenes atferdsintensjoner som "nøytrale" totalvurderinger. Oppsummeringen som presenteres i tabell 8.44, viser et mønster for atferdsintensjonene, som i grove trekk samsvarer med det vi observerte for total-

vurderingene. For respondenter med lav produktkunnskap, forklarer vurderinger av spesifikke produktegenskaper, foretatt med subjektforankret instruksjon og -vurderingsskala, atferdsintensjoner bedre enn spesifikke vurderinger foretatt med objektforankret instruksjon og -vurderingsskala. Dette resultatet samsvarer med hypotesene vi har formulert. Det motsatte er tilfellet for respondenter med høy produktkunnskap, der vi finner at objektforankrete vurderinger forklarer respondentenes atferdsintensjoner bedre enn subjektforankrete vurderinger.

I avhandlingen har vi argumentert for at bruk av subjektforankrete- og objektforankrete instruksjoner og -vurderingsskalaer vil reflektere ulike dimensjoner i kunders produktvurderinger. Med dette som utgangspunkt, bør man ved utvikling av spørsmål som skal benyttes for å måle vurderingene i et spørreskjema, starte med en avklaring av hvilke teoretiske dimensjoner man ønsker å måle. I tillegg til en slik klargjøring av formålet med spørreundersøkelsen, er det basert på diskusjonen ovenfor, en rekke forhold som taler for at man i en gitt situasjon bør velge å benytte den ene instruksjonen og vurderingsskalaen fremfor den andre.

Dersom man er opptatt av spesifikke vurderingers evne til å forklare respondentenes totalvurderinger og atferdsintensjoner bør det benyttes subjektforankret instruksjon og -vurderingsskala for måling av de spesifikke vurderingene når respondentene har lav produktkunnskap. Når respondentene har høy kunnskap om produktene vi ber dem vurdere, har de innsamlede dataene gitt klare indikasjoner på at det bør benyttes objektforankret instruksjon og -vurderingsskala for de spesifikke vurderingsspørsmålene.

I en undersøkelse der et spørreskjema benyttes for å måle produktvurderinger, er det vanskelig å vite noe om respondentenes produktkunnskap før disse har fyllet ut spørreskjemaet. Det kan derfor være vanskelig å identifisere respondenter med ulik produktkunnskap før undersøkelsen gjennomføres. Et unntak, som til en viss grad kan forenkle dette, er når man har mulighet til å benytte et kundekartotek som grunnlag for å velge respondenter som skal delta i kundeundersøkelsen. Med et slikt kundekartotek har man mulighet for å benytte kundeforholdets varighet som grunnlag for å velge mellom å benytte spørreskjema med objektforankret- eller subjektforankret instruksjon og -vurderingsskala. Imidlertid er det viktig å være oppmerksom på at kundeforholdets varighet, som her kan betraktes som en indikator på produktrelatert erfaring, bare er et av flere elementer i kunders produktkunnskap.

Dersom produktvurderingene måles i en undersøkelse der datainnsamlingen gjennomføres ved hjelp av et personlig intervju eller et telefonintervju, kan det også være enklere å ta hensyn til de praktiske implikasjonene vi har konkludert med ovenfor. Her

kan spørsmål som fanger opp respondentenes produktkunnskap stilles før vurderings-spørsmålene i intervjuet. Avhengig av svargivning på produktkunnskapsspørsmålene, kan man så beslutte om respondenten skal foreta objektforankrete eller subjektforankrete produktvurderinger. Ved bruk av edb-baserte intervjuverktøy der svarene registreres samtidig som intervjuet foregår, vil en slik betinget bruk av instruksjon og vurderingsskala for de spesifikke vurderingssørsmålene være relativt lett å gjennomføre.

Ytterligere et forhold som kan diskuteres med utgangspunkt i funnene fra hypotesetestingen, er hvilken instruksjon og vurderingsskala som bør benyttes ved testing av nye produkter før disse lanseres på markedet. Her kan respondentenes produktkunnskap relateres til graden av nyhet som det nye produktet representerer. Dersom det nye produktet representerer noe helt nytt, der det ikke eksisterer etablerte kategorier som grunnlag for vurderinger, kan utfordringen bedriften står overfor sammenlignes med en situasjon der alle respondentene har lav produktkunnskap. Implikasjonene av dette vil være at man bør benytte subjektforankret instruksjon og -vurderingsskala som grunnlag for å måle vurderinger av produktet foretatt av kundene i testmarkedet.

9.1.2 Valg av instruksjon og vurderingsskala, og relativ betydning av immaterielle produktegenskaper

Basert på teoridrøftelsene i kapittel 3 argumenterte vi for at materielle produktegenskaper vil være viktigere, og bety mer for respondentenes totalvurderinger og atferdsintensjoner enn immaterielle egenskaper. Når det benyttes subjektforankret instruksjon og -vurderingsskala i forbindelse med spesifikke vurderingssørsmål i et spørreskjema, antar vi imidlertid at den relative betydningen av immaterielle produktegenskaper vil være større. Resultatene fra hypotesetestingen viser at dette ser ut til å være tilfellet uavhengig av respondentenes produktkunnskap.

Med unntak av funnene for respondenter med lav produktkunnskap, der vi ikke finner forskjeller mellom eksperimentgruppene når objektforankrete totalvurderinger betraktes som avhengig variabel, finner vi at den relative betydningen av immaterielle produktegenskaper er størst, både for totalvurderinger og atferdsintensjoner, når det benyttes subjektforankret instruksjon og -vurderingsskala for å måle de spesifikke produktvurderingene.

Disse funnene støtter sammenligningene vi gjorde mellom preferansevurderinger og subjektforankrede vurderinger, og mellom sammenligningsvurderinger og objektforankrede vurderinger. Preferansevurderinger er generelt blitt betraktet som subjektive, individspesifikke og heterogene på tvers av konsumenter, mens sammenligningsvurderinger betraktes som mer objektive og homogene. Funnene fra hypotesetestingen samsvarer på denne måten med funn fra en studie av Lefkoff-Hagius & Mason (1990) der det ble foretatt en sammenligning av preferansevurderinger og sammenligningsvurderinger. Her ble forskjellene i betydning av materielle- og immaterielle egenskaper ved et produkt ved bruk av de to måtene å foreta vurderinger på fokusert. Blant annet ble det argumentert for at konsumenter benytter hierarkiske informasjonsbehandlingsprosesser når de blir bedt om å foreta sammenligningsvurderinger (Johnson, 1988). Med utgangspunkt i dette vil materielle egenskaper vies mer oppmerksomhet når en konsument blir presentert et produkt med både immaterielle- og materielle egenskaper. Lefkoff-Hagius & Mason (1990) finner imidlertid støtte for at de immaterielle egenskapene blir viktigere ved preferansevurderinger enn ved sammenligningsvurderinger.

I hypotesene som ble formulert på bakgrunn av de teoretiske perspektivene vi benyttet i kapittel 3, ble respondentenes produktkunnskap også her inkludert som et forhold som vi antar vil moderere effektene av instruksjon og vurderingsskala for den relative betydningen av immaterielle produktegenskaper. Når vi gjennom hypotesetestingen finner at subjektforankrede vurderinger fører til større relativ betydning av immaterielle produktegenskaper også for respondenter med høy produktkunnskap, kan dette indikere at eventuelle problemer med å vurdere immaterielle produktegenskaper også finnes for respondenter med høy produktkunnskap. Til tross for at resultatene fra våre tidligere analyser indikerer at disse respondentene lettere kan foreta produktvurderinger, ser det ut som om vurderinger av de immaterielle produktegenskapene får større betydning når det benyttes subjektforankret instruksjon og -vurderingsskala i spørreskjemaet.

Hva er så de praktiske implikasjonene av disse funnene for utvikling av spørreskjema i forbindelse med måling av kunders produktvurderinger? For å svare på dette spørsmålet relaterer vi resultatene fra hypotesetestingen til tjenestemarkedsføringslitteraturen, der en grunnleggende antakelse er at tjenester er vanskeligere å vurdere enn fysiske varer fordi tjenester er immaterielle (se f.eks. Zeithaml (1981)).

I samsvar med Troye (1991) og Levitt (1972) har vi argumentert for at ethvert produkt kan bestå av både materielle- og immaterielle egenskaper. Funnene fra hypotesetestingen gir støtte for at respondentene legger mer vekt på materielle egenskaper enn immaterielle egenskaper når de foretar produktvurderinger. De innsamlete dataene gir imidlertid indikasjoner på at de immaterielle egenskapene ved produktet får større

relativ betydning, både for totalvurderinger og atferdsintensjoner når det benyttes subjektforankret (“tilfreds-tilfreds”) instruksjon og -vurderingsskala for de spesifikke vurderingsspørsmålene i spørreskjemaet. Med utgangspunkt i diskusjonen ovenfor kan det derfor argumenteres for at det bør benyttes subjektforankret instruksjon og -vurderingsskala ved måling av spesifikke produktvurderinger i et spørreskjema, generelt, og spesielt når undersøkelsen foretas i forbindelse med vurderinger av produkter der tjenesteelementet utgjør en vesentlig del, eller kjernen i produktet.

9.1.3 Valg av instruksjon og vurderingsskala, og nivå, konsistens og sikkerhet i vurderinger

Det kan diskuteres hvorvidt det, på bakgrunn av litteraturgjennomgangen vi har foretatt, ville vært mer riktig å betrakte hypotesen om at det ikke vil være forskjell i nivå på subjektforankrete- og objektforankrete vurderinger, som et eksplorativt forsknings-spørsmål. Litteraturgjennomgangen og teoriperspektivene vi har benyttet gir ikke grunnlag for å påstå at vurderinger foretatt med de to instruksjonene og vurderingsskalaene skal være forskjellige. Vi finner imidlertid eksempler på studier der objektforankrete- (“god-dårlig”) og subjektforankrete (“tilfreds-tilfreds”) instruksjoner og -vurderingsskalaer blir benyttet om hverandre (se f.eks. Cadotte, Woodruff, og Jenkins (1987) og Swan og Trawick (1981)). Dette har ledet oss til å stille spørsmål om man oppfatter de to instruksjonene og vurderingsskalaene vi fokuserer på i denne avhandlingen til å være like.

På bakgrunn av noe som i praksis ser ut til å være en “vedtatt sannhet”, har vi valgt å formulere det vi har kalt en “ikke-teori basert” hypotese om at det ikke vil være forskjeller på nivå på vurderinger foretatt med de to instruksjonene og vurderingsskalaene. Med utgangspunkt i teorier for kundetilfredshet (Oliver, 1981; Tse og Wilton, 1988; Yi, 1990) og assimilasjons-kontrast teori (Sherif og Hovland, 1961), har vi nyansert denne generelle hypotesen, og argumentert for at nivå på subjektforankrete- og objektforankrete vurderinger vil være forskjellig avhengig av kvalitetsnivået på produktet som vurderes. Funnene fra hypotesetestingen som oppsummeres i tabell 8.41 og tabell 8.42 viser imidlertid at det ikke finnes grunnlag for å hevde at slike forskjeller i nivå på vurderinger eksisterer.

Det kan være mange årsaker til at vi ikke finner forskjeller i nivå på vurderinger foretatt i spørreskjema med henholdsvis objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer. Den mest nærliggende forklaringen er at slike forskjeller faktisk

ikke eksisterer, og at vi derfor heller ikke finner forskjeller mellom eksperiment-gruppene i de innsamlete datene. Dersom dette er tilfellet kan vi med en viss sikkerhet slå fast at vurderinger, som vi ber respondenter foreta ved å svare på spørsmål i et spørreskjema, er relativt robuste overfor forskjeller i instruksjon og vurderingsskala når det gjelder nivå på vurderingene.

Når vi heller ikke finner forskjeller i nivå på subjektforankrete- og objektforankrete vurderinger, for produkter som representerer ulike kvalitetsnivåer, kan en forklaring være at vi i den gjennomførte studien bare har med vurderinger foretatt av gjester ved 11 hotell fra Inter Nor kjeden. I en tidligere undersøkelse (Troye, Henjesand, og Breivik, 1992) der 31 Inter Nor Hotell deltok, ble de samme hypotesene testet, og det ble her funnet støtte for hypotese 3a og 3c. Disse funnene kan tyde på at vi i denne tidligere undersøkelsen klarte å fange opp nivåforskjeller i kvalitet mellom de 31 hotellene som var store nok til at betegnelsene høykvalitetshotell, gjennomsnittshotell og lavkvalitetshotell er berettiget. Gjennomsnittsvurderingene som er rapportert i tabell 8.4 indikerer at de 11 hotellene som deltar i undersøkelsen som ble gjennomført for å teste hypotesene vi har formulert i avhandlingen, kun er representative for hotell med kvalitet lik-, og under gjennomsnittet i forhold til den tidligere undersøkelsen. Vi mangler med andre ord hotell som representerer kategorien; høykvalitetsprodukter. En sammenligning av de to datasettene viser at dette er tilfellet. Samtlige 11 hotell som deltok i undersøkelsen som ble gjennomført i forbindelse med avhandlingen, ble i den tidligere undersøkelsen kategorisert som lavkvalitetshotell eller gjennomsnittshotell. Dette forholdet kan bidra til å gi en undersøkelsesmetodologisk forklaring på hvorfor vi ikke finner støtte for hypotesene der kvalitetsnivået på produktene som vurderes inngår som modererende variabel.

De innsamlete dataene gir retningsstøtte for implikasjonene av vår teoretiske diskusjon, der vi betrakter subjektforankrete vurderinger som nærmere knyttet til respondentenes egne verdier og selvoppfatning enn objektforankrete vurderinger. Dette antar vi vil føre til lavere konsistens mellom respondentene når de instrueres til å foreta subjektforankrete vurderinger. Når vi instruerer respondenter til å gi uttrykk for sine følelser, eller tilfredshet, ved bruk av subjektforankret instruksjon og -vurderingsskala, ser det ut som om dette innebærer en annen oppgave enn om vi ber dem være "dommer" og foreta en mer objektiv vurdering, ved bruk av objektforankret instruksjon og -vurderingsskala.

Vi har, i teoridelen i avhandlingen, knyttet konsistens i vurderinger til et spørsmål om hvorvidt objektive eller intersubjektive standarder for vurderinger eksisterer. Når vi ber hotellgjester som har overnattet på hotell som tilhører en hotellkjede om å foreta

vurderinger av hotellet de har benyttet, kan det godt tenkes at slike standarder eksisterer, f.eks. som resultat av felles markedsføring, profilering og til en viss grad standardisering av tilbudet ved hotellene i kjeden. I følge våre hypoteser skulle dette nettopp ha ført til det resultatet som vi i dataene har funnet retningsstøtte for. Dersom objektive eller intersubjektive standarder for vurderinger derimot ikke eksisterer, har vi argumentert for at bruk av objektforankret instruksjon og -vurderingsskala kan medføre usikkerhet i vurderingene. Respondentene blir i en slik situasjon instruert til å være "dommer" uten at standarder, eller normer for vurderingene er "gitt" eller kjent. Dette kan føre til lavere konsistens i objektforankrete vurderinger.

Resonnementet ovenfor tilsier altså at tilstedeværelse av objektive eller intersubjektive standarder for vurderinger av et produkt, kan få betydning for konsistens i vurderinger, avhengig av hvilken instruksjon og vurderingsskala som benyttes i et spørreskjema. Hotellene som deltar i undersøkelsen som ble gjennomført for å teste hypotesene, er alle medlemmer av samme hotellkjede, Inter Nor Hotels. Hotell i denne kjeden forsøker å oppnå en felles profil gjennom kjedemedlemskapet, samtidig som de beholder sitt individuelle særpreg. Det kan derfor godt tenkes at objektive eller intersubjektive standarder for vurderinger av hotellene i liten grad er tilstede. Dette kan bidra til å forklare hvorfor vi ikke finner signifikante forskjeller i konsistens mellom eksperiment-gruppene, hverken for de individuelle vurderingene eller for noen av de summerte skalaene.

Vi har i avhandlingen argumentert for at respondentenes sikkerhet i vurderinger av spesifikke egenskaper ved et produkt, vil påvirkes av instruksjonen og vurderingsskalaen som benyttes i spørreskjemaet. Mer spesifikt har vi i hypotese 6 fremsatt en påstand om at respondenter som foretar subjektforankrete vurderinger av et produkt vil være sikrere på at vurderingene er "riktige", enn respondenter som foretar objektforankrete vurderinger. I hypotese 9a og 9b er respondentenes produktkunnskap i tillegg inkludert som et forhold som vi antar vil moderere effektene av instruksjonen og vurderingsskalaen som benyttes i spørreskjemaet. Som illustrert i tabell 8.43 og tabell 8.44, gir ikke resultatene fra hypotesetestingen støtte for at respondenter vil være sikrere i sine produktvurderinger når disse blir foretatt i spørreskjema der det benyttes subjektforankrete instruksjoner og -vurderingsskalaer.

En forklaring på hvorfor vi ikke finner støtte for hypotesen om forskjeller i sikkerhet i vurderinger, kan ligge i tilnærmingen vi har benyttet for å måle sikkerhet. Basert på en gjennomgang av alternative fremgangsmåter for måling av sikkerhet i vurderinger, valgte vi å be respondentene gi uttrykk for sikkerhet ved vurdering av hver produktegenskap.

Rapportering av sikkerhet ble foretatt ved å benytte en fem-punktsskala. Når produktet som vurderes er et hotell, kan det tenkes at de fleste respondentene føler seg relativt sikre på sine vurderinger av produktegenskapene vi ber dem vurdere¹²¹. Selv om hotellproduktet er relativt komplekst, i betydningen at det er mange egenskaper som inngår i vurderingene, er det ikke sikkert at de enkelte vurderingene oppfattes som vanskelige av respondentene. Hotellproduktet er f.eks. i andre empiriske studier karakterisert som et lav-involveringsprodukt (Breivik, 1995). Dette kan ha ført til at sikkerhetsmålene vi har benyttet i spørreskjemaet ikke skiller godt nok mellom respondentene, og at vi derfor ikke finner støtte for hypotesen om at respondenter som foretar subjektforankrete vurderinger vil være sikrere i vurderingene enn respondenter som har foretatt objektforankrete vurderinger.

På bakgrunn av resultatene fra hypotesetestingen og diskusjonen ovenfor, velger vi å konkludere med at de praktiske implikasjonene for valg mellom subjektforankrete- vs. objektforankrete instruksjoner og -vurderingsskalaer er små når vi kun er opptatt av forhold som nivå på vurderinger, konsistens i vurderinger og sikkerhet i vurderinger.

9.2 Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet

I denne hovedblokken retter vi på nytt oppmerksomheten mot problemstillingene vi har reist vedrørende inkludering og alternativ plassering av viktighetsspørsmål i et spørreskjema. Diskusjonen innledes med en oppsummering av resultatene fra hypotesetestingen. Som tidligere reflekterer feltene som ikke er skravert i figur 9.2, funn fra testing av hypotesene.

Kjennskap til respondentenes produktkunnskap får stor betydning, også når man står overfor en situasjon der man skal ta beslutning om inkludering og alternativ plassering av viktighetsspørsmål i et spørreskjema. For respondenter med lav produktkunnskap gir funnene fra hypotesetestingen klare indikasjoner på at rapportering av viktighet før man foretar spesifikke produktvurderinger får betydning for vurderingene.

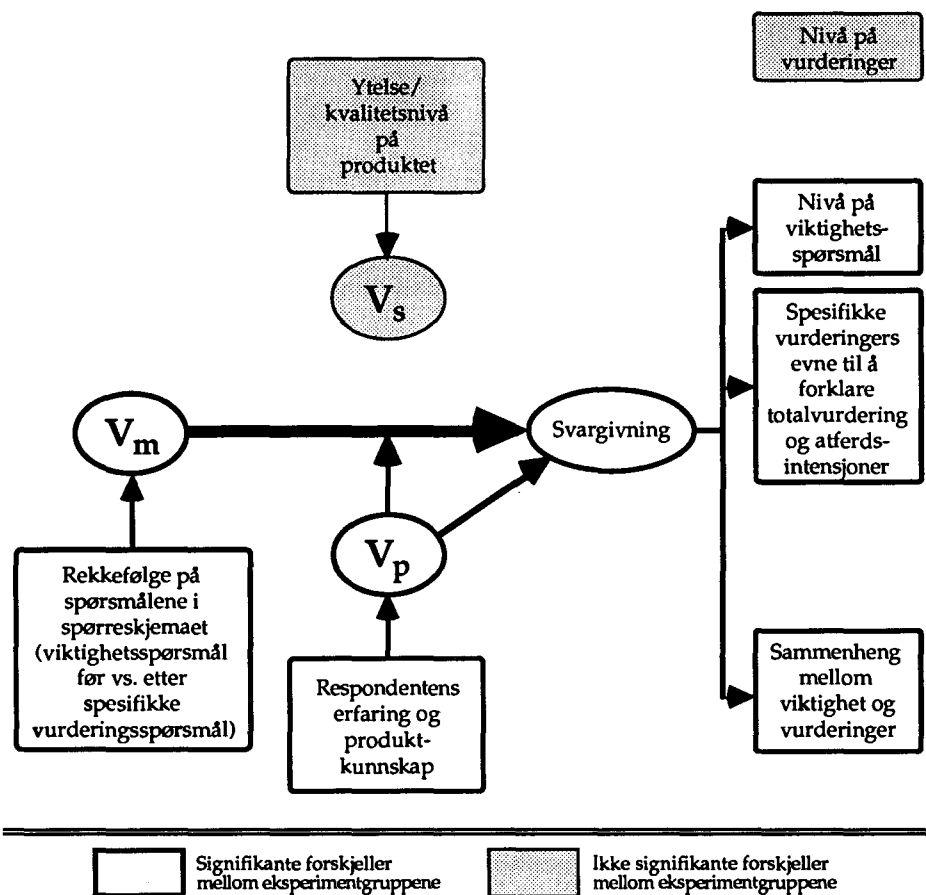
I samsvar med hva vi hadde forventet basert på de teoretiske perspektivene vi har benyttet, finner vi at inkludering og plassering av viktighetsspørsmål før spesifikke

¹²¹Nivået som fremkommer for respondentenes sikkerhet i vurderinger ligger mellom 3,5 og 4,5 på måleskalaen som går fra 1 ("svært usikker") til 5 ("svært sikker").

vurderingspørsmål i spørreskjemaet, ikke får betydning for vurderinger foretatt av respondenter med høy produktkunnskap. Hovedkonklusjonen vi kan trekke ut av disse funnene er at man bør unngå inkludering av viktighetsspørsmål i spørreskjemaet dersom man vet at man står overfor respondenter med lav produktkunnskap.

FIGUR 9.2

Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet - oversikt over funn fra hypotesetestingen



Et funn som leder oss til å konkludere med at man, uavhengig av respondentenes produktkunnskap, bør unngå å måle viktighet ved ulike produkttegenskaper gjennom å inkludere viktighetsspørsmål i et spørreskjema, er et ønske om å måle spesifikke produktvurderinger på en måte som gir høyest mulig prediksjon på totalvurderinger. Funnene fra hypotesetestingen gir klare indikasjoner på en reduksjon i prediktiv evne når viktighetsspørsmål inkluderes i spørreskjemaet.

9.2.1 Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, og sammenheng mellom viktighet og vurderinger

Dersom vi sammenholder resultatene som illustreres i figur 8.5, for nivået på vurderingsspørsmålene, med nivået på viktighetsspørsmålene slik dette er illustrert i figur 8.7, fremkommer et interessant mønster. De produkttegenskapene som vurderes mest positivt ser også ut til å bli rapportert til å være de viktigste for respondentene når de skal vurdere et hotell. Selv om nivå på viktighet og vurderinger representerer gjennomsnitt for samtlige hotellgjester uavhengig av hvilket hotell de har benyttet, er dette mønsteret påfallende. Funnene fra disse analysene samsvarer imidlertid med tidligere studier der man har påvist sammenheng mellom rapportert viktighet og vurderinger av egenskaper ved et produkt (Antil, 1983). Hovedkonklusjonen fra disse studiene er at respondenter tenderer mot å gi mer ekstreme vurderinger til egenskaper som oppfattes som viktige, mens egenskaper som ikke oppfattes som viktige blir gitt vurderinger som ligger nær midten på vurderingsskalaen. Disse funnene forklares med respondentenes engasjement og involvering når de foretar vurderinger.

Når vi ønsker å studere forskjeller i sammenheng mellom rapportert viktighet ved ulike produkttegenskaper og vurdering av de samme egenskapene, avhengig av rekkefølgen på spørsmålene, må dette skje på individnivå. Resultatene fra hypotesetestingen viser at sammenhengen mellom viktighet og vurderinger på individnivå, er sterkest når viktighetsspørsmålene plasseres før vurderingsspørsmålene i spørreskjemaet. Dette gjelder imidlertid kun for respondenter med lav produktkunnskap. For respondenter med høy produktkunnskap finner vi ingen forskjell mellom eksperimentgruppene når det gjelder sammenhenger mellom viktighet og vurderinger. Selv om vi observerer forskjellene som er oppsummert i tabell 8.47 og tabell 8.48, er det verdt å merke seg at sammenhengene mellom viktighet og vurderinger er relativt svake. Z-transformerte gjennomsnittskorrelasjoner varierer fra 0,10 til 0,17. Dette indikerer at sammenhengene vi synes å observere ved å studere nivå på viktighet og vurderinger, ikke er så sterke som disse analysene kan gi inntrykk av.

Når vi ber respondenter vurdere et produkt, er det ønskelig at vurderingene som foretas i størst mulig grad reflekterer sanne vurderinger som ikke påvirkes av måleinstrumentet som benyttes. Vi har tidligere, i kapittel 3, diskutert hvordan viktighets-spørsmål kan påvirke svargivning på vurderingsspørsmål når viktighetsspørsmålene plasseres før vurderingsspørsmålene i spørreskjemaet. På samme måte kan vurderinger påvirke svargivning på viktighetsspørsmål dersom vurderingsspørsmålene plasseres før viktighetsspørsmålene. Feldman og Lynch (1988) benytter den oppfattede diagnose-

verdien av det første spørsmålet for det andre spørsmålet for å forklare hvordan fremkallingseffekter kan påvirke korrelasjoner mellom spørsmålene.

Oppfattet diagnoseverdi av viktighetsspørsmål (vurderingss spørsmål) på etterfølgende vurderinger (viktighetsspørsmål) kan med utgangspunkt i dette defineres som i hvilken grad en respondent oppfatter at svargivning på det første spørsmålet korrekt identifiserer hvordan det andre spørsmålet bør besvares.

Resultatene fra hypotesetestingen indikerer at viktighetsspørsmålene og vurderingss spørsmålene har symmetrisk diagnoseverdi for respondenter med høy produktkunnskap. For disse respondentene vil ikke rekkefølgen på spørsmålene ha betydning for korrelasjonene mellom spørsmålene. Dette innebærer at diagnoseverdien av viktighet for vurderinger er lik diagnoseverdien av vurderinger for rapportering av viktighet.

For respondenter med lav produktkunnskap finner vi at viktighetsspørsmålene og vurderingss spørsmålene har asymmetrisk diagnoseverdi. For disse respondentene blir korrelasjonen mellom viktighetsspørsmål og vurderingss spørsmål påvirket av rekkefølgen på spørsmålene. Resultatene indikerer at rapportering av viktighet har diagnoseverdi for vurderinger når viktighetsspørsmålene er plassert før vurderingss spørsmålene i spørreskjemaet, mens vurderingene ikke har den samme diagnoseverdien for rapportering av viktighet. Disse resultatene gir indikasjoner på at svargivning på viktighetsspørsmål fremkaller kunnskapsstrukturer som er midlertidig tilgjengelige og som påvirker de etterfølgende vurderingene for respondenter med lav produktkunnskap. Når respondenter med lav produktkunnskap svarer på viktighetsspørsmål kan dette føre til at etterfølgende vurderinger ikke foretas gjennom en kognitiv vurderingsprosess, men at de er basert på enklere beslutningsregler der viktighet definerer svargivning på vurderingene (Feldman og Lynch, 1988).

En hovedkonklusjon vi kan trekke ut av diskusjonen vi har foretatt ovenfor, er at inkludering-, og variasjoner i plasseringen av viktighetsspørsmål i et spørreskjema får betydning for respondentenes svargivning. En praktisk implikasjon av dette er at man som forsker bør være forsiktig med å inkludere viktighetsspørsmål før spesifikke vurderingss spørsmål i et spørreskjema der hovedformålet er å måle produktvurderinger. Dette gjelder spesielt dersom man står overfor respondenter med lav produktkunnskap. Måleprosedyren som benyttes kan her påvirke respondentenes svargivning, og dermed måling av de teoretiske begrepene vi er opptatt av.

9.2.2 Inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, og forklaring av totalvurdering og atferdsintensjoner

Før vi diskuterer forskjeller i vurderingsmodellens prediksjonsevne ved plassering av viktighetsspørsmål før vs. etter spesifikke vurderingss spørsmål, vil vi se nærmere på betydningen av å inkludere viktighetsspørsmål i spørreskjemaet. Respondenter tilhørende eksperimentgruppen som er tilordnet spørreskjema der viktighetsspørsmål er utelatt, blir i forbindelse med hypotesene, som omhandler de spesifikke vurderingss spørsmålenes evne til å forklare totalvurdering og atferdsintensjoner, betraktet som en kontrollgruppe. Uavhengig av hvor viktighetsspørsmålene plasseres, gir imidlertid resultatene fra hypotesetestingen klare indikasjoner på at inkludering av viktighetsspørsmål i spørreskjemaet vil redusere de spesifikke vurderingenes prediktive validitet. Disse funnene samsvarer med studier foretatt av Aiello, Czepiel, og Rosenberg (1977) og Sheth og Talarzyk (1972).

Sheth og Talarzyk (1972) undersøker hvorvidt verddiviktighet (*value importance*) eller oppfattet instrumentalitet (*perceived instrumentality*) er dominant for å bestemme holdninger til et produktmerke. Den konsistente reduksjonen i holdningsmodellens prediktive evne når oppfattet instrumentalitet veies med verddiviktighet fremheves, og det konkluderes med at verddiviktighet reduserer oppfattet instrumentalitets evne til å forklare et individs holdninger. Forfatterne gir imidlertid ingen teoretiske forklaringer på funnene. Aiello, Czepiel, og Rosenberg (1977) evaluerer ulike modeller for kundetilfredshet, og rapporterer konsistent høyere prediksjon av totaltilfredshet når det benyttes en enkel regresjonsmodell enn når det benyttes en lineær summert modell. På samme måte som i studien som ble foretatt av Sheth og Talarzyk (1972), konkluderes det også her med at veiing av egenskapsvurderingene med respondentenes selvrapporterte viktighet vil redusere vurderingenes korrelasjon med totaltilfredshet. Det blir heller ikke her gitt teoretiske forklaringer på funnene.

Basert på litteraturgjennomgangen vi har foretatt, har vi ikke funnet teoretiske begrunnelser for at inkludering av viktighetsspørsmål i et spørreskjema skal redusere de spesifikke vurderingenes evne til å forklare respondentenes totalvurdering og atferdsintensjoner. Imidlertid finnes en rekke metodologiske studier der man har undersøkt betydningen av antall spørsmål og lengde på spørreskjema, både for responsrate og datakvalitet (se f.eks. Burchell og Marsh (1992) og Sharp og Frankel (1983)). Hovedkonklusjonen fra disse studiene er at et spørreskjema ikke bør være for langt, da dette kan føre til redusert responsrate og økt feilvarians i målingene.

Inkludering av viktighetsspørsmål vil øke lengden på spørreskjemaet. Dette kan føre til at respondenter blir mindre villige til å svare på hele spørreskjemaet, at de unnlater å svare på enkeltspørsmål, eller at de avslutter svargivningen før spørreskjemaet er ferdig utfyllt. Inkludering av viktighetsspørsmål kan også føre til lavere kvalitet på svarene som blir gitt på spørsmål mot slutten av spørreskjemaet fordi respondentene kan bli utmattet og mindre motivert for å svare¹²². Uavhengig av plasseringen av viktighets-spørsmålene, kan det på bakgrunn av de metodologiske studiene vi har referert ovenfor, argumenteres for at inkludering av viktighetsspørsmål i spørreskjemaet kan bidra til å redusere de spesifikke vurderingenes prediktive validitet. Vi skal nå se nærmere på betydningen av å plassere viktighetsspørsmålene før vs. etter de spesifikke vurderings-spørsmålene i spørreskjemaet.

Uavhengig av om vi inkluderer respondentenes produktkunnskap som modererende variabel i analysene, har vi konkludert med at vurderinger av spesifikke produkt-egenskaper foretatt i spørreskjema der viktighetsspørsmål er plassert før vurderings-spørsmål, forklarer totalvurderinger bedre enn vurderinger av spesifikke egenskaper foretatt i spørreskjema der viktighetsspørsmål er plassert etter de spesifikke vurderings-spørsmålene. For respondentenes atferdsintensjoner finner vi ingen forskjeller mellom eksperimentgruppene.

La oss først se nærmere på mulige forklaringer på funnene fra hypotesetestingen der totalvurderinger betraktes som avhengig variabel. Som nevnt kan inkludering av viktighetsspørsmål i spørreskjemaet føre til reduksjon i kvaliteten på respondentenes svargivning som resultat av en *utmattelseeffekt*. Når viktighetsspørsmålene plasseres før de spesifikke vurderingss spørsmålene, tyder imidlertid funnene fra hypotesetestingen på at denne utmattelseeffekten kan være mindre enn det som er tilfellet dersom viktighetsspørsmålene plasseres etter de spesifikke vurderingss spørsmålene.

En alternativ forklaring på de samme funnene, kan være at plassering av viktighets-spørsmål før vurderingss spørsmål fører til fremkallingseffekter, som bidrar til å gjøre de spesifikke vurderingene mer tilgjengelige for respondentene. Resultatene fra hypotesetestingen samsvarer på denne måten med tidligere studier der tilgjengelighet til vurderinger er funnet å føre til sterkere sammenheng mellom vurderinger av spesifikke produkt-egenskaper og totalvurderinger (Fazio, 1989). Utmattelseeffekten vi trakk frem ovenfor som resultat av å inkludere viktighetsspørsmål i spørreskjemaet, kan imidlertid bidra til å redusere betydningen av fremkallingseffektene. Dette kan være en forklaring

¹²²Det finnes også argumenter for at det å øke lengden på spørreskjemaet kan øke respondentenes oppfatning av betydningen av studien, og at dette kan øke kvaliteten på svargivningen (Bradburn 1977)

på reduksjon i prediksjonsevne ved å inkludere viktighetsspørsmål før spesifikke vurderingsspmårl i spørreskjemaet.

Dersom viktighetsspørsmål plasseres etter de spesifikke vurderingsspmårlene, men før totalvurderingsspmårlene, kan det tenkes at viktighetsspørsmårlene har en *filler effekt* som bidrar til å redusere de spesifikke vurderingenes evne til å predikere totalvurderingene.

Når respondentenes atferdsintensjoner er det man ønsker å predikere ved hjelp av spesifikke produktvurderinger, kan det se ut som disse er mer robuste overfor plassering av viktighetsspørsmårlene i spørreskjemaet. Her finner vi ikke forskjeller mellom eksperimentgruppene. En mulig forklaring på disse funnene kan være at respondentenes intensjoner om å besøke det samme hotellet ved en senere anledning, og sannsynligheten for at de vil anbefale hotellet til andre, i mindre grad enn totalvurderingen av hotellet er relatert til de spesifikke vurderingene av hotellproduktet. Denne forklaringen støttes også av de relativt lave verdiene for forklart varians som rapporteres for atferdsintensjonene¹²³.

Det kan trekkes frem en rekke praktiske implikasjoner for utvikling av spørreskjema av funnene vi har rapportert og drøftet med hensyn på de spesifikke vurderingenes evne til å forklare respondentenes totalvurdering og atferdsintensjoner. Dersom man ønsker å kartlegge viktighet knyttet til produktenskapene som man ber respondenter vurdere i en spørreundesøkelse, kan det som en generell regel konkluderes med at dette bør gjøres på andre måter enn å inkludere viktighetsspørsmål i spørreskjemaet. Resultatene fra hypotesetestingen har vist at inkludering av viktighetsspørsmål reduserer de spesifikke vurderingenes prediksjonsevne. Alternative fremgangsmåter for kartlegging av viktighet er f.eks. bruk av multippel regresjonsanalyse, conjoint analyse, repertory grid og rangering av de ulike produktenskapene.

Dersom viktighetsspørsmål inkluderes i spørreskjemaet, har analysene vi har gjennomført resultert i funn som gir grunnlag for en rekke avveininger når vi skal diskutere implikasjoner for utvikling av spørreskjema.

Et viktig kriterium for å velge plassering av viktighetsspørsmål, vil være å velge den plasseringen av viktighetsspørsmål som fører til at de spesifikke vurderingene best evner å forklare respondentenes totalvurderinger og atferdsintensjoner. Som funnene fra hypotesetestingen viser vil vurderingenes prediktive evne for totalvurderingene være

¹²³Se tabell 8.33 og 8.39.

høyest når viktighetsspørsmålene er plassert før vurderingsspørsmålene i spørreskjemaet. Dette gjelder uavhengig av respondentenes produktkunnskap.

Prediksjonsevnen for respondentenes atferdsintensjoner ser i mindre grad ut til å bli påvirket av spørsmålsrekkefølgen. Atferdsintensjoner ser altså ut til å være mer robuste overfor konteksteffekter enn totalvurderinger. Vi har derfor valgt å konkludere med at man står friere til å velge plassering av viktighetsspørsmål når atferdsintensjoner betraktes som avhengig variabel.

Implikasjonene vi har fremsatt ovenfor, når det gjelder vurderingenes prediktive validitet, må avveies mot funnene vi har rapportert i forbindelse med viktighetsspørsmålenes diagnoseverdi for de spesifikke vurderingene, målt ved gruppegjennomsnitt av korrelasjon mellom viktighet og vurderinger på individnivå. For respondenter med lav produktkunnskap har vi konkludert med at inkludering av viktighetsspørsmål før vurderingsspørsmål i et spørreskjema kan påvirke respondentenes vurderinger, og dermed måling av de teoretiske begrepene vi er opptatt av. Selv om vi finner signifikant forskjellige gjennomsnittskorrelasjoner mellom eksperimentgruppene, er korrelasjonene imidlertid i utgangspunktet små. Dette velger vi å tolke som at problemet med diagnoseverdi av viktighetsspørsmål for vurderinger er lite, men at det kan være noe man bør ta hensyn til for respondenter med lav produktkunnskap¹²⁴.

9.2.3 Inkludering og alternativ plassering av viktighetsspørsmål i et spørreskjema, og nivå på vurderinger og viktighetsspørsmål

Litteraturgjennomgangen vi har foretatt i forbindelse med avhandlingen, har vist at de fleste studiene som er gjennomført med hensyn på effekter av spørsmålsrekkefølge i spørreundersøkelser har fokusert på univariate svarfordelinger (se f.eks. Bickart et al. (1992), Bickart (1991), Bickart (1993), Hippler og Schwarz (1987), Hippler, Schwarz og Sudman (1987), Schuman, Kalton og Ludwig (1983), Schuman og Presser (1981), Schuman, Presser og Ludwig (1981), Schwarz og Bless (1992), Schwarz, Strack og Mai (1991), Schwarz og Sudman (1991)). Til tross for alle disse studiene, er det fortsatt en rekke uklarheter i forbindelse med når konteksteffekter vil oppstå, og når de oppstår hvilken retning og styrke effektene vil ha. Dette har ført til at vi har valgt å betrakte

¹²⁴For en diskusjon av hvordan man kan ta hensyn til respondentenes produktkunnskap ved gjennomføring av en spørreundersøkelse, se kapittel 9.1.1.

problemstillinger i tilknytning til generelle forskjeller i nivå på vurderingsspmårl og viktighetsspmårl som eksplorative.

La oss først se nærmere på resultater fra analyser av nivået på vurderingsspmårlene. De innsamlete dataene gir klare indikasjoner på at inkludering og alternativ plassering av viktighetsspmårl, før vs. etter spesifikke vurderingsspmårl, ikke får betydning for nivå på vurderingene.

Dersom vi antar at manipuleringen vi har foretatt ved å inkludere og variere plassering av viktighetsspmårl i spørreskjemaet har virket må vi konkludere med at nivå på vurderinger ikke påvirkes av svargivning på viktighetsspmårl. En mulig forklaring på at vi ikke finner forskjeller i nivå på vurderingsspmårlene, kan da være at vurderinger av hotellproduktet i stor grad er tilgjengelige for respondentene, og at denne tilgjengeligheten ikke påvirkes i nevneverdig grad gjennom å plassere viktighetsspmårl før vurderingsspmårlene i spørreskjemaet. Dersom dette er tilfellet finner resultatene som rapporteres for nivået på vurderingsspmårlene støtte i litteratur om konteksteffekter i spørreundersøkelser (se f.eks. Tourangeau og Rasinski (1988)). I henhold til denne litteraturen vil konteksteffekter være mindre sannsynlige dersom vurderingene er tilgjengelige for respondentene. I følge den samme litteraturen vil konteksteffekter også være mindre sannsynlige for vurderinger av produkter som respondentene er kjent med. Selv om et hotellprodukt på mange måter kan karakteriseres som komplekst ved at det består av en rekke produktegenskaper, kan vurderinger av de ulike egenskapene likevel være relativt enkle å foreta for respondentene. Dette kan også ha bidratt til at vi ikke finner forskjeller i nivå på vurderingene når viktighetsspmårl inkluderes i spørreskjemaet.

Assimilasjons-kontrastteori kan benyttes for å belyse problemstillingene vi fokuserer på når det gjelder plassering av viktighetsspmårl og vurderingsspmårl i spørreskjemaet. Kontrasteffekter ble i kapittel 3 definert til å oppstå dersom svargivning på viktighetsspmårl aktiverer standarder for vurderinger som er svært forskjellige fra det produktet som respondentene blir bedt om å vurdere i spørreskjemaet. Assimilasjons-effekter kan oppstå dersom svargivning på viktighetsspmårlene påvirker respons på vurderingsspmårlene på en slik måte at vurderinger tilnærmes nivået på den aktiverte standarden. Kvalitetsnivået på produktene som vurderes vil med bakgrunn i disse perspektivene kunne få betydning for hvilke effekter som oppstår når viktighets-spmårl inkluderes og plasseres før de spesifikke vurderings spmårlene i spørreskjemaet.

Når vi heller ikke finner forskjell i nivå på vurderinger av produkter som representerer ulike kvalitetsnivåer, for respondenter som har benyttet spørreskjema der viktighets-spørsmål henholdsvis er ekskludert og inkludert før vurderingsspørsmålene, kan en mulig forklaring være at svargivningen på vurderingsspørsmålene *redigeres* etter at vurderingene er foretatt (Clark og Schober, 1991; Strack og Martin, 1987; Tourangeau og Rasinski, 1988). Når opplevelsen av et produkt avviker fra forventninger i negativ retning kan det argumenteres for at vurderingene av produktet, slik disse f.eks. kommer til uttrykk i et spørreskjema, vil bli oppjustert. Når respondentene vurderer hotell med kvalitet lavere enn gjennomsnittet kan resultatene for testing av hypotese 10a, der vi ikke finner forskjeller i vurderinger mellom eksperimentgruppene, nettopp skyldes individers ønske om å fremstille seg selv på en positiv måte. Kontrasteffekten, som vi antar vil oppstå ved å inkludere viktighetsspørsmålene ved vurdering av lavkvalitetsprodukter, kan dermed bli opphevet av respondentenes ønske om å redusere kognitiv dissonans (Festinger, 1957).

Forsøk på å redusere kognitiv dissonans kan relateres til hvilke årsaker eventuelle positive og negative vurderinger tilskrives av respondentene (Folkes, 1988; Kelley, 1973; Mizersky and Kernan, 1979). Attribusjonsteori kan bidra til å gi en forklaring på hvorfor vi ikke finner støtte for hypotese 10c i de innsamlete dataene. En respondent som har benyttet et høykvalitetshotell antas også ha positive vurderinger av hotellet. Årsaker til positive vurderinger vil ofte være interne, dvs. gjesten tilskriver de positive vurderingene til det at han har gjort et godt valg av hotell. Sammenligningsstandarden vi antar aktiveres av viktighetsspørsmålene kan dermed få mindre betydning for høykvalitetsproduktene.

I forbindelse med diskusjonen av hypotese 3a, 3b og 3c ble en alternativ ikke-teoretisk forklaring på manglende støtte for hypotesene trukket frem. I den gjennomførte studien har vi med vurderinger foretatt av gjester ved 11 hotell fra Inter Nor kjeden. Funnene som ble rapportert for hypotese 3a, 3b og 3c kan tyde på at vi ikke har klart å fange opp nivåforskjeller i kvalitet mellom de 11 hotellene, som er store nok til at betegnelsene høykvalitetshotell, gjennomsnittshotell og lavkvalitetshotell er berettiget. Dette forholdet kan bidra til å gi en undersøkelsesmetodologisk forklaring på hvorfor vi ikke finner støtte for hypotesene der kvalitetsnivå inngår som modererende variabel.

Basert på de innsamlete dataene og diskusjonen ovenfor har vi valgt å konkludere med at inkludering av viktighetsspørsmål i spørreskjemaet ikke påvirker nivået på etterfølgende vurderinger. De praktiske implikasjonene vi kan trekke ut av funnene med hensyn på nivå på vurderinger, for utvikling av spørreskjema til bruk ved måling av

produktvurderinger, er dermed at forskeren står fritt til å velge om viktighetsspørsmål skal inkluderes før vurderingss spørsmål i spørreskjemaet.

Konklusjonene ovenfor er imidlertid kun gyldige dersom kartlegging av nivå på vurderingene er hovedformålet med undersøkelsen¹²⁵. De eksplorative analysene gir klare indikasjoner på at nivået på viktighetsspørsmålene er høyere når disse plasseres etter de spesifikke vurderingene i spørreskjemaet. Analyser som vi ikke har gjengitt i avhandlingen, viser i tillegg at standardavvikene på individnivå i gjennomsnitt er lavere når viktighetsspørsmålene plasseres etter vurderingene¹²⁶. Spesielt gjelder dette for struktur-, backstage- og de kundeavhengige egenskapene. For samhandlingsegenskapene finner vi ikke forskjell i nivå på respondentenes rapportering av viktighet. En praktisk implikasjon av disse funnene kan være at man, i en undersøkelse der hovedformålet er kartlegging av ulike dimensjoner ved service, står friere til å velge hvor man vil plassere viktighetsspørsmål i spørreskjemaet.

Når undersøkelsen også omfatter de øvrige sidene ved produktet blir situasjonen en annen. Høyere nivå på rapportering av viktighet og lavere standardavvik gjør dataene dårligere egnet som grunnlag for segmentering. De praktiske implikasjonene for utvikling av spørreskjema blir i dette tilfellet at viktighetsspørsmålene bør plasseres før de spesifikke vurderingss spørsmålene i spørreskjemaet.

Dersom respondentene svarer på viktighetsspørsmål i et spørreskjema etter at de har foretatt en vurdering av et produkts egenskaper, argumenterte vi for at assimilasjonseffekter kan oppstå som et resultat av at respondentene kan legge mer vekt på egenskaper som vurderes ekstremt positivt eller ekstremt negativt. For å fremstå som konsistent i svargivning kan en egenskap som av respondenten er vurdert som svært bra eller dårlig, kunne bli tillagt høyere vekt enn den ville blitt dersom vurderingen ikke hadde vært foretatt. Dette kan være en mulig forklaring på resultatet vi observerer for nivået på viktighetsspørsmålene når disse plasseres etter de spesifikke vurderingss spørsmålene i spørreskjemaet. Dersom dette er tilfellet vil responseffekten kunne knyttes til *redigering av respons*, slik dette er beskrevet i modellen for respondentenes informasjonsbehandlingsprosess i forbindelse med svargivning på spørsmål i en spørreundersøkelse i kapittel 3 (se figur 3.6). Disse perspektivene kan også forklare hvorfor vi ikke finner forskjeller for nivå på viktigheten av samhandlingsegenskapene. Nivået på viktigheten av samhandlingsegenskaper er høyere

¹²⁵Dersom kartlegging av nivå på vurderinger er hovedformålet med en undersøkelse kan det hevdes at viktighetsspørsmål ikke bør inkluderes i spørreskjemaet. Dette vil gjøre spørreskjemaet kortere og kan bidra til å øke både responsrate og kvalitet på innsamlete data.

¹²⁶Dette kan være resultat av *ceiling* effekter.

enn nivået på de øvrige egenskapene uavhengig av plasseringen av viktighets-spørsmålene. Det er derfor mindre sannsynlig at en assimilasjonseffekt vil oppstå for samhandlingsegenskapene¹²⁷.

En alternativ forklaring på forskjeller i nivå på viktighetsspørsmål kan være at svargivning på spesifikke vurderingsspørsmål i spørreskjemaet aktiverer informasjon som påvirker respondentenes oppfatning av viktighet knyttet til produktegenskapene. Dersom dette er tilfellet vil responseffektene være knyttet til *dannelse av oppfatning* og *vurderingsfasen* i informasjonsbehandlingsprosessen respondenten kan tenkes å gå gjennom ved svargivning på spørsmål i spørreskjemaet. Vi har i kapittel 3 argumentert for at viktigheten av spesifikke produktegenskaper kan knyttes til respondentenes preferanser, mens vurderingene i større grad er knyttet til objektet som vurderes. Dersom viktighetsspørsmålene stilles før vurderingsspørsmålene kan det argumenteres for at rapportering av viktighet forankres i preferansene, og at svargivningen i spørreskjemaet reflekterer disse. Når viktighetsspørsmålene plasseres etter vurderingene kan det tenkes at tillegging av viktighet i større grad reflekterer vurderingene av objektet. Forankringen som benyttes av respondentene ved svargivning på viktighetsspørsmålene kan dermed bli forandret.

Til tross for konklusjonene vi har fremsatt ovenfor, kan det finnes situasjoner der det vil være ønskelig å inkludere viktighetsspørsmål i det samme spørreskjemaet som benyttes for måling av produktvurderingene. Blant annet vil det i produktutviklingssammenheng være ønskelig å vite noe om hvilke egenskaper ved produktet som har betydning for at kunder skal velge å kjøpe produktet. Dette kan f.eks. være tilfellet der vi ønsker å kartlegge hvordan et produkt vurderes forskjellig avhengig av hva kundene legger vekt på. Viktighetsspørsmål kan her benyttes som grunnlag for inndeling av respondentene i segmenter¹²⁸. Kartlegging av viktighet ved bruk av viktighetsspørsmål i spørreskjemaet er en av flere mulige fremgangsmåter for å fremskaffe slik informasjon¹²⁹. Hvilken fremgangsmåte man velger å benytte må vurderes i forhold til formålet med den spesifikke undersøkelsen som skal gjennomføres.

Funnene vi har diskutert ovenfor fører til at vi må revurdere forbeholdene som fremsettes i litteraturen om spørreskjemaforskning (f.eks. Ayidiya og McClendon (1990) og Strack (1991)) angående forskerens kontroll over rekkefølgen spørsmål blir lest og

¹²⁷En alternativ forklaring kan være en *ceiling* effekt.

¹²⁸En slik inndeling kan f.eks. foretas ved hjelp av clusteranalyse

¹²⁹Andre fremgangsmåter for å kartlegge betydning av ulike produktegenskaper for valg av produktalternativ er f.eks. conjoint analyse, repertory grid og rangering av produktegenskaper. Alternativt kan multipl regresjonsanalyse benyttes for å estimere relativ betydning av ulike produktegenskaper for totalvurderinger og atferdsintensjoner.

besvart, som grunnlag for å avdekke rekkefølgeeffekter. Studien som er gjennomført i forbindelse med denne avhandlingen gir indikasjoner på at respondenter kan være mer sekvensielle i sin svargivning enn man tidligere har antatt, og at rekkefølgeeffekter derfor også kan oppstå når spørsmålsrekkefølgen varieres i et skriftlig spørreskjema. Når det finnes grunnlag for å anta at tilstedeværelse av-, og svargivning på spørsmål i et spørreskjema kan påvirke svargivning på etterfølgende spørsmål, bør man som forsker alltid tenke gjennom hvilke effekter spørsmålsrekkefølgen kan få. Dersom man unnlater å inkludere viktighetsspørsmål, vil lengden på spørreskjemaet bli redusert. Vi har tidligere diskutert hvordan dette kan bidra til å øke både responsrate og kvalitet på svarene som blir gitt på spørsmålene i undersøkelsen. Gjennom å ta hensyn til muligheten for at rekkefølgeeffekter kan oppstå ved utforming og testing av spørreskjemaet, før dette benyttes som datainnsamlingsverktøy i en spørreundersøkelse, kan utilsiktete responseffekter unngås.

Som en oppsummering av implikasjonene for utvikling av spørreskjema til bruk ved måling av kunders produktvurderinger, velger vi å konkludere med at viktighets-spørsmål, som en generell regel, ikke bør inkluderes i spørreskjemaet dersom man har mulighet for å kartlegge viktighet knyttet til ulike produktegenskaper på andre måter. Dersom man til tross for dette velger å inkludere viktighetsspørsmål i spørreskjemaet, har vi, basert på en samlet vurdering, valgt å konkludere med at viktighetsspørsmålene bør plasseres før de spesifikke vurderingss spørsmålene.

10

Begrensninger og implikasjoner for videre forskning

Utgangspunktet for denne avhandlingen har vært å fokusere på to valg man står overfor når et spørreskjema skal benyttes for å måle kunders produktvurderinger. For å belyse konsekvenser av disse valgene har vi gjennomført en studie der også vi har stått overfor valg. Først måtte vi velge hvilke perspektiver vi skulle benytte for å belyse problemstillingene vi har reist. Deretter måtte vi velge et undersøkelsesopplegg som gjorde oss i stand til å teste hypotesene som ble formulert basert på de teoretiske perspektivene. Hypotesene er formulert på teoriplan. Som grunnlag for å foreta hypotesetesting var det derfor nødvendig å operasjonalisere en rekke begreper, som f.eks. vurderinger, atferdsintensjoner, viktighet, sikkerhet i vurderinger og produkt-kunnskap. Operasjonalisering innebærer valg med hensyn på hvilke spørsmål i et spørreskjema vi lar representere de teoretiske begrepene. I forbindelse med operasjonaliseringen er det også nødvendig å bestemme hvilken empirisk setting de teoretiske begrepene skal knyttes til. Hypoteser kan ikke testes uten at det samles inn data. Dette innebærer valg, både av vurderingsobjekt/produkt og av subjekter/ respondenter. Hvorvidt de innsamlede dataene er egnet for testing av hypotesene avhenger av valgene vi har foretatt med hensyn på utvalg og utvalgsprosedyrer. Konklusjonene vi trekker på bakgrunn av hypotesetestingen er betinget av at de statistiske analysemetodene vi har valgt er egnet for å gjennomføre en kritisk testing av hypotesene. Dette innebærer blant annet at vi ikke har gjort alvorlige brudd på forutsetningene for å benytte analysemetodene. Til slutt er de praktiske implikasjonene vi har fremmet for utvikling av spørreskjema, et resultat av tolkning av funnene fra hypotesetestingen.

Alle disse valgene har fått betydning for gyldigheten av de konklusjonene vi har trukket, og for de praktiske implikasjonene vi har fremhevet. I dette kapittelet vil vi

diskutere begrensninger i den gjennomførte studien, og fokusere på implikasjoner for videre forskning.

10.1 Valg av perspektiver for å belyse problemstillingene

Selv om de metodologiske valgene vi har fokusert på i avhandlingen er knyttet til utvikling av spørreskjema til bruk ved måling av kunders produktvurderinger, kan valgene betraktes som generelle i den forstand at de er knyttet til ordlyd i spørsmålstilling og svaralternativer og rekkefølge på spørsmålene i et spørreskjema. Når man står overfor slike valg, enten som forsker eller som praktiker, er det naturlig å starte med å søke hjelp i metodelitteraturen. Dette innebærer i første rekke litteratur om måling, ulike skalatyper, spørreskjema utvikling og responseffekter. I kapittel 2 ble en rekke bidrag fra denne litteraturen gjennomgått og drøftet, uten at vi var i stand til å gi spesifikke retningslinjer for de metodologiske valgene. Vi konkluderte med å relatere de metodologiske bidragene til *Stimulus - Organisme - Respons* (S-O-R) modellen. De forskjellige studiene bærer preg av at man har observert forskjeller i svargivning (respons) ved endringer i spørreskjemaet (stimulus). Man har i liten grad relatert endringer i svargivning til respondentens informasjonsbehandlingsprosess (organisme) ved utfylling av spørreskjemaet. En slik kognitiv tilnærming til studier av responseffekter i spørreundersøkelser er imidlertid, i nyere spørreskjemaforskning, belyst av blant annet Hippler, Schwarz, og Sudman (1987) og Tourangeau og Rasinski (1988). Dette ledet oss til å rette oppmerksomheten mot litteratur som beskriver hvordan kunder foretar produktvurderinger som grunnlag for å forstå hvordan valg av ordlyd i spørsmålsstilling, svaralternativer og spørsmålsrekkefølge kan påvirke respondentenes svargivning på vurderingsspørsmål i et spørreskjema.

De teoretiske perspektivene vi har benyttet, har fått konsekvenser for hypotesene som ble formulert for å presisere implikasjoner av de metodologiske valgene. Selv om bruk av alternative teoretiske perspektiver kunne ha ledet oss til å fokusere på andre implikasjoner, vil vi hevde at den teoretiske argumentasjonen vi har benyttet som grunnlag for å formulere hypoteser, representerer en av de sterke sidene ved avhandlingen. Med utgangspunkt i respondentens informasjonsbehandlingsprosess ved svargivning på spørsmål i et spørreskjema, har det vært naturlig å trekke inn bidrag fra holdningsteori og litteratur om kundetilfredshet, i tillegg til en rekke relaterte områder,

for å belyse problemstillingene. Der vi har vært i stand til å formulere empirisk testbare hypoteser har dette dannet et bredt teoretisk grunnlag for hypotesene.

Som nevnt, har man i senere studier av metodologiske valg knyttet til utforming av spørreskjema og gjennomføring av spørreundersøkelser, rettet oppmerksomheten mot respondentens informasjonsbehandlingsprosess ved svargivning, som teoretisk utgangspunkt for å forstå ulike typer responseffekter (se f.eks. Hippler, Schwarz, og Sudman (1987) og Tourangeau og Rasinski (1988)). Denne dreiningen bort fra de ad-hoc pregete kodingene mellom undersøkelsesmetodologi og ulike teoriretninger, som tradisjonelt kjennetegner forskningen som er knyttet til spørreundersøkelser (Groves, 1987), kan være begynnelsen på utvikling av en *teori for gjennomføring av spørreundersøkelser*.

I fremtidige studier bør perspektivene fra informasjonsbehandlingslitteraturen benyttes som grunnlag for å fokusere på de ulike oppgavene en respondent kan tenkes å utføre ved svargivning på spørsmål i en spørreundersøkelse. Responseffekter som følge av metodologiske valg i forbindelse med spørreskjemaautvikling, vil med et slikt utgangspunkt, i større grad kunne relateres til ulike faser i svargivningsprosessen. Dette vil imidlertid kreve andre metodologiske tilnærminger enn de vi har valgt for å belyse problemstillingene vi har tatt opp i avhandlingen. Valg av undersøkelsesopplegg drøftes i neste avsnitt.

10.2 Valg av undersøkelsesopplegg

Som grunnlag for valg av undersøkelsesopplegg har avveininger mellom krav til intern og ekstern validitet (Cook og Campbell, 1979) fått stor betydning. Krav til ekstern validitet tilsier at data bør innsamles på en så realistisk måte som mulig, slik at resultatene kan generaliseres utover en spesifikk undersøkelsessituasjon, mens krav til intern validitet tilsier at datainnsamlingen bør foretas på en slik måte at vi kan utelukke alternative forklaringer. Når vi har valgt å benytte en feltstudie der spørreskjema er utdelt til hotellgjester, favoriseres kravet til ekstern validitet. Undersøkelsessituasjonen er også representativ for en typisk situasjon man som forsker står overfor når et spørreskjema skal benyttes for å måle kundens produktvurderinger¹³⁰.

Når vi har anvendt kognitive perspektiver som grunnlag for å forklare implikasjoner av metodologiske valg ved utvikling av spørreskjema, har vi understreket at de teoretiske

¹³⁰Det valgte undersøkelsesopplegget er også typisk for de fleste metodologiske studiene vi har gjennomgått i kapittel 2.

perspektivene i seg selv ikke underlegges testing. Formålet har vært å studere hvorvidt implikasjonene de kognitive perspektivene forskriver inntreffer. Dette er en viktig presisering i forbindelse med det valgte undersøkelsesopplegget. Kravet til intern validitet kan dermed diskuteres med to forskjellige utgangspunkt.

For det første kan intern validitet diskuteres i forhold til de teoretiske forklaringsvariablene som ble benyttet for å predikere effekter av de metodologiske valgene vi har studert i forbindelse med utvikling av spørreskjemaet. Det kan hevdes at vi ikke kan være sikker på at de teoretiske begrunnelsene vi har benyttet som grunnlag for hypotesene, faktisk er de underliggende årsakene til forskjellene vi observerer i respondentenes svargivning. Selv om vi har formulert hypoteser basert på et relativt bredt teorigrunnlag, kan det finnes alternative teorier med andre variabler som kan foreskrive de samme effektene. Det er derfor vanskelig med grunnlag i den foreliggende studien, å utelukke alternative teoretiske forklaringer der funnene fra hypotesetestingen støtter hypotesene vi har formulert. Dette har heller ikke vært formålet med studien.

En viktig implikasjon av begrensningene vi har påpekt ovenfor, er at man i fremtidige studier av problemstillingene vi har tatt opp i denne avhandlingen, bør velge undersøkelsesopplegg som gjør det mulig å etterspore effekter på ulike trinn i respondentenes informasjonsbehandlingsprosess. Dette kan oppnås gjennom alternative kvalitative undersøkelsesopplegg.

En mulig fremgangsmåte vil være å benytte en *protokoll* tilnærming (se f.eks. Douglas, Craig, og Faivre (1981)). En tradisjonell protokoll tilnærming vil innebære at man ber respondenter *tenke høyt* mens de fyller ut spørreskjemaet. Dette kan gjøre det mulig å tilskrive responseffekter, både i forbindelse med valg av instruksjon og vurderingsskala, og i forbindelse med inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet, til ulike faser i respondentenes informasjonsbehandlingsprosess ved svargivningen. En undersøkelse der man ber respondenter tenke høyt mens de svarer på spørsmålene i et spørreskjema, kan gjennomføres, både som et felteksperiment og som en laboratoriestudie. Hvorvidt man velger felteksperiment fremfor laboratoriestudie, må være et resultat av avveininger mellom krav til intern og ekstern validitet.

Intern validitet kan også diskuteres i forhold til den eksperimentelle manipuleringen som er foretatt i studien vi har gjennomført. I kapittel 5 presiserte vi betydningen av å kunne tilskrive responseffekter til forskjeller i instruksjon og vurderingsskala, og inkludering og alternativ plassering av viktighetsspørsmål i spørreskjemaet. Selv om vi gjennom redegjørelsen for den praktiske gjennomføringen av datainnsamlingen, har

sannsynliggjort at det er tilfeldig hvilke hotellgjester som er tilordnet en bestemt type spørreskjema, kan en laboratorietilnærming, eventuelt et klasseromseksperiment, øke studiens interne validitet. Vi vil da ha større kontroll over randomisering av tilordning av subjekter til de ulike eksperimentrundene.

10.3 Valg av vurderingsobjekter

Det kan diskuteres hvorvidt hotellproduktet er egnet som objekt for en studie der vi ønsker å belyse problemstillingene vi har tatt opp i avhandlingen. En rekke argumenter kan fremsettes både for og i mot dette valget av vurderingsobjekter.

En viktig begrunnelse for valg av hotell fra Inter Nor kjeden som objekter for den gjennomførte studien, er at hotellproduktet tilfredsstillende de kravene hypotesene stiller til vurderingsobjekter bestående av både materielle og immaterielle egenskaper. Videre bidrar gjennomføring av undersøkelsen i en hotellokjede, til muligheten for å studere forskjeller i vurderinger av sammenlignbare produkter med ulikt kvalitetsnivå. Selv om det ikke ser ut til at vi har oppnådd denne målsetningen i den foreliggende studien, kan dette skyldes at det ikke er stor nok variasjon i kvalitetsnivå mellom de 11 hotellene som deltok i studien, mer enn at hotell er valgt som vurderingsobjekter.

Et viktig forhold som taler for at det i en eventuell oppfølgingsstudie bør velges andre produkter som vurderingsobjekt, er at hotellproduktet representerer det vi kan kalle et typisk lav-involveringsprodukt (Breivik, 1995). Resultatene fra hypotesetestingen viser funn som i de fleste tilfeller er svært signifikante, men der de observerte forskjellene mellom eksperimentgruppene er små. Det kan diskuteres hvorvidt de metodologiske valgene vi har studert ville ha fått større betydning dersom vi hadde valgt et mer typisk høy-involveringsprodukt som vurderingsobjekt.

I litteraturen om konteksteffekter i spørreundersøkelser argumenteres det for at høyere involvering, på samme måte som produktkunnskap, vil redusere betydningen av spørsmålsekkefølge (Tourangeau og Rasinski, 1988). For de spesifikke rekkefølgeproblemmstillingene vi har tatt opp i avhandlingen, impliserer dette at de forventede effektene vil være størst, nettopp i den settingen vi har valgt for den gjennomførte studien. Det finnes med andre ord ikke grunnlag for å forvente større effekter dersom det velges vurderingsobjekter der respondentene antas å være mer involvert i vurderingene, for disse problemstillingene. Vi har i kapittel 3 argumentert for at subjektforankrede instruksjoner og -vurderingsskalaer i større grad enn objektforankrede instruksjoner og

vurderingsskalaer involverer respondentene i vurderingene. For problemstillingene i tilknytning til valg av instruksjon og vurderingsskala, kan vi derfor forvente mer markante forskjeller mellom eksperimentgruppene, dersom undersøkelsen gjennomføres i forbindelse med produkter der respondentene er mer involvert i vurderingene.

Et alternativ til hotellproduktet kan f.eks. være kjøp av bil, der det finnes empirisk belegg for at respondentene vil være mer involvert i vurderingene (Breivik 1995). Bilforhandlere som forhandler et bestemt bilmerke vil tilfredsstillende kravene hypotesene stiller, både til et vurderingsobjekt bestående av immaterielle og materielle egenskaper, og til sammenlignbare produkter med ulikt kvalitetsnivå.

10.4 Valg av operasjonaliseringer

I kapittel 6 ble det gjort grundig rede for operasjonalisering av avhandlingens avhengige og uavhengige variabler. I forbindelse med valgene vi har foretatt i tilknytning til operasjonalisering, vil vi her trekke frem tre forhold som har fått betydning for konklusjonene vi har trukket på bakgrunn av de innsamlede dataene; operasjonalisering av subjektforankret- og objektforankret instruksjon og -vurderingsskala, respondentenes produktkunnskap og kvalitetsnivået på produktene som vurderes.

Mono-operasjonalisering ble diskutert som et mulig problem i forbindelse med operasjonalisering av subjektforankrete- og objektforankrete vurderinger. Selv om vi konkluderte med at det ikke finnes noe substitutt for multiple operasjonaliseringer av eksperimentvariablene, ble det besluttet å avgrense studien til å se nærmere på forskjeller mellom "bra-dårlig" og "tilfreds-utilfreds", som operasjonaliseringer av henholdsvis objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer. En slik beslutning har flere årsaker. For det første argumenterte vi for å benytte det vi valgte å kalle ytterpunktene i objektforankrete og subjektforankrete vurderinger. Ytterpunktene ble valgt for å maksimere forskjellene mellom subjektforankrete- og objektforankrete vurderinger. Et annet argument som ble trukket frem for å benytte kun en operasjonalisering av de eksperimentelle årsaksvariablene er begrensninger i mulighetene for å gjennomføre studien innenfor de tids- og kostnadsrammer vi har til rådighet. Økning av antall operasjonaliseringer av objektforankrete- og subjektforankrete instruksjoner og vurderingsskalaer, ville medført behov for en økning av nødvendig antall eksperimentgrupper. Alternative operasjonaliseringer av objektforankrete- og subjektforankrete instruksjoner og -vurderingsskalaer bør imidlertid be-

nyttes i senere studier. En slik tilnærming vil øke generaliserbarheten av funnene i den gjennomførte studien, både over tid og på tvers av studier.

Respondentenes produktkunnskap har stått sentralt som et forhold som kan moderere effektene av de metodologiske valgene vi har tatt opp i avhandlingen. Produktkunnskap er operasjonalisert som et multidimensjonalt begrep, bestående av tre dimensjoner: Objektive mål på kunnskap, subjektive selvrapporterte mål på kunnskap og frekvensmål på tidligere kjøps- og/eller bruksatferd (Brucks, 1985; Brucks, 1986; Cole, Gaeth, og Singh, 1986). Disse tre målene på produktkunnskap danner grunnlag for å dele respondentene inn i tre grupper som representerer hotellgjester med lav, middels høy produktkunnskap. For identifisering av de tre gruppene ble det benyttet clusteranalyse. Det kan diskuteres hvorvidt denne måten å operasjonalisere produktkunnskap har påvirket resultatene fra testing av hypoteser der produktkunnskap inngår som modererende variabel. En indikasjon på godheten av produktkunnskapsmålet, kan være at effektene som fremkommer ved testing av alternative hypoteser der produktkunnskapsmålet inngår, synes å være i samsvar med effekter foreskrevet i de teoretiske perspektivene som danner grunnlag for hypotesene. Produktkunnskapsmålet oppnår dermed en viss gyldighet gjennom teoretisk konsistente resultater på tvers av hypoteser¹³¹.

I tillegg til respondentenes produktkunnskap, er kvalitetsnivået på produktet som vurderes inkludert som modererende variabel i hypotesene. Vi har tidligere påpekt at hotellene som deltok i den gjennomførte studien ikke i tilstrekkelig grad synes å tilfredsstille kravet til sammenlignbare produkter med ulikt kvalitetsnivå. Dette kan ha ført til at vi ikke finner støtte for hypotesene der kvalitetsnivået på produktet som vurderes er inkludert som modererende variabel.

En implikasjon av dette vil være at man, i senere studier av problemstillingene vi har tatt opp i avhandlingen, bør sørge for tilstrekkelig variasjon i kvalitetsnivå på vurderingsobjektene. Et alternativ til å benytte gjennomsnittlige totalvurderinger som grunnlag for å gruppere hotellene i lavkvalitets-, gjennomsnittskvalitets- og høykvalitetshotell, er å benytte ekspertvurderinger. Gruppering av hotellene ved hjelp av ekspertvurderinger kan benyttes, både som grunnlag for å sikre variasjon i kvalitetsnivå før man velger hvilke produkter som skal delta i undersøkelsen, og som selvstendig utgangspunkt for testing av hypoteser der kvalitetsnivå inngår som modererende variabel.

¹³¹Nomologisk validitet

10.5 Valg av utvalg og utvalgsprosedyrer

Utdeling av spørreskjema på hotellene ble foretatt av resepsjonspersonalet. Selv om det ble gitt en grundig instruksjon til disse i forbindelse med utsending av spørreskjema, har vi ikke hatt kontroll med hvordan utdelingen har foregått på de enkelte hotellene. I kapittel 5 gjorde vi grundig rede for fremgangsmåten vi benyttet for å sikre at det er tilfeldig hvilke gjester som fikk en bestemt type spørreskjema. Basert på oppnådde utvalg i de seks eksperimentgruppene argumenterte vi for at tilordning av spørreskjema synes å ha tilfredsstillt kravet om randomisering. Selv om utdelingen av spørreskjemaene ser ut til å ha fungert, kan vi ikke kontrollere for eventuelle problemer knyttet til selvseleksjon. Vi kan således ikke vite om det er spesielt misfornøyde eller fornøyde gjester som har svart.

Implikasjonene av disse begrensningene er at det i videre forskning bør benyttes prosedyrer som sikrer større kontroll over datainnsamlingen. Vi har tidligere drøftet nødvendigheten av å benytte kvalitative tilnærminger som grunnlag for ettersporing av responseffekter til ulike faser i respondentenes informasjonsbehandlingsprosess. Selv om dette vil være nødvendig for å styrke intern validitet knyttet til de teoretiske forklaringsvariablene som benyttes, vil det fortsatt være nødvendig å gjennomføre eksperimenter der spørreskjema benyttes på tilsvarende måte som i studien vi har gjennomført. Data kan som nevnt samles inn ved bruk av spørreskjema i et laboratorieeksperiment¹³². Vi vil da ha større kontroll over utvalg og utvalgsprosedyrer. Avveininger mellom krav til intern og ekstern validitet må imidlertid stå sentralt i forbindelse med slike valg.

Ytterligere en begrensning som kan trekkes frem i forbindelse med datainnsamlingen er bruk av spørreskjema der respondentene selv leser og svarer på spørsmålene. Problemstillingene vi har tatt opp vedrørende inkludering og alternativ plassering av viktighetsspørsmål, før vs. etter spesifikke vurderingss spørsmål, betinger at vi har kontroll, både over rekkefølgen spørsmålene blir presentert, og rekkefølgen på respondentenes svargivning (se f.eks. Ayidiya og McClendon (1990) og Strack (1991)). Dette vil langt på vei være tilfellet i et personlig intervju og i et telefonintervju, men ikke når respondentene svarer på spørsmål ved selv å fylle ut et spørreskjema. I en slik situasjon antar man at respondentene kan lese gjennom hele spørreskjemaet, og til og med lese spørsmålene flere ganger før de markerer sine svar (Bishop & al 1988). Plassering av viktighetsspørsmålene i spørreskjemaet vil dermed ha mindre betydning som forklaring på variasjoner i svargivning. Studien som er gjennomført i forbindelse

¹³²Gjennomføring av eksperimenter i klasserom ved bruk av studenter kan her være en mulighet.

med denne avhandlingen gir imidlertid indikasjoner på at respondenter kan være mer sekvensielle i sin svargivning enn man tidligere har antatt, og at rekkefølge effekter derfor også kan oppstå når spørsmålsrekkefølgen varieres i et skriftlig spørreskjema.

I fremtidige studier vil det være nødvendig å sikre større kontroll, både over rekkefølgen spørsmålene blir presentert, og rekkefølgen på respondentenes svargivning. Innsamling av data gjennom personlig intervju eller telefonintervju vil her være mulige fremgangsmåter for å oppnå dette.

10.6 Valg av statistiske analysemetoder

Variansanalyse ble valgt som analysemetode for testing av de teoribaserte hypotesene. Det relativt høye antall respondenter, i tillegg til den jevne fordelingen av respondenter i eksperimentgruppene, gjør at vi til tross for brudd på forutsetningen om homogenitet i varians for en rekke avhengige variabler, kan forsvare å benytte variansanalyser for å teste de fremsatte hypotesene.

Redegjørelsen vi foretok vedrørende grupperingen av de individuelle vurderingsspmålene i fire produktelementer, avviker på mange måter fra en tradisjonell validitets- og reliabilitetsdrøfting. I kapittel 6 foretok vi et skille mellom formative og reflektive målemodeller som grunnlag for drøfting av begrepsvaliditet. I analysene har vi imidlertid behandlet målene på tradisjonell måte ved å konstruere summerte skalaer. Et alternativ til å benytte variansanalyse ville ha vært å benytte LISREL (Jøreskog og Sørbom, 1993), eller andre analysemetoder som gjør det mulig å skille mellom formative og reflektive målemodeller. I analyser som ikke er inkludert i avhandlingen ble nettopp LISREL benyttet for å sammenligne resultater fra variansanalysene. Resultatene som fremkommer fra disse analysene viser at funnene er de samme uavhengig av hvilken analysemetode vi velger.

Man kan i alle samfunnsvitenskapelige studier stille spørsmål om hvor liten en signifikant effekt skal være før vi velger å se bort fra den. Vi har tidligere drøftet funnene fra hypotesetestingen, som i de fleste tilfeller er kjennetegnet av høye p-verdier og lave effektstørrelser, i relasjon til valg av vurderingsobjekt for den gjennomførte studien. Det er mulig at vi ville funnet sterkere effekter av de metodologiske valgene vi har studert ved valg av andre vurderingsobjekter. Videre forskning er derfor nødvendig for å underbygge de praktiske implikasjonene vi har fremmet for utvikling av spørreskjema.

== Litteraturreferanser ==

- ABELSON, R. P., D. R. KINDER, M. D. PETERS, og S. T. FISKE
1982 Affective and Semantic Components in Political Person Perception. *Journal of Personality and Social Psychology* 42 (No.4) : 616-630.
- AIELLO, A., J. A. CZEPIEL, og L. J. ROSENBERG
1977 Scaling the Heights of Consumer Satisfaction: An Evaluation of Alternative Measures. I *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*. Redigert av R. L. Day. Bloomington, Indiana: Division of Research, School of Business, Indiana University.
- AIKEN, L. S., og S. G. WEST
1991 *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, California: Sage Publications, Inc.
- AKAAH, I. P., og P. K. KORGAONKAR
1983 An empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint, and Hybrid Conjoint Models. *Journal of Marketing Research* XX (May) : 187-197.
- ALBA, J. W., og H. MARMORSTEIN
1986 Frequency Information as a Dimension of Consumer Knowledge. I *Advances in Consumer Research*. Redigert av R. J. Lutz. 446-449. Las Vegas, Nevada: Association for Consumer Research.
- ALBA, J. W., og J. W. HUTCHINSON
1987 Dimensions of Consumer Expertise. *Journal of Consumer Research* 13 (March) : 411-454. X
- ALPERT, M.I.
1971 Identification of Determinant Attributes: A Comparison of Methods. *Journal of Marketing Research* 8 : 184-191.
- ALWIN, D. F.
1991 Research on Survey Quality. *Sociological Methods & Research* 20 (1), 3-29.
- ALWIN, D. F., og J. A. KROSINICK
1985 The Measurement of Values in Surveys: A Comparison of Ratings and Rankings. *Public Opinion Quarterly* 49, 535-552.
- ALWIN, D. F., og J. A. KROSINICK
1991 The Reliability of Survey Attitude Measurement. *Sociological Methods & Research* 20 (1, August), 139-181.

- ANDERSON, N. H.
1981 *Foundations of Information Integration Theory*. New York: Academic Press.
- ANDERSON, R. E.
X 1973 Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance. *Journal of Marketing Research* 10 (February) : 38-44.
- ANDREWS, F. M.
1984 Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach. *Public Opinion Quarterly* 48 : 409-442.
- ANTIL, J. H.
1983 Uses of Response Certainty in Attitude Measurement. I *Advances in Consumer Research*. Redigert av R. Bagozzi og A. Tybout. 409-415. Ann Arbor: Association for Consumer Research.
- AYIDIYA, S. A., og M. K. J. MCCLENDON
1990 Response Effects in Mail Surveys. *Public Opinion Quarterly* 54 : 229-247.
- BAGOZZI, R. P.
1986 *Principles of Marketing Management*. Chicago: Science Research Associates. Inc.
- BECKWITH, N. E., og D. R. LEHMANN
1973 The Importance of Differential Weights in Multiple Attribute Models of Consumer Attitude. *Journal of Marketing Research* X (May) : 141-145.
- BENTON, J. E., og J. L. DALY
1991 A Question Order Effect in a Local Government Survey. *Public Opinion Quarterly* 55 : 640-642.
- BERRY, W. D., og S. FELDMAN
1985 *Multiple Regression in Practice*. Redigert av M. S. Lewis-Beck. Quantitative Applications in the Social Sciences. London New Dehli: Sage Publications, Inc.
- BETTMAN, J. R., og C. W. PARK
1980 Effects of Prior Knowledge and Experience and Phase of The Choice Process on Consumer Decision Processes. *Journal of Consumer Research* 7 (December) : 234-248.
- BICKART, B. A.
1991 Question-Order Effects and Brand Evaluations. I *Context Effects in Social and Psychological Research*. Redigert av N. Schwarz and S. Sudman. 63-79. New York, Berlin, Heidelberg: Springer-Verlag.
- X BICKART, B. A.
1993 Carryover and Backfire Effects in Marketing Research. *Journal of Marketing Research* XXX (February) : 52-62.
- BICKART, B. A., G. MENON, S. SUDMAN, og J. BLAIR
1992 Context Effects in Proxy Judgments. I *Advances in Consumer Research* Redigert av J. F. Sherry, og B. Sternthal. 64-71. Chichago, Illinois: Association for Consumer Research,
- BISHOP, G. F.
1990 Issue Involvement and Response Effects in Public Opinion Surveys. *Public Opinion Quarterly* 54, 209-218.
- BLALOCK, H. M.
1982 *Conceptualization and Measurement in the Social Sciences*. Beverly Hills: Sage Publications.
- BODENHAUSEN, G. V., og R. S. WYER
1987 Social Cognition and Social Reality: Information Acquisition and Use in the Laboratory and the Real World. I *Social Information Processing and Survey Methodology*. Redigert av H.-J. Hippler, N. Schwarz og S. Sudman. 6-41. New York: Springer-Verlag.

- BOLLEN, K. A.
1984 Multiple Indicators: Internal Consistency or No Necessary Relationship? *Quality and Quantity* 18 : 377-385.
- BOLLEN, K. A.
1989 *Structural Equations with Latent Variables*. Wiley series in probability and mathematical statistics. Applied probability and statistics section. New York: John Wiley & Sons, Inc.
- BOLLEN, K., og R. LENNOX
1991 Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin* 110 (2) : 305-314.
- BOOTE, A. S.
1981 Reliability Testing of Psychographic Scales. Five-point or seven-point? Anchored or labeled? *Journal of Advertising Research* 21 (5) : 53-60.
- BRADBURN, N. M., og S. SUDMAN
1981 *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass Publishers.
- BRAY, J. H., og S. E. MAXWELL
1985 *Multivariate Analysis of Variance*. Sage University Paper Series in Quantitative Applications in the Social Sciences. Beverly Hills: Sage Publications.
- BRECKLER, S. J., og E. C. WIGGINS
1989 On Defining Attitude and Attitude Theory: Once More with Feeling. I *Attitude Structure and Function*. Redigert av A. R. Pratkanis, S. J. Breckler og A. G. Greenwald. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- BRIGGS, S. R., og J. M. CHEEK
1986 The Role of Factor Analysis in the Evaluation of Personality Scales. *Journal of Personality* 54 : 106-148.
- BRUCKS, M.
1985 The Effects of Product Class Knowledge on Information Search Behavior. *Journal of Consumer Research* 12 (June) : 1-15.
- BRUCKS, M.
1986 A Typology of Consumer Knowledge Content. I *Advances in Consumer Research*. Redigert av R. J. Lutz. 58-63. Las Vegas, Nevada: Association for Consumer Research.
- CADOTTE, E. R., R. B. WOODRUFF, og R. L. JENKINS
1987 Expectations and Norms in Models of Consumer Satisfaction. *Journal of Marketing Research* XXIV (August) : 305-314. X
- CALDER, B. J., L. W. PHILLIPS, og A. M. TYBOUT
1981 Designing Research for Application. *Journal of Consumer Research* 8 (September) : 197-207.
- CAMPBELL, D.T., og D.W. FISKE
1959 Convergent and Discriminant Validation by the Multitrait - Multimethod Matrix. *Psychological Bulletin* 56 : 81-105.
- CANTRIL, H.
1944 *Gauging Public Opinion*. London: Princeton University Press.
- CATTELL, R.B.
1965 *The Scientific Analysis of Personality*. New York: Penguin Books.
- CHURCHILL, G. A.
1979 A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research* 16 (1) : 64-73.

- CHURCHILL, G. A.
1991 *Marketing Research: Methodological Foundations*. 5 ed. Chicago: The Dryden Press.
- X CHURCHILL, G. A., og C. SURPRENANT
1982 An Investigation Into the Determinants of Customer Satisfaction. *Journal of Marketing Research* 19 (November) : 491-504.
- CICCHETTI, D. V., D. SHOWALTER, og P. J. TYRER
1985 The effect of Number of Rating Scale Categories on Levels of Interrater Reliability: A Monte Carlo Investigation. *Applied Psychological Measurement* 9 (1, March) : 31-36.
- CLARK, H. H., og M. F. SCHOBEL
1991 Asking Questions and Influencing Answers. I *Questions about Questions, Inquiries into the Cognitive Bases of Surveys*. Redigert av J. M. Tanur. 15-48. New York: Russel Sage Foundation.
- COLE, C. A., G. GAETH, og S. N. SINGH
1986 Measuring Prior Knowledge. I *Advances in Consumer Research*. Redigert av R. J. Lutz. 64-66. Las Vegas, Nevada: Association for Consumer Research.
- COOK, T. D., og D. T. CAMPBELL
1979 *Quasi-Experimentation, Design & Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company.
- COOPER, W. H.
1981 Ubiquitous Halo. *Psychological Bulletin* 90 : 218-244.
- COX, E. P.
1980 The Optimal Number of Response Alternatives for a Scale: A Review. *Journal of Marketing Research* XVII (November) : 407-422.
- COX, E. P.
1982 *A Functional Classification of Scales and the Reliability Criteria for their Evaluation*. Working Paper, The university og Texas at Austin, Department of Marketing Administration.
- CRESPI, I., og D. MORRIS
1984 Question Order Effect and the Measurement of Candidate Preference in the 1982 Connecticut Elections. *Public Opinion Quarterly* 48 : 578-591.
- DARBY, M. R., og E. KARNI
1973 Free Competition and the Optimal Amount of Fraud. *Journal of Law and Economics* 16 (April) : 67-88.
- DIXON, P. N., M. BOBO, og R. A. STEVICK
1984 Response Differences and Preferences for All-Category-defined and End-defined Likert Formats. *Educational and Psychological Measurement* 44
- DOUGLAS, S. P., S. C. CRAIG, og J.-P. FAIVRE
1981 Protocols in Consumer Research: Problems, Methods and Uses. I *Research in Marketing*. Redigert av J. N. Sheth. 29-58. JAI Press.
- DUVAL, S., og R. A. WICKLUND
1972 *A Theory of Self-Awareness*. New York: Academic Press.
- EIGLIER, P., E. LANGEARD, J.E.G. BATESON, og C. LOVELOCK
1979 Consumers' Participation in the Service System: Concept Definition and Measurement,. I *Proceedings, 6th International Research Seminar in Marketing at Senanque*. Aix-en-Provence, France, Institut d' Administration des Enterprises: Institut d' Administration des Enterprises.
- ENGEL, J. F., og R. D. BLACKWELL
1982 *Consumer Behavior*. 4 ed. Chicago: The Dryden Press.

- FAZIO, R. H.
1989 On the Power and Functionality of Attitudes: The Role of Attitude Accessibility. I *Attitude Structure and Function*. Redigert av A. R. Pratkanis, S. J. Breckler og A. G. Greenwald. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- FAZIO, R. H.
1990 A Practical Guide to the Use of Response Latency in Social and Psychological Research. I *Research Methods in Personality and Social Psychology*. Redigert av C. Hendrick and M. S. Clark. London: Sage Publications.
- FAZIO, R. H., og M. P. ZANNA
1981 Direct Experience and Attitude-Behavior Consistency. I *Advances in Experimental Social Psychology*. Redigert av L. Berkowitz. 161-202. New York: Academic Press.
- FELDMAN, J. M., og J. G. LYNCH jr.
1988 Self-Generated Validity and other Effects of Measurement on Belief, Attitude, Intention and Behavior. *Journal of Applied Psychology* 73 : 421-435. X
- FERGUSON, G. A., og Y. TAKANE
1989 *Statistical Analysis in Psychology and Education*. 6 ed. Psychology Series. New York: McGraw-Hill Book Company.
- FESTINGER, L.
1957 *A Theory of Cognitive Dissonance*. Palo Alto, California: Stanford University Press.
- FISHBEIN, M.
1963 An investigation of the relationship between beliefs about an object and the attitude toward that object. *Human Relations* 16 : 233-240.
- FISHBEIN, M.
1972 The Search for Attitudinal- Behavioral Consistency. I *Science Foundations in Consumer Behavior*. Redigert av J. S. Cohen. New York: the Free Press, a division of Macmillan Publishing Co., Inc.
- FISHBEIN, M., og I. AJZEN
1975 *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading: MA: Addison-Wesley.
- FISKE, S. T., og S. E. TAYLOR
1991 *Social Cognition*. Redigert av P. G. Zimbardo. McGraw-Hill Series in Social Psychology. New York: McGraw-Hill, Inc.
- FOLKES, V. S.
1988 Recent Attribution Research in Consumer Behavior: A Review and New Directions. *Journal of Consumer Research* 14 (March) : 548-565. X
- GERBING, D. W., og J. C. ANDERSON
1988 An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment. *Journal of Marketing Research* XXV (May) : 186-192.
- GHISELLI, E. E.
1963 Moderating Effects and Differential Reliability and Validity. *Journal of Applied Psychology* 47 (April) : 81-86.
- GHISELLI, E. E., J. P. CAMPBELL, og S. ZEDECK
1981 *Measurement Theory for the Behavioral Sciences*. San Francisco: W. H. Freeman and Company.
- GLASS, G. V., P. D. PECKHAM, og J. R. SANDERS
1972 Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analysis of Variance and Covariance. *Review of Educational Research* 42 (2) : 237-288.
- GRICE, H.
1975. Logic and Conversation. I *Syntax and Semantics*. Redigert av P. Cole og T. Morgan. New York: Seminar Press.

- GROVES, R. M.
1987 Research on Survey Data Quality. *Public Opinion Quarterly* 51 : 156-172.
- GRÜNERT, K. G.
1989 Attributes, Attribute Values and their Characteristics: A Unifying Approach and an Example Involving a Complex Household Investment. *Journal of Economic Psychology* 10 : 229-251.
- GRØNHAUG, K.
1985 Problemer i empirisk forskning. I *Metoder og perspektiver i økonomisk-administrativ forskning*. Redigert av A. Kinserdal. 7-18. Oslo: Universitetsforlaget.
- GUILFORD, J. P.
1954 *Psychometric methods*. New York: McGraw-Hill.
- HAIR JR., J. F., R. E. ANDRESON, og R. L. TATHAM
1987 *Multivariate Data Analysis*. Second ed. New York: Macmillian Publishing Company.
- HARRIS, R.
1978 Improving Patient Satisfaction Through Action Research. *Journal of Applied Behavioral Science* 14 : 382-399.
- HASTIE, R.
1987 Information Processing Theory for the Survey Researcher. I *Social Information Processing and Survey Methodology*. Redigert av H.-J. Hippler, N. Schwarz og S. Sudman. 42-70. New York: Springer-Verlag.
- HEIDE, M.
1990 *Effects of Mood and Information Content on Product Evaluations*. Doktor-avhandling, Norges Handelshøyskole, Bergen.
- HENJESAND, I. J.
1991a *Inter Nor Hotels kvalitetsprogram, Delrapport V, Gjesteundersøkelse 1991, Hvordan vurderer feriegjester Inter Nor hotellene?* Stiftelsen for samfunns- og næringslivsforskning, Bergen. NTIS, 87.
- HENJESAND, I. J.
1991b *Inter Nor Hotels kvalitetsprogram, Delrapport II, Måling og analyse av kvalitet, gjesteundersøkelse*. Stiftelsen for samfunns- og næringslivsforskning, Bergen. NTIS, 14.
- HENJESAND, I. J., og S. V. TROYE
1991 *Inter Nor Hotels kvalitetsprogram, Kvalitetsikring og kvalitetstyring, En ramme for å forstå problemer og muligheter*. Stiftelsen for samfunns- og næringslivsforskning, Bergen. NTIS, 13.
- HIGGINS, E.T., og J. A. BARGH
1987 Social Cognition and Social Perception. *Annual Review of Psychology* 38 : 369-425.
- HIPPLER, H.-J., og N. SCHWARZ
1987 Response Effects in Surveys. I *Social Information Processing and Survey Methodology*. Redigert av H.-J. Hippler, N. Schwarz og S. Sudman. 102-122. New York: Springer-Verlag.
- HIPPLER, H.-J., N. SCHWARZ, og S. SUDMAN
1987 *Social Information Processing and Survey Methodology*. New York: Springer-Verlag.
- HOGARTH, R. M.
1982 *Question Framing and Response Consistency*. San Francisco: Jossey-Bass Inc. Publishers.
- HOWARD, J. A., og J. N. SHETH
1969 *The Theory of Buyer Behavior*. New York: John Wiley and Sons, Inc.
- HUBER, J., og J. MCCANN
1982 The Impact of Inferential Beliefs on Product Evaluations. *Journal of Marketing Research* XIX (August) : 324-333.

- HUTCHINSON, J. W.
 1983 On The Locus of Range Effects in Judgment and Choice. I *Advances in Consumer Research*. Redigert av R. Bagozzi and A. Tybout. 305-308. Ann Arbor: Association for Consumer Research.
- JACCARD, J., D. BRINBERG, og L. J. ACKERMAN
 1986 Assessing Attribute Importance: A Comparison of Six Methods. *Journal of Consumer Research* 12 (March) : 463-468.
- JACOBY, J., J.C. OLSON, og R.A. HADDOCK
 1971 Price , Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality. *Journal of Applied Psychology* 55 : 570-579. X
- JACOBY, J., T. TROUTMAN, A. KUSS, og D. MAZURSKY
 1983 Experience and Expertise in Complex Decision Making. I *Advances in Consumer Research*. Redigert av R. J. Lutz. Las Vegas, Nevada: Association for Consumer Research.
- JOHNSON, M. D.
 1988 Comparability and Hierarchical Processing in Multialternative Choice. *Journal of Consumer Research* 15 (December) : 303-314.
- JONES, W. H., og J. R. LANG
 1982 Reliability and Validity Effects under Mail Survey Conditions. *Journal of Business Research* 10 : 339-353.
- JØRESKOG, K. G., og D. SØRBOM
 1993 LISREL 8: *Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International, Inc.
- KELLEY, H.
 1973 The Processes of Causal Attributions. *American Psychologist* 28 : 107-128.
- KEPPEL, G.
 1982 *Design and Analysis: A Researcher's Handbook*. Englewood Cliffs, N.J.: Prentice-Hall.
- KIDDER, L. H., og C. M. JUDD
 1986 *Research Methods in Social Relations*. 5 ed. HRW International Editions. New York: CBS College publishing.
- KOTLER, P.
 1984 *Marketing Management: Analysis, Planning and Control*. 5 ed. The Prentice-Hall Series in Marketing. London: Prentice/Hall International Inc.
- KOTLER, P.
 1986 The Prosumer Movement: A New Challenge for Marketers. I *Advances in Consumer Research*. Redigert av R. J. Lutz. 510-513. Las Vegas, Nevada: Association for Consumer Research.
- KROSNICK, J. A.
 1989 The Polls - A Review. Question Wording and Reports of Survey Results: The Case of Louis Harris and Associates and Aetna Life and Casualty. *Public Opinion Quarterly* 53 : 107-113.
- KROSNICK, J. A., og D. F. ALWIN
 1987 An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. *Public Opinion Quarterly* 51 : 201-219.
- LANCASTER, K.
 1966 A New Approach to Consumer Theory. *Journal of Political Economy* LXXIV (No.2 (April)) : 132-157.
- LANCASTER, K.
 1971 *Consumer Demand, A New Approach*. New York: Columbia University Press.

- LATOURE, S. A.; og N. C. PEAT
 1979 Conceptual and Methodological Issues in Consumer Satisfaction Research. I *Advances in Consumer Research*. Redigert av W. L. Wilkie. 431-437. Ann Arbor: Association for Consumer Research.
- LEFKOFF-HAGIUS, R., OG C. H. MASON
 1990 The Role of Tangible and Intangible Attributes in Similarity and Preference Judgments. I *Advances in Consumer Research*. Redigert av M. E. Goldberg, G. Gorn og R. W. Pollay. 135-143. New Orleans, Louisiana: Association for Consumer Research.
- LEVITT, T.
 1972 Production-line approach to service. *Harvard Business Review* (September-October) : 41-52.
- LEVITT, T.
 1981 Marketing Intangible Products and Product Intangibles. *Harvard Business Review* (May-June)
- LEWIS-BECK, M. S.
 1980 *Applied Regression: An introduction*. Redigert av J. L. Sullivan. Quantitative Applications in the Social Sciences. Beverly Hills/London: Sage Publications.
- LORD, F.M., og M.R. NOVICK
 1968 *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.
- LOVELOCK, C. H.
 1988 *Managing Services Marketing, Operations, and Human Resources*. Englewood Cliffs, New Jersey: Prentice-Hall.
- LUTZ, R. J.
 1991 The Role of Attitude Theory in Marketing. I *Perspectives in Consumer Behavior*. Redigert av H. H. Kassarian og T. S. Robbertson. Illinois: Prentice-Hall International
- LUTZ, R. J., og J. R. BETTMAN
 1977 Multiattribute Models in Marketing: A Bicentennial Review. I *Consumer and Industrial Buying Behavior*. Redigert av A. G. Woodside, J. N. Sheth og P. D. Bennett. New York: North-Holland.
- LYNCH, J. G. Jr.
 1982 On the External Validity of Experiments in Consumer Research. *Journal of Consumer Research* 9 (3) : 225-239.
- LYNCH, J. G. Jr.
 1983 The Role of External Validity in Theoretical Research. *Journal of Consumer Research* 10 (1) : 109-111.
- LYNCH, J. G. Jr.
 1985 Uniqueness Issues in the Decompositional Modeling of Multiattribute Overall Evaluations: An Information Integration Perspective. *Journal of Marketing Research* XXII (February) : 1-19.
- LYNCH, J. G. Jr., D. CHAKRAVARTI, og A. MITRA
 1991 Contrast Effects in Consumer Judgments: Changes In Mental Representations or in the Anchoring of Rating Scales? *Journal of Consumer Research* 18 (3 (December)) : 248-298.
- MARKS, L. J., og M. A. KAMINS
 1988 The Use of Product Sampling and Advertising: Effects of Sequence of Exposure and Degree of Advertising Claim Exaggeration on Consumers' Belief Strength, Belief Confidence, and Attitudes. *Journal of Marketing Research* XXV (August) : 266-281.

- MAZIS, M. B., O. T. AHTOLA, og R. E. KLIPPEL
 1975 A Comparison of Four Multi-Attribute Models in the Prediction of Consumer Attitudes. *Journal of Consumer Research* 2 (June) : 38-52.
- MAZURSKY, D., og A. GEVA
 1989 Temporal Decay in Satisfaction - Purchase Intention Relationship. *Psychology & Marketing* 6 (3) (Fall) : 211-227.
- MCCLENDON, M. K. J., og D. J. O'BRIEN
 1988 Question-Order Effects on the Determinants of Subjective Well-Being. *Public Opinion Quarterly* 52 : 351-364.
- MCGRATH, J. E., og D. BRINBERG
 1983 External Validity and the Research Process: A Comment on the Calder/Lynch Dialogue. *Journal of Consumer Research* 10 (June) : 115-125.
- MEEK, G. E., H. L. TAYLOR, K. A. DUNNING, og K. A. KLAFEIN
 1987 *Business Statistics*. Newton, Massachusetts: Allyn and Bacon Inc.
- MERVIS, C. B., og E. ROSCH
 1981 Categorization of Natural Objects. *Review of Psychology* 35 : 113-135.
- MILLER, P. V.
 1984 Alternative Question Forms for Attitude Scale Questions in Telephone Interviews. *Public Opinion Quarterly* 48 : 766-778.
- MITCHELL, T. R.
 1985 An Evaluation of the Validity of Correlational Research Conducted in Organizations. *Academy of Management Review* 10 (No. 2) : 192-205.
- MIZERSKY, R.W., og J. KERNAN
 1979 The Attribution Process in Consumer Decision Making. *Journal of Consumer Research* 6 (September) : 123-140.
- MORDAL, T. L.
 1989 *Som man spør, får man svar. Arbeid med survey-opplegg*. TANO.
- MURPHY, K. R., og D. H. REYNOLDS
 1988 Does True Halo Affect Observed Halo? *Journal of Applied Psychology* 73 (2) : 235-238.
- MURPHY, K. R., og R. L. ANHALT
 1992 Is Halo Error a Property of the Rater, Ratees, or the Specific Behaviors Observed. *Journal of Applied Psychology* 77 (4) : 494-500.
- NACHMIAS, C., og D. NACHMIAS
 1985 *Research Methods in Social Sciences*. 2 ed. London: Edward Arnold (Publishers) Ltd.
- NELSON, P.
 1974 Advertising as Information. *Journal of Political Economy* 82 (4 (July/August)) : 729-754.
- NESLIN, S. A.
 1981 Linking Product Features to Perceptions: Self-Stated Versus Statistically Revealed Importance Weights. *Journal of Marketing Research* XVIII (February) : 80-86.
- NEUMANN, L.
 1983 Effects of Scale Length on Means and Correlation Coefficients. *Quality and Quantity* 17 : 405-408.
- NEUMANN, L., og Y. NEUMANN
 1981 Comparison of Six Lengths of Rating Scales: Students' Attitudes Toward Instruction. *Psychological Reports* 48 : 399-404.
- NUNNALLY, J. C.
 1978 *Psychometric Theory*. 2 ed. New York: McGraw-Hill.

- NUNNALLY, J. C., og I. H. BERNSTEIN
 1994 *Psychometric Theory*. 3rd ed. McGraw-Hill Series in Psychology. New York: McGraw-Hill.
- OLIVER, R. L.
 X 1980 A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research* XVII (November) : 460-469.
- OLIVER, R. L.
 X 1981 Measurement and Evaluation of Satisfaction Process in Retail Setting. *Journal of Retailing* 57 (Fall) : 25-48.
- OLIVER, R. L., og W. S. DESARBO
 X 1988 Response Determinants in Satisfaction Judgments. *Journal of Consumer Research* 14 (March) : 495-507.
- OLSHAVSKY, R. W., og D. H. GRANBOIS
 1979 Consumer Decision making - Fact of Fiction? *Journal of Consumer Research* 6 (September) : 93-100.
- OLSON, J. C.
 1977 Price as an Informational Cue: Effects on Product Evaluations. I *Consumer and Industrial Buying Behavior*. Redigert av A. G. Woodside, J. N. Sheth og P. D. Bennett. 267-286. New York: North-Holland.
- OLSON, J. C., og J. JACOBY
 1972 Cue Utilization in the Quality Perception Process. I *Advances in Consumer Research*. Redigert av M. Venkatesan. 167-179. Iowa City: Association for Consumer Research
- OTTATI, V.C., E.J. RIGGLE, R.S. WYER JR., N. SCHWARZ, og J. KUKLINSKI
 1989 Cognitive and affective bases of opinion survey responses. *Journal of Personality and Social Psychology* 57 : 404-415.
- PARASURAMAN, A.
 1986 *Marketing Research*. Reading: Addison-Wesley.
- PARASURAMAN, A., V. A. ZEITHAML, og L. L. BERRY
 1985 A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing* 49 (Fall) : 41-50.
- PARASURAMAN, A., V. A. ZEITHAML, og L. L. BERRY
 1988 SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing* 64 (Spring) : 12-40.
- PETERSON, R. A., og W. R. WILSON
 1992 Measuring Customer Satisfaction: Fact and Artifact. *Journal of the Academy of Marketing Science* 20 (1) : 61-71.
- PETTY, R. E., G. A. RENNIER, og J. T. CACIOPPO
 1987 Assertion Versus Interrogation Format in Opinion Surveys. *Public Opinion Quarterly* 51 : 481-494.
- POE, G. S., I. SEEMAN, J. MCLAUGHLIN, E. MEHL, og M. DIETZ
 1988 "Don't Know" Boxes in Factual Questions in a mail Questionnaire, Effects on Level and Quality of Response. *Public Opinion Quarterly* 52 : 212-222.
- RADEN, D.
 1985 Strength-Related Attitude Dimensions. *Social Psychology Quarterly* 48 (No.4) : 312-330.
- RAJU, P. S.
 1977 Product Familiarity, Brand Name, and Price Influences on Product Evaluation. I *Advances in Consumer Research*. Redigert av W. D. Perrault JR. 64-71. Atlanta: Association for Consumer Research.

- RAO, A. R., og K. B. MONROE
 1988 The Moderating Effect of Prior Knowledge on Cue Utilization in Product Evaluations. *Journal of Consumer Research* 15 (September) : 253-264.
- RASINSKI, K. A.
 1989 The Effect of Question Wording on Public Support for Government Spending. *Public Opinion Quarterly* 53 : 388-394.
- ROSCH, E.
 1988 Principles of Categorization. I *Cognition and Categorization*. Redigert av E. Rosch og B. B. Lloyd. 27-48. Hillsdale, New Jersey: Erlbaum.
- ROSENBERG, M.J.
 1956 Cognitive Structure and Attitudinal Effect. *Journal of Abnormal and Social Psychology* 53 : 367-372.
- RUST, R. T., A. J. ZAHORIK, og T. L. KEININGHAM
 1994 *Return on Quality: Measuring the Financial Impact of Your Company's Quest for Quality*. Chicago, Illinois: Probus Publishing Company.
- SAMPLE, J., og R. WARLAND
 1973 Attitude and Prediction of Behavior. *Social Forces* 51 (March) : 292-304.
- SAWYER, A. G., og A. D. BALL
 1981 Statistical Power and Effect Size in Marketing Research. *Journal of Marketing Research* XVIII (August) : 275-90.
- SCHUL, Y., og M. SCHIFF
 1993 Measuring Satisfaction With Organizations. Predictions From Information Accessibility. *Public Opinion Quarterly* 57 : 536-551.
- SCHUMAN, H., G. KALTON, og J. LUDWIG
 1983 Context and Contiguity in Survey Questionnaires. *Public Opinion Quarterly* 47 : 112-115.
- SCHUMAN, H., og S. PRESSER
 1981 *Questions and Answers in Attitude Surveys*. San Diego, California: Academic Press.
- SCHUMAN, H., S. PRESSER, og J. LUDWIG
 1981 Context effects on survey responses to questions about abortion. *Public Opinion Quarterly* 45 : 216-223.
- SCHWARZ, N.
 1990 Assessing Frequency Behaviors of Mundane Behaviors. Contributions of Cognitive Psychology to Questionnaire Construction. I *Research Methods in Personality and Social Psychology*. Redigert av C. Hendrick og M. Clark S. 98-119. London: Sage Publications.
- SCHWARZ, N., B. KNÄUPER, H.-J. HIPPLER, E. NOELLE-NEUMANN, og L. CLARK
 1991 Rating Scales: Numeric Values May Change the Meaning of Scale Labels. *Public Opinion Quarterly* 55 : 570-582.
- SCHWARZ, N., F. STRACK, og H.-P. MAI
 1991 Assimilation and Contrast Effects in Part-Whole Question Sequences: A Conversational Logic Analysis. *Public Opinion Quarterly* 55 : 3-23.
- SCHWARZ, N., H.-J. HIPPLER, B. DEUTSCH, og F. STRACK
 1985 Response Scales: Effects of Category Range on Reported Behavior and Comparative Judgments. *Public Opinion Quarterly* 49 : 388-395.
- SCHWARZ, N., og H. BLESS
 1992 Assimilation and Contrast Effects in Attitude Measurement: An Inclusion/Exclusion Model. I *Advances in Consumer Research*. Redigert av J. F. Sherry og B. Sternthal. 72-77. Chicago, Illinois: Association for Consumer Research.

- SCHWARZ, N., og H.-J. HIPPLER
 1987 What Response Scales May Tell your Respondents: Informative Functions of Response Alternatives. I *Social Information Processing and Survey Methodology*. Redigert av H.-J. Hippler, N. Schwarz og S. Sudman. 163-178. New York: Springer-Verlag.
- SCHWARZ, N., og S. SUDMAN
 1991 *Context Effects in Social and Psychological Research*. New York: Springer-Verlag.
- SELLITZ, C., L.S. WRIGHTSMAN, og S.W. COOK.
 1976 *Research Methods in Social Relations*. 3 ed. New York: Holt, Rinehart & Winston.
- SELNES, F.
 1986 *Effekten av produktkunnskap på informasjonssøking ved forbrukerkjøp*. Dr. avhandling, Norges Handelshøyskole, Bergen.
- SELNES, F., og K. GRØNHAUG
 1986 Subjective and Objective Measures of Product Knowledge Contrasted. I *Advances in Consumer Research*. Redigert av R. J. Lutz. 67-71. Las Vegas, Nevada: Association for Consumer Research.
- SELNES, F., og S. V. TROYE
 1989 Buying Expertise, Information Search, and Problem Solving. *Journal of Economic Psychology* 10 (3 (November)): 411-428.
- SHARMA, S., R. M. DURAND, og O. GUR-ARIE
 1981 Identification and Analysis of Moderator Variables. *Journal of Marketing Research* XVIII (August): 291-300.
- SHERIF, M., og C.I. HOVLAND
 1961 *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven: Yale University Press.
- SHOSTACK, G. L.
 1977 Breaking Free from Product Marketing. *Journal of Marketing* April: 73-80.
- SRULL, T.K.
 1983 The Role of Prior Knowledge in the Acquisition, Retention, and use of New Information. I *Advances in Consumer Research*. Redigert av R. Bagozzi og A. Tybout. 572-576. Ann Arbor, MI: Association for Consumer Research.
- STEENKAMP, J.-B. E. M.
 1989 *Product Quality*. Assen/maastricht, The Netherlands: Van Gorcum.
- STINCHCOMBE, A. L.
 1968 *Constructing Social Theories*. Chicago: The University of Chicago Press.
- STOKMANS, M.
 1991 The Relation Between the Characteristics of a Product Used by the Management in New Product Development and Those Used by the Consumer in Making a Choice. *European Marketing Academy Conference (EMAC)* Dublin: European Marketing Academy.
- STRACK, F.
 1991 "Order Effects" in Survey Research: Activation and Information Functions of Preceding Questions. I *Context Effects in Social and Psychological Research*. Redigert av N. Schwarz og S. Sudman. 23-34. New York: Springer-Verlag.
- STRACK, F., og L. L. MARTIN
 1987 Thinking, Judging, and Communicating: A Process Account of Context Effects in Attitude Surveys. I *Social Information Processing and Survey Methodology*. Redigert av H.-J. Hippler, N. Schwarz og S. Sudman. 123-148. New York: Springer-Verlag.

- SUDMAN, S., og N. M. BRADBURN
 1974 *Response effects i surveys: A review and synthesis*. Chicago: Aldine Publishing Company.
- SUDMAN, S., og N. M. BRADBURN
 1983 *Asking Questions*. The Jossey-Bass Series in Social and Behavioral Sciences. san Francisco: Jossey-Bass Publishers.
- SUJAN, M.
 1985 Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments. *Journal of Consumer Research* 12 (June) : 31-46. X
- SWAN, J. E., og I. F. TRAWICK
 1981 Disconfirmation of Expectations and Satisfaction with a Retail Service. *Journal of Retailing* 57 (Fall) : 49-67.
- SZYBILLO, G. J., og J. JACOBY
 1974 Intrinsic Versus Extrinsic Cues as Determinants of Perceived Product Quality. *Journal of Applied Psychology* 59 (1) : 74-78.
- TANUR, J. M.
 1991 *Questions about Questions. Inquiries into the Cognitive Bases of Surveys*. New York: Russel Sage Foundation.
- TOFFLER, A.
 1980 *The Third Wave*. Willam Morrow.
- TORGERSON, W.S.
 1958 *Theory and Methods of Scaling*. New York: Wiley.
- TOURANGEAU, R.
 1991 Context Effects on Responses to Attitude Questions: Attitudes as Memory Structures. I *Context Effects in Social and Psychological Research*. Redigert av N. Schwarz og S. Sudman. 35-47. New York: Springer-Verlag.
- TOURANGEAU, R., og K. A. RASINSKI
 1988 Cognitive Processes Underlying Context Effects in Attitude Measurement. *Psychological Bulletin* 103 (3) : 299-314.
- TOURANGEAU, R., K. A. RASINSKI, og N. BRADBURN
 1991 Measuring Happiness in Surveys: A Test of the Subtraction Hypothesis. *Public Opinion Quarterly* 55, 255-266.
- TOURANGEAU, R., K. A. RASINSKI, N. BRADBURN, og R. D'ANDRADE
 1989a Carryover Effects in Attitude Surveys. *Public Opinion Quarterly* 53 : 495-524.
- TOURANGEAU, R., K. A. RASINSKI, N. BRADBURN, og R. D'ANDRADE
 1989b Belief Accessibility and Context Effects in Attitude Measurement. *Journal of Experimental Social Psychology* 25 : 401-421.
- TROYE, S. V.
 1990 *Markedsføring: Styling av kvalitet*. 1 ed. Oslo: Universitetsforlaget.
- TROYE, S. V.
 1992 *Perspektiv og ramme for mlling av subjektiv kvalitet*. Norges Handelshøyskole, Bergen.
- TROYE, S. V., I. J. HENJESAND, og E. BREIVIK
 1992 Does Labelling Matter? Some Empirical Differences Between Subject and Objekt Anchored Response Scales in Customer Evaluations. *IAREP/GEW Conference Proceedings* : Frankfurt.
- TROYE, S. V., T. ØGAARD, og I. J. HENJESAND
 1995 The Triple Importance of Service: An Alternative Perspective of Product and Service Quality. I *Marketing Today and for the 21st Century, the 24th Annual Conference of the European Marketing Academy* (Proceedings) Redigert av M. Bergadaà. Cergy-Pontoise, France: The European Marketing Academy.

- TSE, D. K., og P. C. WILTON
 1988 Models of Consumer Satisfaction Formation: An Extension. *Journal of Marketing Research* XXV (May) : 204-212.
- UHL, K. P., og G. D. UPAH
 1983 The Marketing of Services: why and how it is different? *Research in Marketing* 6 : 231-257.
- URBAN, G.L., og J.R. HAUSER
 1980 *Design and Marketing of New Products*. London: Prentice Hall International.
- WESTBROOK, R. A.
 1980 A Rating Scale for Measuring Product/Service Satisfaction. *Journal of Marketing* 44 (Fall) : 68-72.
- WESTBROOK, R. A.
 1987 Product/Consumption-Based Affective Responses and Postpurchase Processes. *Journal of Marketing Research* XXIV (August) : 258-270.
- WILCOX, R.R.
 1987 New Designs in Analysis of Variance. I *Annual Review of Psychology*. Redigert av M. R. Rosenweig and L. W. Porter. 29-60. Palo Alto, Ca: Annual Reviews. Inc.
- WILDT, A. R., og M. B. MAZIS
 1978 Determinants of Scale Response: Label Versus Position. *Journal of Marketing Research* XV (May) : 261-267.
- WILSON, T. D.
 1990 Self-Persuasion Via Self-Reflection. I *Self-Inference Processes: The Ontario Symposium*. Redigert av J. M. Olson and M. P. Zanna. 43-67. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- WIND, Y.J.
 1982 *Product Policy: Concepts, Methods and Strategy*. London: Addison-Wesley Publishing Company.
- WINER, B.J., D.R. BROWN, og K.M. MICHELS
 1991 *Statistical Principles in Experimental Design*. New York: McGraw-Hill.
- WOODRUFF, R. B., E. R. CADOTTE, og R. L. JENKINS
 1983 Modeling Consumer Satisfaction Processes Using Experience-Based Norms. *Journal of Marketing Research* XX (August) : 296-304.
- WYER, R. S. Jr., og T. K. SRULL
 1986 Human Cognition in its Social Context. *Psychological Review* 93 : 322-359.
- YI, YOUJAE.
 1990 *A Critical Reveiw of Consumer Satisfaction*. I *Review of Marketing*. Redigert av V. A. Zeithaml. Chicago Il.: American Marketing Association.
- ZALTMAN, G., C. R. A. PINSON, og R. ANGELMAR
 1972 *Metatheory and Consumer Research*. Hisdale, Illinois: Druden Press.
- ZANNA, M. P., J. M. OLSON, og R. H. FAZIO
 1980 Attitude-Behavior Consistency: An Individual Difference Perspective. *Journal of Personality and Social Psychology* 38 (3) : 432-440.
- ZEDECK, S.
 1971 Problems with the Use of "Moderator" Variables. *Psychological Bulletin* 76 (October) : 295-310.
- ZEITHAML, V. A.
 1981 How Consumer Evaluation Processes Differ between Goods and Services. I *Marketing of Services*. Redigert av J. H. Donnelly and W. R. George. Chicago: American Marketing Association.

ZEITHAML, V. A., A. PARASURAMAN, og L. L. BERRY

1985 Problems and Strategies in Services Marketing. *Journal of Marketing* 49 (spring) : 33-46.

ZELLER, R. A., og E. G. CARMINES

1980 *Measurement in The Social Sciences: The Link Between Theory and Data*. Cambridge: Cambridge University Press.

== APPENDIKS ==

A

**Empiriske studier:
ordlyd i spørsmålsstilling,
svarskala,
og
rekkefølge på spørsmål**

Tabell A.1

Empiriske studier: Problemstillinger i tilknytning til ordlyd i spørsmålstillning i spørreundersøkelser

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|------------------------------|---|-----------------------------------|--|---|
| Andrews (1984) | Effekt av lengde på introduksjon og på etterfølgende spørsmål på data kvalitet, og effekt av absolutte vs. komparative vurderings-spørsmål | Data kvalitet Begrepsvaliditet | 6 forskjellige surveys og strukturert modellering med LISREL | Introduksjon på 16-64 ord med etterfølgende spørsmål på 16-24, og 25+ ord gir høyest validitet, lavest residualeffekt og lavest metodeeffekt. Anbefaler bruk av komparative spørsmåls formuleringer da disse gir mer presise svar pga. forrankingspunkter |
| Miller (1984) | Sammenligning av ulike måter å stille 7- punktsskala holdningsspørsmål på i en helseundersøkelse 1 = one step scale 2 = two step scale | Ad-hoc, mangler teori | Telefontintervju eksperiment med to grupper der en gruppe svarer på spørsmålet i ett trinn (one step scale) ved å velge ett tall, mens den andre gruppen først skal gi en generell, og deretter en spesiell vurdering på tallskalen (two step scale) | Ingen forskjeller mhp. frekvens distribusjoner One step scale, mindre positiv responssekvenshet One step scale, mindre missing data One step scale, høyere interkorrelasjon |
| Alwin og Krosnick (1985) | Sammenligning av "rating" og "ranking" ved måling av verdier ved foreldres verdsettning av barns kvaliteter | Ad-hoc, mangler teori | 3 forskjellige spørsmålstyper: Standard reduced ranking (he) Standard reduced ranking (child) Fem punkt "rating" skala Data fra "full probability cluster sample" | "Rating" og "ranking" like mhp. å få frem relativ viktighet for verdvaig, men ulike mhp. latent struktur og korrelasjon med relevante prediksjonsvariabler. "Ranking" gir best prediksjon. |
| Petty, Rennier et al. (1987) | Undersøke effekten av å presentere holdningsspørsmål i "påstand"-format (assertion) vs. "spørre"-format (interrogation) | Social psychology | Eksperiment ved gjennomføring av survey for barnehøveler og kalkulatorer. Manipulering av spørsmålsformat (påstand- vs. spørreformat). 7 punkt "enig- uenig"-skala. Randomisert til- deling til 91 studenter. | Sterkere polarisering i svargivning ved bruk av "spørre" format (interrogation) enn ved bruk av "påstand" format (assertion) Mer informasjonsbehandling ved "spørre" format enn ved "påstands" format "Spørre" format anbefales |
| Rasinski (1989) | Undersøke effekt av ordlyd i spørsmål for måling av vurderinger av offentlig forbruk innenfor ulike områder. | Ad-hoc mangler teori | "Spill-ballo" eksperimenter i 1984, 1985 og 1986 i "General Social Survey". Bruk av multinomial logit analyse. | Signifikante effekter av endringer i ordlyd i spørsmålstillning i løpet av de tre årene |
| Krosnick (1989) | Undersøke effekten av å benytte "støtte" (support), "akseptabel" (acceptable) eller "foretrekke" (favor) ved måling av holdninger til rettsystemet i USA. | Ad-hoc mangler teori | Eksperiment med ulike spørsmåls- formulering ved gjennomføring av telefonintervjuer | Signifikante effekter at ordlyd i spørsmål- stilling. Bruk av "akseptabel" og "foretrekker" (acceptable og favor) alternativene gir høyere positiv respons enn ved bruk av "støtte" alternativet (support) |

Tabell A.2

Empiriske studier: Problemstillinger i tilknytning til svarskala i spørreundersøkelser

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|---------------------------|---|---|---|---|
| Wildt, & Mazis (1978) | Undersøke betydningen av å bruke benevnelse på alle verdier på en vurderingsskala vs. benevnelse bare på endepunktene på skalaen | Skaleringsteori | Seks forskjellige spørreskjema med ulike grad av benevnelse på skalaverdier utviklet for vurdering av ulike produkter (bla. TV-, show, og biler). Randomisert tildeling til studenter ved University of Florida | Respons på en skala er en funksjon av både benevnelse og plassering av benevnelsen på tallskalaen. Finner ingen signifikante forskjeller mellom skala med tilnærmet intervall benevnelse på samtlige skalaverdier, skaler med benevnelse bare på endepunktene |
| Cox III (1980) | Review over forskning på optimalt antall responsalternativer på en skala. | Skaleringsteori (Stimuluscentrert, Subjektcentrert og Responsstilnærming) Informasjonsteori | Litteraturreview | Finnes ikke "riktig" antall responsalternativer. Anbefalinger: (i) skalaer med 2-3 svar alternativer gir lite informasjon og frustrerer respondentene (ii) skalaer med mer enn 9 svaralternativer gir små og marginale forbedringer (iii) skaler med midtpunkt anbefales der respondentene antas å ha en nøytral posisjon (iv) forståelig instruksjon og benevnelse på svaralternativene er nødvendig (v) Subjektcentrert skala- opp til 5 svaralt. Stimuluscentrert skala- opp til 9 svaralt. |
| Boote (1981) | Undersøke i hvilken grad forskjeller i reliabilitet i psykografiske skalaer kan tilskrives variasjon i antall verdier på skalaen og forskjeller i forankring /benevnelse på skalavertiene | Reliabilitet | Fire forskjellige spørreskjema med ulikt antall verdier på vurderingsskalaene og ulik grad av benevnelse på vertiene (5- og 7 punktiskaler med benevnelse på alle verdiene og kun på endepunkt) Randomisert tilordning til 600 husholdninger fra et panel. Gjentaatt måling etter seks uker med samme spørreskjema til til samme respondenter | Korrelasjon mellom vurderinger på tidspunkt 1 og 2 gir ingen forskjell i reliabilitet basert på antall verdier på skalaen, mens skale med benevnelse på samtlige skalaverdier er mer reliabile enn skala med benevnelse kun på endepunktene. Andel konsistente vurderinger (test- retest) gir høyere reliabilitet for 5- punkt skala enn for 7-punkt skala og ingen forskjell i reliabilitet mellom skala med benevnelse på alle verdier og skaler med benevnelse kun på endepunktene. Ekstremt viktige/ uviktige egenskaper gir høyere reliabilitet enn middels viktige egenskaper. Gjennomsnittsvurdering er lavere for skaler med benevnelse på alle verdiene. |
| Neumann, & Neumann (1981) | Sammenligning av seks vurderingsskalaer med ulik lengde (2,3,4,5,7 og 10 verdier) ved måling av studenters holdninger til instruksjon. | Litteraturreview Ad-hoc ingen teoriebruk | Seks forskjellige spørreskjema med ulikt antall verdier på vurderingsskalaen (2,3,4,5,7 og 10 verdier) ble randomisert tildelt 500 studenter. | Alle fordelinger er skjeve i negativ retning. På de lange skalaene ungår man ekstremt verdiene Korrelasjon mellom variablene øker med skallengde. Eta-koeffisienter (grad av reduksjon i varians) øker med skallengde. Anbefaler bruk av 5-punktiskala i holdningsundersøkelser |

Tabell A.2 - Fortsetter

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|---|--|--|--|--|
| Neumann (1983) | Sammenligning av tre vurderingskalaer med ulik lengde (3,5 og 7 verdier) ved måling av studenters holdninger til instruksjon. | Ad-hoc, ingen teorbrikk Replikasjon av | 57 studenter fylte ut spørreskjema vedr. tilfredshet med kurs og professorer. Samtlige studenter fylte ut spørreskjema tre ganger med ulikt antall verdier på vurderingsskalaene | Alle fordelinger er skjeve i negativ retning. På de lange skalaene ungår man ekstrem korrelasjon mellom variablene øker med skållengde. Eta-koeffisienter (grad av reduksjon i varians) øker med skållengde. Anbefaler bruk av 5-punktskala i holdningsundersøkelser |
| Dixon, Bobo, og Stevick (1984) | Sammenligning av to Likert-type skåler. Den ene med benevnelse på alle skålpunkter, den andre med benevnelse bare på endepunktene | Litteraturgjennomgang Ad-hoc, mangler teori | 121 subjektive svarte alle på halvparten av spørsmålene med benevnelse på alle skålpunkter og den andre halvparten med benevnelse bare på endepunktene | Ingen signifikante forskjeller mellom skalaene Skållengde med benevnelse kun på endepunktene viste større varians enn skalaene med benevnelse på alle skålpunktene. |
| Andrews (1984) | Effekt av antall svaralternativer, bruk av "vet ikke" alternativ og lengde på item batteri på data kvalitet | Data kvalitet Begrepsvaliditet | 6 forskjellige surveys og strukturert modellering med LISREL | Data kvalitet øker med økende antall skålkategorier. Bruk av eksplisitt "vet ikke" alternativ gir høyere data kvalitet Data kvalitet reduseres med økende antall item i itembatteriet, selv om det er fordeler forbundet med å benytte felles introduksjon og svarskåll knyttet til itembatteriet |
| Cicchetti, Showalter et al. (1985) | Undersøke hvorvidt "interrater" reliabilitet for en klinisk måleskala er avhengig av antall kategorier eller punkt på skalaen | Reliabilitet | Monte Carlo simulering | Reliabilitet øker jevnt opp til 7 skålpunkter og flater deretter ut (selv om antall skålpunkter økes til 100) |
| Schwarz, Hippler, Deutsch, og Strack (1985) | Undersøke effekter av rekkevidde (range) på responskategorier i lukkede svaralternativer i spørsmål for måling av rapportert atferd (daglig bruk av TV) for vurdering av tilfredshet med fritiden. | Litteraturgjennomgang Informasjonsbehandlingsperspektiv | Ekspertimenter med to ulike skållengder og en variant med åpen respons for måling av TV-tidsbruk i spørreskjema for måling av tilfredshet med fritiden. | Indikasjon på at respondenter som blir presentert en skala med små kategorilengde velger kategorier på midten av listen, mens respondenter som blir presentert en skala med vidt kategorilengde velger den første kategorien i skalaen. Respondenter som presenteres en skala med vidt kategorilengde rapporterer at de er mer tilfreds med fritid enn respondenter som er presentert en skala med små kategorilengde. Anbefaling: bruke åpne svaralternativer ved måling av atferdsrekvenser og bruksrekvenser |
| Bishop (1987) | Undersøke effekter av å inkludere eller utelukke midtpunkt (middle response alternative) på en | Litteraturgjennomgang Ad-hoc, mangler teori | Ekspertiment gjennomført som del av en survey om holdninger til sosial trygghet, forsvars- | 1) Signifikant større sannsynlighet for å velge midtpunktet (middle response alternative) når dette er eksplisitt med som del av spørsmålet. |

Tabell A.2 - Fortsetter

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|--|---|--|---|--|
| | vurderingsskala. | | budsjett og atomkraft. Randomisert tilordning av spørreskjema med eller uten midtpunkt (middle response alternative) | 2) Ved å nevne midtpunktet i innledningen til spørsmålet økes sannsynligheten for at dette veiges slett om det ikke er gitt som et eksplisitt alternativ. 3) Plassering av midtalternativet (i midten eller til slutt) har betydning 4) Respondenter som velger midtalternativet når dette er tilgjengelig, vil ikke nødvendigvis svare på spørsmålet på samme måte som andre respondenter når de blir tvunget til å velge side |
| Krosnick, og Alwin (1987) | Teste en kognitiv teori for å forklare effekter av rekkefølge på responsalternativer i spørreundersøkelser ved måling av foreldres rangering av hvilke verdier det er viktig at barn har | Kognitiv teori "primacy"- og "recency" effekter Litteratur-gjenomgang Psykologi | "Split-ballot" eksperiment der rekkefølgen på responsalternativer ble variert ved bruk av "show-cards" | Plassering av et responsalternativ blant de tre første på en liste, øker sannsynligheten for at det rangeres som et av de tre viktigste. Respondenter som er mindre kognitivt sofistikert blir i sterkere grad påvirket av endring i responsalternativ rekkefølge |
| Poe, Seeman, McLaughlin, Mehl, og Dietz (1988) | Måling av effekt av å inkludere "vet ikke" alternativ på response-rate ved fakaspørsmål i en postal survey | Ad-hoc, mangler teori | Eksperiment med spørreskjema med og uten "vet ikke" alternativ gjennomført i undersøkelse der man skulle finne ut hva man husket fra en avdød i nær families siste leveår | Selv-administrert spørreskjema uten "vet ikke" anbefales fordi: (i) høyere item response rate (ii) mindre behov for korrigering av missing, og dermed lavere sannsynlighet for å introdusere en feilkilde i dataene. (iii) bruk av "vet ikke" alternativ fører til få store substansielle endringer i dataene (iv) ingen vesentlige forskjeller i responsefeilrate når "vet ikke" alternativ utelates (v) utelatelse av "vet ikke" alternativer påvirker total response rate (vi) Spørreskjemaene blir kortere og enklere å fylle ut |
| Ayidiya, og McClendon (1990) | Undersøke om funn vedrørende rekkefølge på spørsmål, rekkefølge på svaralternativer, "vet ikke"-filter, midtpunkt på skala og føyelighet i personlige intervju undersøkelser også gjelder i postale spørreundersøkelser | Literaturgjennomgang Ad-hoc teoribruk "Recency"- og "primacy" effekter | Eksperiment med spørreskjema i survey med 532 respondenter. Randomisert tildeling av to spørreskjema varianter i postal spørreundersøkelse | "Recency" effekter oppstår ikke i postale skriftlige spørreundersøkelser. Finnes en svakt signifikant "primacy" effekt der midtalt. "fortsette som før" veiges oftere der dette er plassert i midten i forhold til der det er plassert til slutt. |
| Bishop (1990) | Forsøk på å teste antakelsene om at responseffekter knyttet til variasjoner i spørsmålform, | Literaturgjennomgang | Meta-analyse over 15 spørreskjema eksperimenter | Effekter av spørsmålrekkefølge og responsalternativ rekkefølge ser ut til å være upåvirket av respondentenes grad av involvering i temaet |

Tabell A.2 - Fortsetter

| STUDIE | FORMÅL | PERSEKTIV | METODE | FUNN/BESLUTNINGER |
|--|--|--|--|--|
| Alwin, og Krosnick (1991) | ordlyd eller kontekst vil være sterkere for de respondentene som er minst involvert i temaet det spørres om. Effekter av karakteristika ved spørsmålsstilling og respondent på reliabilitet i målinger i spørre-skjemaundersøkelser | Datakvalitet Reliabilitet | Data fra 5 ulike paneldatasett med minimum 200 resp. i hver. Reliabilitetsanalyse for attityde-spørsmål mhp. (i) antall responskategorier (ii) bruk av midtpunkt på skala (iii) benevnelse på svarålt. (iv) bruk av "vet ikke" alternativ | for undersøkelsen. Responseeffikter knyttet til bruk av balansert argumentasjon (counter arguments) og midtalternativer er sterkere for respondenter med lavere engasjement i temaet det spørres om (i) Reliabilitet øker når responsalternativene på en vurderingsskala øker fra 2 til 7 og flater deretter ut. For enig -uenig skalaer finnes ikke signifikante forskjeller i reliabilitet for 2 - og 5 punkts skaler (ii) Bruk av midtpunkt på skala reduserer reliabilitet, men kan fungere som forankringspunkt på lange skalaer (f.eks. mer enn 7) (iii) Benevnelse på alle skalaverdier gir høyere reliabilitet enn benevnelse bare på endepkt. (iv) Bruk av "vet ikke" alternativ reduserer reliabilitet |
| Schwarz, Knauper, Hippler, Noelle-Neumann, og Clark (1991) | Effekt av tallskala (+/+ eller -/+) på respondenters tolking av benevnelser på endepunktene på en vurderingsskala | Ad-hoc teori- bruk i diskusjon av resultater | Tre eksperimenter: Eksp. 1: måling av tilfredshet med livet. Bruk av 0 - 10 skala og -5 - 5 skala, 1032 resp. i nasjonal survey Eksp. 2: replikasjon av eksp. 1 med 101 studenter ved tysk universitet Eksp. 3: nøyaktighet i inferens om andre basert på minimumt med informasjon 0 - 10 skala og -5 - 5 skala, 22 studenter | Respondenter kan bruke tallverdiene på vurderingsskalaen for å gi mening til benevnelser på skalen (i) 0 - 10 skala: Unipolar konseptualisering, grad av tilstedeværelse av egenskapen som vurderes. (ii) - 5 - 5 skala: Bipolar, konseptualisering tilstedeværelse = 0, negativ verdi indikerer fravær av egenskapene Anbefalinger: (a) Semantisk differensial -> negativ til positiv skala (b) Informasjon/tilstedeværelse av en bestemt egenskap -> 0 til positiv skala (c) Valg av skala baseres på kunnskap om responsdistribusjon for relevante svar |

Tabell A.3

Empiriske studier. Problemstillinger i tilknytning til rekkefølge på spørsmål i spørreundersøkelser

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|------------------------------------|--|--|--|--|
| Jones, og Lang (1982) | Fokuserer blant annet på rekkefølge på sammenlignings-spørsmål og attributtvurderings-spørsmål ved vurderinger av sitt eget nabolag | Reliabilitet og validitet. "Group mean" og "individual-level" valideringstilnærming | Feltestperiment i en survey der subjekter var 2926 personer som hadde kjøpt nytt sted å bo det siste året. Manipulering av sponsor, følgebrev, rekkefølge på spørsmål og kontaktmåte. Randomisert tilordning av subjekter | Pretesting indikerer at attributtvurderings-spørsmål er lettere å svare på enn sammenlignings-spørsmål med attributtvurderings-spørsmål for sammenligningspersoner gj høyere reliabilitet og ingen effekt på validitet i tillegg til en positiv effekt på responsrate sammenlignet med spørreskjema der sammenlignings-spørsmålene stilles først. |
| Schuman, Kalton, og Ludwig (1983) | Undersøke om konteksteffekter kan reduseres ved å legge inn nøytrale spørsmål mellom spørsmål som man antar, eller vet vil påvirke hverandre når det gjelder svargivning | Ad-hoc, ingen teoribruk | Eksperiment ved måling av amerikanske holdninger til åttitaamerikanske reportere til Sovjet (Amer) vs. tillate russiske reportere til USA (Com). Manipulering av rekkefølge: Com/Amer, Amer/Com, Amer/17 nøytrale spm./Com | Finner signifikant kontekst effekt: 70% ja ved Amer/Com 45% ja ved Com/Amer 66% ja ved Amer/17 nøytrale spm./Com Indikasjon på at det ikke reduserer kontekst-effektene ved å innføre nøytrale mellomspørsmål |
| Crespi, og Morris (1984) | Effekt av spørsmålsrekkefølge ved meningsmåling i sammenheng med senats- og guvernørvalg | Ad-hoc, ingen teoribruk | Survey med 763 respondenter der halvparten ble først spurt om hvordan de ville stemme ved Senat-valget og deretter hvordan de ville stemme ved valg av guvernør. Den andre halvparten ble først spurt om hvordan de ville stemme ved valg av guvernør. | Moisatte preferanser for to ulike guvernør-kandidater ved bruk av de to alternative rekke-følgene på spørsmålene. |
| McClendon, og O'Brien (1988) | Undersøke effekten av å plassere et sett av spesifikke tilfredshets-spørsmål før eller etter et generelt tilfredshets spørsmål. | Literaturgjennomgang Informasjons-behandlings-teori | "Split-ballot" eksperiment med manipulering av rekkefølge på generelle og spesifikke spørsmål i spørreskjema | Ingen signifikant effekt av rekkefølge på spørsmål for gjennomsettene på de spesifikke spørsmålene. Signifikante effekter av rekkefølge på regressjonsligningen for gifte respondenter. Ekteskap og helse hadde høyere betydning når de spesifikke spørsmålene kom først, mens for hold ved arbeidet her hadde mindre betydning. |
| Tourangeau, Rasinski et al. (1989) | Undersøke om overføringseffekter kan generaliseres over ulike holdningstaner. Undersøke effektene knyttet til bruk av "buffer" spørsmål mellom kontekst og hovedspørsmål. Undersøke effekter av holdningskonflikter og | Literaturgjennomgang Informasjons-behandlingsteori Holdningsteori | Ti versjoner av et spørreskjema ble benyttet i telefonintervjuer. Randomisert tildelt 2629 hushold. Manipulering av rekkefølge på spørsmål | Faktorer som påvirker betydning av kontekst: Respondenter med blandede holdninger eller med holdninger i konflikt, men som er sentrale, er spesielt påvirkelige av konteksteffektene. Tilstedeværelse av nøytrale/irrelevante "buffer" spørsmål mellom kontekst- og hoved-spørsmålene gav ingen klare effekter. |

Tabell A.3 - Fortsetter

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|------------------------------------|--|--|---|--|
| | grad av sentralitet i holdningene. | | | Retning på korrelasjoner mellom respons på hovedspørsmålene påvirker retningen på konteksteffektene. Konteksteffektene var ikke avhengig av kjennskap til tema eller respons på kontekstspørsmålene. Det samme gjelder tughet med kontekstspørsmålene. |
| Tourangeau, Rasinski et al. (1991) | Undersøke om det er slik at dersom et generell tilfredshetsspørsmål følger etter et spesifikt spørsmål om tilfredshet med ekteskapet, så vil respondentsene tolke det generelle spørsmålet til å gjelde forhold som kommer i tillegg til forhold ved ekteskapet. Dette kalles "subtraksjons hypotesen" | Literaturgjennomgang av studier vedrørende konteksteffekter | Fire versjoner av spørreskjema i et "split-ballot" eksperiment med ulik rekkefølge og ordlyd på det generelle spørsmålet. | Versjonen der det generelle spørsmålet eksplisitt hadde med "i tillegg til ekteskapet", gav respons som var lik versjonen der dette ikke var tatt med når det generelle spørsmålet kom etter det spesifikke. Versjonen der det generelle spørsmålet eksplisitt hadde med "inkludert ekteskapet", gav respons som var lik versjonen der dette var utelatt når det generelle spørsmålet kom før det spesielle. Subtraksjonshypotesen blir ikke støttet. |
| Benton, og Daly (1991) | Undersøke effekter av rekkefølge på spesifikke- vs. generelle spørsmål i en undersøkelse om kvaliteten ved offentlige tjenester | Literaturgjennomgang Replikasjon av studiet | To versjoner av spørreskjema med manipulering av rekkefølge på generelle og spesifikke vurderingsspørsmål. Telefonintervjuer av 808 respondenter | Ingen effekt av rekkefølge på spørsmålene Responsene varierer som en funksjon av utdanning. Spørsmålsrekkefølgeeffektene større for respondenter med lav utdanning |
| Schwarz, Strack, og Mai (1991) | Undersøke om effekter av rekkefølge på generelle og spesifikke spørsmål om tilfredshet med livet påvirkes av i hvilken grad spørsmålene oppfattes som å høre sammen (conversational logic) | Literaturgjennomgang Informasjons- behandlings teori Kognitiv teori | Ni forskjellige spørreskjema med ulik rekkefølge og instruksjon på spørsmålene i tillegg til ulikt antall spesifikke spørsmål: Gen./Spes. - 3 spes. Spes./Gen. - 1/3 spes. Spes./Gen./felles inl. - 1/3 spes. Spes./Gen. /eksp. inkl - 1/3 spes. Spes./Gen. /eksp. ekskl - 1/3 spes. | Sammenhenger mellom spesifikke og generelle spørsmål: Sterkest sammenheng når Spes./Gen. - 1 spes. og Spes./Gen. /eksp. inkl - 1 spes. Indikerer at respondentsene benytter den sist aktive informasjonen når de svarer på det generelle spørsmålet. Flere spesifikke spørsmål øker tilgjengeligheten til et variert sett av potensiell relevant informasjon og reduserer dermed sammenhengen mellom et spesifikt spørsmål (relasjons tilfredshet) og det generelle spørsmålet. Effekten av konversasjons norm operasjonalisert ved felles introduksjon til de spesifikke og generelle spørsmålene bidrar ikke til økning i korrelasjon mellom de spesifikke- og det generelle spørsmålet. Forskjeller i gjennomsnitt: Respondenter som rapporterte en høy relasjons-tilfredshet, rapporterte også en høyere generell tilfredshet med livet. |

| STUDIE | FORMÅL | PERSPEKTIV | METODE | FUNN/BESLUTNINGER |
|---|--|---|---|---|
| Bickart, Menon, Sudman, og Blair (1992) | Undersøke hvordan kontekstuelle faktorer, som spørreskjema-format, påvirker respondenter bruk av selvrapportering når de blir bedt om å vurdere ektefellen | Ad-hoc preget men teori basert litteraturgjennomgang og hypoteser | Eksperiment ved gjennomføring av survey for å måle tilfredshet med ektefellen relatert til forhold som TV-seing, alkoholforbruk, helse og holdninger til offentlig alkoholpolitikk. Manipulering av tilgjengelighet til selvrapport ved å variere avstanden mellom selv- og andre(proxy) spørsmål. a) selvrapporteringer for proxy b) selvrapportering a, proxy a, selvrapportering b, proxy b, osv. Sammenligningsbasis manipulert ved å benytte et "priming" spørsmål som aktiverer en bestemt sammenlikningsstandard. | Avhengig variabel: vektlegging av selvrapport ved vurdering av andre (proxy) Når selvrapportering og proxy vurdering er skilt fra hverandre med en serie spørsmål, tillegges mer vekt til selvrapportering når likheter mellom ektefellene er primet i et spørsmål som kommer før proxy vurderingen. Vekt på selvrapport reduseres når selvrapport og ektefelle (proxy) rapportering er sekvensiell og likhet mellom ektefellene er primet |
| Schul, og Schiff (1993) | Undersøker effekt av rekkefølge på generelle- og spesifikke tilfredshets spørsmål ved måling av kunders tilfredshet med service organisasjoner. | Informasjons-behandlings-teori, Assimilations- og kontrassteori | "Split-ballot" eksperiment med to utvalg av kunder hos the Israel Jewish Telephone Company. (881 og 904 kunder) Telefonintervjuer. Manipulering av rekkefølge på generell tilfredshet spørsmål og spesifikke tilfredshet spørsmål. | Korrelasjoner mellom generell tilfredshet og attributtspesifikke tilfredshets spørsmål, høyere når de attributtspesifikke spørsmålene kommer først i spørreskjemaet Når generell tilfredshet spørsmål kommer etter de attributtspesifikke tilfredshets spørsmålene i spørreskjemaet økes tilgjengeligheten til både positiv og negativ informasjon. Når de generelle tilfredshets spørsmålene kommer først blir negativ informasjon mer tilgjengelig enn positiv informasjon |

B

Eksempel på spørreskjema benyttet for innsamling av data

HOTELLUNDERSØKELSE

Kjære gjest!

Vi håper De kan avsette tid til å svare på en del spørsmål i forbindelse med Deres opphold på hotellet hvor De fikk utdelt dette spørreskjemaet. Deres hjelp vil gi oss en mulighet til å utvikle og forbedre tilbudet av hotellovernattinger i Norge.

Undersøkelsen gjennomføres som ledd i et større reiseliussprosjekt ved Norges Handelshøyskole.

*De som leverer inn skjemaet i utfylt stand, vil delta i trekning av en gavesjekk, **verdi: kr. 5000,-**.*

Dersom De ønsker å delta i denne trekningen ber vi Dem oppgi navn og adresse på siste side.

Alle opplysninger vil bli behandlet konfidensielt.

Vær vennlig å levere skjemaet i resepsjonen før De forlater hotellet, eller legg det i vedlagte svarkonvolutt og putt det i nærmeste postkasse.

Vi takker på forhånd for Deres hjelp og ønsker Dem et fortsatt hyggelig ferieopphold!

1. **Hva er navnet på hotellet hvor De fikk utdelt dette spørreskjemaet?** Oppgi navn på hotell og sted.

.....

2. **Hvilken dato ankom De og Deres reisefølge dette hotellet ?**

.....

3. **Hvor mange netter har De overnattet på dette hotellet hittil på denne ferieturen?**

.....overnattinger

4. **Benytter De Hotellpass i forbindelse med Deres hotellovernattinger i denne ferien?**

Sett sirkel rundt det tallet som angir deres svar.

1. Ja 2. Nei

5. Hva er Deres nasjonalitet?

.....

6. Hva er Deres alder?

.....år.

7. **Kjønn** Sett sirkel rundt det tallet som angir Deres svar.

1. Kvinne 2. Mann

8. **Hvor lenge planlegger De å feriere i Norge?**

..... dager.

9. **La oss anta at De skal velge hotell som overnattingssted for neste feriereise. Hvilke egenskaper ved hotellene ville De ha benyttet for å skille mellom de ulike alternativene basert på Deres nåværende kunnskaper?**

| | |
|--------|---------|
| 1..... | 10..... |
| 2..... | 11..... |
| 3..... | 12..... |
| 4..... | 13..... |
| 5..... | 14..... |
| 6..... | 15..... |
| 7..... | 16..... |
| 8..... | 17..... |
| 9..... | 18..... |

10. **Hvordan tror De Deres venner vurderer Deres kunnskaper når De skal vurdere et hotell?**

Sett en sirkel rundt tallet som uttrykker Deres svar.

Svært dårlige kunnskaper

Svært gode kunnskaper

1 2 3 4 5 6 7

11. **I hvilken grad er De enig eller uenig i følgende utsagn: "Når jeg skal vurdere et hotell vet jeg hvilke egenskaper som er viktige, og som må være tilstede for at jeg skal bli fornøyd med oppholdet"**

Sett en sirkel rundt tallet som uttrykker Deres svar.

Svært uenig

Svært enig

1 2 3 4 5 6 7

12. **Hvor mange ganger har De overnattet på hotell hittil på denne ferieturen?**

.....overnattinger

13. Hvor mange ganger har De overnattet på hotell i løpet av de siste 3 årene? Sett sirkel rundt det tallet som angir Deres svar.

- | | |
|--------------------------|---------------------------------|
| 1. 0 - 5 overnattinger | 4. 16 - 25 overnattinger |
| 2. 6 - 10 overnattinger | 5. 26 - 35 overnattinger |
| 3. 11 - 15 overnattinger | 6. 36 overnattinger eller flere |

14. Hvor mange ganger har De overnattet på dette hotellet i løpet av de siste 3 årene? Sett sirkel rundt det tallet som angir Deres svar.

- | | |
|--------------------------|---------------------------------|
| 1. 0 - 5 overnattinger | 4. 16 - 25 overnattinger |
| 2. 6 - 10 overnattinger | 5. 26 - 35 overnattinger |
| 3. 11 - 15 overnattinger | 6. 36 overnattinger eller flere |

15. Når bestemte De og Deres reisefølge å legge ferien til Norge? Sett sirkel rundt det tallet som angir Deres svar.

- | | | |
|-------------------------------|--------------------------------|----------------------------------|
| 1. Mer enn 3 mnd. før avreise | 2. Mellom 1-3 mnd. før avreise | 3. Mindre enn 1 mnd. før avreise |
|-------------------------------|--------------------------------|----------------------------------|

16. Hvem reiser De sammen med? Sett sirkel rundt det tallet som angir Deres svar.

- | | | |
|---------------------------------|--|-------------------------|
| 1. Alene | 2. Venner | 3. Ektefelle |
| 4. Familie med barn under 15 år | 5. Familie uten barn, eller med barn 15 år eller eldre | 6. Gruppe/Selskapsreise |
| 7. Andre | | |

17. Hvor mange i Deres reisefølge (inkludert barn) er:

- | | |
|-------------------------|-------------------------------|
| Under 7 år.....personer | 25 - 34 år.....personer |
| 7 - 14 år.....personer | 35 - 44 år.....personer |
| 15 - 19 år.....personer | 45 - 66 år.....personer |
| 20 - 24 år.....personer | 67 år eller mer:.....personer |

18. Hvordan ble Deres ferietur organisert? Sett sirkel rundt det tallet som angir Deres svar.

1. Selskaps/grupperreise arrangert av reisebyrå, reisearrangør, forening, turoperatør etc.
2. Individuelt opplagt tur for Dem og Deres reisefølge helt eller delvis bestilt gjennom reisebyrå/reisearrangør
3. Individuelt opplagt tur arrangert av Dem selv og Deres reisefølge uten forhåndsbestilling gjennom reisebyrå/reisearrangør

19. Hvor viktig mener De følgende faktorer er for et vellykket hotell-opphold? Sett en sirkel rundt tallet som uttrykker Deres oppfatning.

| | <i>Helt uvesentlig</i> | | | | | <i>Svært vesentlig</i> | | |
|--|------------------------|---|---|---|---|------------------------|---|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 1. Personalets imøtekommenhet ved innsjekking på hotellet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 2. Informasjon om aktivitetstilbud på stedet via brosjyrer, oppslagstavler o.l på hotellet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 3. Hotellets beliggenhet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 4. Aktivitetstilbudet på hotellet (f.eks. svømming, trim, minigolf, solbad,...) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 5. Personalets høflighet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 6. Maten på hotellet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 7. Hotellrommenes utseende/Innendørsarkitektur | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 8. Kultur og underholdning på stedet (f.eks. konserter, revy, teater,.....) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 9. Personalets serviceinnstilling | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 10. Åpningstidene i bar og restaurant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 11. Baderommet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 12. Turmulighetene på stedet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 13. Personalets evne til å løse problemer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 14. Menyens variasjon | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 15. Sengens komfort | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 16. Natteliv og underholdningstilbudet om kvelden | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 17. Personalets evne til å informere om aktivitetstilbud på stedet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 18. Rutine/prosedyrer for betaling (f.eks. for opphold, minibar, restaurantbesøk) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 19. Rengjøring av rommet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 20. Shoppingtilbudet på stedet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 21. Personalets evne til å yte rask service | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 22. Ventilasjon på rommet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 23. Hotellets utseende/arkitektur | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 24. Informasjon om sikkerhetsrutiner (f.eks branninstruks) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 25. Hotellets utendørsanlegg (f.eks hotellhage) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |

Hvor viktig forts.....

| | Helt uvesentlig | | | | | Svært vesentlig | | |
|--|-----------------|---|---|---|---|-----------------|---|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 26. Parkeringsmulighetene | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| 27. Mulighet for rolig og fredelig opphold | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |

20. Nedenfor ber vi Dem gi en vurdering av det hotellet hvor De fikk utlevert dette spørreskjemaet. Hvor tilfreds eller utilfreds er De når det gjelder de egenskapene som er nevnt nedenfor? I tillegg ber vi Dem gi uttrykk for hvor sikker De er i Deres vurdering.
 Sett en sirkel rundt tallet som uttrykker Deres svar.

Et eksempel:

| | Svært utilfreds | | | | | | | | | | Svært tilfreds | | | | | | | | | | Hvor sikker er De i Deres vurdering ? | | | | | | |
|-----------------------------|-----------------|----|----|----|----|---|---|---|---|---|----------------|----|----|----|----|----|---|---|---|---|---------------------------------------|---|---|---|---|---|---|
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 1. Informasjon om utflukter | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |

Denne gjesten er hverken tilfreds eller utilfreds med hotellets informasjon om utflukter. Rubrikken til høyre viser at gjesten er nokså sikker i denne vurderingen.

| | Svært utilfreds | | | | | | | | | | Svært tilfreds | | | | | | | | | | Hvor sikker er De i Deres vurdering ? | | | | | | |
|---|-----------------|----|----|----|----|---|---|---|---|---|----------------|----|----|----|----|----|---|---|---|---|---------------------------------------|---|---|---|---|---|---|
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 1. Personalets imøtekommenhet ved innsjekking på hotellet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 2. Informasjon om aktivitets-tilbud på stedet via brosjyrer, oppslagstavler o.l på hotellet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 3. Hotellets beliggenhet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 4. Aktivitetstilbudet på hotellet (f.eks. svømming, trim, minigolf, solbad,...) | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 5. Personalets høflighet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 6. Maten på hotellet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 7. Hotellrommenes utseende/innendørsarkitektur | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 8. Kultur og underholdning på stedet (f.eks. konserter, revy, teater,.....) | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 9. Personalets serviceinnstilling | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |

Hvor tilfreds eller utilfreds er De, forts.....

**Hvor sikker er De i
Deres vurdering ?**

| | Svært utilfreds | | | | | | | | | | | Svært tilfreds | | | | | | | | | | | Svært usikker | | | | | Svært sikker | | | | |
|---|----------------------------|----|----|----|----|---|---|---|---|---|---|---------------------------|----|----|----|----|---|---|---|---|---|---|--------------------------|---|---|---|---|-------------------------|---|---|---|---|
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 10. Åpningstidene i bar og restaurant | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 11. Baderommet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 12. Turmulighetene på stedet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 13. Personalets evne til å løse problemer | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 14. Menyens variasjon | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 15. Sengens komfort | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 16. Natteliv og underholdnings-tilbudet om kvelden | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 17. Personalets evne til å informere om aktivitets-tilbud <u>på stedet</u> | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 18. Rutine/prosedyrer for betaling (f.eks. for opphold, minibar, restaurantbesøk) | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 19. Rengjøring av rommet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 20. Shoppingtilbud på stedet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 21. Personalets evne til å yte rask service | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 4 | 1 | 2 | 3 | 4 | 5 |
| 22. Ventilasjon på rommet | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 23. Hotelllets utseende/arkitektur | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 24. Informasjon om sikkerhetsrutiner (f.eks branninstruks) | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 25. Hotelllets utendørsanlegg (f.eks hotellhage) | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 26. Parkeringsmulighetene | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 27. Mulighet for rolig og fredelig opphold | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |

21. Hva vil De eventuelt trekke frem som spesielt positivt eller negativt ved hotellet?

Spesielt positivt:.....

.....

28. For mange er valg av hotell i ferien en viktig beslutning. Nedenfor har vi listet opp en del informasjonskilder som ofte benyttes før valg av hotell. Hvordan vi De vurdere Deres bruk av de ulike informasjonskildene når De skulle velge det hotellet hvor De fikk utdelt dette spørreskjemaet? Sett sirkel rundt det tallet som angir Deres svar.

| | <i>Svært lite benyttet</i> | | | | | | | <i>Svært mye benyttet</i> | |
|--|----------------------------|---|---|---|---|---|---|---------------------------|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| 1. Informasjon fra reisebyrå/turoperatør | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| 2. Reiseberetninger i ukeblader/ferie og fritidsmagasiner/TV-programmer/kino | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| 3. Anbefaling fra familie/venner/ arbeidskollegaer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| 4. Brosjyre med informasjon om hotellet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| 5. Feriekatalog | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |
| 6. Erfaring fra tidligere besøk på dette hotellet | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | |

29. Vurderte De flere hotell enn dette som alternative overnattingssteder under forberedelsene til denne ferien? Sett en sirkel rundt tallet som uttrykker Deres svar.

1. Nei 2. Ja Dersom Ja, kan de huske hvor mange? Antall hotell:.....

30. Hvor viktig er det for Dem at De har gjort et godt valg av hotell som overnattingssted for ferien?

Sett en sirkel rundt tallet som uttrykker Deres svar.

Svært uenig

Svært enig

1 2 3 4 5 6 7

*Takk for at De tok Dem tid til å fylle ut skjemaet.
Dersom De ønsker å delta i trekningen av en gavesjekk på **kr.5000,-**,
ber vi Dem oppgi navn og adresse nedenfor*

Navn:.....

Adresse:.....

Vær vennlig å levere skjemaet i resepsjonen før De forlater hotellet, eller legg det i vedlagte svarkonvolutt og putt det i nærmeste postkasse

C

Korrelasjonsmatrise og operasjonaliseringer

Tabell C.1

Korrelasjonsmatrise for vurderings spørsmål, totalvurderinger og atferdsintensjoner

| | VU 1 | VU 5 | VU 9 | VU 13 | VU 17 | VU 21 | VU 2 | VU 6 | VU 10 | VU 14 | VU 18 | VU 24 | VU 3 | VU 7 | VU 11 | VU 15 | VU 19 | VU 22 | VU 23 | VU 25 | VU 26 | VU 27 | VU 4 | VU 8 | VU 12 | VU 16 | VU 20 | S2 | S3 | S4 | S5 | S6 | | | | |
|------|------|------|------|-------|-------|-------|------|------|-------|-------|-------|-------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|-------|-------|-------|-----|-----|-----|-----|----|--|--|--|--|
| VU1 | .60 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU5 | .59 | .73 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU9 | .37 | .34 | .41 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU13 | .35 | .38 | .37 | .52 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU17 | .37 | .43 | .51 | .61 | .48 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU21 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU2 | .34 | .37 | .31 | .33 | .55 | .26 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU6 | .35 | .48 | .45 | .29 | .35 | .34 | .33 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU10 | .32 | .37 | .39 | .40 | .39 | .39 | .26 | .41 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU14 | .32 | .35 | .37 | .46 | .44 | .34 | .56 | .45 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU18 | .40 | .46 | .48 | .47 | .48 | .54 | .40 | .40 | .48 | .41 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU24 | .28 | .35 | .33 | .37 | .42 | .38 | .35 | .37 | .31 | .38 | .43 | | | | | | | | | | | | | | | | | | | | | | | | | |
| VU3 | .28 | .44 | .41 | .31 | .24 | .26 | .26 | .34 | .28 | .30 | .36 | .27 | | | | | | | | | | | | | | | | | | | | | | | | |
| VU7 | .34 | .37 | .44 | .31 | .26 | .38 | .27 | .36 | .30 | .31 | .33 | .27 | .31 | | | | | | | | | | | | | | | | | | | | | | | |
| VU11 | .32 | .36 | .39 | .36 | .29 | .37 | .24 | .28 | .31 | .26 | .27 | .22 | .21 | .60 | | | | | | | | | | | | | | | | | | | | | | |
| VU15 | .35 | .41 | .39 | .31 | .30 | .40 | .23 | .36 | .31 | .31 | .32 | .26 | .30 | .46 | .48 | | | | | | | | | | | | | | | | | | | | | |
| VU19 | .37 | .41 | .49 | .38 | .37 | .42 | .31 | .40 | .36 | .28 | .50 | .36 | .27 | .42 | .43 | .36 | | | | | | | | | | | | | | | | | | | | |
| VU22 | .35 | .40 | .38 | .42 | .36 | .44 | .30 | .36 | .34 | .35 | .35 | .39 | .29 | .45 | .46 | .42 | .37 | | | | | | | | | | | | | | | | | | | |
| VU23 | .30 | .35 | .36 | .35 | .30 | .38 | .30 | .37 | .37 | .38 | .38 | .43 | .33 | .46 | .34 | .32 | .37 | .50 | | | | | | | | | | | | | | | | | | |
| VU25 | .23 | .23 | .24 | .27 | .34 | .23 | .41 | .34 | .31 | .36 | .32 | .33 | .33 | .42 | .21 | .12 | .20 | .17 | .33 | .38 | | | | | | | | | | | | | | | | |
| VU26 | .29 | .24 | .26 | .16 | .24 | .25 | .21 | .32 | .27 | .22 | .26 | .31 | .18 | .26 | .21 | .27 | .28 | .28 | .33 | .38 | | | | | | | | | | | | | | | | |
| VU27 | .24 | .35 | .28 | .21 | .28 | .29 | .33 | .31 | .33 | .27 | .32 | .37 | .39 | .25 | .18 | .30 | .31 | .46 | .33 | .35 | .40 | | | | | | | | | | | | | | | |
| VU4 | .15 | .22 | .21 | .24 | .29 | .22 | .45 | .30 | .25 | .33 | .31 | .22 | .30 | .17 | .08* | .14 | .19 | .18 | .18 | .51 | .17 | .34 | | | | | | | | | | | | | | |
| VU8 | .23 | .29 | .35 | .30 | .30 | .31 | .30 | .22 | .32 | .24 | .25 | .17 | .29 | .34 | .32 | .23 | .27 | .27 | .25 | .13 | .13 | .20 | .20 | | | | | | | | | | | | | |
| VU12 | .18 | .31 | .25 | .34 | .32 | .26 | .38 | .24 | .33 | .36 | .36 | .28 | .32 | .16 | .20 | .19 | .30 | .18 | .25 | .27 | .19 | .32 | .37 | .32 | | | | | | | | | | | | |
| VU16 | .26 | .34 | .31 | .40 | .44 | .40 | .29 | .33 | .56 | .37 | .42 | .32 | .26 | .29 | .30 | .32 | .35 | .29 | .20 | .26 | .30 | .32 | .25 | .40 | .37 | | | | | | | | | | | |
| VU20 | .27 | .30 | .28 | .34 | .34 | .39 | .30 | .20 | .31 | .30 | .36 | .23 | .29 | .29 | .30 | .24 | .33 | .27 | .29 | .17 | .17 | .17 | .07 | .43 | .34 | .40 | | | | | | | | | | |
| S2 | .34 | .39 | .44 | .40 | .33 | .44 | .28 | .37 | .33 | .31 | .34 | .35 | .29 | .55 | .45 | .41 | .41 | .51 | .45 | .30 | .29 | .34 | .20 | .26 | .20 | .32 | .23 | | | | | | | | | |
| S3 | .33 | .41 | .50 | .38 | .32 | .45 | .29 | .40 | .30 | .31 | .31 | .33 | .27 | .56 | .49 | .45 | .44 | .50 | .42 | .23 | .27 | .30 | .21 | .26 | .18 | .27 | .20 | .84 | | | | | | | | |
| S4 | .39 | .44 | .49 | .38 | .38 | .46 | .29 | .43 | .32 | .33 | .36 | .34 | .37 | .54 | .48 | .48 | .46 | .53 | .43 | .28 | .31 | .38 | .22 | .30 | .22 | .33 | .24 | .73 | .78 | | | | | | | |
| S5 | .31 | .38 | .44 | .28 | .31 | .38 | .28 | .34 | .26 | .30 | .27 | .36 | .39 | .27 | .31 | .32 | .41 | .32 | .28 | .17 | .33 | .26 | .20 | .16 | .30 | .14 | .63 | .61 | .65 | | | | | | | |
| S6 | .31 | .40 | .43 | .28 | .36 | .38 | .32 | .38 | .30 | .30 | .33 | .27 | .36 | .46 | .31 | .33 | .36 | .41 | .33 | .35 | .25 | .35 | .29 | .20 | .18 | .33 | .16 | .66 | .66 | .70 | .90 | | | | | |

Alle korrelasjonskoeffisienter signifikante, $p < 0.05$ (2-halet), med unntak av *, $n = 608$ (korrelasjonsmatrisen er basert på pairwise deletion av missing variable)

Tabell C.2
Backstage skalaen med spørsmål og statistiske egenskaper

| SPØRSMÅL* | GJ.SN. KORR. MED ØVRIGE SPM. INNENF. SKALAEN | GJ.SN. KORR. MED SPM. UTENFOR SKALAEN |
|--|---|--|
| VU2 Informasjon om aktivitetstilbudet på stedet via brosjyrer, oppslagstavler o.l. på hotellet | 0,33 | 0,32 |
| VU6 Maten på hotellet | 0,41 | 0,33 |
| VU10 Åpningstidene i bar og restaurant | 0,38 | 0,34 |
| VU14 Menyens variasjon | 0,43 | 0,33 |
| VU18 Rutine prosedyre for betaling | 0,42 | 0,38 |
| VU24 Informasjon om sikkerhetsrutiner | 0,37 | 0,31 |
| Gjennomsnittskorrelasjon | <u>0,39</u> | <u>0,34</u> |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendix B

Tabell C.3
Kundeavhengig skalaen med spørsmål og statistiske egenskaper

| SPØRSMÅL* | GJ.SN. KORR. MED ØVRIGE SPM. INNENF. SKALAEN | GJ.SN. KORR. MED SPM. UTENFOR SKALAEN |
|--|---|--|
| VU4 Aktivitetstilbudet på hotellet | 0,22 | 0,24 |
| VU8 Kultur og underholdning på stedet | 0,34 | 0,26 |
| VU12 Turmulighetene på stedet | 0,35 | 0,27 |
| VU16 Natteliv og underholdningstilbudet om kvelden | 0,36 | 0,33 |
| VU20 Shoppingtilbud på stedet | 0,31 | 0,28 |
| Gjennomsnittskorrelasjon | <u>0,32</u> | <u>0,28</u> |

* Spørsmålene i tabellen finnes i spørreskjemaet som er gjengitt i appendix B

D

Beskrivende statistikk

Tabell D.1
Beskrivende statistikk for hele utvalget

| VARIABEL | GJ. SNITT. | ST. AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|----------------------------------|---------------|--------------|-------|-------|------|------|-----|
| <i>Samhandling vurderinger</i> | | | | | | | |
| VU1 Imøtekommenhet | 4,27 | 1,22 | 10,84 | -2,86 | -4 | 5 | 606 |
| VU5 Høflighet | 4,32 | 1,11 | 10,54 | -2,71 | -4 | 5 | 594 |
| VU9 Serviceinnstilling | 4,09 | 1,31 | 7,60 | -2,32 | -5 | 5 | 590 |
| VU13 Problemløsningsevne | 2,61 | 2,09 | -1,60 | -0,45 | -4 | 5 | 512 |
| VU17 Evne til å informere | 1,99 | 2,28 | 0,15 | -0,63 | -5 | 5 | 531 |
| VU21 Evne til rask service | 3,31 | 1,85 | -0,02 | -1,04 | -4 | 5 | 537 |
| <i>Struktur vurderinger</i> | | | | | | | |
| VU3 Hotelllets beliggenhet | 3,71 | 1,76 | 4,85 | -2,04 | -5 | 5 | 595 |
| VU7 Hotellrom utseende | 3,23 | 1,91 | 2,68 | -1,63 | -4 | 5 | 606 |
| VU11 Baderommet | 3,25 | 2,13 | 2,76 | -1,71 | -5 | 5 | 578 |
| VU15 Sengens komfort | 3,45 | 1,87 | 5,24 | -2,10 | -5 | 5 | 591 |
| VU19 Rengjøring av rommet | 3,78 | 1,76 | 5,13 | -2,15 | -5 | 5 | 576 |
| VU22 Ventilasjon | 2,61 | 2,51 | 1,14 | -1,36 | -5 | 5 | 588 |
| VU23 Hotelllets utseende | 2,52 | 2,13 | 0,99 | -1,12 | -5 | 5 | 585 |
| VU25 Utendørsanlegg | 2,16 | 2,42 | 0,59 | -0,96 | -5 | 5 | 558 |
| VU26 Parkering | 3,69 | 1,94 | 6,25 | -2,38 | -5 | 5 | 587 |
| VU27 Rolig og fredelig | 3,44 | 2,18 | 3,39 | -1,91 | -5 | 5 | 589 |
| <i>Backstage vurderinger</i> | | | | | | | |
| VU2 Informasjon | 3,03 | 2,08 | 1,63 | -1,32 | -5 | 5 | 584 |
| VU6 Maten på hotellet | 3,82 | 1,53 | 4,55 | -1,95 | -4 | 5 | 574 |
| VU10 Åpningstider | 3,18 | 2,03 | 1,25 | -1,25 | -5 | 5 | 548 |
| VU14 Menyens variasjon | 2,75 | 1,96 | 1,24 | -1,10 | -5 | 5 | 547 |
| VU18 Rutine for betaling | 3,56 | 1,75 | 1,52 | -1,42 | -5 | 5 | 547 |
| VU24 Info. om sikkerhetsrut. | 2,55 | 2,28 | 0,87 | -1,07 | -5 | 5 | 574 |
| <i>Kundeavhengig vurderinger</i> | | | | | | | |
| VU4 Aktivitetstilbudet | 2,23 | 2,71 | 0,14 | -0,93 | -5 | 5 | 553 |
| VU8 Kultur og underholdn. | 1,39 | 2,10 | -0,05 | -0,09 | -5 | 5 | 530 |
| VU12 Turmulighetene | 3,04 | 2,09 | 0,58 | -1,06 | -5 | 5 | 553 |
| VU16 Natteliv | 2,27 | 2,18 | 0,11 | -0,72 | -5 | 5 | 534 |
| VU20 Shoppingtilbud | 2,93 | 2,06 | -0,33 | -0,81 | -4 | 5 | 551 |
| <i>Totalvurderinger</i> | | | | | | | |
| S2 Sammenlign. m. andre | 3,13 | 1,83 | 3,18 | -1,58 | -5 | 5 | 592 |
| S3 I forhold til pris | 3,33 | 1,76 | 3,25 | -1,68 | -5 | 5 | 592 |
| S4 Total tilfredshet | 3,59 | 1,42 | 5,92 | -1,87 | -5 | 5 | 590 |
| <i>Atferdsintensjoner</i> | | | | | | | |
| S5 Gjengkjøp | 9,07 | 2,31 | 2,00 | -1,53 | 1 | 11 | 595 |
| S6 Anbefaling til andre | 9,12 | 2,27 | 1,82 | -1,47 | 1 | 11 | 596 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.2

Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med objektforankret instruksjon og -vurderingsskala

| VARIABEL* | GJ. SNITT. | ST. AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|--------------------------|------------|-----------|-------|-------|-------|-------|-----|
| Summert gj.sn. vurdering | 3,21 | 1,07 | -0,14 | -0,64 | -0,62 | 4,96 | 316 |
| Samhandling vurdering | 3,61 | 1,14 | 0,43 | -0,86 | -1,00 | 5,00 | 316 |
| Struktur vurdering | 3,27 | 1,21 | 0,98 | -1,03 | -2,00 | 5,00 | 316 |
| Backstage vurdering | 3,20 | 1,30 | 0,24 | -0,84 | -1,67 | 5,00 | 316 |
| Kundeavhengig vurdering | 2,51 | 1,56 | 0,23 | -0,59 | -4,00 | 5,00 | 313 |
| Samhandling viktighet | 5,91 | 0,73 | 1,63 | -0,86 | 2,67 | 7,00 | 199 |
| Struktur viktighet | 5,69 | 0,61 | -0,23 | -0,30 | 3,80 | 7,00 | 199 |
| Backstage viktighet | 5,28 | 0,80 | 0,86 | -0,50 | 2,17 | 7,00 | 199 |
| Kundeavhengig viktighet | 4,18 | 1,10 | 0,50 | -0,24 | 1,20 | 7,00 | 199 |
| Totalvurdering | | | | | | | |
| Objektforankret | 3,18 | 1,79 | 4,25 | -1,81 | -4,50 | 5,00 | 307 |
| Subjektforankret | 3,62 | 1,79 | 7,71 | -2,04 | -5,00 | 5,00 | 305 |
| Atferdsintensjoner | 9,12 | 2,22 | 2,56 | -1,65 | 1,00 | 11,00 | 309 |
| Summert gj.sn. sikkerhet | 4,14 | 0,67 | 0,87 | -0,81 | 1,04 | 5,00 | 305 |
| Samhandling sikkerhet | 4,11 | 0,81 | -0,02 | -0,76 | 1,00 | 5,00 | 305 |
| Struktur sikkerhet | 4,39 | 0,59 | 3,77 | -1,44 | 1,00 | 5,00 | 301 |
| Backstage sikkerhet | 4,08 | 0,80 | 0,32 | -0,82 | 1,00 | 5,00 | 297 |
| Kundeavhengig sikkerhet | 3,74 | 1,01 | -0,49 | -0,59 | 1,00 | 5,00 | 295 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.3

Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med subjektforankret instruksjon og -vurderingsskala

| VARIABEL* | GJ. SNITT. | ST. AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|--------------------------|------------|-----------|-------|-------|-------|-------|-----|
| Summert gj.sn. vurdering | 3,09 | 1,25 | 2,56 | -1,17 | -3,69 | 4,92 | 292 |
| Samhandling vurdering | 3,46 | 1,35 | 0,75 | -1,00 | -2,17 | 5,00 | 291 |
| Struktur vurdering | 3,15 | 1,39 | 2,95 | -1,40 | -4,10 | 5,00 | 292 |
| Backstage vurdering | 3,13 | 1,45 | 2,98 | -1,31 | -4,50 | 5,00 | 291 |
| Kundeavhengig vurdering | 2,36 | 1,59 | 1,09 | -0,73 | -5,00 | 5,00 | 289 |
| Samhandling viktighet | 5,94 | 0,83 | 0,63 | -0,81 | 2,67 | 7,00 | 182 |
| Struktur viktighet | 5,72 | 0,75 | 2,57 | -1,10 | 2,30 | 7,00 | 181 |
| Backstage viktighet | 5,30 | 0,94 | 0,46 | -0,65 | 2,00 | 7,00 | 181 |
| Kundeavhengig viktighet | 4,23 | 1,09 | 0,20 | -0,32 | 1,00 | 7,00 | 179 |
| Totalvurdering | | | | | | | |
| Objektforankret | 3,29 | 1,63 | 1,80 | -1,34 | -4,00 | 5,00 | 307 |
| Subjektforankret | 3,55 | 1,43 | 4,33 | -1,71 | -3,00 | 5,00 | 307 |
| Atferdsintensjoner | 9,04 | 2,25 | 1,49 | -1,39 | 1,00 | 11,00 | 288 |
| Summert gj.sn. sikkerhet | 4,18 | 0,69 | 1,26 | -0,96 | 1,00 | 5,00 | 277 |
| Samhandling sikkerhet | 4,15 | 0,82 | 0,91 | -1,04 | 1,00 | 5,00 | 277 |
| Struktur sikkerhet | 4,44 | 0,55 | 0,39 | -0,97 | 2,70 | 5,00 | 272 |
| Backstage sikkerhet | 4,14 | 0,78 | 0,52 | -0,88 | 1,00 | 5,00 | 273 |
| Kundeavhengig sikkerhet | 3,74 | 1,07 | -0,31 | -0,73 | 1,00 | 5,00 | 269 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.4

Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema uten viktighetsspørsmål

| VARIABEL* | GJ. SNITT. | ST. AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|--------------------------|---------------|--------------|-------|-------|-------|-------|-----|
| Summert gj.sn. vurdering | 3,13 | 1,13 | -0,43 | -0,61 | 0,04 | 4,96 | 220 |
| Samhandling vurdering | 3,45 | 1,26 | 0,48 | -0,94 | -1,00 | 5,00 | 220 |
| Struktur vurdering | 3,18 | 1,23 | -0,29 | -0,71 | -0,50 | 5,00 | 220 |
| Backstage vurdering | 3,16 | 1,34 | -0,18 | -0,74 | -0,83 | 5,00 | 220 |
| Kundeavhengig vurdering | 2,54 | 1,54 | -0,34 | -0,50 | -2,40 | 5,00 | 218 |
| Totalvurdering | 3,09 | 1,77 | 1,42 | -1,29 | -3,50 | 5,00 | 219 |
| Atferdsintensjoner | 9,05 | 2,32 | 1,29 | -1,41 | 1,00 | 11,00 | 220 |
| Summert gj.sn. sikkerhet | 4,12 | 0,73 | 0,91 | -0,91 | 1,04 | 5,00 | 206 |
| Samhandling sikkerhet | 4,03 | 0,87 | 0,25 | -0,84 | 1,00 | 5,00 | 206 |
| Struktur sikkerhet | 4,38 | 0,63 | 4,19 | -1,56 | 1,00 | 5,00 | 202 |
| Backstage sikkerhet | 4,07 | 0,84 | 0,34 | -0,87 | 1,00 | 5,00 | 202 |
| Kundeavhengig sikkerhet | 3,73 | 1,07 | -0,24 | -0,75 | 1,00 | 5,00 | 199 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.5

Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med viktighetsspørsmål for spesifikke vurderingsspørsmål

| VARIABEL* | GJ. SNITT. | ST. AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|--------------------------|---------------|--------------|-------|-------|-------|-------|-----|
| Summert gj.sn. vurdering | 3,15 | 1,13 | 0,07 | -0,68 | -0,73 | 4,94 | 204 |
| Samhandling vurdering | 3,61 | 1,20 | 0,22 | -0,82 | -0,67 | 5,00 | 204 |
| Struktur vurdering | 3,28 | 1,28 | 1,71 | -1,19 | -2,00 | 5,00 | 204 |
| Backstage vurdering | 3,14 | 1,36 | 0,19 | -0,79 | -1,67 | 5,00 | 202 |
| Kundeavhengig vurdering | 2,21 | 1,62 | 0,97 | -0,56 | -5,00 | 5,00 | 198 |
| Samhandling viktighet | 5,86 | 0,80 | 0,73 | -0,78 | 2,67 | 7,00 | 203 |
| Struktur viktighet | 5,59 | 0,70 | 2,53 | -1,03 | 2,30 | 7,00 | 203 |
| Backstage viktighet | 5,11 | 0,90 | 0,68 | -0,59 | 2,00 | 7,00 | 203 |
| Kundeavhengig viktighet | 3,98 | 1,11 | 0,08 | -0,31 | 1,00 | 7,00 | 203 |
| Totalvurdering | 3,23 | 1,66 | 3,97 | -1,66 | -4,50 | 5,00 | 198 |
| Atferdsintensjoner | 9,03 | 2,25 | 1,91 | -1,47 | 1,00 | 11,00 | 199 |
| Summert gj.sn. sikkerhet | 4,15 | 0,69 | 1,73 | -1,05 | 1,00 | 5,00 | 200 |
| Samhandling sikkerhet | 4,20 | 0,80 | 0,98 | -1,10 | 1,00 | 5,00 | 200 |
| Struktur sikkerhet | 4,41 | 0,56 | 0,08 | -0,87 | 2,78 | 5,00 | 197 |
| Backstage sikkerhet | 4,06 | 0,80 | 0,22 | -0,79 | 1,00 | 5,00 | 196 |
| Kundeavhengig sikkerhet | 3,72 | 1,02 | -0,40 | -0,61 | 1,00 | 5,00 | 193 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.6

Beskrivende statistikk for eksperimentgruppen som er tilordnet spørreskjema med vi
hetsspørsmål etter spesifikke vurderingss spørsmål

| VARIABEL | GJ. SNITT. | ST. AVVIK | KURT. | SKEW. | MIN. | MAX. | N |
|--------------------------|---------------|--------------|-------|-------|-------|-------|-----|
| Summert gj.sn. vurdering | 3,19 | 1,23 | 5,15 | -1,61 | -3,69 | 4,88 | 184 |
| Samhandling vurdering | 3,57 | 1,26 | 1,80 | -1,18 | -2,17 | 5,00 | 183 |
| Struktur vurdering | 3,18 | 1,40 | 4,64 | -1,79 | -4,10 | 5,00 | 184 |
| Backstage vurdering | 3,22 | 1,42 | 5,83 | -1,84 | -4,50 | 5,00 | 182 |
| Kundeavhengig vurdering | 2,56 | 1,56 | 1,56 | -0,93 | -4,00 | 5,00 | 182 |
| Samhandling viktighet | 6,00 | 0,75 | 1,57 | -0,88 | 2,67 | 7,00 | 178 |
| Struktur viktighet | 5,83 | 0,63 | -0,26 | -0,38 | 4,20 | 7,00 | 177 |
| Backstage viktighet | 5,49 | 0,79 | 0,29 | -0,46 | 2,83 | 7,00 | 177 |
| Kundeavhengig viktighet | 4,46 | 1,01 | 0,58 | -0,13 | 1,20 | 7,00 | 175 |
| Totalvurdering | 3,40 | 1,70 | 6,16 | -2,10 | -4,50 | 5,00 | 178 |
| Atferdsintensjoner | 9,20 | 2,12 | 3,51 | -1,75 | 1,00 | 11,00 | 178 |
| Summert gj.sn. sikkerhet | 4,22 | 0,59 | -0,94 | -0,37 | 2,85 | 5,00 | 176 |
| Samhandling sikkerhet | 4,16 | 0,74 | -0,36 | -0,61 | 1,83 | 5,00 | 176 |
| Struktur sikkerhet | 4,46 | 0,50 | 0,51 | -0,98 | 2,90 | 5,00 | 174 |
| Backstage sikkerhet | 4,20 | 0,71 | 0,49 | -0,79 | 1,67 | 5,00 | 172 |
| Kundeavhengig sikkerhet | 3,76 | 1,03 | -0,56 | -0,62 | 1,00 | 5,00 | 172 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.7

Resultater av tester for univariat homogenitet i varians for H1 - H9b

| VARIABEL | COCHRANS C | BARTLETT- BOX F |
|--------------------------|-----------------------------|----------------------------------|
| Summert gj.sn. vurdering | C(303,2) = ,57804 p = 0,006 | F(1,1097107) = 7,47919 p = 0,006 |
| Samhandling vurdering | C(303,2) = ,58540 p = 0,003 | F(1,1093084) = 8,96739 p = 0,003 |
| Struktur vurdering | C(303,2) = ,56630 p = 0,021 | F(1,1097107) = 5,37626 p = 0,021 |
| Backstage vurdering | C(301,2) = ,55461 p = 0,058 | F(1,1083345) = 3,61059 p = 0,058 |
| Kundeavhengig vurdering | C(298,2) = ,50893 p = 0,758 | F(1,1062452) = 0,09475 p = 0,758 |
| Samhandling viktighet | C(190,2) = ,56306 p = 0,081 | F(1,428616) = 3,03609 p = 0,082 |
| Struktur viktighet | C(189,2) = ,59613 p = 0,008 | F(1,426066) = 7,12699 p = 0,008 |
| Backstage viktighet | C(189,2) = ,58301 p = 0,022 | F(1,426066) = 5,28508 p = 0,022 |
| Kundeavhengig viktighet | C(188,2) = ,50419 p = 0,909 | F(1,420937) = 0,01314 p = 0,909 |
| Totalvurdering | C(297,2) = ,54727 p = 0,103 | F(1,1052062) = 2,64895 p = 0,104 |
| Atferdsintensjoner | C(298,2) = ,50648 p = 0,823 | F(1,1058551) = 0,04987 p = 0,823 |
| Summert gj.sn. sikkerhet | C(290,2) = ,51558 p = 0,596 | F(1,1002943) = 0,28101 p = 0,596 |
| Samhandling sikkerhet | C(290,2) = ,50988 p = 0,737 | F(1,1002943) = 0,11286 p = 0,737 |
| Struktur sikkerhet | C(286,2) = ,53589 p = 0,225 | F(1,971412) = 1,46470 p = 0,226 |
| Backstage sikkerhet | C(284,2) = ,51163 p = 0,695 | F(1,963272) = 0,15306 p = 0,696 |
| Kundeavhengig sikkerhet | C(281,2) = ,53038 p = 0,308 | F(1,942136) = 1,03746 p = 0,309 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

Tabell D.8
Resultater av tester for univariat homogenitet i varians for H10a - H14b

| VARIABEL | COCHRANS C | | BARTLETT- BOX F | |
|--------------------------|-------------------|-----------|-----------------------|-----------|
| Summert gj.sn. vurdering | C(202,3) = ,37191 | p = 0,237 | F(2,813587) = 0,93290 | p = 0,394 |
| Samhandling vurdering | C(201,3) = ,34430 | p = 1,000 | F(2,810260) = 0,30647 | p = 0,736 |
| Struktur vurdering | C(202,3) = ,38416 | p = 0,097 | F(2,813587) = 1,76000 | p = 0,172 |
| Backstage vurdering | C(200,3) = ,35816 | p = 0,541 | F(2,801685) = 0,40261 | p = 0,669 |
| Kundeavhengig vurdering | C(198,3) = ,35453 | p = 0,653 | F(2,786783) = 0,33167 | p = 0,718 |
| Samhandling viktighet | C(190,2) = ,53486 | p = 0,337 | F(1,425945) = 0,91410 | p = 0,339 |
| Struktur viktighet | C(189,2) = ,55001 | p = 0,169 | F(1,423270) = 1,87752 | p = 0,171 |
| Backstage viktighet | C(189,2) = ,56672 | p = 0,066 | F(1,423270) = 3,35045 | p = 0,067 |
| Kundeavhengig viktighet | C(188,2) = ,54693 | p = 0,198 | F(1,417891) = 1,64218 | p = 0,200 |
| Totalvurdering | C(197,3) = ,35613 | p = 0,605 | F(2,775826) = 0,41026 | p = 0,664 |
| Atferdsintensjoner | C(198,3) = ,36013 | p = 0,490 | F(2,780505) = 0,76434 | p = 0,466 |
| Summert gj.sn. sikkerhet | C(193,3) = ,39164 | p = 0,058 | F(2,746340) = 4,04914 | p = 0,018 |
| Samhandling sikkerhet | C(193,3) = ,38744 | p = 0,082 | F(2,746340) = 2,34447 | p = 0,096 |
| Struktur sikkerhet | C(190,3) = ,40942 | p = 0,012 | F(2,723989) = 4,44205 | p = 0,012 |
| Backstage sikkerhet | C(189,3) = ,38005 | p = 0,149 | F(2,715391) = 2,72972 | p = 0,065 |
| Kundeavhengig sikkerhet | C(187,3) = ,35144 | p = 0,772 | F(2,701808) = 0,21109 | p = 0,810 |

Samhandling = immaterielle egenskaper, Struktur = materielle egenskaper

E

Hypotesetesting: tabeller og figurer

Tabell E.1

F-tester gjennomført separat for hvert hotell for høyere standardavvik knyttet til spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala

| ANTALL STANDARDAVVIK HØYEST I EKSPERIMENTGRUPPEN MED SUBJEKTFORANKRET INSTRUKSJON OG -VURDERINGSSKALA | | |
|--|---|-----------------------------------|
| HOTELL | a < 0,05 - EN-HALE TEST | a < 0,10 - EN-HALE TEST |
| Solborg Hotel, Mandal | VU8, 12, 20 | VU5, 22, 23, 10, 18 |
| Hotel Røros, Røros | VU1, 5, 9, 13, 21, 3, 19, 22, 25, 27, 6, 24, 4, 12 | VU17, 11 |
| Grand Hotel, Larvik | VU15 | VU8 |
| Grand Hotel Olav, Trondheim | VU1, 5, 3, 11, 22, 20 | VU14 |
| Hotel Scandinavie, Ålesund | VU11, 19, 8, 12 | VU5, 9, 13, 21 |
| Ulstein Hotel, Ulsteinvik | VU13, 26, 2 | VU18, 24 |
| Sunnfjord Hotell, Førde | VU5 | VU21, 11, 22, 27 |
| Hotel Prinsen, Trondheim | VU11, 19, 26, 10, 20 | VU22, 25 |
| Hotel Residence, Sandnes | VU1, 19, 27 | VU22 |
| Highland Hotel, Geilo | VU1, 27, 2, 24, 12 | VU17, 7, 16 |
| Grand Hotel, Kristiansand | VU7, 14, 12 | VU5, 9, 15, 6, 10 |

Tabell E.2

ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|---------------------------|-----------|------------------------|----------|-----------------------|
| Within+residual | 244,40 | 188 | 1,30 | | |
| Konstant | 3,35 | 1 | 3,35 | 2,58 | 0,110 |
| Samtlige 27 vurderinger | 204,99 | 1 | 204,99 | 157,69 | < 0,001 |
| Instruksjon og vurderingsskala | 1,66 | 1 | 1,66 | 1,27 | 0,260 ^a |
| Samtlige 27 vurderinger by instruksjon og vurderingsskala | 1,26 | 1 | 1,26 | 0,97 | 0,326 ^b |
| (Model) | 220,76 | 3 | 73,59 | 56,61 | < 0,001 |
| (Total) | 465,17 | 191 | 2,44 | | |
| R-squared = | 0,475 | | | | |
| Adjusted R-squared = | 0,466 | | | | |

^a p = 0,130 for en-hale test

^b p = 0,163 for en-hale test

Tabell E.3

ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|-------|--------------------|
| Within+residual | 290,04 | 147 | 1,97 | | |
| Konstant | 2,50 | 1 | 2,50 | 1,27 | 0,262 |
| Samtlige 27 vurderinger | 153,99 | 1 | 153,99 | 78,05 | < 0,001 |
| Instruksjon og vurderingsskala | 6,13 | 1 | 6,13 | 3,10 | 0,080 ^a |
| Samtlige 27 vurderinger x instruksjon og vurderingsskala | 2,16 | 1 | 2,16 | 1,09 | 0,297 ^b |
| (Model) | 165,33 | 3 | 55,11 | 27,93 | < 0,001 |
| (Total) | 455,37 | 150 | 3,04 | | |
| R-squared = | 0,363 | | | | |
| Adjusted R-squared = | 0,350 | | | | |

^a p = 0,004 for en-hale test

^b p = 0,149 for en-hale test

Tabell E.4

ANOVA, OBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 533,00 | 244 | 2,18 | | |
| Konstant | 3,76 | 1 | 3,76 | 1,72 | 0,191 |
| Samtlige 27 vurderinger | 291,39 | 1 | 291,39 | 133,40 | < 0,001 |
| Instruksjon og vurderingsskala | 20,09 | 1 | 20,09 | 9,20 | 0,003 ^a |
| Samtlige 27 vurderinger x instruksjon og vurderingsskala | 16,41 | 1 | 16,41 | 7,51 | 0,007 ^b |
| (Model) | 292,97 | 3 | 97,66 | 44,71 | < 0,001 |
| (Total) | 825,96 | 247 | 3,34 | | |
| R-squared = | 0,355 | | | | |
| Adjusted R-squared = | 0,347 | | | | |

^a p = 0,001 for en-hale test

^b p = 0,003 for en-hale test

Tabell E.5

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 191,16 | 184 | 1,04 | | |
| Konstant | 37,17 | 1 | 37,17 | 35,77 | < 0,001 |
| Samtlige 27 vurderinger | 130,00 | 1 | 130,00 | 125,13 | < 0,001 |
| Instruksjon og vurderingsskala | 6,78 | 1 | 6,78 | 6,53 | 0,011 ^a |
| Samtlige 27 vurderinger by instruksjon og vurderingsskala | 4,90 | 1 | 4,90 | 4,72 | 0,031 ^b |
| (Model) | 153,31 | 3 | 51,10 | 49,19 | < 0,001 |
| (Total) | 344,47 | 187 | 1,84 | | |
| R-squared = | 0,445 | | | | |
| Adjusted R-squared = | 0,436 | | | | |

^a p = 0,005 for en-hale test

^b p = 0,015 for en-hale test

Tabell E.6

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|-------|--------------------|
| Within+residual | 188,38 | 147 | 1,28 | | |
| Konstant | 33,71 | 1 | 33,71 | 26,30 | < 0,001 |
| Samtlige 27 vurderinger | 102,74 | 1 | 102,74 | 80,17 | < 0,001 |
| Instruksjon og vurderingsskala | 2,23 | 1 | 2,23 | 1,74 | 0,190 ^a |
| Samtlige 27 vurderinger x instruksjon og vurderingsskala | 1,16 | 1 | 1,16 | 0,91 | 0,343 ^b |
| (Model) | 106,06 | 3 | 35,35 | 27,59 | < 0,001 |
| (Total) | 294,44 | 150 | 1,96 | | |
| R-squared = | 0,360 | | | | |
| Adjusted R-squared = | 0,347 | | | | |

^a p = 0,095 for en-hale test

^b p = 0,172 for en-hale test

Tabell E.7

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 318,97 | 243 | 1,31 | | |
| Konstant | 33,59 | 1 | 33,59 | 25,59 | < 0,001 |
| Samtlige 27 vurderinger | 214,47 | 1 | 214,47 | 163,39 | < 0,001 |
| Instruksjon og vurderingsskala | 8,19 | 1 | 8,19 | 6,24 | 0,013 ^a |
| Samtlige 27 vurderinger x instruksjon og vurderingsskala | 7,32 | 1 | 7,32 | 5,58 | 0,019 ^b |
| (Model) | 214,76 | 3 | 71,59 | 54,54 | < 0,001 |
| (Total) | 533,72 | 246 | 2,17 | | |
| R-squared = | 0,402 | | | | |
| Adjusted R-squared = | 0,395 | | | | |

^a p = 0,006 for en-hale test

^b p = 0,008 for en-hale test

Tabell E.8

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 505,11 | 188 | 2,70 | | |
| Konstant | 778,71 | 1 | 778,71 | 288,12 | < 0,001 |
| Samtlige 27 vurderinger | 257,35 | 1 | 257,35 | 95,22 | < 0,001 |
| Instruksjon og vurderingsskala | 18,41 | 1 | 18,41 | 6,81 | 0,010 ^a |
| Samtlige 27 vurderinger by instruksjon og vurderingsskala | 16,04 | 1 | 16,04 | 5,93 | 0,016 ^b |
| (Model) | 311,26 | 3 | 103,75 | 38,39 | < 0,001 |
| (Total) | 819,37 | 191 | 4,29 | | |
| R-squared = | 0,380 | | | | |
| Adjusted R-squared = | 0,370 | | | | |

^a p = 0,005 for 1-hale test

^b p = 0,008 for 1-hale test

Tabell E.9

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 626,69 | 147 | 4,26 | | |
| Konstant | 489,81 | 1 | 489,81 | 114,89 | < 0,001 |
| Samtlige 27 vurderinger | 237,23 | 1 | 237,23 | 55,65 | < 0,001 |
| Instruksjon og vurderingsskala | 1,83 | 1 | 1,83 | 0,43 | 0,513 ^a |
| Samtlige 27 vurderinger x instruksjon og vurderingsskala | 1,65 | 1 | 1,65 | 0,39 | 0,535 ^b |
| (Model) | 240,98 | 3 | 80,33 | 18,84 | < 0,001 |
| (Total) | 867,67 | 150 | 5,78 | | |
| R-squared = | 0,278 | | | | |
| Adjusted R-squared = | 0,263 | | | | |

^a p = 0,257 for 1-hale test

^b p = 0,268 for 1-hale test

Tabell E.10

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--|----------------|-----|-------------|--------|--------------------|
| Within+residual | 938,60 | 246 | 3,82 | | |
| Konstant | 1143,92 | 1 | 1143,92 | 299,81 | < 0,001 |
| Samtlige 27 vurderinger | 333,55 | 1 | 333,55 | 87,42 | < 0,001 |
| Instruksjon og vurderingsskala | 41,61 | 1 | 41,61 | 10,91 | 0,001 ^a |
| Samtlige 27 vurderinger x instruksjon og vurderingsskala | 39,01 | 1 | 39,01 | 30,11 | 0,002 ^b |
| (Model) | 344,65 | 3 | 114,88 | 30,11 | < 0,001 |
| (Total) | 1283,26 | 249 | 5,15 | | |
| R-squared = | 0,269 | | | | |
| Adjusted R-squared = | 0,260 | | | | |

^a p < 0,001 for 1-hale test

^b p = 0,001 for 1-hale test

Tabell E.11

ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegnaker foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 213,50 | 185 | 1,15 | | |
| Konstant | ,28 | 1 | 0,28 | 0,25 | 0,621 |
| Immaterielle vurderinger | 13,65 | 1 | 13,65 | 11,83 | 0,001 |
| Materielle vurderinger | 75,29 | 1 | 75,29 | 65,24 | < 0,001 |
| Instruksjonog vurderingsskala | 1,51 | 1 | 1,51 | 1,31 | 0,255 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 1,43 | 1 | 1,43 | 1,24 | 0,267 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | ,15 | 1 | 0,15 | 0,13 | 0,723 ^c |
| (Model) | 251,04 | 5 | 50,21 | 43,50 | < 0,001 |
| (Total) | 464,54 | 190 | 2,44 | | |
| R-squared = | 0,540 | | | | |
| Adjusted R-squared = | 0,528 | | | | |

^a p = 0,128 for en-hale test

^b p = 0,134 for en-hale test

^c p = 0,362 for en-hale test

Tabell E.12

ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegnaker foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 270,89 | 145 | 1,87 | | |
| Konstant | 1,86 | 1 | 1,86 | 1,00 | 0,320 |
| Immaterielle vurderinger | ,46 | 1 | ,46 | ,25 | 0,620 |
| Materielle vurderinger | 80,54 | 1 | 80,54 | 43,11 | < 0,001 |
| Instruksjonog vurderingsskala | 6,62 | 1 | 6,62 | 3,54 | 0,062 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 0,00 | 1 | 0,00 | 0,00 | 0,970 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | 2,61 | 1 | 2,61 | 1,39 | 0,240 ^c |
| (Model) | 184,48 | 5 | 36,90 | 19,75 | < 0,001 |
| (Total) | 455,37 | 150 | 3,04 | | |
| R-squared = | 0,405 | | | | |
| Adjusted R-squared = | 0,385 | | | | |

^a p = 0,031 for en-hale test

^b p = 0,485 for en-hale test

^c p = 0,120 for en-hale test

Tabell E.13

ANOVA, OBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 464,05 | 242 | 1,92 | | |
| Konstant | 0,00 | 1 | 0,00 | 0,00 | 0,997 |
| Immaterielle vurderinger | 23,86 | 1 | 23,86 | 12,44 | 0,001 |
| Materielle vurderinger | 96,99 | 1 | 96,99 | 50,58 | < 0,001 |
| Instruksjonog vurderingsskala | 22,67 | 1 | 22,67 | 11,82 | 0,001 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 1,63 | 1 | 1,63 | 0,85 | 0,358 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | 4,85 | 1 | 4,85 | 2,53 | 0,113 ^c |
| (Model) | 361,92 | 5 | 72,38 | 37,75 | < 0,001 |
| (Total) | 825,96 | 247 | 3,34 | | |
| R-squared = | 0,438 | | | | |
| Adjusted R-squared = | 0,427 | | | | |

^a p < 0,001 for en-hale test

^b p = 0,179 for en-hale test

^c p = 0,057 for en-hale test

Tabell E.14

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 166,21 | 182 | 0,91 | | |
| Konstant | 19,83 | 1 | 19,83 | 21,71 | < 0,001 |
| Immaterielle vurderinger | 19,25 | 1 | 19,25 | 21,08 | < 0,001 |
| Materielle vurderinger | 32,63 | 1 | 32,63 | 35,73 | < 0,001 |
| Instruksjonog vurderingsskala | 7,01 | 1 | 7,01 | 7,67 | 0,006 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 5,86 | 1 | 5,86 | 6,42 | 0,012 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | 0,42 | 1 | 0,42 | 0,46 | 0,497 ^c |
| (Model) | 178,26 | 5 | 35,65 | 39,04 | < 0,001 |
| (Total) | 344,47 | 187 | 1,84 | | |
| R-squared = | 0,518 | | | | |
| Adjusted R-squared = | 0,504 | | | | |

^a p = 0,003 for en-hale test

^b p = 0,006 for en-hale test

^c p = 0,249 for en-hale test

Tabell E.15

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 162,67 | 145 | 1,12 | | |
| Konstant | 21,64 | 1 | 21,64 | 19,29 | < 0,001 |
| Immaterielle vurderinger | 0,03 | 1 | 0,03 | 0,02 | 0,877 |
| Materielle vurderinger | 63,29 | 1 | 63,29 | 56,41 | < 0,001 |
| Instruksjonog vurderingsskala | 2,51 | 1 | 2,51 | 3,85 | 0,137 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 0,42 | 1 | 0,42 | 0,38 | 0,540 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | 4,32 | 1 | 4,32 | 3,85 | 0,052 ^c |
| (Model) | 131,77 | 5 | 26,35 | 23,49 | < 0,001 |
| (Total) | 294,44 | 150 | 1,96 | | |
| R-squared = | 0,448 | | | | |
| Adjusted R-squared = | 0,428 | | | | |

^a p = 0,069 for en-hale test

^b p = 0,270 for en-hale test

^c p = 0,026 for en-hale test

Tabell E.16

ANOVA, SUBJEKTFORANKRET totalvurdering forklart av vurderinger av materielle og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 280,60 | 241 | 1,16 | | |
| Konstant | 24,16 | 1 | 24,16 | 20,75 | < 0,001 |
| Immaterielle vurderinger | 4,70 | 1 | 4,70 | 4,04 | 0,046 |
| Materielle vurderinger | 95,24 | 1 | 95,24 | 81,80 | < 0,001 |
| Instruksjonog vurderingsskala | 7,16 | 1 | 7,16 | 6,15 | 0,014 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 0,06 | 1 | 0,06 | 0,06 | 0,815 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | 3,33 | 1 | 3,33 | 2,86 | 0,092 ^c |
| (Model) | 253,13 | 5 | 50,63 | 43,48 | < 0,001 |
| (Total) | 533,72 | 246 | 2,17 | | |
| R-squared = | 0,474 | | | | |
| Adjusted R-squared = | 0,463 | | | | |

^a p < 0,007 for en-hale test

^b p = 0,408 for en-hale test

^c p = 0,046 for en-hale test

Tabell E.17

ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--------------------------------|----------------|-----|-------------|--------|--------------------|
| Within+residual | 473,59 | 185 | 2,56 | | |
| Konstant | 602,16 | 1 | 602,16 | 235,23 | < 0,001 |
| Immaterielle vurderinger | 35,11 | 1 | 35,11 | 13,72 | < 0,001 |
| Materielle vurderinger | 66,16 | 1 | 66,16 | 25,84 | < 0,001 |
| Instruksjonog vurderingsskala | 12,61 | 1 | 12,61 | 4,93 | 0,028 ^a |
| Immaterielle vurderinger x | | | | | |
| instruksjon og vurderingsskala | 9,61 | 1 | 9,61 | 3,75 | 0,054 ^b |
| Materielle vurderinger x | | | | | |
| instruksjon og vurderingsskala | 0,29 | 1 | 0,29 | 0,11 | 0,736 ^c |
| (Model) | 344,73 | 5 | 68,95 | 26,93 | < 0,001 |
| (Total) | 818,31 | 190 | 4,31 | | |
| R-squared = | 0,421 | | | | |
| Adjusted R-squared = | 0,406 | | | | |

^a p = 0,014 for en-hale test

^b p = 0,027 for en-hale test

^c p = 0,368 for en-hale test

Tabell E.18

ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, MIDDELS PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|--------------------------------|----------------|-----|-------------|-------|--------------------|
| Within+residual | 629,12 | 145 | 4,34 | | |
| Konstant | 290,05 | 1 | 290,05 | 66,85 | < 0,001 |
| Immaterielle vurderinger | 20,84 | 1 | 20,84 | 4,80 | 0,030 |
| Materielle vurderinger | 50,53 | 1 | 50,53 | 11,65 | 0,001 |
| Instruksjonog vurderingsskala | 2,13 | 1 | 2,13 | 0,49 | 0,484 ^a |
| Immaterielle vurderinger x | | | | | |
| instruksjon og vurderingsskala | 1,37 | 1 | 1,37 | 0,31 | 0,576 ^b |
| Materielle vurderinger x | | | | | |
| instruksjon og vurderingsskala | 0,08 | 1 | 0,08 | 0,02 | 0,893 ^c |
| (Model) | 238,55 | 5 | 47,71 | 11,00 | < 0,001 |
| (Total) | 867,67 | 150 | 5,78 | | |
| R-squared = | 0,275 | | | | |
| Adjusted R-squared = | 0,250 | | | | |

^a p = 0,242 for en-hale test

^b p = 0,288 for en-hale test

^c p = 0,447 for en-hale test

Tabell E.19

ANOVA, ATFERDSINTENSJONER forklart av vurderinger av materielle- og immaterielle produkttegenskaper foretatt i spørreskjema med objektforankret- og subjektforankret instruksjon og -vurderingsskala, HØY PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 886,73 | 244 | 3,63 | | |
| Konstant | 971,38 | 1 | 971,38 | 267,29 | < 0,001 |
| Immaterielle vurderinger | 20,16 | 1 | 20,16 | 5,55 | 0,019 |
| Materielle vurderinger | 110,39 | 1 | 110,39 | 30,37 | < 0,001 |
| Instruksjon og vurderingsskala | 44,02 | 1 | 44,02 | 12,11 | 0,001 ^a |
| Immaterielle vurderinger x instruksjon og vurderingsskala | 2,11 | 1 | 2,11 | 0,58 | 0,447 ^b |
| Materielle vurderinger x instruksjon og vurderingsskala | 14,20 | 1 | 14,20 | 3,91 | 0,049 ^c |
| (Model) | 369,52 | 5 | 79,30 | 21,82 | < 0,001 |
| (Total) | 1283,26 | 249 | 5,15 | | |
| R-squared = | 0,309 | | | | |
| Adjusted R-squared = | 0,295 | | | | |

^a p < 0,001 for en-hale test

^b p = 0,224 for en-hale test

^c p = 0,025 for en-hale test

Tabell E.20

ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål for spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, LAV PRODUKTKUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 206,23 | 186 | 1,11 | | |
| Konstant | 22,94 | 1 | 22,94 | 20,69 | > 0,001 |
| Samtlige 27 vurderinger | 142,58 | 1 | 142,58 | 128,59 | > 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 10,37 | 2 | 5,18 | 4,68 | 0,010 ^a |
| Samtlige 27 vurderinger x Tilstedeværelse og plassering av viktighetsspørsmål | 7,01 | 2 | 3,50 | 3,16 | 0,045 ^b |
| (Model) | 177,29 | 5 | 35,46 | 31,98 | > 0,001 |
| (Total) | 383,52 | 191 | 2,01 | | |
| R-squared = | 0,462 | | | | |
| Adjusted R-squared = | 0,448 | | | | |

^a p = 0,005 for en-hale test

^b p = 0,022 for en-hale test

Tabell E.21

ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, **MIDDELS PRODUKTUNNSKAP**

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|-------|--------------------|
| Within+residual | 225,71 | 146 | 1,55 | | |
| Konstant | 4,93 | 1 | 4,93 | 3,19 | 0,076 |
| Samtlige 27 vurderinger | 143,85 | 1 | 143,85 | 93,05 | > 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 1,61 | 2 | 0,80 | 0,52 | 0,596 ^a |
| Samtlige 27 vurderinger x Tilstedeværelse og plassering av viktighetsspørsmål | 1,74 | 2 | 0,87 | 0,56 | 0,570 ^b |
| (Model) | 146,68 | 5 | 29,34 | 18,98 | > 0,001 |
| (Total) | 372,39 | 151 | 2,47 | | |
| R-squared = | 0,394 | | | | |
| Adjusted R-squared = | 0,373 | | | | |

^a p = 0,298 for en-hale test

^b p = 0,285 for en-hale test

Tabell E.22

ANOVA, TOTALVURDERING forklart av spesifikke vurderinger foretatt i spørreskjema spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, **HØY PRODUKTUNNSKAP**

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 392,52 | 244 | 1,61 | | |
| Konstant | 14,30 | 1 | 14,30 | 8,89 | 0,003 |
| Samtlige 27 vurderinger | 255,86 | 1 | 255,86 | 159,05 | > 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 16,80 | 2 | 8,40 | 5,22 | 0,006 ^a |
| Samtlige 27 vurderinger x Tilstedeværelse og plassering av viktighetsspørsmål | 11,10 | 2 | 5,55 | 3,45 | 0,033 ^b |
| (Model) | 267,64 | 5 | 53,53 | 33,27 | > 0,001 |
| (Total) | 660,16 | 249 | 2,65 | | |
| R-squared = | 0,405 | | | | |
| Adjusted R-squared = | 0,393 | | | | |

^a p = 0,003 for en-hale test

^b p = 0,016 for en-hale test

Tabell E.23

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, LAV PRODUKT-KUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 564,11 | 186 | 3,03 | | |
| Konstant | 838,65 | 1 | 838,65 | 276,65 | > 0,001 |
| Samtlige 27 vurderinger | 231,69 | 1 | 231,69 | 76,39 | > 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 2,75 | 2 | 1,37 | 0,45 | 0,636 ^a |
| Samtlige 27 vurderinger x Tilstedeværelse og plassering av viktighetsspørsmål | 0,89 | 2 | 0,45 | 0,15 | 0,864 ^b |
| (Model) | 255,26 | 5 | 51,05 | 16,83 | > 0,001 |
| (Total) | 819,37 | 191 | 4,29 | | |
| R-squared = | 0,312 | | | | |
| Adjusted R-squared = | 0,293 | | | | |

^a p = 0,438 for en-hale test

^b p = 0,424 for en-hale test

Tabell E.24

ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, MIDDELS PRODUKT-KUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 640,23 | 145 | 4,42 | | |
| Konstant | 454,31 | 1 | 454,31 | 102,89 | > 0,001 |
| Samtlige 27 vurderinger | 217,33 | 1 | 217,33 | 49,22 | > 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 2,11 | 2 | 1,06 | 0,24 | 0,788 ^a |
| Samtlige 27 vurderinger x Tilstedeværelse og plassering av viktighetsspørsmål | 2,64 | 2 | 1,32 | 0,30 | 0,742 ^b |
| (Model) | 227,44 | 5 | 45,49 | 10,30 | > 0,001 |
| (Total) | 867,67 | 150 | 5,78 | | |
| R-squared = | 0,262 | | | | |
| Adjusted R-squared = | 0,237 | | | | |

^a p = 0,394 for en-hale test

^b p = 0,371 for en-hale test

Tabell E.25

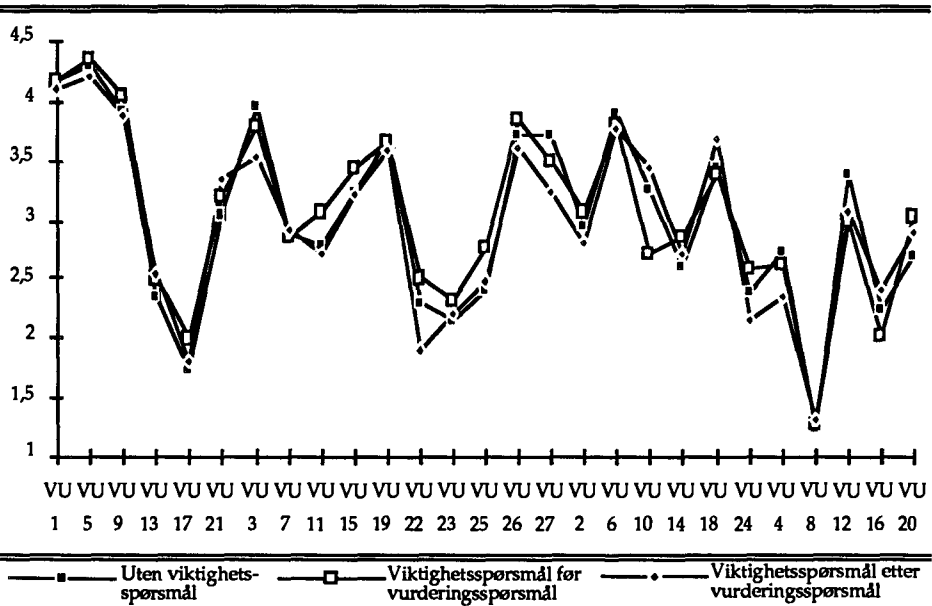
ANOVA, ATFERDSINTENSJONER forklart av spesifikke vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, HØY PRODUKT-KUNNSKAP.

| SOURCE OF VARIATION | SUM OF SQUARES | DF | MEAN SQUARE | F | SIGN. OF F |
|---|----------------|-----|-------------|--------|--------------------|
| Within+residual | 952,22 | 244 | 3,90 | | |
| Konstant | 1207,72 | 1 | 1207,72 | 309,47 | > 0,001 |
| Samtlige 27 vurderinger | 317,15 | 1 | 317,15 | 81,27 | > 0,001 |
| Tilstedeværelse og plassering av viktighetsspørsmål | 20,99 | 2 | 10,50 | 2,69 | 0,070 ^a |
| Samtlige 27 vurderinger x Tilstedeværelse og plassering av viktighetsspørsmål | 15,40 | 2 | 7,70 | 1,97 | 0,141 ^b |
| (Model) | 331,04 | 5 | 66,21 | 16,97 | > 0,001 |
| (Total) | 1283,26 | 249 | 5,15 | | |
| R-squared = | 0,258 | | | | |
| Adjusted R-squared = | 0,243 | | | | |

^a p = 0,035 for en-hale test
^b p = 0,070 for en-hale test

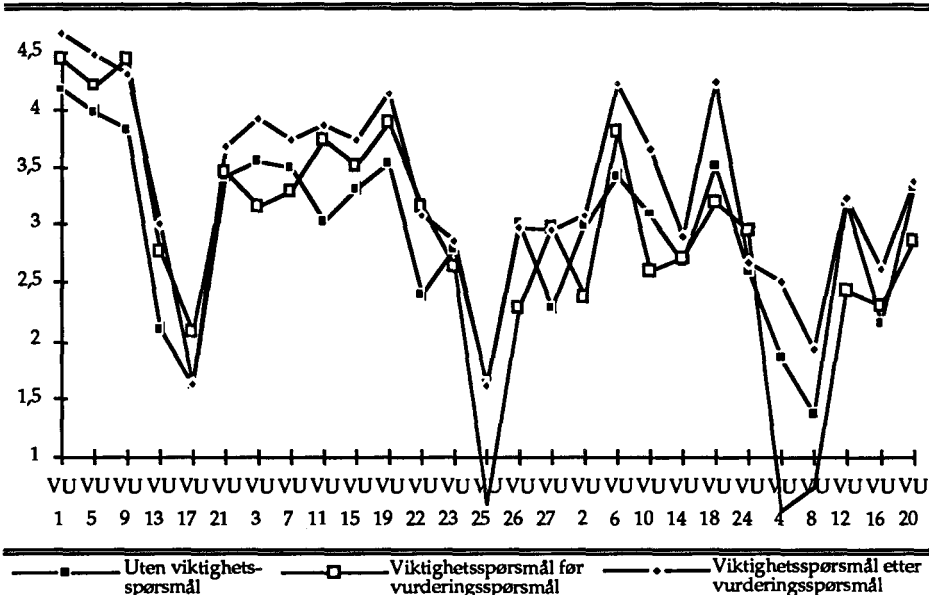
Figur E.1a

Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, lavkvalitetshotell



Figur E.1b

Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, gjennomsnittshotell



Figur E.1c

Nivå på vurderinger foretatt i spørreskjema uten viktighetsspørsmål, med viktighetsspørsmål før spesifikke vurderingsspørsmål og med viktighetsspørsmål etter spesifikke vurderingsspørsmål, høykvalitetshotell

