



Visualisering av Big Data i revisjon

En eksperimentell studie

Johnny Vu Tran & Lars Michelsen

Veileder: Jonas Gaudernack

Masteroppgave i regnskap og revisjon

&

Masteroppgave i økonomi og administrasjon

Hovedprofil: Business Analytics

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i regnskap- og revisjon, og økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer inntår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

Sammendrag

Revisjonsbransjen har et økende fokus på innovative metoder for å analysere data. Det er konsensus om at bruk av stadig større og mer komplekse datamaterialer vil være sentralt for morgendagens digitale revisjon. Denne studien bidrar til litteraturen ved å undersøke hvordan revisor påvirkes av moderne visualiseringsteknikker. Vi stiller følgende spørsmål: «*Blir revisors profesjonelle skjønn påvirket av informasjonsrekkefølge og presentasjonsformat, ved bruk av Big Data i risikovurderingsprosessen?*»

For å svare på dette, utførte vi et eksperiment med 64 deltakere bestående av praktiserende revisorer og revisorstudenter. De ble delt inn i fire grupper og bedt om å vurdere iboende risiko. Vurderingene baserte seg på et informasjonssett bestående av fire ulike nettsider med datavisualiseringer. Deltakerne vurderte hver nettside individuelt, i tillegg til å gjøre en avsluttende totalvurdering. To nettsider kommuniserte positiv informasjon (lavere risiko), og to nettsider kommuniserte negativ informasjon (høyere risiko). Halvparten av respondentene fikk informasjonen i rekkefølgen positiv-negativ (PPNN) og den andre halvparten fikk den i motsatt rekkefølge (NNPP). I tillegg ble det laget to versjoner av alle nettsidene. Halvparten fikk se et statisk format, og den andre halvparten fikk den samme informasjonen med interaktive funksjoner.

Vi finner en statistisk signifikant formateffekt (5 % nivå) for nettsidene med positiv informasjon. En toveis variansanalyse viser at gruppene som fikk se et statisk format, vurderte risiko høyere enn de som fikk informasjonen med interaktive funksjoner. Det var lignende signaler for nettsidene med negativ informasjon og for den avsluttende totalvurderingen, men disse forskjellene var ikke signifikante. Videre var det ingen tegn til at informasjonsrekkefølge påvirket deltakerne i noen retning. Det var heller ingen interaksjonseffekt, og vi konkluderer på følgende måte: Resultatene fra studien er interessante, men ikke sterke nok til å kunne argumentere for at revisor påvirkes av informasjonsrekkefølge eller presentasjonsformat, ved vurdering av iboende risiko.

Nøkkelord – Big Data, Datavisualisering, Risikovurdering, ISA 315, Iboende risiko

Forord

Denne masteroppgaven ble skrevet våren 2022 ved Norges Handelshøyskole; for Johnny som en del av masterstudiet i regnskap og revisjon, og for Lars som en del av masterstudiet i økonomi og administrasjon. Vår felles interesse for den teknologiske utviklingen i revisjonsbransjen, ga oss inspirasjon til å skrive en oppgave hvor vi krysser Business Analytics med regnskap og revisjon.

Det har vært veldig lærerikt å benytte et eksperimentelt forskningsdesign. Veien fra idéutvikling til drøftelse av resultater, har vært både omfattende og givende. Gjennom prosessen har vi fått bedre forståelse for revisjon, og videreutviklet våre tekniske og analytiske ferdigheter. Det har vært særlig spennende å utvikle webapplikasjoner med moderne visualiseringsteknikker i en revisjonskontekst. Vi håper dette vil være nyttige erfaringer som vi kan ta med oss videre.

Først og fremst ønsker vi å takke vår veileder Jonas Gaudernack for en ryddig prosess med gode diskusjoner. En stor takk går også til statistiker John Pavia for gode innspill rundt det tekniske, og lektor Ole Jakob Michelsen for korrekturlesing. Til slutt ønsker vi å takke pilottesterne som hjalp oss med prestudien, og alle deltakerne som tok seg tid til å gjennomføre undersøkelsen.

Norges Handelshøyskole

Bergen, mai 2022

Johnny Vu Tran

Lars Michelsen

Innhold

Sammendrag	i
Forord	ii
Figurer og tabeller	v
Forkortelser	vi
1 Innledning	1
2 Teori	2
2.1 Big Data	2
2.1.1 Hva er Big Data?	2
2.1.2 Big Data i regnskap og revisjon	3
2.1.3 Visualisering av Big Data	4
2.2 Beslutningsvariasjon	5
2.2.1 Linsemodellen	5
2.2.2 Skjønnsmessige vurderinger i revisjon	7
2.2.3 Risikovurdering i revisjon	7
2.3 Informasjonsrekkefølge	9
2.4 Presentasjonsformat	10
3 Forskningsmodell og hypoteser	13
3.1 Validitetsramme for forskningsmodellen	13
3.2 Avhengig variabel	14
3.3 Uavhengige variabler	14
3.3.1 Informasjonsrekkefølge	15
3.3.2 Presentasjonsformat	15
3.4 Kontrollvariabler	15
3.5 Hypoteseutvikling	16
3.5.1 Informasjonsrekkefølge og revisors profesjonelle skjønn	16
3.5.2 Presentasjonsformat og revisors profesjonelle skjønn	17
3.5.3 Interaksjonseffekt mellom format og rekkefølge	18
4 Metode	19
4.1 Forskningsdesign	19
4.2 Design av undersøkelsen	20
4.2.1 Del 1: Innledning og kontekst	21
4.2.2 Del 2: Webapplikasjoner	21
4.2.3 Del 3: Totalvurdering og kontrollspørsmål	25
4.3 Prestudie	25
4.4 Innsamling av data	26
4.4.1 Utvalg	26
4.4.2 Gjennomføring	27
4.4.3 Forskningsetikk	27
4.5 Analyse av data	28
4.5.1 Konstruksjon av variabler	28
4.5.2 Variansanalyser	29
4.6 Validitet og reliabilitet	30

4.6.1	Validitet	30
4.6.2	Reliabilitet	32
5	Resultater	33
5.1	Datagrunnlag	33
5.1.1	Demografi	33
5.2	Deskriptiv statistikk	35
5.2.1	Utvalget under ett	35
5.2.2	Fordelt per gruppe	36
5.3	Hypotesetesting	38
5.3.1	Hovedhypoteser	38
5.3.2	Tilleggshypoteser	39
6	Diskusjon og konklusjon	41
6.1	Informasjonsrekkefølge og profesjonelt skjønn	41
6.2	Presentasjonsformat og profesjonelt skjønn	43
6.3	Interaksjon mellom rekkefølge og format	44
6.4	Andre funn	44
6.4.1	Formateffekt for spesifikke signaltyper	44
6.4.2	Alternativ tolkning av avhengig variabel	45
6.5	Betydning for revisjonspraksis	45
6.6	Begrensninger	46
6.7	Forslag til videre forskning	47
	Referanser	48
	Appendiks	54
A1	Spørreundersøkelse	54
A2	Webapplikasjoner	63
A3	Forutsetninger for variansanalyser	70
A4	Alternativ tolkning av avhengig variabel	71

Figurliste

2.1	Brunswiks linsemodell	6
2.2	Keltons beslutningsrammeverk	11
3.1	Studiens validitetsrammeverk	13
4.1	Studiens 2×2 faktorielle design	20
5.1	Demografisk fordeling	34
5.2	Fordelinger av risikovurderinger for hele utvalget	35
5.3	Fordelinger av risikovurderinger per gruppe	37
5.4	Profildiagram for hver gruppe	37
A1.1	Spørreundersøkelse - Innledning 1	54
A1.2	Spørreundersøkelse - Innledning 2	55
A1.3	Spørreundersøkelse - Innledning 3	56
A1.4	Spørreundersøkelse - Vurdering av tidsserier	57
A1.5	Spørreundersøkelse - Vurdering av ordsky	58
A1.6	Spørreundersøkelse - Vurdering av flytskjema	59
A1.7	Spørreundersøkelse - Vurdering av kart	60
A1.8	Spørreundersøkelse - Samlet vurdering og kontrollspørsmål 1	61
A1.9	Spørreundersøkelse - Samlet vurdering og kontrollspørsmål 2	62
A2.1	Webapplikasjon - Tidsserier (statisk versjon)	63
A2.2	Webapplikasjon - Tidsserier (interaktiv versjon)	64
A2.3	Webapplikasjon - Ordsky (statisk versjon)	65
A2.4	Webapplikasjon - Ordsky (interaktiv versjon)	65
A2.5	Webapplikasjon - Flytskjema (statisk versjon)	66
A2.6	Webapplikasjon - Flytskjema (interaktiv versjon)	67
A2.7	Webapplikasjon - Kart (statisk versjon)	68
A2.8	Webapplikasjon - Kart (interaktiv versjon)	69

Tabelliste

5.1	Respondenter i utvalget	33
5.2	Deskriptiv statistikk	36
5.3	Anova-tabell for den samlede risikovurderingen	38
5.4	Anova-tabell for risikovurdering av positive cues	39
5.5	Anova-tabell for risikovurdering av negative cues	40
A3.1	Shapiro-Wilks test	70
A3.2	Levenes test	70
A4.1	Anova-tabell for samlet vurdering (alternativ)	71
A4.2	Anova-tabell for positive cues (alternativ)	72
A4.3	Anova-tabell for negative cues (alternativ)	72

Forkortelser

ADA	–	Audit Data Analytics
AICPA	–	American Institute of Certified Public Accountants
ANCOVA	–	Analysis of Covariance
ANOVA	–	Analysis of Variance
IAASB	–	International Audit and Assurance Standards Board
ISA	–	International Standard on Auditing
MRR	–	Master i regnskap og revisjon
NHH	–	Norges Handelshøyskole
NSD	–	Norsk senter for forskningsdata
RVF	–	Risiko for vesentlig feilinformasjon
ØKAD	–	Økonomi og administrasjon

1 Innledning

Nærmest all aktivitet på digitale plattformer genererer informasjon i form av data. Det er flere som spår at morgendagens næringsliv vil domineres av de som har størst tilgang til data og de sterkeste algoritmene (Fry, 2018; Goodwin, 2020). Revisors primære oppgave er å kvalitetssikre finansielle rapporter, og en stor del av dette arbeidet består av å gjennomgå både finansiell og ikke-finansiell data. Med bakgrunn i dataen bruker revisor sitt profesjonelle skjønn til å trekke slutninger om rapporteringen er gjort i henhold til gjeldende rammeverk.

Revisjonsbransjen har et økende fokus på innovative metoder for å analysere data. *Audit Data Analytics*¹ (ADA) er et satsingsområde hos alle de store revisjonsselskapene i Norge, men det er foreløpig begrenset hvor mye det brukes i praksis (Eilifsen et al., 2020). Selv om det per i dag er begrensninger i den praktiske bruken, virker det å være konsensus om at analyser av stadig større og mer komplekse datamaterialer, vil være sentralt for morgendagens digitale revisjon. I fagmiljøet har det vært uttrykt bekymringer rundt revisors manglende forståelse av moderne dataanalyser, og villigheten til å tilegne seg nødvendig kunnskap (Alles, 2015; Perkhofer et al., 2019; IDEA, 2020). Som allmennhetens tillitsperson skal revisor øke tilliten til regnskapsrapporteringen, og manglende kjennskap til moderne analyseverktøy kan være en dempende faktor på den overordnede revisjonskvaliteten.

I den forbindelse ønsker vi å rette fokuset mot prosessen hvor enorme og ofte rotete datamaterialer leses inn, struktureres og visualiseres. Dette er en prosess som kan stille større krav til teknologisk kompetanse. Når slike visualiseringer innhentes og presenteres for revisor, er det nødvendig å være obs på eventuelle faktorer som kan påvirke revisors skjønnsmessige vurderinger. Vi ønsker derfor å svare på forskningsspørsmålet:

Blir revisors profesjonelle skjønn påvirket av informasjonsrekkefølge og presentasjonsformat, ved bruk av Big Data i risikovurderingsprosessen?

¹En forenklet oversettelse av definisjonen til AICPA (2018):

«Kunsten å analysere datamaterialer for å identifisere avvik og annen informasjon som kan være relevant for revisjonsarbeidet.»

2 Teori

I dette kapittelet vil vi gjøre rede for det teoretiske fundamentet bak vår studie. Vi presenterer Big Data som begrep og hvordan det brukes i revisjon. Videre blir det redegjort for teori rundt beslutningstaking, og hvordan informasjonsrekkefølge og presentasjonsformat kan påvirke revisor i en risikovurderingsprosess.

2.1 Big Data

Big Data som konsept har blitt spådd å forandre samfunnet, ettersom loggføring av nesten all menneskelig aktivitet på digitale plattformer åpner en ny verden av muligheter (Pedersen og Wilkinson, 2019). Det skrives om hvordan analyser av stadig større datamengder vil kunne avdekke informasjon på en ny måte. Samfunnets sterke fokus på kausalitet kan bli byttet ut med enkle korrelasjoner. *Hvorfor* vil kunne byttes ut med *hva*, noe som kan utfordre etablert praksis rundt beslutningstaking (Mayer-Schönberger og Cukier, 2013, s.7).

2.1.1 Hva er Big Data?

Det er utfordrende å skulle spore seg tilbake til opprinnelsen og betydningen av begrepet *Big Data*. Francis Diebold hevder at det sannsynligvis oppstod i tekniske miljøer på midten av 90-tallet, og han finner de første sporene av begrepet i akademisk litteratur mot slutten av 90-tallet (Diebold, 2012). Videre skulle det ta mer enn ti nye år før bruken av begrepet virkelig spredte seg fra 2010 og utover (Gandomi og Haider, 2015).

Douglas Laney et al. introduserte i en artikkel fra 2001 de tre konseptene *volum*, *hastighet* og *variasjon*² i tilknytning til data (Laney et al., 2001), og dette har blitt en av de mest siterte definisjonene på hva Big Data er, selv om artikkelen ikke nevner de faktiske ordene «Big Data» (Ward og Barker, 2013). Laney's tilnærming virker å ha festet seg, og mange tar utgangspunkt i det som ofte omtales som *The Three V's of Big Data* i sine definisjoner av begrepet³. Volum er relatert til størrelsen, hastighet til hvor raskt ny data genereres, og

²Oversatt fra de engelske begrepene *volume*, *velocity* og *variety*.

³En populær definisjon er fra Gartner (2022):

«Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation»

variasjon er knyttet til den strukturelle heterogeniteten (Russom et al., 2011; Gandomi og Haider, 2015; De Mauro et al., 2016). Begrepet kan defineres på forskjellige måter, og det benyttes ulikt på tvers av fagfelt, noe som bidrar til økt uklarhet (Ward og Barker, 2013).

I denne studien tar vi som utgangspunkt at all data som gjøres tilgjengelig for revisor, særlig i større revisjonsoppdrag, er av en slik art og størrelse at det går under paraplyen Big Data. Det kan være snakk om enorme mengder transaksjonsdata, oversikt over hvem som har logget seg på datasystemene, historikk fra nøkkelbrikker, kommunikasjon med kunder gjennom sosiale medier, interne og eksterne uttalelser etc. Når dette akkumuleres over tid vil revisor måtte ta stilling til nye og innovative metoder for å tilrettelegge en målrettet og effektiv revisjon.

2.1.2 Big Data i regnskap og revisjon

I senere år har det vært et økende fokus på bruk av Big Data i de store revisjonsselskapene (Insights og KPMG, 2020; PwC, 2013). I akademia er det også mye litteratur som fokuserer på bruken av Big Data i revisjon. Revisor møter på komplekse utfordringer ved å skulle samle, vurdere og gjøre mening ut av stadig større datamengder fra ulike kilder. Overbelastning av informasjon, informasjonsrelevans, mønstergjenkjenning og tvetydighet er områder som kan svekke kvaliteten på revisors vurderinger og beslutninger (Brown-Liburd et al., 2015). Samtidig argumenteres det for hvordan verktøy som ekspertsystemer⁴ og prediktive modeller kan bidra til økt treffsikkerhet og effektivitet (Brown-Liburd og Vasarhelyi, 2015).

Generelt åpner Big Data nye dører, men revisor må vise aktsomhet ved implementering av slike nye verktøy. Korrelasjoner og avvik som kommer til syne gjennom analyser av store datamaterialer, vil i seg selv ikke være tilstrekkelig som revisjonsbevis. Revisor må fortsatt gjøre tilstrekkelige revisjonshandlinger for å finne bakenforliggende årsaker til det som eventuelt finnes (Brown-Liburd et al., 2015). Med Internet of Things (IoT), GPS-sporing, nøkkelbrikker etc. vil det genereres enorme mengder data som revisor må ta stilling til. Tradisjonell revisjonsmetodikk for å innhente bevis vil ikke nødvendigvis være like tilstrekkelig som før (Brown-Liburd og Vasarhelyi, 2015). I tillegg ser man et skifte i regnskapsrapporteringen, hvor det i økende grad rapporteres om ikke-finansiell

⁴Fra Store norske leksikon (2022):

«Ekspertsystem er et dataprogram som benytter kunstig intelligens til å løse eller bidra til å løse problemer innenfor et spesialisert fagområde.»

informasjon (IMPACT, 2020). Økt kompleksitet og mangfold av data medfører et økt behov for moderne analyseverktøy for å gjøre mening ut av dataene (Brown-Liburd og Vasarhelyi, 2015; Vasarhelyi et al., 2015).

Selv med den eksponentielle veksten i Big Data og analyseverktøy, har revisjonsprofesjonen vært treg med å tilpasse seg ny teknologi (Alles, 2015). Alles stiller spørsmål ved om moderne teknologi vil integreres i revisjonsarbeidet som følge av økt etterspørsel fra klienter, og ikke nødvendigvis som et resultat av revisors aksept for nyere verktøy (Alles, 2015). Mangelen på kunnskap om moderne visualiseringsteknikker, i tillegg til det sterke fokuset på Microsoft Excel, er hindrende faktorer for bruken av ny og moderne visualiseringssoftware (Perkhofer et al., 2019). Ettersom bruken av Big Data -relaterte verktøy blir stadig mer utbredt hos revisjonsklientene, bør revisor også mestre disse verktøyene for å ikke risikere svekkelse av troverdighet og integritet (Alles, 2015).

Selv om disse verktøyene blir stadig mer integrert, og tilgangen og størrelsen på datamaterialene stadig vokser, vil målet med et regnskap være det samme – å rapportere beslutningsnyttig informasjon til interne og eksterne regnskapsbrukere (Janvrin og Watson, 2017). Revisjonens formål vil også være som før – å øke tilliten til regnskapsrapporteringen.

2.1.3 Visualisering av Big Data

Veksten i Big Data og stadig økende datakompleksitet har skapt et større behov for visualiseringsteknikker som kan komprimere større mengder data på en oversiktlig og hensiktsmessig måte (Alawadhi, 2015). En revisor vil i større grad være avhengig av avanserte visualiseringsteknikker, da tradisjonelle analysemetoder blir mindre effektive i en Big Data -verden. Strategiske datavisualiseringer kan bidra til å øke revisors forståelse av informasjon, samt forbedre beslutningstakingen og revisjonskvaliteten (Alawadhi, 2015).

På generelt grunnlag skilles det mellom to hovedtyper av datavisualisering; forklarende og utforskende. I forklarende visualiseringer er informasjonen i datamaterialet kjent og det visuelle blir brukt til å kommunisere et budskap. Utforskende visualiseringer brukes for å søke etter sammenhenger og interessant informasjon i datamaterialet (Iliinsky og Steele, 2011). I en revisjonskontekst kan utforskende visualiseringer være nyttig blant annet i planleggingsfasen av et oppdrag, hvor det potensielt er mye data som må gjennomgås for

å avklare videre revisjonshandlinger. Forklarende visualiseringer kan være nyttig i flere kontekster, blant annet kan det brukes som supplerende revisjonsbevis, for å kommunisere resultater til andre i revisjonsteamet eller til klient.

Bruken av datavisualiseringer medfører ikke alltid økt forståelse, og visualiseringen må utarbeides strategisk for å virke hensiktsmessig. For at datavisualiseringer skal virke effektivt, må den tilpasses formålet og brukerne av visualiseringen. Feil bruk av datavisualiseringer kan i noen tilfeller medføre flere ulemper enn fordeler. Tidligere forskning på datavisualiseringer har som regel vært rettet mot enkle visualiseringer som tabeller og grafer, og det er et behov for forskning på mer moderne teknikker. For revisjonsforskning vil det være særlig aktuelt å undersøke hvordan visualiseringer av Big Data kan effektiviseres for å forbedre revisors beslutningstaking ([Alawadhi, 2015](#)).

2.2 Beslutningsvariasjon

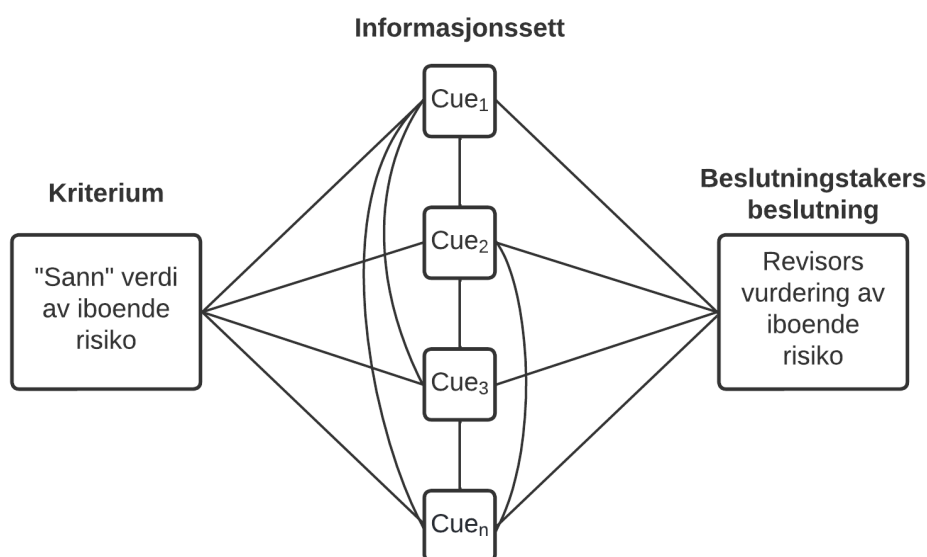
Denne studien vil falle inn under *judgement and decision making research*. [Bonner \(2008\)](#) definerer dette som:

«Research that focuses on something about judgments or decisions as either the dependent variable or independent variable.»

Dette kan deles inn i to hovedbolker: (1) variasjon i vurderingsprosesser og beslutninger, og (2) kvaliteten av vurderingsprosesser og beslutninger ([Bonner, 2008](#), s.2). Denne studien fokuserer på beslutningsvariasjonen med formål om å bidra med videre forståelse av hvordan rekkefølgeeffekt og formateffekt påvirker revisors beslutninger.

2.2.1 Linsemodellen

The lens model er blant de mer dominante modellene for hvordan revisors kognitive prosess kan studeres ([Bonner, 2008](#), s.116). Modellen ble presentert av [Frenkel-Brunswik \(1952\)](#) og har blitt brukt i både eldre og nyere studier som et teoretisk grunnlag ([Hursch et al., 1964](#); [Majid et al., 2001](#)). Vi vil framover i oppgaven referere til modellen som linsemodellen.



Figur 2.1: Forenklet tolkning av Brunswiks linsemodell.

Denne studien benytter en forenklet linsemodell utarbeidet av [Hammond et al. \(1964\)](#) som består av tre hovedkomponenter: (1) kriterium, (2) informasjonssett, og (3) beslutningstakers beslutning. Kriteriet defineres som den sanne beslutningen, altså det «riktige» svaret til en gitt oppgave ([Bonner, 2008](#), s.116). Informasjonssettet består av ulike relevante *cues*⁵ som beslutningstakeren blir presentert for før en beslutning gjøres. Det er viktig å merke seg at en cue har evnen til å påvirke senere cues som presenteres. Den siste komponenten er beslutningstakerens beslutning, som tas på grunnlag av presenterte cues.

Linsemodellen benyttes i denne studien som et teoretisk utgangspunkt for å forstå den kognitive prosessen til beslutningstakeren. Siden studien fokuserer på beslutningsvariasjon, vil kriteriet «sann» verdi ikke være et relevant element. Fokuset vil være på hvordan beslutningstakeren benytter et informasjonssett som utgangspunkt for sin beslutningstaking. Modellen brukes sammen med relevant teori om informasjonsrekkefølge og presentasjonsformat.

⁵En *cue* kan beskrives som et objekt med en substansiell og ikke-substansiell dimensjon ([Savolainen, 2009](#)). En cue skal trigge kognitive mekanismer og forekommer ofte som en del av noe større. Vi finner ingen god norsk oversettelse og benytter det engelske begrepet cue (entall) og cues (flertall) i denne masteroppgaven.

2.2.2 Skjønnsmessige vurderinger i revisjon

Revisjonsarbeid er i stor grad preget av skjønnsmessige vurderinger. Utøvelsen av profesjonelt skjønn nevnes som et sentralt attributt i IAASBs⁶ rammeverk for god revisjonskvalitet. I norsk revisjon forankres profesjonelt skjønn i revisorloven § 9-1 som omtaler revisor som allmennhetens tillitsperson. Revisor er pliktig til å utøve sitt yrke med integritet, objektivitet og aktsomhet ([Revisorloven, 2021](#)). ISA⁷ 200 definerer profesjonelt skjønn som:

«Anvendelsen av relevant opplæring, kunnskap og erfaring innenfor rammen av standarder for revisjon, regnskap og etikk når det fattes informerte beslutninger om hvilke handlingsplaner som er hensiktsmessige etter omstendighetene ved revisjonsoppdraget.».

Definisjonen innebærer at revisor anvender seg av kunnskap og kompetanse opparbeidet gjennom opplæring, studier og kurs, samt erfaring og ferdigheter fra tidligere arbeid. Profesjonelt skjønn nevnes i tillegg som en av de fem prinsippene for god revisjonsskikk ([IESBA, 2018](#)). Som følge av forankringen i lovgivningen samt IAASBs rammeverk er utøvelsen av profesjonelt skjønn et sentralt konsept. Dette vises også i den norske revisjonspraksisen gjennom søkelyset rettet mot det profesjonelle skjønnen. De fire store revisjonshusene i Norge arbeider aktivt for å opprettholde god skikk, aktsomhet og integritet i revisjonstjenestene deres ([Deloitte, 2021](#); [EY, 2021](#); [KPMG, 2021](#); [PwC, 2021](#)).

2.2.3 Risikovurdering i revisjon

Skjønnsmessige vurderinger gjøres i hele revisjonsprosessen ([IAASB, 2013](#)). Denne studien retter fokuset mot planleggingsfasen hvor revisor gjør seg kjent med klienten. I denne prosessen blir det gjort skjønnsmessige vurderinger av risikoen for vesentlig feilinformasjon. Dette er en risikovurdering som følger av revisorloven § 9-4 bokstav c.

⁶International Auditing and Assurance Standards Board

⁷International Standards on Auditing

ISA 200 definerer det overordnede målet for revisjon som:

«Formålet med revisjon er å øke de tiltenkte brukernes tillit til regnskapet. Dette oppnås ved at revisor gir uttrykk for en mening om hvorvidt regnskapet i det alt vesentlige er utarbeidet i samsvar med det gjeldende rammeverket for finansiell rapportering.»

Videre i standarden står det at revisors konklusjon på regnskapet skal bygges på en oppnådd betryggende sikkerhet for at regnskapet ikke inneholder vesentlig feilinformasjon, verken på grunn av misligheter eller feil. Betryggende sikkerhet forekommer når revisor har redusert revisjonsrisikoen til et tilfredsstillende lavt nivå. Revisjonsrisikoen jfr ISA 200 blir definert som:

«Risikoen for at revisor gir uttrykk for en uriktig mening i revisjonsberetningen når regnskapet inneholder vesentlig feilinformasjon. Revisjonsrisiko er en funksjon av risikoene for vesentlig feilinformasjon og oppdagelsesrisiko.»

Revisjonsrisikoen består av tre komponenter: Iboende risiko, kontrollrisiko og oppdagelsesrisiko (IAASB, 2009). Iboende- og kontrollrisiko er komponentene av interesse, da dette er risikofaktorer som revisor må identifisere og vurdere (Messier et al., 2006, s.81). ISA 200 punkt 13 bokstav n viser til risikoen for vesentlig feilinformasjon som et samlebegrep for disse to komponentene. Risikoen for vesentlig feilinformasjon (heretter RVF) defineres som:

«Risikoen for at regnskapet inneholder vesentlig feilinformasjon før det revideres.»

Vurderingen av RVF er en viktig prosess i revisjonsarbeidet da den vil styre planleggingen av videre revisjonshandlinger (Gulden, 2016, s.120). Det er derfor viktig for revisor å utøve profesjonelt skjønn når RVF fastsettes, ettersom feilaktig vurdering kan trekke ned revisjonskvaliteten og -effektiviteten (Messier et al., 2006, s.82). I praksis kan det være utfordrende å måle RVF. Etter ISA 200 punkt A42 kan det gjøres både kvalitativt og kvantitativt. De to komponentene, iboende risiko og kontrollrisiko, kan måles enten samlet eller separat. Messier et al. (2006) argumenterer for å måle iboende risiko og kontrollrisiko separat, da f.eks. en lav kontrollrisiko tilsier at revisjonsomfanget knyttet til rutiner og

internkontrollen kan reduseres. Derimot argumenterer [Gulden \(2016\)](#) for å måle risikoene samlet under ett mål for RVF, da de to risikoelementene er nært knyttet til hverandre. I praksis vil revisor være mer opptatt av sannsynligheten for at vesentlig feilinformasjon forekommer, og kilden for feilinformasjonen vil være av mindre betydning.

Revidert ISA 315

I 2019 ble den reviderte ISA 315 fremlagt med økt fokus på identifisering og vurdering av risikoer for vesentlig feilinformasjon. Den reviderte standarden trer i kraft for regnskapsperioder som starter f.o.m. 15.12.2021. Blant endringene er et nytt krav om å gjøre en separat vurdering av iboende- og kontrollrisiko når RVF skal fastsettes ([ICAEW, ud; Olsen, 2020](#)). I tillegg har den reviderte standarden et økt fokus på profesjonelt skjønn og en ny veiledning for måling av iboende risiko. Risikoen skal nå måles ved å vurdere skjæringspunktet mellom sannsynligheten og konsekvensen av mulig feilinformasjon på påstandsnivå.

2.3 Informasjonsrekkefølge

Revisjonsarbeid er en høykompleks prosess preget av usikkerhet, skjønnsmessige estimater og informasjonmengder utenfor revisors kontroll ([Yankova, 2015](#), s.73). Gjennom revisjonen blir det ofte gjort sekvensielle vurderinger ettersom ny informasjon blir tilgjengelig ([Gibbins, 1984](#)). Dette åpner opp for at rekkefølgen av informasjon kan være med å påvirke revisors vurderinger, noe som kan være et problem i praksis ([Bell et al., 2005](#), s.34).

Et sentralt konsept knyttet til rekkefølge er *primacy/recency effect* ([Bonner, 2008](#), s.184). Denne effekten tilsier at informasjon som blir presentert enten på en tidlig (primacy) eller sen (recency) fase blir mer vektlagt i den kognitive vurderingen. Disse to effektene har hvert sitt eget navn og utløses på ulike tidspunkter, men brukes samlet i begrepet *rekkefølgeeffekt* da de på et konseptuelt nivå viser til det samme fenomenet. Et eksempel på rekkefølgeeffekt er når to personer med det samme informasjonssettet, kommer til ulike konklusjoner, dersom informasjonen blir presentert i ulike rekkefølger ([Yankova, 2015](#), s.25).

Hogarth og Einhorn (1992) bygger videre på ideen om rekkefølgeeffekt og har utarbeidet *The belief-adjustment model*. Modellen beskriver hvordan en rekkefølgeeffekt kan forekomme som følge av en *anchoring-and-adjustment* strategi. I vurderingsprosessen dannes et anker, og kommende informasjon vurderes i henhold til dette ankeret. Modellen definerer et sett med prediktive hypoteser, og er en dominerende modell i litteratur som omhandler informasjonsrekkefølge i en revisjonskontekst (Bonner, 2008, s.185).

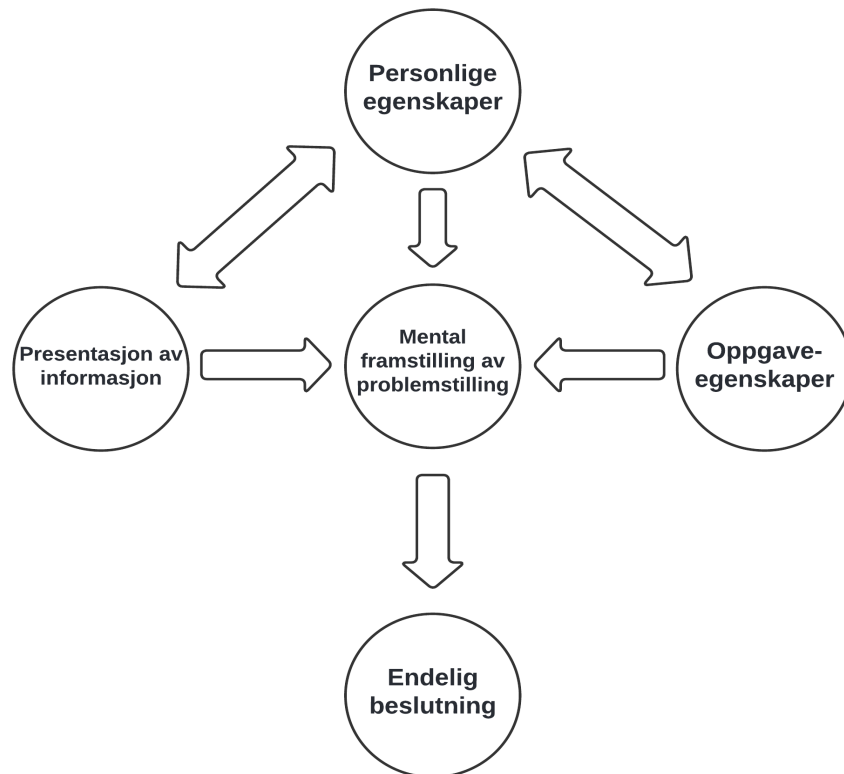
Studiene i dette forskningsområdet er for det meste konsistente med denne modellen (Bonner, 2008, s.188), men det er sprikende resultater i litteraturen. Trotman og Wright (1996) undersøkte revisors vurdering av fortsatt drift og oppdaget en rekkefølgeeffekt som ble moderert av erfaringsnivå. Derimot fant studiene til Asare (1992) og Ahlawat (1999) at rekkefølgeeffekten ikke ble moderert av erfaringsnivå under de samme omstendighetene.

Det argumenteres for at modellen til Hogarth og Einhorn kan utløse sprikende funn avhengig av oppgavekompleksitet og erfaring. Sistnevnte kan potensielt redusere rekkefølgeeffekten fordi økt erfaring reduserer den kognitive belastningen (Trotman og Wright, 2000). Studien til Monroe og Ng (2000) undersøkte revisors vurdering av iboende risiko, og fant ingen rekkefølgeeffekt. De hevder at under bestemte omstendigheter vil ulike oppgavekarakteristika kunne motvirke rekkefølgeeffekten.

2.4 Presentasjonsformat

Med presentasjonsformat menes det i hvilken utforming informasjon blir gitt til en beslutningstaker i revisjonsprosessen. For eksempel kan informasjon om en spesifikk regnskapslinje presenteres både grafisk og tabellarisk (Kelton et al., 2010). Litteraturen viser til at formateffekt kan forekomme ved at en beslutningstaker gjør ulike vurderinger avhengig av hvordan informasjon blir presentert (Bonner, 2008, s.189).

Det er mangel på teori som kan generalisere hvordan revisor påvirkes av ulike presentasjonsformater (Bonner, 2008, s.189). Flere studier har derfor benyttet seg av teori fra andre forskningsområder som utgangspunkt (Dilla et al., 2010; Kelton et al., 2010). En mye brukt modell er *The cognitive fit model* fra studien til Shaft og Vessey (2006), som inkluderer presentasjonsformat som en beslutningsvariabel. Inspirert av denne, utarbeidet Kelton et al. (2010) en kognitiv modell tilpasset revisjonsforskning.



Figur 2.2: Keltons beslutningsrammeverk for presentasjonsformat

Modellen i Figur 2.2 viser at presentasjonsformatet er en faktor i den mentale framstillingen av en problemstilling. Sammen med egenskaper ved både person og oppgave, vil dette kunne påvirke hvordan en beslutning blir tatt (Kelton et al., 2010). Det kan påvirke beslutningen på to ulike måter; enten direkte gjennom beslutningstakerens forståelse av problemstillingen, eller indirekte ved å trigge visse personlige egenskaper hos beslutningstakeren (Kelton et al., 2010). Et eksempel på en indirekte effekt kan være at et format trigger skeptiske holdninger hos de med lite kjennskap til formattypen.

I denne studien er statiske og interaktive formater to sentrale konsepter, hvor vi ønsker å undersøke om bruken av disse vil påvirke revisors beslutninger. Vi tar utgangspunkt i Yi et al. (2007) sin definisjon av interaktivitet hvor de viser til egenskaper som bl.a: (1) filtrere data etter visse kriterier, (2) markere dataelementer av interesse, (3) rearrangere format, og (4) fremheve data i mer eller mindre detalj. Interaktive formater er en kontrast til statiske formater, hvor datavisualiseringen er absolutt og ikke kan påvirkes.

[Kelton et al. \(2010\)](#) gjorde en litteraturgjennomgang hvor fokuset var rettet mot statiske presentasjonsformater. Formålet med studien var å identifisere trender samt hvilke statiske formater (f.eks. grafisk og tabellarisk) som virket mer effektivt. De oppsummerer funnene i dette forskningsområdet som blandet og at det er vanskelig å identifisere trender.

[Dilla et al. \(2010\)](#), gjorde en litteraturgjennomgang av studier rettet mot interaktive formater. De finner også sprikende resultater. Blant annet viser studien til [Hodge \(2001\)](#) at bruken av interaktive formater øker oppgavekompleksiteten og reduserer nøyaktigheten av revisors vurderinger. Derimot finner [Kumar og Benbasat \(2004\)](#) og [Schuff et al. \(2006\)](#) at bruken av interaktive formater øker effektiviteten og forbedrer beslutninger. Det kan argumenteres for at interaktive egenskaper kan redusere belastningen på den kognitive prosessen, da man blant annet kan filtrere og selektivt velge ut relevant informasjon. Men for at interaktive formater skal virke effektivt, må designet være hensiktsmessig for oppgaven ([Alawadhi, 2015](#)).

Som følge av utviklingen i visualiseringsteknikker i nyere tid kan det være relevant å undersøke formateffekten videre med moderne visualiseringer ([Alawadhi, 2015](#)). Det er tilsynelatende få studier som direkte undersøker hvordan statisk vs interaktivt presentasjonsformat påvirker beslutninger, men det finnes likevel noe litteratur på området. Eksempelvis er det en studie fra [Lurie og Mason \(2007\)](#) som undersøkte hvordan disse formatene påvirket beslutningene til markedsføringsledere. De fant at bruken av interaktive visualiseringsverktøy bidrar til bedre beslutninger, samtidig advarer de mot å generalisere disse funnene. Vi håper å kunne bidra til litteraturen ved å undersøke dette i en revisjonskontekst.

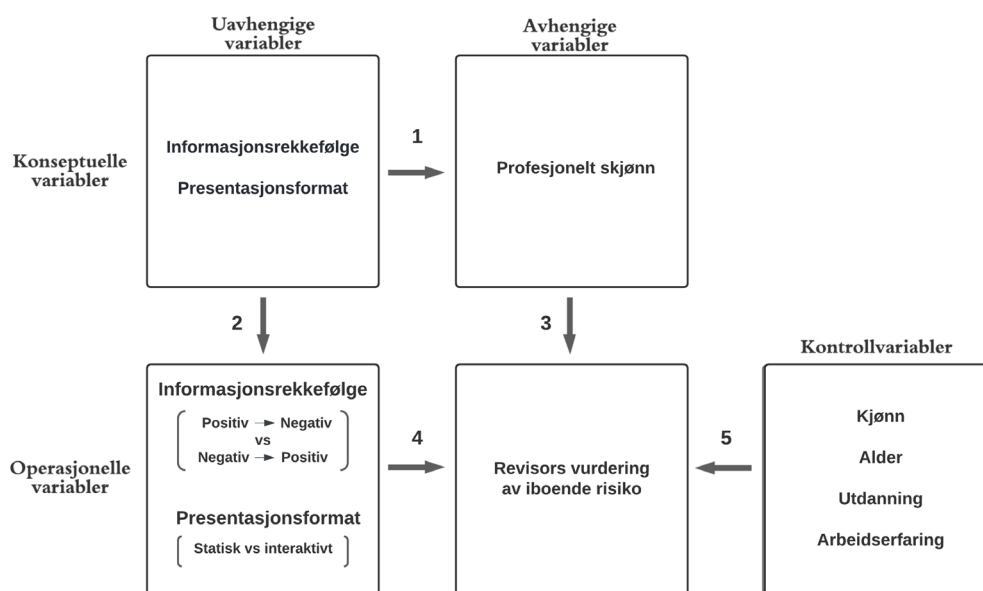
3 Forskningsmodell og hypoteser

Vi vil i det følgende kapittelet redegjøre for studiens forskningsmodell og validitetsrammeverk både på et konseptuelt og et operasjonelt nivå. Det blir også redegjort for studiens variabler og hypoteser.

3.1 Validitetsramme for forskningsmodellen

Med utgangspunkt i forskningsspørsmålet, har et validitetsrammeverk blitt utviklet med inspirasjon fra Libby (1981). Rammeverket gir et godt overblikk over hovedkomponentene, og veien fra litteratur til observasjon. Dette brukes som et prediktivt verktøy for å kartlegge den interne og eksterne validiteten, hvor Libby et al. (2002) påpeker:

«An experiment is effective if it provides evidence of sufficient internal validity that readers should believe the results of hypothesis tests, while being of sufficient external validity that it bears on a significant part of the financial accounting issue of interest.»



Figur 3.1: Validitetsrammeverk for forskningsmodellen

Figur 3.1 viser grafisk hvordan rammeverket langs den ene dimensjonen består av et konseptuelt og operasjonelt nivå. Det konseptuelle nivået tar utgangspunkt i litteraturen og tidligere teori om fenomenene som er av interesse. Det operasjonelle nivået viser hvordan fenomenene observeres og måles (Libby og Thorne, 2020). Langs den andre dimensjonen deles rammeverket opp i uavhengige variabler, avhengige variabler og kontrollvariabler.

3.2 Avhengig variabel

Den avhengige variabelen er den variabelen som vi ønsker å måle. Mer konkret ønsker vi å undersøke om den *avhenger* av andre variabler, altså om den kan påvirkes gjennom manipulasjon av en eller flere uavhengige variabler (Saunders et al., 2019, s.179). Link 3 i forskningsmodellen vår (Figur 3.1) viser hvordan vi har operasjonalisert den avhengige variabelen. Vi tar utgangspunkt i planleggingsfasen av et nytt revisjonsoppdrag, og bruker vurdering av iboende risiko som et mål på revisors profesjonelle skjønn.

Vi har valgt å fokusere på den iboende risikoen, ettersom revisors vurdering av dette er skjønnsmessig av natur jfr ISA 315 punkt A208. Risikovurderingene vil gjennomføres i tråd med den reviderte ISA 315 punkt 31 jfr A205 som stiller krav til at revisor tar stilling til både *sannsynligheten* for at det foreligger feilinformasjon i regnskapet, og *konsekvensen* av den mulige feilinformasjonen. Det er skjæringspunktet mellom de to som vil utgjøre den relevante iboende risikoen. I denne studien blir respondentene bedt om å vurdere sannsynlighet og konsekvens for fire individuelle cues i tillegg til en avsluttende og samlet totalvurdering. Designet tar utgangspunkt i studiene fra Ashton og Ashton (1988) og Monroe og Ng (2000). Videre detaljer om designet blir utredet i Kapittel 4 Metode.

3.3 Uavhengige variabler

Link 2 i forskningsmodellen (Figur 3.1) viser operasjonaliseringen av de uavhengige variablene. En uavhengig variabel kan beskrives som en variabel som manipuleres eller endres på, for å måle dens effekt og påvirkning på den avhengige variabelen (Saunders et al., 2019, s.179). Modellen vår består av to konseptuelle uavhengige variabler; informasjonsrekkefølge og presentasjonsformat. Vi ønsker å undersøke om det er en sammenheng mellom disse og vurdering av iboende risiko.

3.3.1 Informasjonsrekkefølge

Informasjonsrekkefølgen operasjonaliseres ved at vi utarbeider fire cues som blir presentert etter hverandre. Mer konkret er det snakk om fire separate nettsider/webapplikasjoner⁸ som inneholder datavisualiseringer og tilleggsinformasjon. Disse webapplikasjonene inneholder informasjon som kan knyttes til iboende risikofaktorer i den reviderte ISA 315 Vedlegg 2 punkt 5 (IAASB, 2019). To webapplikasjoner skal kommunisere positive signaler (lavere risiko) og to webapplikasjoner skal kommunisere negative signaler (høyere risiko). Rekkefølgen manipuleres ved at ulike grupper av respondenter får presentert nettsidene i motsatt rekkefølge.

3.3.2 Presentasjonsformat

Presentasjonsformatet operasjonaliseres ved at det utarbeides to versjoner av de fire webapplikasjonene. Hver cue vil da finnes i et statisk og et interaktivt format. De statiske versjonene vil ha et absolutt presentasjonsformat som betyr at respondentene ikke har mulighet til å påvirke utseendet eller funksjonaliteten på det som blir visualisert. De interaktive versjonene vil derimot være programmert med funksjoner som krever at brukeren aktivt må interagere med applikasjonene, for å filtrere og utheve ønsket data. For å isolere effekten til presentasjonsformatet, vil de statiske versjonene inneholde nøyaktig den samme informasjonen som de interaktive. Ulike grupper av respondenter får presentert applikasjonene i de to formatene.

3.4 Kontrollvariabler

Link 5 i validitetsrammeverket (Figur 3.1) omfatter de ulike kontrollvariablene som inkluderes i modellen. Et eksperiment opprettholder intern validitet dersom all variasjon i den avhengige variabelen kan forklares gjennom de uavhengige variablene (Campbell og Riecken, 1968). Trotman (2001) refererer til denne påstanden og beskriver mangel på kontrollvariabler som en av de større utfordringene i revisjonsforskning. Den interne validiteten svekkes dersom forklaringskraften på den avhengige variabelen stammer fra variabler som ikke er fanget opp i modelldesignet. Kontrollvariabler brukes derfor for å fange opp og identifisere eksternt støy.

⁸Heretter vil disse begrepene bli brukt om hverandre.

Studien til [Chung og Monroe \(2001\)](#) viste at kjønn påvirket beslutningsvariasjonen i lavkomplekse revisjonsoppgaver der menn viste lavere beslutningsvariasjon sammenlignet med kvinner. Vi tar utgangspunkt i disse funnene og inkluderer kjønn som en av kontrollvariablene i forskningsmodellen. Videre har det vært empiriske funn i tidligere forskningsartikler som viser til at erfaring og ekspertise påvirker revisors beslutninger ([Messier Jr og Tubbs, 1994](#); [Reckers og Schultz, 1993](#); [Ashton og Ashton, 1995](#)). Med bakgrunn i dette er derfor alder, utdanning og arbeidserfaring også inkludert som kontrollvariabler.

3.5 Hypoteseutvikling

Ved bruk av eksperimenter må hypoteser utredes. De skal gjenspeile forskningsspørsmålet og det er disse hypotesene som skal testes og analyseres ([Saunders et al., 2019](#), s.178). Hypoteser skal vise studiens prediksjoner om hvilke sammenhenger som finnes mellom de variablene som skal testes ([Trochim et al., 2016](#), s.17).

De deles ofte inn i en nullhypotese H_0 og en alternativhypotese H_A ([Saunders et al., 2019](#), s.179). Nullhypotesen er prediksjon om ingen årsakssammenheng (*A har ingen effekt på B*) og alternativhypotesen er forskerens prediksjoner om årsakssammenheng (*A har følgende effekt på B*). Gjennom eksperimentet og datainnsamlingen skal hypotesene testes. Dersom det er grunnlag for det, vil forskeren kunne forkaste H_0 og argumentere for en årsakssammenheng mellom variabel A og B .

3.5.1 Informasjonsrekkefølge og revisors profesjonelle skjønn

Vi tar utgangspunkt i studien⁹ til [Hogarth og Einhorn \(1992\)](#), og vil undersøke om vi finner de samme resultatene som deres modell predikerer. Eksperimentet vårt er konstruert som en enkel evalueringsoppgave¹⁰, hvor respondentene bes om å gjøre en samlet vurdering etter eksponering av fire cues med blandede signaler (både positive og negative).

⁹Se Kapittel 2.3 Informasjonsrekkefølge for redegjørelse av modellen til Hogarth et al.

¹⁰[Monroe og Ng \(2000\)](#) beskriver vurdering av iboende risiko som en evalueringsoppgave.

Med disse forutsetningene forventes det fra modellen en *primacy effect*. Dette betyr at dersom man skal gjøre en vurdering etter å ha sett gjennom en rekke med informasjon, så påvirkes man mest av den informasjonen man ser først. Våre hypoteser blir derfor følgende:

$H_0^{Rekkefølge}$: *Etter eksponering av fire cues med blandede signaler, vurderes iboende risiko likt uavhengig av hvilken signaltipe som presenteres først.*

$H_A^{Rekkefølge}$: *Etter eksponering av fire cues med blandede signaler, vurderes iboende risiko lavere ved presentasjon av positive cues før negative cues, og høyere ved presentasjon av negative cues før positive cues.*

3.5.2 Presentasjonsformat og revisors profesjonelle skjønn

Vi finner ikke tidligere forskning som spesifikt undersøker om statiske vs. interaktive presentasjonsformater påvirker revisors vurdering av iboende risiko. Vår studie kan forhåpentligvis bidra til å fylle dette rommet i litteraturen. Det er dog gjort studier som sammenligner statiske og interaktive presentasjonsformater i kontekst av mer generell beslutningstaking. Av disse peker flere mot at implementering av interaktive formater øker variasjonen i beslutningene. De argumenterer samtidig for at resultatene vanskelig kan generaliseres, og at oppgaveegenskaper spiller en stor rolle på resultatene. (Huang og Windsor, 1998; Huang, 2003; Wheeler og Arunachalam, 2008). Med mangelfull og sprikende litteratur er det vanskelig å predikere hvordan presentasjonsformatet vil påvirke respondentenes vurdering av iboende risiko, etter å ha sett gjennom fire cues. Vi er derfor retningsnøytrale og hypotesene blir følgende:

$H_0^{Format^1}$: *Ved samlet risikovurdering etter fire cues, blir ikke revisor påvirket av presentasjonsformatet*

$H_A^{Format^1}$: *Ved samlet risikovurdering etter fire cues blir revisor påvirket av presentasjonsformatet*

I vårt eksperiment gjør respondentene også en risikovurdering av fire individuelle cues, og vi har da muligheten til å undersøke formateffekt for spesifikke signaltyper (positiv og negativ). Monroe og Ng (2000) retter oppmerksomhet mot hvordan revisor vektlegger

høyrisikoinformasjon annerledes enn lavrisikoinformasjon. Denne skjevheten kan skyldes at revisorer relativt sett er mer sensitive og kritiske til informasjon som peker mot høyere risiko (Yankova, 2015, s.79). Studien til Monroe og Ng er gjort i kontekst av rekkefølgeeffekt, men vi mener det er interessant å undersøke dette også i konteksten av en mulig formateffekt og inkluderer følgende tilleggshypoteser:

$H_0^{Format^2}$: Ved risikovurdering av to positive cues blir ikke revisor påvirket av presentasjonsformatet

$H_A^{Format^2}$: Ved risikovurdering av to positive cues blir revisor påvirket av presentasjonsformatet

$H_0^{Format^3}$: Ved risikovurdering av to negative cues blir ikke revisor påvirket av presentasjonsformatet

$H_A^{Format^3}$: Ved risikovurdering av to negative cues blir revisor påvirket av presentasjonsformatet

3.5.3 Interaksjonseffekt mellom format og rekkefølge

Vi finner lite forskning på interaksjonseffekt mellom format og rekkefølge i en revisjonskontekst. Det er dog en studie som fant slik interaksjonseffekt utenfor en revisjonskontekst. Her eksperimenterte de med numeriske og verbale formater kombinert med manipulasjon av rekkefølge. Funnene viste at formatene trigget ulike kognitive prosesser som påvirket styrken av rekkefølgeeffekten (Shen og Hue, 2007).

Alawadhi argumenterer i sin doktoravhandling for at interaktive formater kan redusere den kognitive belastningen (Alawadhi, 2015). Andre studier argumenterer derimot for at interaktive formater øker kompleksiteten og kan føre til kognitiv overbelastning (Huang og Windsor, 1998; Huang, 2003; Wheeler og Arunachalam, 2008). Med bakgrunn i dette blir det utfordrende å predikere hvordan og i hvilken grad presentasjonsformatet kan påvirke rekkefølgeeffekten. Hypotesene våre blir dermed:

$H_0^{Interaksjon}$: Ved samlet risikovurdering etter fire cues vil presentasjonsformatet ikke påvirke rekkefølgeeffekten.

$H_A^{Interaksjon}$: Ved samlet risikovurdering etter fire cues vil presentasjonsformatet påvirke rekkefølgeeffekten.

4 Metode

I dette kapitlet vil vi redegjøre for forskningsdesignet og metoden. Forskningsmetoden kan beskrives som en strategisk tilnærming for å innhente og analysere data for å besvare forskningsspørsmålet (Johannesen og Christoffersen, 2016, s.26).

4.1 Forskningsdesign

Forskningsdesignet har som funksjon å binde alle delkomponentene i et forskningsprosjekt sammen (Trochim et al., 2016). Overordnet kan man skille mellom tre ulike typer forskningsdesign; eksplorativt, beskrivende og kausalt (Trochim et al., 2016). Denne studien fokuserer på sistnevnte.

Kausalitet kan defineres som sammenhengen mellom årsak og virkning. Årsaken kan anses som endringer i en eller flere uavhengige variabler. Virkning er effekten som dette har på den avhengige variabelen (Trochim et al., 2016). Denne studien forsøker å avdekke en slik årsakssammenheng, hvor formålet er å teste om manipulasjon av informasjonsrekkefølge og presentasjonsformat vil påvirke revisors vurdering av iboende risiko. Det vil da være hensiktsmessig med et eksperimentelt design. Dette er en metode som kjennetegnes ved at det samles inn data fra sammenlignbare grupper (Trochim et al., 2016). Gruppene vil være sammenlignbare ved at alle kjennetegnene er like, med unntak av de manipulerede tilstandene som skal undersøkes (Johannesen et al., 2020, s.263). Eksperimentelle design er også den mest forekommende metoden i studier rettet mot revisors atferd (Trotman, 2001).

Vår studie har to uavhengige variabler som manipuleres for å avdekke en mulig årsakssammenheng. Med flere uavhengige variabler kan et faktorielt design på eksperimentet være fordelaktig (Trochim et al., 2016, s.237). Et faktorielt design retter fokuset mot manipulasjonsfaktorene og de ulike tilstandene som disse faktorene kan ha. Det blir da mulig å fastslå om de ulike faktorene/tilstandene faktisk påvirker den avhengige variabelen. De ulike faktorene kan påvirke den avhengige variabelen enten enkeltvis, eller samlet gjennom en interaksjonseffekt (Trochim et al., 2016, s.238).

I slike design deles utvalget i grupper som eksponeres for ulike tilstander enten *within-subjects*¹¹ eller *between-subjects*¹². (Trotman, 2001). Dette kan framstilles i en oversiktlig matrise som viser hvordan faktorene og tilstandene fordeles mellom gruppene. Figur 4.1 viser studiens faktorielle design med to tilstander under hver av de to faktorene (2×2). For å øke studiens validitet er det benyttet «between-subjects» -fordeling, som betyr at hver gruppe eksponeres for en unik kombinasjon av tilstander.

		Informasjonsrekkefølge	
		Lav risiko → Høy risiko (PPNN)	Høy risiko → Lav risiko (NNPP)
Presentasjonsformat	Statisk	Gruppe 1	Gruppe 3
	Interaktivt	Gruppe 2	Gruppe 4

Figur 4.1: Studiens 2×2 faktorielle design. To grupper eksponeres for samme presentasjonsformat, men motsatt informasjonsrekkefølge. To grupper eksponeres for samme informasjonsrekkefølge, men ulikt presentasjonsformat.

4.2 Design av undersøkelsen

Ved konstruksjon av undersøkelsen ble Qualtrics' løsning¹³ for spørreskjemaer benyttet. Gjennom NHH sin lisens fikk vi tilgang til alle nødvendige funksjonaliteter, blant annet randomisering av grupper, linking til eksterne nettsider og eksport av data i csv-format for fleksibel implementering i vårt analyseskript. Nettsidene ble hovedsakelig skrevet i programmeringsspråket R og satt i produksjon gjennom RStudios løsning¹⁴ for drift av webapplikasjoner. Selve undersøkelsen ble bygget opp på følgende måte: (1) introduksjon med praktisk informasjon og kontekst, (2) hoveddel med webapplikasjoner og individuelle risikovurderinger, og (3) en samlet og avsluttende risikovurdering, hvor det også ble svart på noen kontrollspørsmål.

¹¹En gruppe eksponeres for alle tilstandene i en faktor.

¹²En gruppe eksponeres for kun én tilstand per faktor.

¹³<https://www.qualtrics.com/>

¹⁴<https://www.rstudio.com/products/shinyapps/>

4.2.1 Del 1: Innledning og kontekst

Først blir deltakerne gitt informasjon om hva som er formålet med undersøkelsen, i tillegg til praktisk informasjon om anonymitet og gjennomføring. Dersom respondentene svarer bekreftende på at de ønsker å delta, blir de presentert for det fiktive caset og konteksten rundt det de skal gjøre.

Deltakerne ble bedt om å se for seg at de jobbet i revisjonsselskapet *Revisjon AS* og at de skulle gjøre en innledende vurdering av iboende risiko i regnskapet til deres nye kunde *Produksjonsselskapet AS*. Kunden driver med produksjon og salg til industrien, og det var én spesifikk produktlinje som skulle være i fokus. For å redusere eventuelle bias, ble det ikke gitt informasjon om hva bedriften produserte eller hvilken bransje de opererer i. Videre ble det gitt en kort definisjon av hva iboende risiko er. Som en kontrollsjekk ble deltakerne bedt om å gjengi denne definisjonen før de gikk videre til nettsidene.

4.2.2 Del 2: Webapplikasjoner

Hoveddelen av undersøkelsen bestod av fire separate webapplikasjoner som inneholdt datavisualiseringer og tilleggsinformasjon. Hver nettside inneholdt én spesifikk type datavisualisering, satt i en revisjonskontekst. De ulike typene av visualiseringer var; tidsserier, ordsky, flytskjema, og kart. Deltakerne ble bedt om å bruke ca to minutter på hver nettside til å bli kjent med innholdet og gjøre seg opp en mening om den iboende risikoen.

Etter hver nettside ble de bedt om å vurdere iboende risiko basert på det de hadde sett, ved å tallfeste både *sannsynligheten* for feilinformasjon og *konsekvensen* av den potensielle feilinformasjonen. Begge to ble målt på skalaer fra 0-100, hvor 0 representerer hhv. ingen risiko for feil og null konsekvens, og 100 betyr at det helt sikkert er feil og maks konsekvens. Etter ISA 315 punkt 31 jfr A205 er det skjæringspunktet mellom sannsynlighet og konsekvens som utgjør den relevante iboende risikoen. Vi har operasjonalisert dette til å være gjennomsnittet av sannsynlighet og konsekvens.¹⁵

¹⁵En alternativ definisjon av variablene har også blitt utforsket. Se 6.4.2 Alternativ tolkning av avhengig variabel for detaljer.

Manipulasjonsfaktorene som vist i Figur 4.1 er informasjonsrekkefølge og presentasjonsformat. Totalt ble det laget åtte webapplikasjoner, ettersom fire cues (to positive og to negative) ble designet i to ulike versjoner; statisk og interaktiv. Respondentene ble så delt inn i fire grupper som fikk presentert nettsidene på følgende måte:

Gruppe 1 - $\frac{PPNN}{Statisk}$ $Tidsserier(P) \rightarrow Ordsky(P) \rightarrow Flytskjema(N) \rightarrow Kart(N)$

Gruppe 2 - $\frac{PPNN}{Interaktiv}$ $Tidsserier(P) \rightarrow Ordsky(P) \rightarrow Flytskjema(N) \rightarrow Kart(N)$

Gruppe 3 - $\frac{NNPP}{Statisk}$ $Kart(N) \rightarrow Flytskjema(N) \rightarrow Ordsky(P) \rightarrow Tidsserier(P)$

Gruppe 4 - $\frac{NNPP}{Interaktiv}$ $Kart(N) \rightarrow Flytskjema(N) \rightarrow Ordsky(P) \rightarrow Tidsserier(P)$

Tidsserier - Positiv cue

For denne nettsiden ble respondentene bedt om å vurdere risiko knyttet til bransjeendringer og forutsetningen for fortsatt drift. Nettsiden som respondentene ble presentert for inkluderte tre ulike tidsserier¹⁶, alle med årlig intervall:

1. Antall registrerte selskaper fra 2012 til 2021.
2. Bransjens totale omsetning knyttet til produktlinjen, vist i MNOK fra 2012 til 2021.
3. Klientens totale omsetning knyttet til produktlinjen, vist i MNOK fra 2012 til 2021.

I tillegg til historisk data, ble fremtidsprediksjoner vist for de neste tre årene (2022 - 2024). Intensjonen var å vise at Produksjonsselskapet AS har gjort det godt i markedet og at det er en positiv utvikling i etterspørselen etter det de selger. Disse faktorene skal tale i retning av lavere risiko.

Den statiske versjonen av nettsiden (se Figur A2.1) presenterte alle tre tidsseriene ved siden av hverandre. Den interaktive versjonen (se Figur A2.2) viste kun én tidsserie av gangen, uten fremtidsprediksjoner. Den krevde at respondentene aktivt filtrerte mellom de ulike figurene, i tillegg til å måtte trykke på en knapp for å få se fremtidsprediksjonene.

¹⁶En tidsserie er data som observeres sekvensielt over tid. Eksempler på tidsserier er årlig avkastning, kvartalsvise salgstill eller månedlig regnfall (Hyndman og Athanasopoulos, 2021).

Ordsky - Positiv cue

For denne nettsiden ble respondentene bedt om å vurdere risiko knyttet til interne forhold i selskapet, hvor de primære risikofaktorene var kompetanse, integritet og organisasjonsholdninger. Konteksten var at alle epost-utvekslinger mellom regnskapskontoret og ledelsen hadde blitt samlet sammen. Basert på dette var det laget en ordsky¹⁷ som viste de mest brukte ordene i epost-utvekslingene. Intensjonen var at ordskyen skulle ha en overvekt av positivt ladete ord, som kommuniserer en god tone mellom regnskapskontoret og ledelsen. Dette skal tale i retning av lavere risiko.

Ordene som ble vist tok utgangspunkt i prosjektet til [Barnes et al. \(2019\)](#) hvor et omfattende engelsk sentiment-leksikon ble oversatt til norsk. Det ble i alt plukket ut 100 ord, hvorav de fleste var positivt ladet. Det ble også inkludert noen negative og noen mer nøytrale ord som forventes i en regnskapskontekst, slik som *bilag*, *transaksjon* og *beløp*.

Den statiske versjonen av nettsiden (se Figur A2.3) presenterte ordskyen uten mulighet for filtrering og endring. Den interaktive versjonen (se Figur A2.4) presenterte også ordskyen med alle 100 ordene, men respondentene fikk her muligheten til å filtrere antall ord som ble vist, i tillegg til mulighet for å endre på den visuelle framtoningen med ulike former og farger.

Flytskjema - Negativ cue

For denne nettsiden ble respondentene bedt om å vurdere risiko knyttet til bokføring av salgstransaksjoner, hvor den primære risikofaktoren var misligheter. Prosessen rundt bokføring av salgstransaksjoner ble beskrevet i korte trekk, hvor det kom fram at de fleste transaksjoner ble matchet automatisk og så ført i hovedbok. Men det kom også fram at selgere hadde mulighet til å manuelt overstyre eventuelle avvik.

Denne prosessen ble visualisert gjennom et *Sankey-diagram*, som er et flytskjema¹⁸ hvor bredden på nodene er proporsjonal med mengden som flyter gjennom de. Diagrammets første node representerte det totale salget. Derfra ble det vist hvor mye som «fløt» til automatisk eller manuell registrering, og videre til 3-way match med eller uten avvik. De

¹⁷Fra [NAOB](#): «Billedlig fremstilling av ordene fra en gitt tekst, hvor ordene vises med større bokstaver jo hyppigere de forekommer i teksten.»

¹⁸Fra [NAOB](#): «Grafisk fremstilling som viser gangen i noe, f.eks. enkeltoperasjoner i en prosess eller en serie operasjoner på datamaskin.»

siste nodene viste hvor mye avvik i kroner, som var manuelt overstyrt av selger eller sendt til regnskapskontoret for videre sjekk.

Nettsiden hadde totalt tre slike diagrammer, ett for hvert av de siste tre årene. De kunne vise at det var en negativ utvikling, hvor både mengden som var overstyrt av selger, i tillegg til de totale avvikene, hadde økt. En slik utvikling skulle tale i retning av høyere risiko. Den statiske versjonen (se Figur A2.5) viste alle tre diagrammene under hverandre, hvor det ikke var muligheter for filtrering eller interaksjon. Den interaktive versjonen (se Figur A2.6) viste kun ett diagram, hvor brukeren aktivt måtte filtrere på år. I tillegg fikk brukeren mulighet til å utheve avvikene i rødt, samt muligheten til å fjerne alt som ikke var avvik fra skjemaet.

Kart - Negativ cue

For denne nettsiden ble respondentene bedt om å vurdere risiko knyttet til leverandørendringer. Konteksten var at Produksjonsbedriften AS i løpet av de siste tre årene hadde skiftet ut leverandørene som var knyttet til den aktuelle produktlinjen. Dette ble visualisert gjennom tabeller og kart, med ulik finansiell og ikke-finansiell informasjon. Det finansielle var fokusert rundt transaksjonsbeløp og kredittider. Den ikke-finansielle informasjonen inkluderte bl.a. informasjon om hvor utsatt de ulike leverandørlandene var for hvitvasking og korrupsjon.

Gjennom årene har klienten byttet ut leverandører lokalisert i høykostnadsland som Norge, Sverige og Tyskland. Her har de hatt et langt samarbeid og gode betingelser (lang kredittid). Disse er erstattet med leverandører lokalisert i lavkostnadsland som Kina, India og Bangladesh, hvor de har kuttet kostnader, men også fått dårligere betingelser. Slike endringer skulle peke i retning av høyere risiko.

Den statiske versjonen (se Figur A2.7) viste all tilgjengelig informasjon på én og samme side. Det var et kart hvor alle leverandørene var markert, og tabeller med informasjon for de siste tre årene. Den interaktive versjonen (se Figur A2.8) krevde at respondentene aktivt filtrerte på hvilket år de ville se. Kartet var programmert slik at det automatisk zoomet inn på de aktuelle leverandørene for det aktuelle året. Videre var kartet programmert interaktivt, slik at brukeren kunne zoome inn/ut og trykke på markørene for å avdekke informasjon.

4.2.3 Del 3: Totalvurdering og kontrollspørsmål

Etter å ha sett gjennom de fire nettsidene blir deltakerne bedt om å gjøre en samlet risikovurdering basert på alle cues. Denne måles på samme måte som for de individuelle målingene, altså med separat vurdering av sannsynlighet og konsekvens. Etter den siste risikovurderingen blir deltakerne bedt om å svare på spørsmål om kjønn, alder, utdanning og arbeidserfaring innen regnskap og revisjon. Avslutningsvis blir respondentene spurt om de hadde samarbeidet med noen andre under gjennomføringen av undersøkelsen. Se Appendiks A1 for skjermtdump av spørreundersøkelsen.

4.3 Prestudie

Prestudier brukes for å få tilbakemelding på utformingen av undersøkelsen ([Johannesen et al., 2020](#), s.295). For å teste om undersøkelsen kunne gi oss svar på forskningsspørsmålet, ønsket vi tilbakemeldinger fra deltakerne i prestudien om hvorvidt:

- Det var klart og tydelig hvordan undersøkelsen skulle gjennomføres
- Det var et realistisk case med relevant informasjon
- Nettsidene kommuniserte tiltenkt signaltipe (positiv og negativ)
- De interaktive funksjonalitetene var intuitive, og om de faktisk ble brukt
- Nettsidene fungerte som tiltenkt på tvers av ulike skjermer og enheter

Prestudien hadde fire deltakere, som alle hadde kjennskap til revisjon (to MRR-studenter og to praktiserende revisorer). Hver deltaker ble plassert i en av de fire gruppene slik at alle gruppene ble dekket. For å minimere eventuelle bias ble informasjon om studiens manipulasjonsfaktorer holdt skjult. Fra prestudien fikk vi gode tilbakemeldinger om at oppgaven var tydelig og at nettsidene virket relevante og realistiske for praksis. Ved å inkludere revisorstudenter i prestudien fikk vi også bekreftet at oppgaven var forståelig for de uten arbeidserfaring.

En viktig faktor for å kunne undersøke rekkefølgeeffekt er hvorvidt nettsidene faktisk kommuniserer tiltenkt signaltipe. Etter å ha gjennomført undersøkelsen, informerte vi deltakerne i prestudien om studiens formål og manipulasjonsfaktorer. Det var enighet om at skillet mellom positive og negative signaler fungerte som tiltenkt. Videre var det enighet om at de interaktive funksjonalitetene virket intuitive og at de faktisk ble brukt.

Vi ble dog informert om at noen funksjoner ikke hadde fungert som tiltenkt på skjermer med lavere oppløsning. Vi gjorde derfor funksjonene mer fleksible på tvers av skjermer, i tillegg til andre små endringer relatert til informasjonsmengde og generell estetikk.

4.4 Innsamling av data

For å observere fenomenene som studien undersøker må det innhentes data ([Johannesen et al., 2020](#), s.25). Det er i den forbindelse nødvendig å definere hvem man skal innhente data fra og hvordan man skal nå fram til disse. I tillegg er det forskningsetiske retningslinjer å ta hensyn til.

4.4.1 Utvalg

I de fleste eksperimentelle studier ønsker forskeren å vite noe om en populasjon¹⁹ ([Johannesen et al., 2020](#), s.268). Forskningsspørsmålet vårt viser til revisors profesjonelle skjønn, og det er derfor praktiserende revisorer på alle nivåer som blir den teoretiske populasjonen i studien. Å gjennomføre et eksperiment med hele denne populasjonen er både vanskelig og unødvendig ([Johannesen et al., 2020](#), s.269). I stedet er det hensiktsmessig å trekke ut et mindre utvalg som skal være representativt for alle norske revisorer. Eksperimentet gjennomføres på utvalget, og resultatene kan forhåpentligvis generaliseres til hele populasjonen.

Når et representativt utvalg skal trekkes, er det ulike strategier som kan brukes. Ved trekking av utvalget har vi i stor grad vært begrenset til å bruke eget nettverk. Dette kan anses som en bekvemmelighetsutvelgelse, altså et ikke-sannsynlighetsutvalg basert på hvem som er enklest å nå ut til ([Johannesen et al., 2020](#), s.68). En slik utvelgelse kan være til fordel når en studie har begrenset med tid og ressurser ([Taherdoost, 2016](#)).

For at studiens utvalg skal være representativt bør det inkludere et tilstrekkelig stort antall revisorer. Med studiens begrensninger ble det tidlig klart at dette kom til å bli en utfordring. Vi valgte derfor å inkludere revisorstudenter uten praktisk erfaring som en del av utvalget og designet undersøkelsen med hensyn til dette. Vi antar at studentenes vurderinger kan likestilles med vurderingene som gjøres av praktiserende revisorer. Fra

¹⁹Målgruppen som en studie ønsker å vite noe om ([Johannesen et al., 2020](#), s.268)

litteraturen er det delte meninger om bruken av studenter, og vi kommer tilbake til dette i Kapittel 4.6.1 Validitet.

4.4.2 Gjennomføring

Som nevnt ble spørreskjemaet utformet i Qualtrics, og ved publisering ble det laget en link som sendte respondentene randomisert til én av de fire ulike versjonene/gruppene. Undersøkelsen ble markedsført til revisorstudenter ved NHH, BI, OsloMet og UiA gjennom Facebook-grupper, Canvas, og personlige henvendelser. For å nå ut til praktiserende revisorer ble den markedsført gjennom LinkedIn og personlige henvendelser til vårt eget nettverk. Det ble da spurt om de kunne videresende undersøkelsen til kollegaer og andre med utdanning og/eller erfaring innenfor revisjon. Undersøkelsen ble publisert 20. april 2022 og var aktiv fram til 9. mai 2022.

4.4.3 Forskningsetikk

Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH²⁰) har vedtatt forskningsetiske retningslinjer som veiledning for hva en forsker må ta hensyn til ved innsamling av data. Disse retningslinjene kan oppsummeres i tre punkter (Johannesen et al., 2020, s.45): (1) informantens rett til selvbestemmelse og autonomi, (2) forskerens plikt til å respektere informantens privatliv og (3) forskerens ansvar for å unngå skade.

Etter introduksjonen i undersøkelsen blir deltakeren spurt om hen ønsker å delta i undersøkelsen. Selvbestemmelse og autonomi opprettholdes ved at deltakeren blir gitt valgfrihet til å delta. Under gjennomføringen av undersøkelsen er det også mulighet til å trekke seg dersom deltakeren ikke lenger ønsker å delta.

Deltakerens personvern ivaretas ved at undersøkelsen gjennomføres anonymt. Det innhentes indirekte personopplysninger som kjønn, alder, utdanning og arbeidserfaring. Men disse opplysningene, enten enkeltvis eller kombinert, gjør det ikke mulig å identifisere en deltaker. Det samles heller ikke inn IP-adresser eller annen informasjon som gjør det mulig å identifisere respondentene. Det har derfor ikke vært nødvendig å melde studien til NSD²¹.

²⁰<https://www.forskningsetikk.no/retningslinjer/hum-sam/>

²¹Norsk senter for forskningsdata

Forskerens ansvar for å unngå skade er mer aktuelt i medisinsk forskning, men kan også være av betydning for samfunnsvitenskapelig forskning (Johannesen et al., 2020, s.46). Undersøkelsen vår tar for seg en vurdering av iboende risiko, og den vil ikke på noen måte berøre sensitive temaer eller områder.

4.5 Analyse av data

For å analysere datamaterialet ble det laget et skript i R. I dette skriptet ble rådata rensset, variabler konstruert og hypotesetesting utført med variansanalyser.

4.5.1 Konstruksjon av variabler

Respondentene ble bedt om å vurdere iboende risiko ved å kvantifisere sannsynlighet og konsekvens for hver cue. I tillegg gjorde de en avsluttende og samlet totalvurdering. Dette resulterer i ti unike risikomålinger fra hver respondent. Basert på disse målingene ble det konstruert tre avhengige variabler til bruk i modellene: (1) gjennomsnittet av fire målinger for positive cues, (2) gjennomsnittet av fire målinger for negative cues, og (3) gjennomsnittet av to målinger for den avsluttende samlede vurderingen²². Algebraisk kan dette vises som

$$\begin{aligned}
 y_{Pos} &= \frac{1}{4}(T_{Sann} + T_{Kons} + O_{Sann} + O_{Kons}) = \mathbf{P-Cues}_{Risk} \in [0, 100] \\
 y_{Neg} &= \frac{1}{4}(F_{Sann} + F_{Kons} + K_{Sann} + K_{Kons}) = \mathbf{N-Cues}_{Risk} \in [0, 100] \\
 y_{Sam} &= \frac{1}{2}(S_{Sann} + S_{Kons}) = \mathbf{Samlet}_{Risk} \in [0, 100]
 \end{aligned}$$

hvor

T_{Sann} og T_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med tidsserier.

O_{Sann} og O_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med ordsky.

F_{Sann} og F_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med flytskjema.

K_{Sann} og K_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med kart.

S_{Sann} og S_{Kons} er måling av sannsynlighet og konsekvens for den samlede vurderingen.

²²En alternativ definisjon av variablene har også blitt utforsket. Se 6.4.2 Alternativ tolkning av avhengig variabel for detaljer.

De uavhengige variablene; informasjonsrekkefølge og presentasjonsformat, kodes i modellene som dummy, hvor

$$z_{Rekkefølge} = \begin{cases} 0 & \text{hvis rekkefølgen er PPNN} \\ 1 & \text{hvis rekkefølgen er NNPP} \end{cases}$$
$$z_{Format} = \begin{cases} 0 & \text{hvis formatet er statisk} \\ 1 & \text{hvis formatet er interaktivt} \end{cases}$$

Kontrollvariablene; kjønn, alder, utdanning og erfaring inkluderes i modellene som hhv. $x_{Kjønn}$, x_{Alder} , $x_{Utdanning}$, og $x_{Erfaring}$. Også her blir de kategoriske variablene (kjønn, utdanning og erfaring), kodet som dummy.

4.5.2 Variansanalyser

Studiens faktorielle design gjør at en toveis variansanalyse er hensiktsmessig for å teste hypotesene (Sirkin, 2006, s.349). Dette er en analysemetode som ligner mye på lineær regresjon, hvor effektene fra de ulike variablene testes på kryss og tvers. Gjennom en slik analyse vil vi kunne si noe om både hovedeffekten til de uavhengige variablene, og interaksjonseffekten mellom dem. For å styrke validiteten blir det i tillegg til vanlig ANOVA gjennomført en ANCOVA. Sistnevnte inkluderer kontrollvariabler, og benyttes for å fjerne støy som antas å påvirke modellen (Montgomery, 2012, s.656). Sentrale forutsetninger for en slik variansanalyse er blant annet antakelse om en normalfordelt avhengig variabel, og homogenitet i variansen mellom gruppene. Normalitet kan testes med Shapiro-Wilk -test, og homogenitet kan testes med Levenes test (Lærd, 2022).

For tilleggshypotesene om formateffekt inkluderes de samme uavhengige variablene som for hovedhypotesene, ettersom man bør være forsiktig med ta vekk interaksjonsleddet i slike tofaktor-modeller (Montgomery, 2012, s.202). For tilleggshypotesene vil det dog være hovedeffekten til presentasjonsformatet som er av interesse, ettersom studien ikke er designet for å teste rekkefølgeeffekten til individuelle cues.

4.6 Validitet og reliabilitet

Validitet og reliabilitet er to kvalitetsegenskaper som ser på studiens gyldighet og pålitelighet (Johannesen et al., 2020). De er nært knyttet til hverandre, da det vil være lite hensiktsmessig å se på en studies validitet dersom reliabiliteten er svak, og motsatt. For å opprettholde god kvalitet på studien må både reliabilitet og validitet være til stede (Grenness, 2004; Saunders et al., 2019).

4.6.1 Validitet

Validitet viser i hvilken grad en undersøkelse måler det den er ment til å måle (Grenness, 2004; Johannesen et al., 2020). Det skilles mellom *begrepsvaliditet*, *intern validitet* og *ekstern validitet*. I denne studien er det nødvendig å stille spørsmål ved om undersøkelsen gir en troverdig måling av revisors profesjonelle skjønn.

Begrepsvaliditet

Det dreier seg om i hvilken grad dataene som innhentes kan knyttes opp mot det fenomenet som studeres (Johannesen et al., 2020, s.44). Studiens formål er å undersøke revisors profesjonelle skjønn og dette operasjonaliseres gjennom en vurdering av iboende risiko. En slik vurdering er skjønnsmessig av natur, og må anses å falle inn under hvordan IAASB beskriver utøvelsen av *profesjonelt skjønn* i deres rammeverk for god revisjonskvalitet. For å opprettholde begrepsvaliditeten i studien er det nødvendig at respondentene har forståelse for begrepet *iboende risiko*. For å legge til rette for dette, blir en definisjon av begrepet først presentert, før deltakerne blir bedt om å svare på hva iboende risiko er.

Intern validitet

Intern validitet viser til hvordan en undersøkelse er egnet til å påvise en kausal sammenheng mellom de uavhengige og avhengige variablene (Johannesen et al., 2020, s.407). Denne studien har et faktorielt design med randomisert fordeling av tilstander mellom gruppene. Gruppene testes «between-subjects» for å redusere potensielle undersøkelseeffekter²³ og skjevheter blant gruppene. For å påvise årsakssammenheng mellom informasjonsrekkefølge

²³En effekt der deltakere som blir eksponert for flere tilstander i en faktor klarer å identifisere manipulasjonsvariabelen.

og profesjonelt skjønn, er det nødvendig at våre cues kommuniserer tiltenkt signaltipe (positivt eller negativt). De individuelle risikovurderingene som gjøres for hver cue brukes som en validitetssjekk for å teste tiltenkt signaltipe. Videre inkluderes relevante kontrollvariabler for å redusere eksternt støy og muligheten for spuriøse²⁴ sammenhenger.

Et mulig tiltak for å ytterligere styrke intern validitet er å være til stede under gjennomføringen av eksperimentet (Trotman, 2001). Studiens ressursbegrensninger gjorde det ikke mulig for oss å være fysisk til stede, for å overvåke deltakernes forståelse og gjennomføring av undersøkelsen. Den digitale gjennomføringen av undersøkelsen kan derfor bidra til å svekke den interne validiteten.

Ekstern validitet

Ekstern validitet handler om hvordan resultatene fra studiens utvalg kan generaliseres til å gjelde for hele populasjonen. En forutsetning for dette er at utvalget er representativt (Trochim et al., 2016, s.83). Representativitet oppnås dersom sammensetningen av relevante egenskaper i utvalget tilsvarer sammensetningen i populasjonen. En vanlig strategi for å oppnå dette er å benytte utvalg av en viss størrelse som er tilfeldig trukket fra populasjonen (Johannesen et al., 2020, s.269). Med ressursbegrensningene i denne studien var det ikke mulig å trekke et utvalg på denne måten. Som et alternativ har vi benyttet en strategisk utvelgelse hvor vi har nådd ut til revisorer gjennom vårt eget nettverk. Dette kan være med på å svekke studiens eksterne validitet.

For å øke utvalgsstørrelsen har vi inkludert revisorstudenter. Noen hevder at dette kan redusere sannsynligheten for at utvalget er representativt, og differanser i personlighetstrekk mellom studenter og revisorer kan være en svakhet (Sears, 1986). Andre oppfordrer derimot til å inkludere studenter. Deres kunnskapsnivå kan hensyntas i designet av eksperimentet, og de bidrar til et mer ressursvennlig utvalg (Peecher og Solomon, 2001). For å motvirke potensielle svekkelser ble det inkludert studenter i pilottestingen. Ved utforming av undersøkelsen ble det tatt hensyn til studentenes manglende arbeidserfaring. Oppgavene holder seg innenfor grunnleggende revisjonspensum, og nødvendig informasjon blir forklart i detalj.

²⁴En falsk eller tilsynelatende sammenheng.

4.6.2 Reliabilitet

Reliabilitet handler om nøyaktigheten av datainnsamlingen og påliteligheten av resultatene (Grenness, 2004; Johannesen et al., 2020). Reliabiliteten sier altså i hvilken grad resultatene gjenspeiler utvalget. For å styrke reliabiliteten har studien implementert anonymitet i undersøkelsen. Dette er gjort for å motivere deltakerne til å være ærlige og føle seg trygge på besvarelsene sine. Spørsmålene i undersøkelsen ble formulert på en konkret og nøytral måte for å redusere potensielle bias. En pilottesting ble gjennomført for å identifisere uklarheter på forhånd før publiseringen av undersøkelsen.

5 Resultater

I dette kapittelet blir det redegjort for hvordan datamaterialet ble filtrert og analysert. Videre blir resultatene fra analysene presentert og vurdert i lys av studiens hypoteser.

5.1 Datagrunnlag

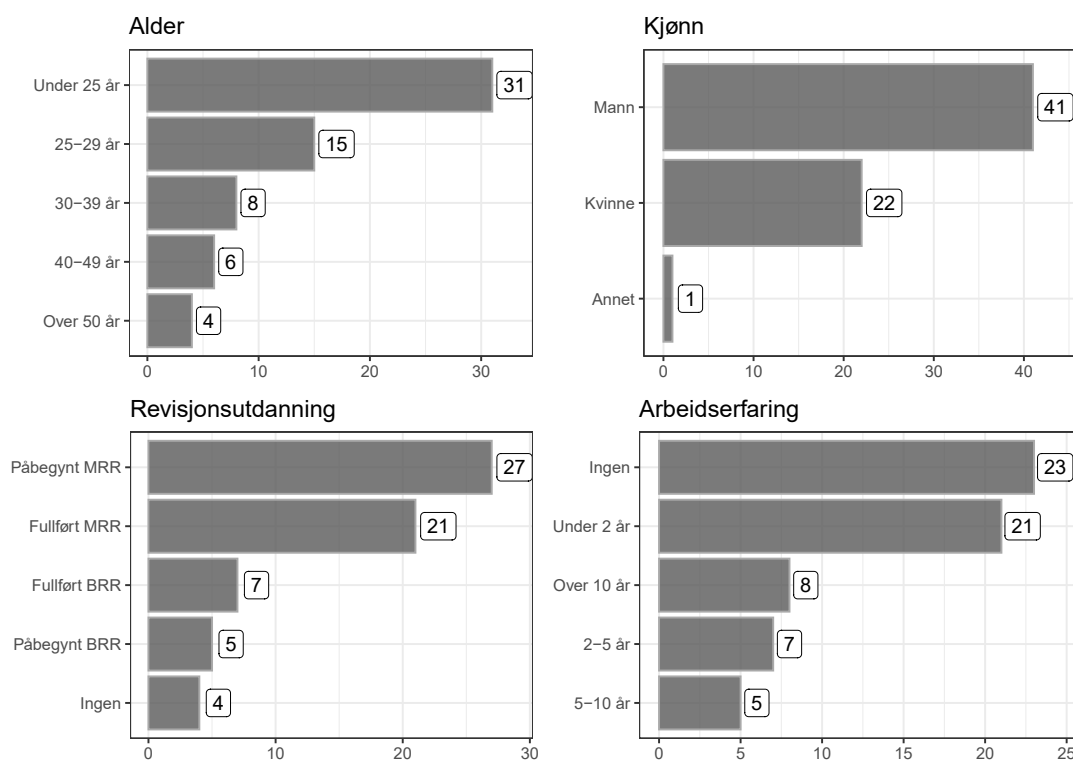
Det var 82 respondenter som fullførte hele undersøkelsen. Blant disse var det 12 respondenter som verken hadde utdanning eller erfaring innen revisjon. Vi antar at disse 12 vil redusere sannsynligheten for et representativt utvalg og de ble filtrert ut av datasettet. I tillegg var det ytterligere ni respondenter med varierende grad av kompetanse som svarte feil på det innledende spørsmålet «*Hva er iboende risiko?*». Blant disse var det tre respondenter med fullført mastergrad i revisjon og lang arbeidserfaring som revisor. Vi antar at disse har god forståelse for konseptet og inkluderte dem likevel i datasettet. De resterende seks som svarte feil på spørsmålet, ble fjernet. Etter filtrering var det 64 respondenter, fordelt relativt jevnt mellom gruppene, som dannet grunnlaget for analysen.

Gruppe	Rekkefølge	Format	Antall før filter	Filtrert	Antall etter filter
Gruppe 1	PPNN	Statisk	23	-7	16
Gruppe 2	PPNN	Interaktiv	20	-4	16
Gruppe 3	NNPP	Statisk	19	-4	15
Gruppe 4	NNPP	Interaktiv	20	-3	17
<i>Sum:</i>			82	-18	64

Tabell 5.1: Antall respondenter i utvalget fordelt per gruppe før og etter filtrering.

5.1.1 Demografi

I undersøkelsen ble det innhentet informasjon om alder, kjønn, utdanning og arbeidserfaring. Denne informasjonen blir brukt for å få en bedre oversikt over respondentene i utvalget, samt at det brukes som kontrollvariabler i analysen. Figur 5.1 viser grafisk hvordan utvalget på 64 respondenter fordelte seg.



Figur 5.1: Utvalget på 64 respondenter fordelt på demografiske variabler.

Alder: I utvalget på 64 personer er det en overvekt av unge personer hvor nesten halvparten er under 25 år. Dette har en naturlig sammenheng med antallet studenter i utvalget, og at de praktiserende revisorer som vi har nådd gjennom vårt nettverk i hovedsak er unge personer.

Kjønn: Det er en skjevhet i utvalget hvor nesten to tredeler er menn. Det er vanskelig å spekulere i hvorfor det er slik ettersom vi har markedsført undersøkelsen til det vi oppfatter som et jevnt antall av kvinner og menn.

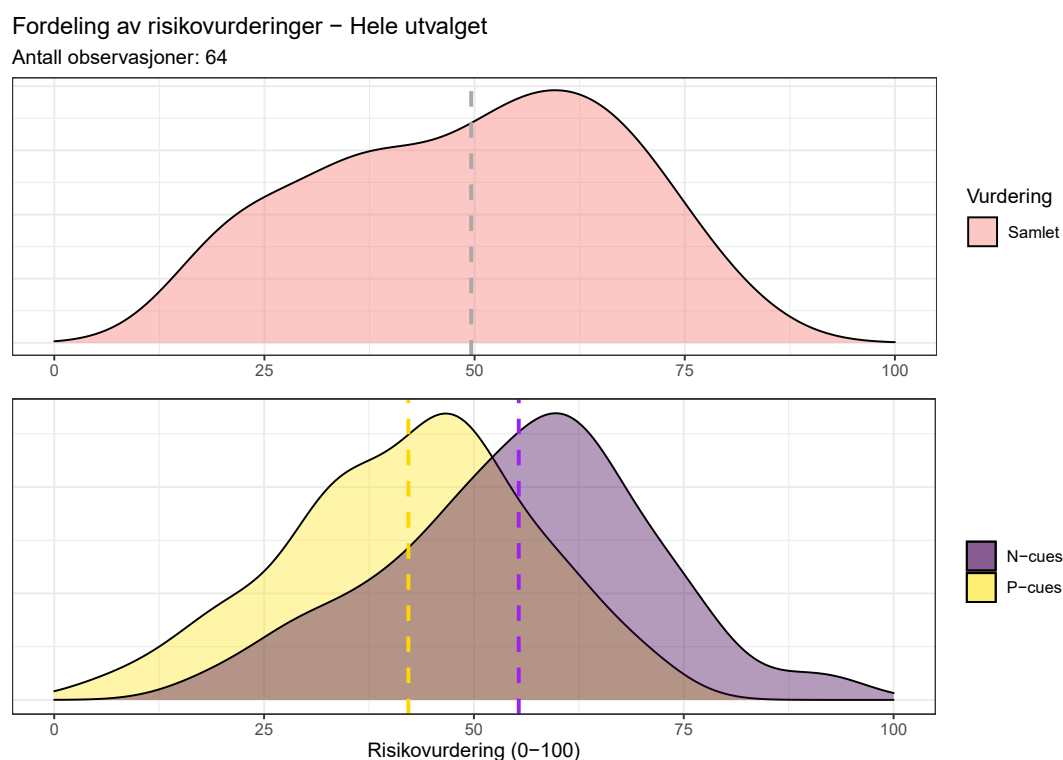
Revisjonsutdanning: Utvalget har en ganske jevn fordeling hvor ca halvparten har fullført en revisjonsgrad enten på master -eller bachelornivå. Av de som har fullført eller har påbegynt en grad, er det en klar overvekt på masternivå, noe som må anses som naturlig.

Arbeidserfaring: Fra oversikten ser vi at en drøy tredel av utvalget består av revisorstudenter uten arbeidserfaring. Det er nesten like mange som har under to års erfaring og resten er fordelt på ulike nivåer.

5.2 Deskriptiv statistikk

5.2.1 Utvalget under ett

Respondentene gjorde en risikovurdering for hver informasjons-cue i tillegg til en avsluttende og samlet totalvurdering. Det er sistnevnte som blir brukt til å teste hovedhypotesene $H_0^{Rekkefølge}$, $H_0^{Format^1}$ og $H_0^{Interaksjon}$. De individuelle vurderingene blir gruppert i positive og negative cues og viser hvordan respondentene har reagert på informasjon som peker mot hhv. lavere og høyere risiko. Dette brukes til å teste tilleggshypotesene $H_0^{Format^2}$ og $H_0^{Format^3}$, i tillegg til å fungere som validitetssjekk for å bekrefte at designet av undersøkelsen har fungert som tiltenkt.



Figur 5.2: Fordeling og gjennomsnitt samlet for hele utvalget. De skyggelagte områdene viser fordelingen, hvor større bredde betyr større variasjon. De stiplete linjene viser gjennomsnittsverdier for de ulike risikovurderingene.

Figur 5.2 viser hvordan de tre målene for risikovurdering fordeler seg for hele utvalget sett under ett. Utvalget har vurdert den samlede risikoen med et gjennomsnitt på 49.6 og et standardavvik på 17.6. For at designet skal fungere som tiltenkt er det nødvendig at risiko blir vurdert lavere for den positive informasjonen enn for den negative, og motsatt. Nedre

halvdel av Figur 5.2 viser tydelig forskjellen, hvor positive cues (*gjennomsnitt* = 42.2, *standardavvik* = 14.6) ligger lenger til venstre enn negative cues (*gjennomsnitt* = 55.3, *standardavvik* = 15.6). En t-test viser at dette er en signifikant forskjell på alle relevante nivåer; $t(125.4) = -4.9, p < .001$

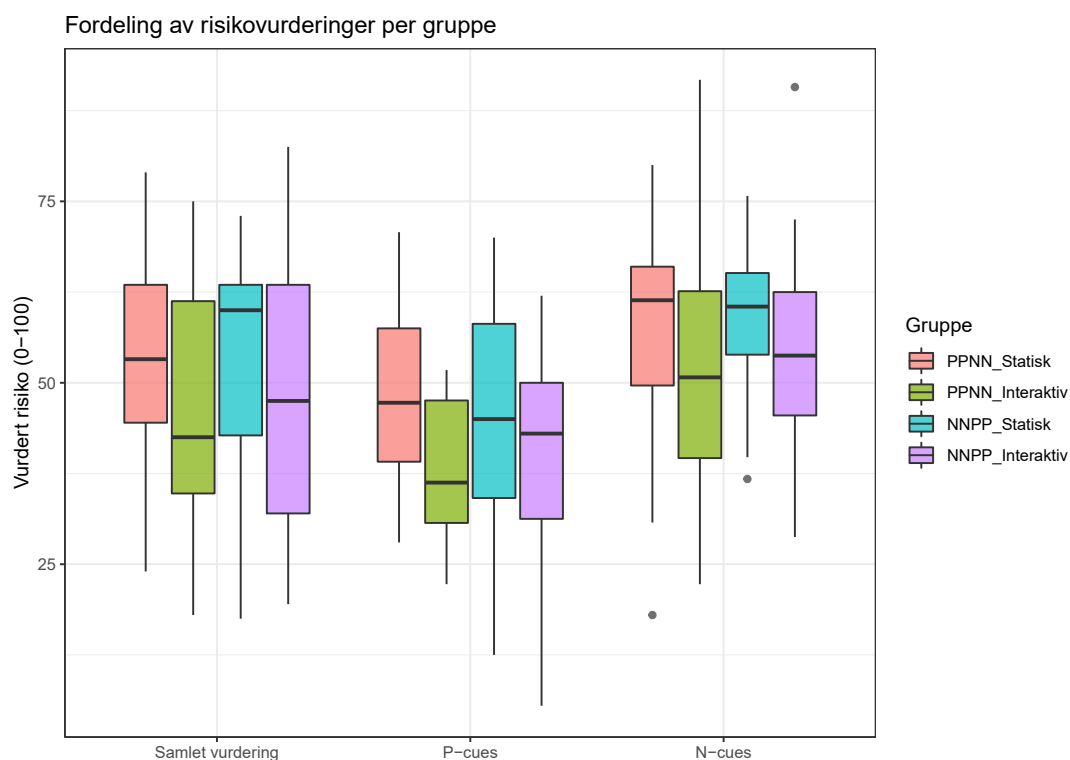
5.2.2 Fordelt per gruppe

Tabell 5.2 viser en oversikt over utvalgte deskriptive mål for gruppene risikovurderinger. Ut fra tallene er det vanskelig å se noen klar forskjell mellom de som har fått informasjonen i rekkefølgen PPNN (Gruppe 1 & 2), og de som har fått den i rekkefølgen NNPP (Gruppe 3 & 4). Derimot ser det ut som en trend i dataen hvor de som har sett statiske visualiseringer (Gruppe 1 & 3) generelt vurderer risiko noe høyere enn de som har sett interaktive visualiseringer (Gruppe 2 & 4). Dette visualiseres i Figur 5.3, hvor medianverdien (sort horisontal strek) til de statiske gruppene ligger noe høyere enn for de interaktive gruppene.

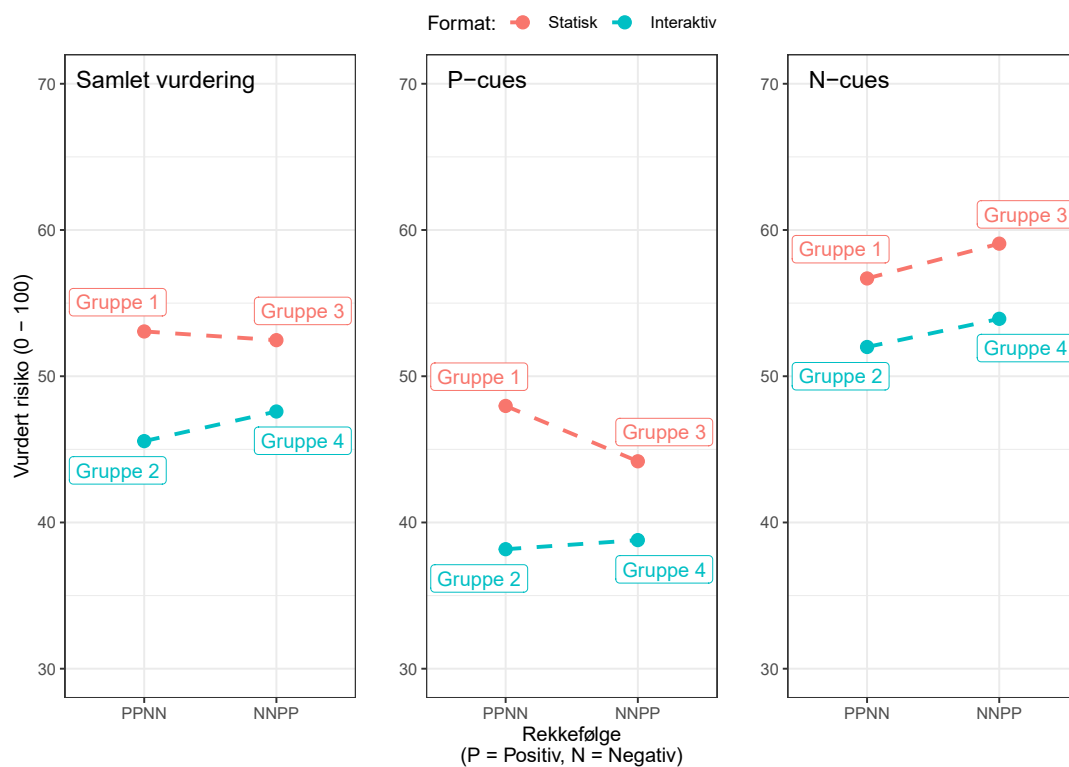
Gruppe	Rekkefølge	Format	Vurdering	Antall	Min	Max	Median	Gj.snitt	Std.avvik
Gruppe 1	PPNN	Statisk	Samlet	16	24.0	79.0	53.2	53.1	16.0
Gruppe 2	PPNN	Interaktiv	Samlet	16	18.0	75.0	42.5	45.6	18.6
Gruppe 3	NNPP	Statisk	Samlet	15	17.5	73.0	60.0	52.5	16.0
Gruppe 4	NNPP	Interaktiv	Samlet	17	19.5	82.5	47.5	47.6	19.7
Gruppe 1	PPNN	Statisk	P-cues	16	28.0	70.8	47.2	48.0	12.2
Gruppe 2	PPNN	Interaktiv	P-cues	16	22.2	51.8	36.2	38.2	10.0
Gruppe 3	NNPP	Statisk	P-cues	15	12.5	70.0	45.0	44.2	17.4
Gruppe 4	NNPP	Interaktiv	P-cues	17	5.5	62.0	43.0	38.8	16.5
Gruppe 1	PPNN	Statisk	N-cues	16	18.0	80.0	61.4	56.7	16.2
Gruppe 2	PPNN	Interaktiv	N-cues	16	22.2	91.8	50.8	52.0	19.0
Gruppe 3	NNPP	Statisk	N-cues	15	36.8	75.8	60.5	59.1	11.2
Gruppe 4	NNPP	Interaktiv	N-cues	17	28.8	90.8	53.8	53.9	15.4

Tabell 5.2: Deskriptiv statistikk som viser hvordan gruppene har vurdert den samlede risikoen, i tillegg til hvordan de har vurdert positive og negative cues individuelt.

En annen måte å visualisere dette på er å se isolert på gjennomsnittsverdiene til gruppene. Dette vises gjennom et profildiagram i Figur 5.4 hvor de statiske gruppene i utvalget har konsekvent vurdert risikoen høyere enn de interaktive gruppene. Når gjennomsnittet isoleres, blir det også lettere å identifisere mulige forskjeller i gruppene med ulik informasjonsrekkefølge, samt for eventuelle interaksjoner mellom rekkefølge og format. Fra grafen til venstre i Figur 5.4 er det vanskelig å se noen tydelige signaler på forskjeller mellom informasjonsekkefølgene for den samlede totalvurderingen. Det er heller ingen tydelige tegn på en interaksjonseffekt.



Figur 5.3: Boxplot som for hver gruppe viser minste verdi, første kvartil, median, tredje kvartil og største verdi. Figuren viser verdier for både den samlede risikovurderingen, i tillegg til individuelle vurderinger av positive og negative cues.



Figur 5.4: Profildiagram som viser hver gruppes gjennomsnittlige risikovurdering for både den samlede risikovurderingen, i tillegg til individuelle vurderinger av positive og negative cues.

5.3 Hypotesetesting

For å teste hvilken effekt informasjonsrekkefølge og presentasjonsformat har på revisors vurdering av iboende risiko, har det blitt gjennomført toveis variansanalyser. Disse er gjennomført både uten kontrollvariabler (ANOVA) og med kontrollvariabler (ANCOVA). Residualanalyser ble gjort for å avgjøre om forutsetningene for å gjennomføre slike variansanalyser var til stede. For alle variansanalyser ble uteliggere avdekket gjennom en boxplot-metode. Antakelsen om normalfordelte residualer ble begrunnet med Shapiro-Wilks test og antakelsen om homogenitet i variansen mellom gruppene ble begrunnet med Levenes test. Det ble ikke funnet noen relevante uteliggere, residualene var normalfordelte (se Tabell A3.1), og det var homogenitet i variansen (se Tabell A3.2).

5.3.1 Hovedhypoteser

Etter eksponering av fire cues ble respondentene bedt om å gjøre en samlet vurdering av iboende risiko, basert på all informasjon som de hadde fått se. Det er denne vurderingen som danner grunnlaget for å teste studiens tre hovedhypoteser om rekkefølgeeffekt, formateffekt og interaksjon mellom de to.

<u>Avhengig variabel: Samlet vurdering</u>									
	<i>Toveis ANOVA</i>					<i>Toveis ANCOVA</i>			
	Sum Sq	Df	F value	Pr(>F)	Sum Sq	Df	F value	Pr(>F)	
Rekkefølge	9.1	1	0.0290	0.8653	0.5	1	0.0014	0.9705	
Format	612.2	1	1.9445	0.1683	650.8	1	1.8378	0.1814	
Rekkefølge:Format	27.4	1	0.0871	0.7689	66.2	1	0.1869	0.6674	
Alder					477.0	1	1.3468	0.2515	
Kjønn					14.2	2	0.0201	0.9801	
Utdanning					865.6	4	0.6111	0.6566	
Erfaring					195.4	4	0.1379	0.9674	
Residualer	18,890.2	60			17,353.2	49			

*p<0.1; **p<0.05; ***p<0.01

Tabell 5.3: Anova-tabell (Type II tester) for den avsluttende og samlede risikovurderingen.

Rekkefølgeeffekt

Det ble ikke funnet en statistisk signifikant forskjell mellom gruppene som fikk ulik informasjonsrekkefølge, verken etter en ANOVA; $F(1, 60) = 0.03$, $p > .05$, eller en ANCOVA; $F(1, 49) = 0.00$, $p > .05$. Vi beholder derfor $H_0^{Rekkefølge}$.

Formateffekt

Det ble ikke funnet en statistisk signifikant forskjell mellom gruppene som fikk ulikt presentasjonsformat, verken etter en ANOVA; $F(1, 60) = 1.94$, $p > .05$, eller en ANCOVA; $F(1, 49) = 1.84$, $p > .05$. Vi beholder derfor $H_0^{Format^1}$.

Interaksjonseffekt

Det ble ikke funnet en statistisk signifikant interaksjon mellom rekkefølge og format, verken etter en ANOVA; $F(1, 60) = 0.09$, $p > .05$, eller en ANCOVA; $F(1, 49) = 0.19$, $p > .05$. Vi beholder derfor $H_0^{Interaksjon}$.

5.3.2 Tilleggshypoteser

Utover de tre hovedhypotesene ble det gjennomført tilsvarende variansanalyser for å teste tilleggshypotesene om formateffekt. Grunnlaget for disse testene var at de individuelle risikovurderingene ble gruppert til én verdi for positive cues og én verdi for negative cues.

Formateffekt for positive cues

En toveis variansanalyse resulterte i en hovedeffekt for presentasjonsformatet, hvor gruppene som fikk se statisk informasjon vurderte risiko for positive cues annerledes enn gruppene som fikk se interaktiv informasjon. Dette vises både gjennom en ANOVA; $F(1, 60) = 4.48$, $p < .05$, og en ANCOVA; $F(1, 49) = 6.16$, $p < .05$. Med en signifikant hovedeffekt ble det videre gjennomført parvise sammenligninger med Bonferroni-korreksjon av p-verdiene. Dette resulterte fortsatt i signifikante verdier på 5 % nivå og vi forkaster derfor $H_0^{Format^2}$.

<u>Avhengig variabel: P-cues</u>									
	<i>Toveis ANOVA</i>					<i>Toveis ANCOVA</i>			
	Sum Sq	Df	F value	Pr(>F)	Sum Sq	Df	F value	Pr(>F)	
Rekkefølge	36.6	1	0.1778	0.6748	58.4	1	0.2719	0.6044	
Format	921.7	1	4.4819	0.0384**	1,322.1	1	6.1599	0.0165**	
Rekkefølge:Format	77.6	1	0.3771	0.5415	3.8	1	0.0178	0.8944	
Alder					166.0	1	0.7732	0.3835	
Kjønn					124.9	2	0.2910	0.7488	
Utdanning					411.4	4	0.4792	0.7508	
Erfaring					904.9	4	1.0540	0.3893	
Residualer	12,339.2	60			10,517.2	49			

*p<0.1; **p<0.05; ***p<0.01

Tabell 5.4: Anova-tabell (Type II tester) for risikovurdering av positive cues.

Formateffekt for negative cues

Det ble ikke funnet en statistisk signifikant forskjell mellom gruppene som fikk ulikt presentasjonsformat, verken etter en ANOVA; $F(1, 60) = 1.55, p > .05$, eller en ANCOVA; $F(1, 49) = 1.06, p > .05$. Vi beholder derfor $H_0^{Format^3}$.

<u>Avhengig variabel: N-cues</u>								
	<i>Toveis ANOVA</i>				<i>Toveis ANCOVA</i>			
	Sum Sq	Df	F value	Pr(>F)	Sum Sq	Df	F value	Pr(>F)
Rekkefølge	73.6	1	0.2962	0.5883	39.2	1	0.1357	0.7142
Format	385.5	1	1.5517	0.2177	305.4	1	1.0562	0.3091
Rekkefølge:Format	0.8	1	0.0033	0.9544	18.9	1	0.0654	0.7993
Alder					7.1	1	0.0246	0.8759
Kjønn					29.4	2	0.0508	0.9505
Utdanning					420.6	4	0.3637	0.8333
Erfaring					253.7	4	0.2194	0.9264
Residualer	14906.7	60			14167.5	49		

*p<0.1; **p<0.05; ***p<0.01

Tabell 5.5: Anova-tabell (Type II tester) for risikovurdering av negative cues.

6 Diskusjon og konklusjon

For å besvare forskningsspørsmålet skal vi nå drøfte resultatene i lys av studiens tre hovedhypoteser og to tilleggshypoteser. Det vil også redegjøres for bl.a. studiens begrensninger og betydningen for praksis.

6.1 Informasjonsrekkefølge og profesjonelt skjønn

Formålet med å teste hypotesen $H_0^{Rekkefølge}$ var å avdekke hvorvidt informasjonsrekkefølge påvirket revisors vurdering av iboende risiko. Med bakgrunn i studien til [Hogarth og Einhorn \(1992\)](#) forventet vi å finne en primacy effect, altså at den informasjonen som gis først, blir styrende for den totale vurderingen. Modellen har mye støtte i litteraturen, men det er likevel motstridende signaler fra studier som tar utgangspunkt i modellen ([Trotman og Wright, 2000](#)).

Resultatene våre viser ingen tegn til primacy effect, når fire informasjons-cues med blandede signaler danner grunnlaget for en vurdering av iboende risiko. Det er ingen signifikant forskjell i noen retninger, verken primacy effect eller recency effect. Fra middelverdiene i den deskriptive statistikken er det heller ingen tegn til at rekkefølgen av cues har påvirket respondentene i utvalget vårt. Når resultatene er så fraværende, skal vi være forsiktige med å spekulere for mye i årsaker, men vi tillater oss likevel å drøfte et par punkter knyttet til: (1) oppgaveegenskaper, (2) styrken av signalene, og (3) utforming av cues.

Oppgaveegenskaper, altså ulike karakteristika ved en konkret oppgave, kan bidra til å moderere rekkefølgeeffekten ([Trotman og Wright, 2000](#)). Studien vår skiller seg fra lignende studier ved at respondentene gjør en simultan vurdering av den iboende risikoen. Altså presenteres de først for fire cues og gjør deretter en samlet vurdering. Det kan argumenteres for at det er vanskeligere å identifisere en rekkefølgeeffekt gjennom simultane vurderinger ([Trotman og Wright, 2000](#); [Yankova, 2015](#)). Andre studier har undersøkt vurderinger av iboende risiko sekvensielt, hvor respondentene oppdaterer sin samlede vurdering etter hver cue. Studien til [Monroe og Ng \(2000\)](#) finner da ingen rekkefølgeeffekt, men studien til [Anderson og Maletta \(1999\)](#) finner primacy effect.

Videre hevdes det at rekkefølgeeffekten kan modereres ved vurdering av iboende risiko, da dette er en oppgave som er svært betydningsfull for den videre revisjonen (Monroe og Ng, 2000). Ved slike vurderinger kan revisor være mer bevisst på å opprettholde sitt profesjonelle skjønn da feilaktige vurderinger kan få store konsekvenser (Smith og Kida, 1991). Dette er faktorer som kan være med på å redusere en eventuell rekkefølgeeffekt, og vi stiller derfor spørsmål ved om oppgaveegenskapene i vårt eksperiment kan ha påvirket denne.

Signalstyrke er en annen faktor som muligens kan påvirke rekkefølgeeffekten. Nedre halvdel av Figur 5.2 viser en signifikant forskjell mellom de ulike signaltypene, hvor lavrisiko-cues og høyrisiko-cues i gjennomsnitt ble vurdert til hhv. 42.2 og 55.3 på en risikoskala fra 0-100. Tidligere studier som undersøker rekkefølgeeffekt i kontekst av risikovurderinger skiller som regel mellom høyrisiko og lavrisiko-cues uten å spesifisere styrken av signalene (Ahlawat, 1999; Monroe og Ng, 2000). For å unngå at en cue blir overstyrende i den totale risikovurderingen har vi bevisst forsøkt å designe cues med en moderat signalstyrke. Vi stiller spørsmål ved om dette kan ha vært med på å moderere rekkefølgeeffekten. Altså om større forskjell mellom lavrisiko-cues og høyrisiko-cues vil gi andre resultater. Vi finner dog ingen støtte for dette i litteraturen og poengterer at det i høyeste grad er spekulativt.

En annen faktor som kan være minst like relevant er hvordan de fire webapplikasjonene er designet. Det er ulike visualiseringsteknikker som benyttes i hver cue, og det kan stilles spørsmål ved om dette er en faktor som er med på å viske ut eventuelle effekter. I tillegg til forskjellige visualiseringsteknikker er det også variasjon i hvilke risikofaktorer som respondentene skal vurdere. Dette er potensielt støyende faktorer som vi ikke har kunnet kontrollere for.

Med utgangspunkt i resultatene fra Kapittel 5 og drøftelsen ovenfor, finner vi ikke noe grunnlag til å argumentere for at informasjonsrekkefølge har påvirket revisors profesjonelle skjønn ved vurdering av iboende risiko.

6.2 Presentasjonsformat og profesjonelt skjønn

Formålet med å teste hypotesen $H_0^{Format^1}$ var å avdekke hvorvidt revisors vurdering av iboende risiko ble påvirket av hvilket presentasjonsformat informasjon ble presentert i. Vi testet to ulike formattyper; interaktivt og statisk. Modellen til [Kelton et al. \(2010\)](#) (se Figur 2.2) predikerer at presentasjonsformat vil ha en direkte påvirkning på revisors beslutninger, og vi forventet derfor å se en effekt.

Resultatene våre viser at gruppene som fikk cues i et statisk format i gjennomsnitt vurderte den totale risikoen noe høyere enn gruppene som fikk informasjonen i et interaktivt format. Variansanalysene viser dog at dette ikke var en signifikant forskjell for den samlede totalvurderingen, og vi skal være forsiktige med å spekulere for mye i eventuelle årsaker. Vi ønsker likevel å nevne et par punkter som kan være relevante.

På lik linje med rekkefølgeeffekten, kan miksen av ulike visualiseringsteknikker og risikofaktorer trekkes inn som en mulig årsak til manglende formateffekt. I tillegg er det variasjon i hvordan de interaktive funksjonalitetene er utformet på tvers av de ulike webapplikasjonene. Slike faktorer kan være med å påvirke resultatene.

Tidligere litteratur som støtter seg på modellen til [Kelton et al.](#) viser enighet i at format kan påvirke beslutninger, men at oppgaveegenskaper kan moderere en slik formateffekt ([Dilla et al., 2010](#); [Alawadhi, 2015](#)). Som nevnt i diskusjonen for rekkefølgeeffekt, er risikovurderinger en viktig oppgave i revisjon hvor konsekvenser av eventuelle feilvurderinger kan være av stor betydning. Dette kan ha bidratt til å skjerpe revisors atferd og muligens visket ut en potensiell formateffekt. Videre hevdes det at forskningen på presentasjonsformat er preget av en «*prøve-og-feile*»-metodikk ([Bonner, 2008](#), s.193). Manglende teori og mangfoldet av ulike presentasjonsformater gjør det vanskelig å generalisere tidligere funn, som skaper utfordringer til hvilke formater å teste og hvilken sammenheng formatene bør testes i.

Med utgangspunkt i resultatene fra Kapittel 5 og drøftelsen ovenfor, finner vi ikke noe grunnlag til å argumentere for at presentasjonsformat har påvirket revisors profesjonelle skjønn ved en samlet vurdering av iboende risiko.

6.3 Interaksjon mellom rekkefølge og format

Formålet med å teste hypotesen $H_0^{Interaksjon}$ var å avdekke hvorvidt ulike presentasjonsformater påvirket rekkefølgeeffekten. Resultatene viser ingen signifikant rekkefølge- eller formateffekt for den totale risikovurderingen, og det vil heller ikke være interaksjoner mellom disse som er aktuelt å drøfte videre.

6.4 Andre funn

6.4.1 Formateffekt for spesifikke signaltyper

Utover studiens tre hovedhypoteser testet vi også to tilleggshypoteser som er knyttet til presentasjonsformatet. Hypotesen $H_0^{Format^2}$ ble testet for å undersøke om presentasjonsformatet påvirket revisor ved vurdering av *positive cues*, og hypotesen $H_0^{Format^3}$ testet det samme for *negative cues*. For de to tilleggshypotesene forventet vi også å finne en effekt som predikert av modellen til [Kelton et al. \(2010\)](#).

Resultatene viser en signifikant formateffekt ved risikovurdering av positive cues, både med og uten kontrollvariabler. Derimot er det ikke signifikant formateffekt ved risikovurdering av negative cues. Figur 5.3 tegner et bilde av variasjonen innad i de ulike gruppene. Her vises det et klart mønster mellom de statiske og interaktive gruppene ved risikovurderingen av positive cues. Dette skillet vises også tydelig i Figur 5.4, der de statiske gruppene vurderer risikoen høyere enn de interaktive. Et spørsmål som reiser seg fra diagrammene er hvorfor *P-cues* gir signifikant forskjell, mens *N-cues* ikke gjør det.

Studier som har undersøkt lignende risikovurderinger argumenterer for at revisor er mer sensitiv til høyrisikoinformasjon, sammenlignet med lavrisikoinformasjon ([Reckers og Schultz, 1993](#); [Monroe og Ng, 2000](#)). Dette kan tyde på en konservativ atferd, hvor revisor stiller seg mer kritisk og er mer skjerpet, når høyrisikoinformasjon skal vurderes ([Yankova, 2015](#), s.79). Spørsmålet vi da stiller er om revisor blir mindre skjerpet ved vurdering av positive cues, og om dette er en årsak til at vi her finner en formateffekt.

En annen mulighet er at resultatene skyldes tilfeldigheter. Som vi har vært inne på i tidligere drøftelser har webapplikasjonene blitt designet med en miks av ulike

visualiseringsteknikker og risikofaktorer. Dette, sammen med et lite utvalg som ikke nødvendigvis er representativt, gjør at vi skal være forsiktige med å trekke sterke slutninger ut fra resultatene. Med bakgrunn i resultatene fra Kapittel 5 og drøftelsen ovenfor så konkluderer vi på følgende måte: Det er signaler som peker mot at presentasjonsformat har vært en påvirkende faktor, når respondentene i vår studie vurderte iboende risiko. Det er dog ikke sterke nok signaler til at vi kan generalisere dette.

6.4.2 Alternativ tolkning av avhengig variabel

Ordlyden i ISA 315 sier at det er *skjæringspunktet* mellom sannsynlighet og konsekvens som utgjør den relevante iboende risikoen. I denne studien har vi operasjonalisert dette som gjennomsnittet av sannsynlighet og konsekvens. En alternativ tolkning som kan være vel så relevant, er å ta utgangspunkt i risikomodellen *Loss Given Default* (LGD) som brukes mye i bank og finans. I vårt design vil «loss» være synonymt med konsekvens, og det vil være produktet av sannsynlighet og konsekvens som utgjør den relevante risikoen. Dette betyr i praksis at dersom det er helt sikkert feil (100 på skalaen), men feilen har ingen konsekvens for regnskapet (0 på skalaen), så vil den relevante iboende risikoen være lik null.

Vi gjennomførte de samme variansanalysene med denne tolkningen og kommer fram til samme konklusjon som ovenfor. Det er kun positive cues som gir signifikante resultater. Ettersom vi får de samme resultatene har vi ikke viet mer plass til denne alternative tolkningen, men algebraisk definisjon av variabler og resultater fra variansanalysene finnes i Appendiks A4.

6.5 Betydning for revisjonspraksis

Tidligere studier viser at revisjonsbransjen er treg med å tilpasse seg nye og innovative teknikker. Denne studien retter fokuset mot bruken av ADA, og spesifikt hvordan bruk og visualisering av Big Data kan brukes i revisjonen. Den tar i bruk innovative teknikker for å presentere både finansiell og ikke-finansiell informasjon i en revisjonskontekst. Disse teknikkene kan forhåpentligvis være til inspirasjon når morgendagens revisjonsmetodikk skal utformes.

Studien finner noen signaler som peker mot at revisor vurderer iboende risiko lavere dersom visualiseringer presenteres med et interaktivt format. Hvis dette er en realitet, er det noe som praktiserende revisorer bør være bevisst på, ettersom det kan utnyttes for å manipulere revisor til å vurdere risiko i en gitt retning.

6.6 Begrensninger

Deltakerne i denne studien er plassert i en konstruert situasjon, som til en viss grad vil avvike fra virkeligheten. Dette er en begrensende faktor ved gjennomføring av eksperimenter (Trochim et al., 2016, s.252). Deltakerne opplever ikke nødvendigvis det samme presset ved en fiktiv risikovurdering, som de ville gjort i en reell situasjon. I vårt eksperiment vil grove feilvurderinger ikke føre til noen konsekvenser for deltakerne, noe som kan bidra til å svekke reliabiliteten av resultatene.

I virkeligheten er risikovurderingene i planleggingsfasen av stor betydning, da det er styrende for videre revisjonshandlinger. Det er i praksis et omfattende arbeid og eksperimentet vårt tar for seg en forenklet risikovurdering for å redusere omfanget. I tillegg gis det minimalt med tilleggsinformasjon for å redusere bias. Forenklingen av oppgaven og lav informasjonsmengde bidrar til å svekke sammenligningsstyrken mellom eksperimentet og virkeligheten.

For rekkefølgeeffekten, er det essensielt at informasjonen blir presentert sekvensielt før det gjøres en vurdering. I undersøkelsen er det hyperlinker som sender deltakerne til én webapplikasjon om gangen. Det blir gitt beskjed om å lukke de enkelte applikasjonene når de har prosessert informasjonen før de går videre til neste. Dette er ikke noe vi har kunnet kontrollere for ved gjennomføringen av undersøkelsen. Risikoen for at respondentene ikke krysser ut webapplikasjonene og bruker tidligere cues mens de gjør den samlede risikovurderingen kan moderere rekkefølgeeffekten som de utsettes for.

En annen begrensning i studien er utvalget, som i stor grad består av revisorer og revisorstudenter fra vårt eget nettverk. Gruppene i eksperimentet har en gjennomsnittlig størrelse på 16 deltakere, noe som må anses å være lavt. Med et lite utvalg blir sannsynligheten for representativitet svekket, i tillegg vil unormale observasjoner få større påvirkningskraft på resultatene. Bruken av bekvemmelighetsutvalg supplert med

revisjonsstudenter kan også svekke representativiteten til utvalget. Med bakgrunn i dette skal vi være forsiktige med å generalisere resultatene fra studien til å gjelde for hele populasjonen av revisorer.

6.7 Forslag til videre forskning

Først og fremst må studiens utvalg anses som en svakhet, og det kan derfor være interessant å gjennomføre en lignende studie med et større utvalg av praktiserende revisorer. Utvalget bør da trekkes på en hensiktsmessig måte for å øke sannsynligheten for representativitet. Det kan tenkes at en slik studie vil gi andre resultater enn det vi har funnet.

Videre har vår studie handlet om beslutningsvariasjon, og vi kan ikke si noe om hva «riktig» utøvelse av profesjonelt skjønn innebærer. Vi identifiserer avvik og variasjon i det profesjonelle skjønnnet, men kan ikke vise til hvordan de aktuelle variablene skal manipuleres for å gi bedre beslutninger. Å bygge videre på denne studien ved å måle revisors beslutningskvalitet kan derfor være av interesse. Ved å inkludere kriteriet «*sann*» verdi av *iboende risiko* i linsemodellen fra [Frenkel-Brunswik \(1952\)](#) (se Figur 2.1), kan studien utvides. Avhengig variabel kan måles på samme måte med sannsynlighet og konsekvens, men i stedet for å teste avvik mellom gruppene, rettes fokuset mot hvorvidt gruppene avviker fra et fasitsvar. Dette kan testes på flere måter, eksempelvis ved å måle konsensus innad i en gruppe og/eller konsensus med en ekspertgruppe ([Bonner, 2008](#), s.29).

Resultatene våre viser ingen tegn til at informasjonsrekkefølge påvirker vurderingen av iboende risiko. Det er dog noen signaler som peker i retning av at presentasjonsformat kan ha en effekt, og det kan være interessant å se videre på dette. En ny studie kan eksempelvis fokusere utelukkende på statiske vs. interaktive presentasjonsformater og hvordan det påvirker revisors vurderinger.

Referanser

- Ahlawat, S. S. (1999). Order effects and memory for evidence in individual versus group decision making in auditing. *Journal of Behavioral Decision Making*, 12(1):71–88.
- AICPA (2018). Guide to Audit Data Analytics an overview. Hentet fra: <https://www.aicpa.org/resources/article/guide-to-audit-data-analytics-an-overview>. Dato hentet: 26.05.2022.
- Alawadhi, A. (2015). *The application of data visualization in auditing*. PhD thesis, Rutgers University.
- Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons*, 29(2):439–449.
- Anderson, B. H. og Maletta, M. J. (1999). Primacy effects and the role of risk in auditor belief-revision processes. *Auditing: A Journal of Practice & Theory*, 18(1):75–89.
- Asare, S. K. (1992). The auditor's going-concern decision: Interaction of task variables and the sequential processing of evidence. *The Accounting Review*, 67(2):379–393.
- Ashton, A. H. og Ashton, R. H. (1988). Sequential belief revision in auditing. *Accounting Review*, sider 623–641.
- Ashton, R. H. og Ashton, A. H. (1995). *Judgment and decision-making research in accounting and auditing*. Cambridge University Press.
- Barnes, J., Touileb, S., Øvreid, L., og Velldal, E. (2019). Lexicon information in neural sentiment analysis: a multi-task learning approach. I *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, Turku, Finland.
- Bell, T. B., Peecher, M. E., og Solomon, I. (2005). *The 21st Century Public Company Audit - Conceptual Elements of KPMG's Global Audit Methodology*. KPMG International.
- Bonner, S. E. (2008). *Judgment and decision making in accounting*. Pearson Prentice Hall, Upper Saddle River, N.J.
- Brown-Liburud, H., Issa, H., og Lombardi, D. (2015). Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions. *Accounting Horizons*, 29(2):451–468.
- Brown-Liburud, H. og Vasarhelyi, M. A. (2015). Big Data and audit evidence. *Journal of Emerging Technologies in Accounting*, 12(1):1–16.
- Campbell, D. T. og Riecken, H. (1968). Quasi-experimental design. *International encyclopedia of the social sciences*, 5(3):259–263.
- Chung, J. og Monroe, G. S. (2001). A research note on the effects of gender and task complexity on an audit judgment. *Behavioral Research in Accounting*, 13(1):111–125.
- De Mauro, A., Greco, M., og Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*.
- Deloitte (2021). Åpenhetsrapport 2021. Hentet fra: <https://nor.deloitte.com/rs/712-CNF-326/images/Apenhetsrapporten-2021.pdf>. Dato hentet: 01.02.22.

- Diebold, F. X. (2012). On the origin (s) and development of the term “Big Data”. *PIER Working Paper*.
- Dilla, W., Janvrin, D. J., og Raschke, R. (2010). Interactive data visualization: New directions for accounting information systems research. *Journal of Information Systems*, 24(2):1–37.
- Eilifsen, A., Kinserdal, F., Messier, W. F., og McKee, T. E. (2020). An exploratory study into the use of audit data analytics on audit engagements. *Accounting Horizons*, 34(4):75–103.
- EY (2021). Åpenhetsrapport 2021. Hentet fra: https://assets.ey.com/content/dam/ey-sites/ey-com/no_no/home-index/ey-norge-aapenhetsrapport-2021.pdf. Dato hentet: 01.02.22.
- Frenkel-Brunswik, E. (1952). Interaction of psychological and sociological factors in political behavior. *American Political Science Review*, 46(1):44–65.
- Fry, H. (2018). *Hello World: How to be Human in the Age of the Machine*. Random House.
- Gandomi, A. og Haider, M. (2015). Beyond the hype: Big Data concepts, methods, and analytics. *International journal of information management*, 35(2):137–144.
- Gartner (2022). Information technology glossary. Hentet fra: <https://www.gartner.com/en/information-technology/glossary/big-data>. Dato hentet: 10.02.22.
- Gibbins, M. (1984). Propositions about the psychology of professional judgment in public accounting. *Journal of Accounting Research*, sider 103–125.
- Goodwin, M. (2020). *AI: Myten om maskinene*. Humanist forlag.
- Grenness, T. (2004). *Hvordan kan du vite om noe er sant? : veiviser i forsknings- og utredningsarbeid for studenter, ledere, konsulenter og journalister*. Cappelen akademisk forl.
- Gulden, B. P. (2016). *Revisjon - teori og metode*. Cappelen Damm.
- Hammond, K. R., Hursch, C. J., og Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review*, 71(6):438 – 456.
- Hodge, F. D. (2001). Hyperlinking unaudited information to audited financial statements: Effects on investor judgments. *The Accounting Review*, 76(4):675–691.
- Hogarth, R. M. og Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24(1):1–55.
- Huang, A. H. (2003). Effects of multimedia on document browsing and navigation: an exploratory empirical investigation. *Information & management*, 41(2):189–198.
- Huang, A. H. og Windsor, J. C. (1998). An empirical assessment of a multimedia executive support system. *Information & Management*, 33(5):251–262.
- Hursch, C. J., Hammond, K. R., og Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological review*, 71(1):42.
- Hyndman, R. og Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*. OTexts, Australia, 3rd edition.

- IAASB (2009). International Standard on Auditing 200. Hentet fra: <https://www.ifac.org/system/files/downloads/a008-2010-iaasb-handbook-isa-200.pdf>. Dato hentet: 02.02.22.
- IAASB (2013). A framework for audit quality. Hentet fra: <https://www.ifac.org/system/files/publications/files/A%20Framework%20for%20Audit%20Quality.pdf>. Dato hentet: 02.02.22.
- IAASB (2019). International Standard on Auditing 315 (revised 2019). Hentet fra: <https://www.ifac.org/system/files/publications/files/ISA-315-Full-Standard-and-Conforming-Amendments-2019-.pdf>. Dato hentet: 09.05.22.
- ICAEW (ud). Revised ISA 315 for 2022 audits. Hentet fra: <https://www.icaew.com/technical/audit-and-assurance/audit/risk-assessment-internal-control-and-response/5-revised-isa-315-for-2022-audits>. Dato hentet: 19.04.22.
- IDEA (2020). IAASB discusses challenges with the use and adoption of data analytics. Hentet fra: <https://idea.caseware.com/blog/iaasb-discusses-challenges-with-the-use-and-adoption-of-data-analytics>. Dato hentet: 23.05.2022.
- IESBA (2018). Handbook of the international code of ethics for professional accountants. Hentet fra: <https://www.ifac.org/system/files/publications/files/IESBA-Handbook-Code-of-Ethics-2018.pdf>. Dato hentet: 19.04.22.
- Iliinsky, N. og Steele, J. (2011). *Designing data visualizations: Representing informational Relationships*. "O'Reilly Media, Inc."
- IMPACT, K. (2020). The KPMG survey of sustainability reporting 2020. Hentet fra: <https://assets.kpmg/content/dam/kpmg/xx/pdf/2020/11/the-time-has-come.pdf>. Dato hentet: 16.05.22.
- Insights, F. og KPMG (2020). A focus on change. Hentet fra: <https://images.forbes.com/forbesinsights/StudyPDFs/KPMG-AFocusOnChange-REPORT.pdf>. Dato hentet: 24.02.22.
- Janvrin, D. J. og Watson, M. W. (2017). "Big Data": A new twist to accounting. *Journal of Accounting Education*, 38:3–8.
- Johannesen, A. og Christoffersen, L. (2016). *Introduksjon til Samfunnsvitenskapelig Metode*. Abstrakt forlag.
- Johannesen, A., Christoffersen, L., og Tufte, P. A. (2020). *Forskningsmetode for Økonomisk-administrative Fag*. Abstrakt forlag.
- Kelton, A. S., Pennington, R. R., og Tuttle, B. M. (2010). The effects of information presentation format on judgment and decision making: A review of the information systems research. *Journal of Information Systems*, 24(2):79–105.
- KPMG (2021). Åpenhetsrapport 2021. Hentet fra: <https://home.kpmg/content/dam/kpmg/no/pdf/2022/01/Aopenhetsrapport2021.pdf>. Dato hentet: 01.02.22.
- Kumar, N. og Benbasat, I. (2004). The effect of relationship encoding, task type, and complexity on information representation: An empirical evaluation of 2d and 3d line graphs. *Mis Quarterly*, sider 255–281.

- Laney, D. et al. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70):1.
- Libby, R. (1981). *Accounting and human information processing : theory and applications*. Englewood Cliffs, N.J. : Prentice-Hall, ©1981.
- Libby, R., Bloomfield, R., og Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, organizations and society*, 27(8):775–810.
- Libby, T. og Thorne, L. (2020). *The Routledge Companion to Behavioural Accounting Research*. Routledge.
- Lurie, N. H. og Mason, C. H. (2007). Visual representation: Implications for decision making. *Journal of marketing*, 71(1):160–177.
- Lærd, S. (2022). Two-way ANOVA in SPSS Statistics. Hentet fra: <https://statistics.laerd.com/spss-tutorials/two-way-anova-using-spss-statistics.php>. Dato hentet: 12.05.2022.
- Majid, A., Gul, F. A., og Tsui, J. S. (2001). An analysis of hong kong auditors' perceptions of the importance of selected red flag factors in risk assessment. *Journal of Business Ethics*, 32(3):263–274.
- Mayer-Schönberger, V. og Cukier, K. (2013). *Big Data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Messier, W. F., Glover, S. M., og Prawitt, D. F. (2006). *Auditing & assurance services: A systematic approach 4th edition*. McGraw-Hill Irwin New York, NY.
- Messier Jr, W. F. og Tubbs, R. M. (1994). Recency effects in belief revision: The impact of audit experience and the review process. *Auditing*, 13(1):57.
- Monroe, G. og Ng, J. (2000). An examination of order effects in auditors' inherent risk assessments. *Accounting & Finance*, 40(2):153–167.
- Montgomery, D. (2012). *Design and Analysis of Experiments, 8th Edition*. John Wiley & Sons, Incorporated.
- NAOB (2022a). Flytskjema. Hentet fra: <https://naob.no/ordbok/ordsky>. Dato hentet: 03.05.2022.
- NAOB (2022b). Ordsky. Hentet fra: <https://naob.no/ordbok/ordsky>. Dato hentet: 03.05.2022.
- Olsen, A. B. (2020). Ny oppdatert ISA 315. Hentet fra: <https://www.nkrf.no/nyheter/2020/06/30/ny-oppdateret-isa-315>. Dato hentet: 19.04.22.
- Pedersen, J. S. og Wilkinson, A. (2019). *Big Data : promise, application and pitfalls*. Edward Elgar Publishing.
- Peecher, M. E. og Solomon, I. (2001). Theory and experimentation in studies of audit judgments and decisions: Avoiding common research traps. *International Journal of Auditing*, 5(3):193–203.
- Perkhofer, L. M., Hofer, P., Walchshofer, C., Plank, T., og Jetter, H.-C. (2019). Interactive visualization of Big Data in the field of accounting: A survey of current practice and potential barriers for adoption. *Journal of Applied Accounting Research*.

- PwC (2013). Breaking through Big Data's barriers. Hentet fra: <https://www.pwc.com/gx/en/services/advisory/assets/digital-iq-big-data.pdf>. Dato hentet: 24.02.22.
- PwC (2021). Åpenhetsrapport 2021. Hentet fra: <https://www.pwc.no/no/om-oss/apenhetsrapporten-2021/aapenhetsrapport-2021.pdf>. Dato hentet: 01.02.22.
- Reckers, P. M. J. og Schultz, J. J. J. (1993). The effects of fraud signals, evidence order, and group-assisted counsel on independent auditor judgment. *Behavioral research in accounting*, 5:124.
- Revisorloven (2021). Lov om revisjon og revisorer (LOV-2020-11-20-128). <https://lovdata.no/dokument/NL/lov/2020-11-20-128?q=revisorloven>.
- Russom, P. et al. (2011). Big Data analytics. *TDWI best practices report, fourth quarter*, 19(4):1–34.
- Saunders, M., Lewis, P., og Thornhill, A. (2019). *Research Methods for Business Students*. Pearson.
- Savolainen, R. (2009). Interpreting informational cues: An explorative study on information use among prospective homebuyers. *Journal of the American Society for Information Science and Technology*, 60(11):2244–2254.
- Schuff, D., Turetken, O., og D'Arcy, J. (2006). A multi-attribute, multi-weight clustering approach to managing “e-mail overload”. *Decision Support Systems*, 42(3):1350–1365.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of personality and social psychology*, 51(3):515.
- Shaft, T. M. og Vessey, I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *Mis Quarterly*, sider 29–55.
- Shen, Y.-C. og Hue, C.-W. (2007). The role of information presentation formats in belief updating. *International Journal of Psychology*, 42(3):189–199.
- Sirkin, R. M. (2006). *Statistics for the social sciences*. Sage.
- Smith, J. F. og Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological bulletin*, 109(3):472.
- Store norske leksikon (2022). Ekspertsystem. Hentet fra: <https://snl.no/ekspertsystem>. Dato hentet: 11.02.22.
- Taherdoost, H. (2016). Sampling methods in research methodology; how to choose a sampling technique for research. *How to Choose a Sampling Technique for Research (April 10, 2016)*.
- Trochim, W. M., Kanika, A., og Donnelly, J. P. (2016). *Research Methods: The Essential Knowledge Base*. Cengage Learning.
- Trotman, K. T. (2001). Design issues in audit JDM experiments. *International Journal of Auditing*, 5(3):181–192.
- Trotman, K. T. og Wright, A. (1996). Recency effects: Task complexity, decision mode, and task-specific experience. *Behavioral Research in Accounting*, 8:175 – 193.

- Trotman, K. T. og Wright, A. (2000). Order effects and recency: where do we go from here? *Accounting & Finance*, 40(2):169 – 182.
- Vasarhelyi, M. A., Kogan, A., og Tuttle, B. M. (2015). Big Data in accounting: An overview. *Accounting Horizons*, 29(2):381–396.
- Ward, J. S. og Barker, A. (2013). Undefined by data: a survey of Big Data definitions. *arXiv preprint arXiv:1309.5821*.
- Wheeler, P. R. og Arunachalam, V. (2008). The effects of decision aid design on the information search strategies and confirmation bias of tax professionals. *Behavioral Research in Accounting*, 20(1):131–145.
- Yankova, K. (2015). *The Influence of Information Order Effects and Trait Professional Skepticism on Auditors' Belief Revisions*. Springer.
- Yi, J. S., ah Kang, Y., Stasko, J., og Jacko, J. A. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics*, 13(6):1224–1231.

Appendiks

A1 Spørreundersøkelse

Figur A1.1: Skjermdump av spørreundersøkelse: Innledning 1

NHH



Denne undersøkelsen er en del av vår masteroppgave ved NHH om digital revisjon.

Den handler om bruk og visualisering av *Big Data*, og gjennomføres på følgende måte:

Du får presentert fire ulike nettsider som inneholder datavisualiseringer og tilleggsinformasjon. Du bes om å gjøre en risikovurdering for hver enkelt nettside, i tillegg til en avsluttende totalvurdering. Det er ingen tidsbegrensning, men det er tiltenkt at du bruker ca 2 minutter på hver nettside/vurdering. Undersøkelsen vil i så fall ta i underkant av 10 minutter.

Vi ber deg om å lese all informasjon nøye før du besvarer spørsmålene og går videre. Det er ikke mulig å navigere tilbake til tidligere spørsmål.

- Undersøkelsen er ikke optimalisert for mobilskjermer og bør helst gjennomføres på PC eller nettbrett.
- Undersøkelsen skal gjennomføres alene, og det er viktig at informasjon vedrørende spørreundersøkelsen ikke spres videre.
- Undersøkelsen er anonym, og data som innhentes vil kun brukes i sammenheng med oppgaven.

Vi setter stor pris på at du tar deg tid til å bidra til masteroppgaven vår!

Mvh

Lars Michelsen og Johnny Vu Tran

Ønsker du å delta på spørreundersøkelsen?

Ja

Nei



Figur A1.2: Skjermdump av spørreundersøkelse: Innledning 2

NHH



Du jobber for revisjonsselskapet Revisjon AS og dere har fått en ny klient - Produksjonsbedriften AS. Du er i planleggingsfasen og skal bli kjent med selskapet som driver med produksjon og salg til industrien. De har én spesifikk produktlinje som utgjør en vesentlig andel av deres omsetning. Oppgaven din blir å gjøre en vurdering av den iboende risikoen som skriver seg fra aktiviteter knyttet til denne produktlinjen.

Iboende risiko kan defineres som:

"Muligheten for at en påstand om en regnskapslinje inneholder feilinformasjon som enkeltvis eller samlet kan være vesentlig, før eventuelle kontroller tas i betraktning."

Risikovurderingen gjennomføres i tråd med den reviderte ISA 315. Standarden stiller krav til at revisor vurderer iboende risiko ved å vurdere sannsynligheten for og konsekvensen av mulig feilinformasjon.

Det er ikke meningen at du skal gjøre en fullstendig og tidkrevende risikovurdering. Hensikten er at du danner deg et overblikk og gjør en kort vurdering basert på det som blir presentert.



Figur A1.3: Skjermdump av spørreundersøkelse: Innledning 3

NHH



Hva er iboende risiko?

Muligheten for at en påstand om en regnskapslinje inneholder feilinformasjon som enkeltvis eller samlet kan være vesentlig, før eventuelle kontroller tas i betraktning.

Risikoen for at revisor gir uttrykk for en uriktig mening i revisjonsberetningen når regnskapet inneholder vesentlig feilinformasjon.


Risikoen for at en påstand om en transaksjonsklasse, kontosaldo eller tilleggsopplysning inneholder feilinformasjon som ikke oppdages eller korrigeres av tilhørende kontroller.

→

Lever av Qualtrics 

Figur A1.4: Skjermdump av spørreundersøkelse: Vurdering av tidsserier

NHH




Du skal nå vurdere risiko knyttet til utvikling i bransjen.

[Trykk her for å åpne nettsiden.](#)

Basert på informasjonen som er gitt om utvikling i bransjen, hvordan vurderer du sannsynligheten for at det er feilinformasjon i regnskapet?

Lav sannsynlighet Moderat sannsynlighet Høy sannsynlighet
0 10 20 30 40 50 60 70 80 90 100


Sannsynlighet





Basert på informasjonen som er gitt om utvikling i bransjen, hvordan vurderer du konsekvensen av mulig feilinformasjon i regnskapet?

Lav konsekvens Moderat konsekvens Høy konsekvens
0 10 20 30 40 50 60 70 80 90 100

Konsekvens





Levert av Qualtrics 

Figur A1.5: Skjermdump av spørreundersøkelse: Vurdering av ordsky

NHH



Du skal nå vurdere risiko knyttet til interne forhold i selskapet.

[Trykk her for å åpne nettsiden.](#)

Basert på informasjonen som er gitt om interne forhold i selskapet, hvordan vurderer du sannsynligheten for at det er feilinformasjon i regnskapet?

Lav sannsynlighet 0 10 20 30 Moderat sannsynlighet 40 50 60 70 Høy sannsynlighet 80 90 100

Sannsynlighet

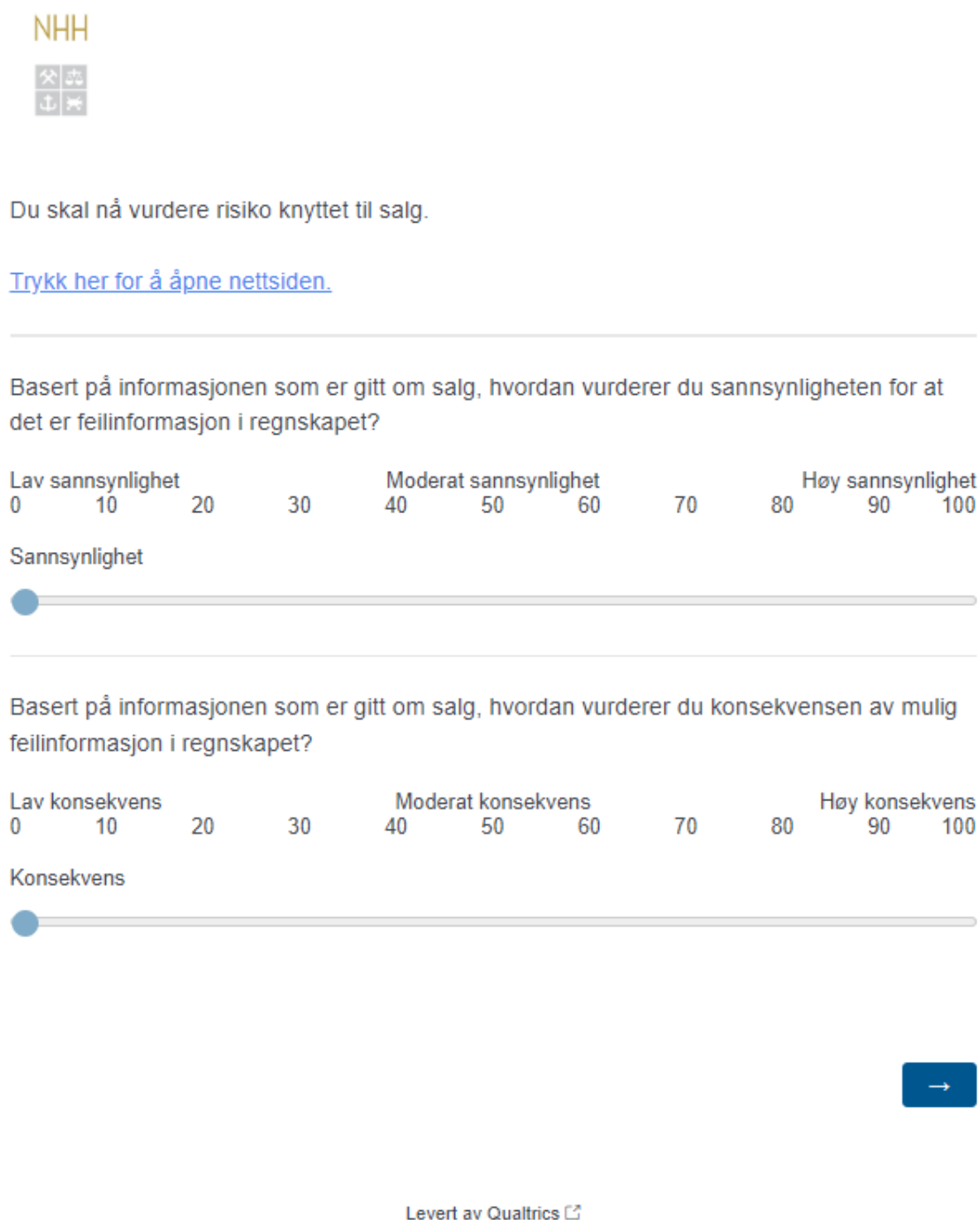


Basert på informasjonen som er gitt om interne forhold i selskapet, hvordan vurderer du konsekvensen av mulig feilinformasjon i regnskapet?

Lav konsekvens 0 10 20 30 Moderat konsekvens 40 50 60 70 Høy konsekvens 80 90 100

Konsekvens



Figur A1.6: Skjermdump av spørreundersøkelse: Vurdering av flytskjema

NHH

⌕ ⌕
⬇ ⌕

Du skal nå vurdere risiko knyttet til salg.

[Trykk her for å åpne nettsiden.](#)

Basert på informasjonen som er gitt om salg, hvordan vurderer du sannsynligheten for at det er feilinformasjon i regnskapet?

Lav sannsynlighet			Moderat sannsynlighet				Høy sannsynlighet			
0	10	20	30	40	50	60	70	80	90	100

Sannsynlighet

●

Basert på informasjonen som er gitt om salg, hvordan vurderer du konsekvensen av mulig feilinformasjon i regnskapet?

Lav konsekvens			Moderat konsekvens				Høy konsekvens			
0	10	20	30	40	50	60	70	80	90	100

Konsekvens


●

→

Levert av Qualtrics

Figur A1.7: Skjermdump av spørreundersøkelse: Vurdering av kart

NHH




Du skal nå vurdere risiko knyttet til leverandører.

[Trykk her for å åpne nettsiden.](#)

Basert på informasjonen som er gitt om leverandører, hvordan vurderer du sannsynligheten for at det er feilinformasjon i regnskapet?

Lav sannsynlighet Moderat sannsynlighet Høy sannsynlighet
0 10 20 30 40 50 60 70 80 90 100


Sannsynlighet




Basert på informasjonen som er gitt om leverandører, hvordan vurderer du konsekvensen av mulig feilinformasjon i regnskapet?

Lav konsekvens Moderat konsekvens Høy konsekvens
0 10 20 30 40 50 60 70 80 90 100


Konsekvens





Figur A1.8: Skjermdump av spørreundersøkelse: Samlet vurdering og kontrollspørsmål 1

NHH




Du er nå ferdig med nettsidene og skal gjøre en samlet vurdering.

Basert på all informasjon som er gitt, hvordan vurderer du den samlede sannsynligheten for at det er feilinformasjon i regnskapet?

Lav sannsynlighet Moderat sannsynlighet Høy sannsynlighet

0 10 20 30 40 50 60 70 80 90 100

Sannsynlighet




Basert på all informasjon som er gitt, hvordan vurderer du den samlede konsekvensen av mulig feilinformasjon i regnskapet?

Lav konsekvens Moderat konsekvens Høy konsekvens

0 10 20 30 40 50 60 70 80 90 100

Konsekvens



Hva er ditt kjønn?

Mann

Kvinne

Annet/ønsker ikke å svare

Hva er din alder?

Figur A1.9: Skjermdump av spørreundersøkelse: Samlet vurdering og kontrollspørsmål 2

Studerer du, eller har du studert Regnskap og Revisjon?

Nei

Har påbegynt bachelor

Har fullført bachelor

Har påbegynt master

Har fullført master

Har du arbeidserfaring innen revisjon?

Ingen erfaring

0-2 års erfaring

2-5 års erfaring

5-10 års erfaring

Over 10 års erfaring

Samarbeidet du med noen under spørreundersøkelsen?

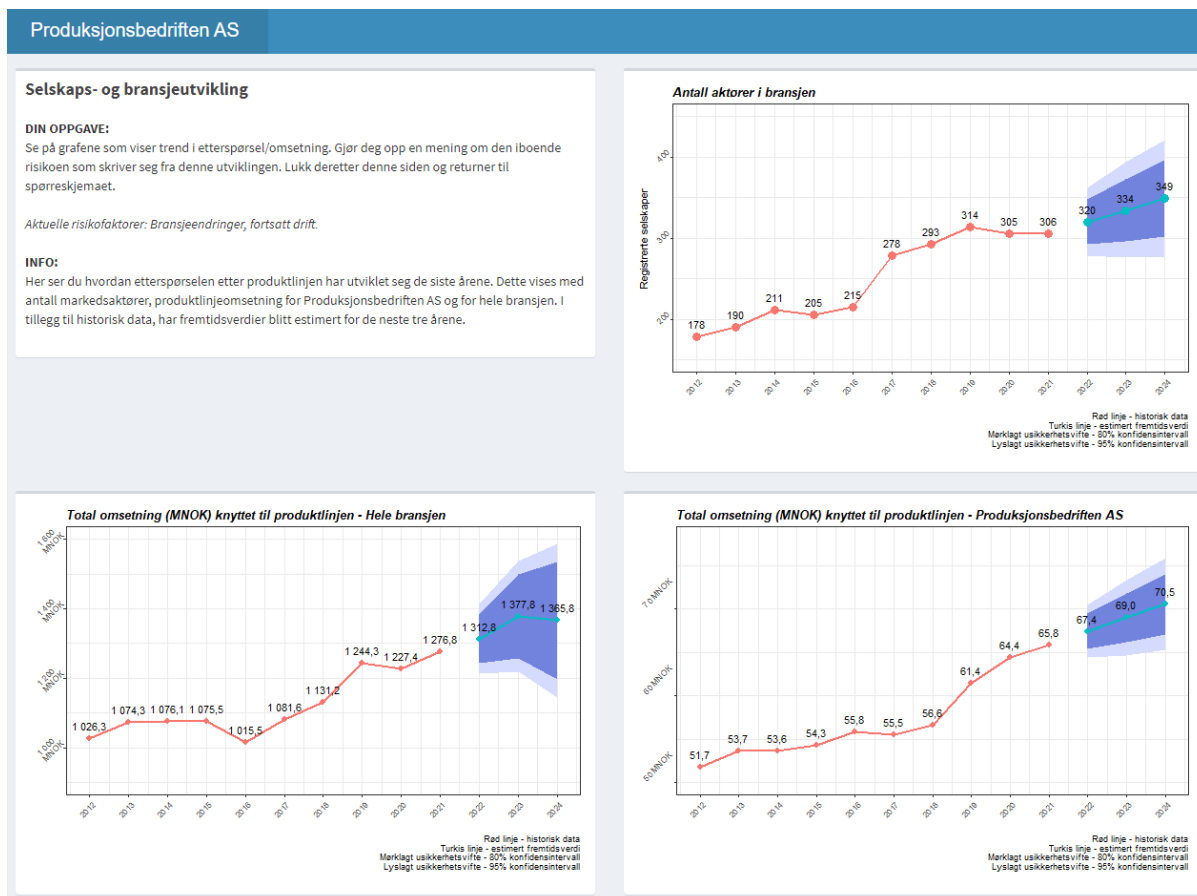
Ja

Nei

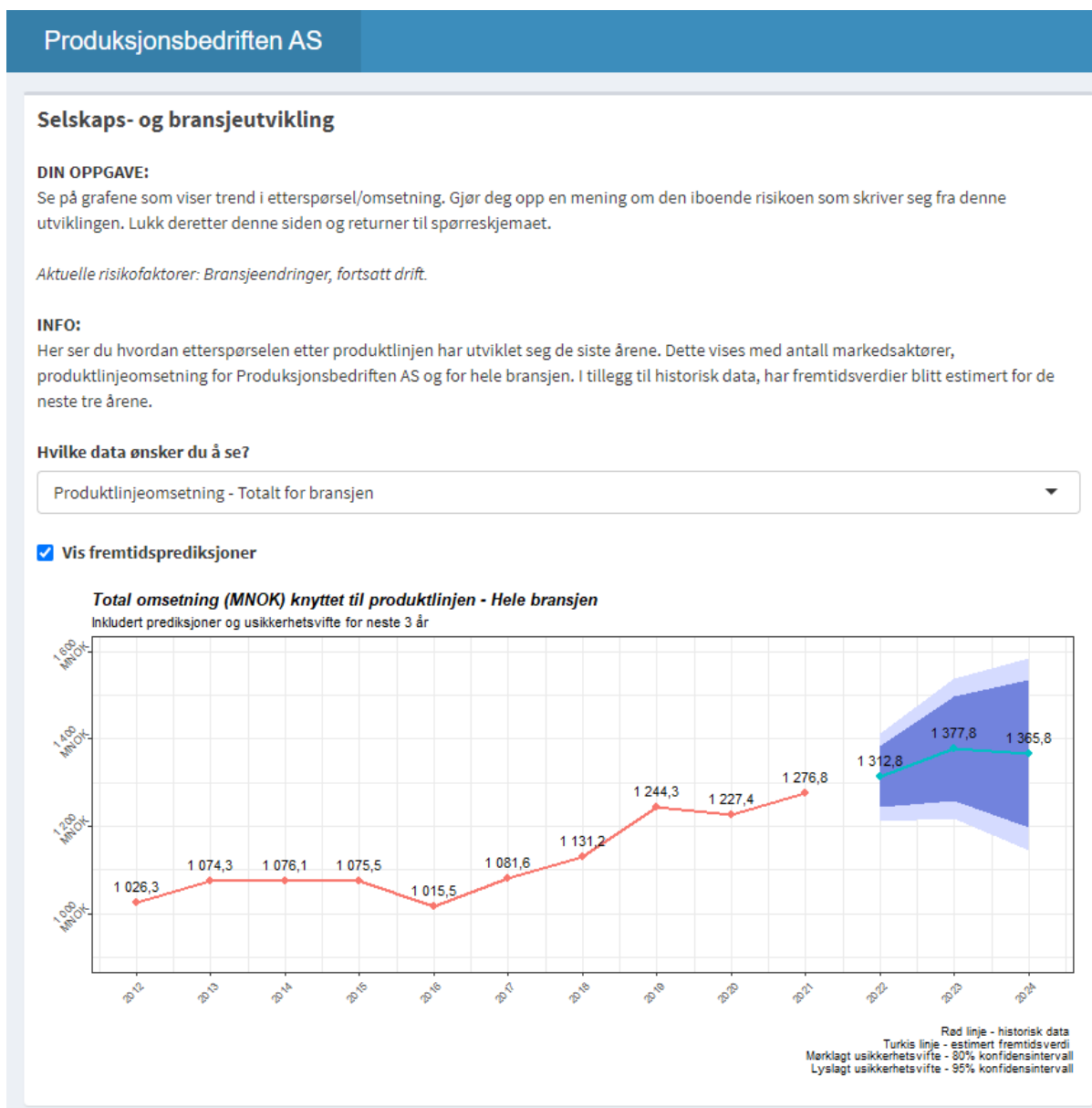


A2 Webapplikasjoner

Figur A2.1: Skjermdump av webapplikasjon: Tidsserier - (Positiv cue / Lavere risiko) - Statisk versjon



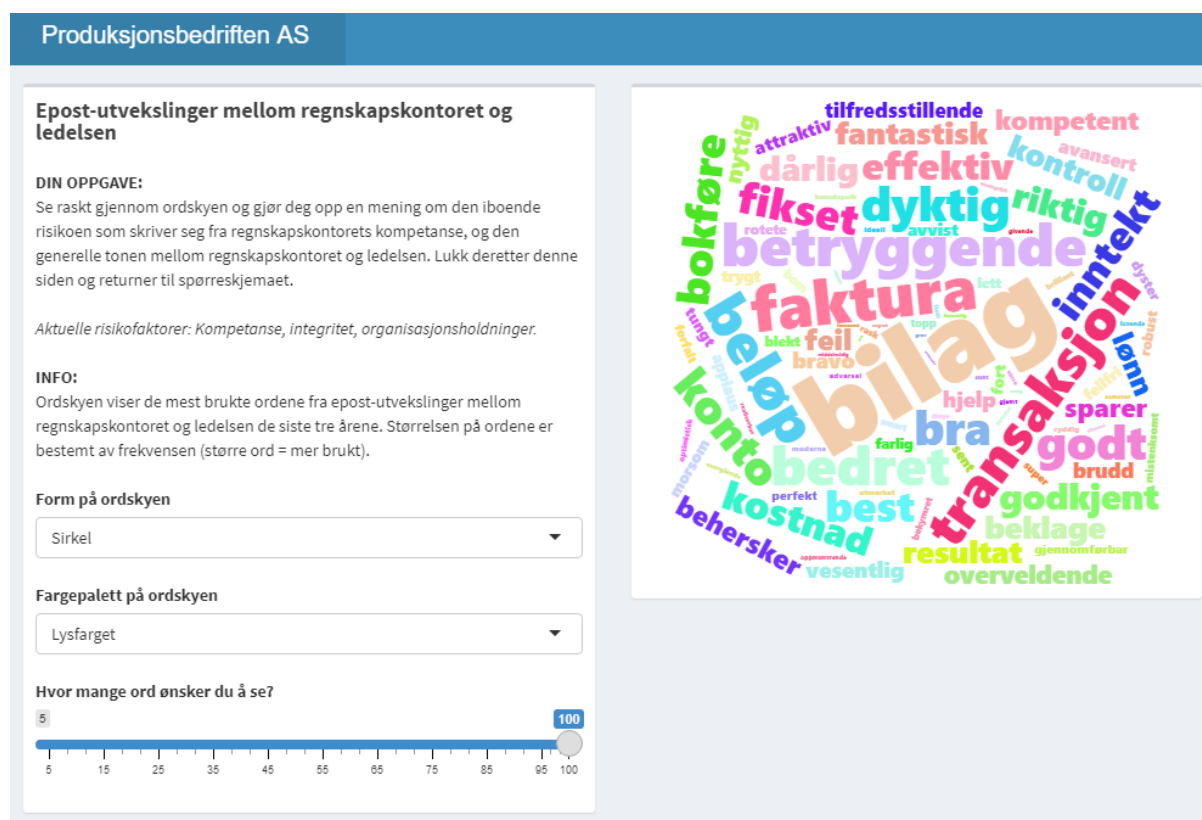
Figur A2.2: Skjermdump av webapplikasjon:
Tidsserier - (Positiv cue / Lavere risiko) - Interaktiv versjon



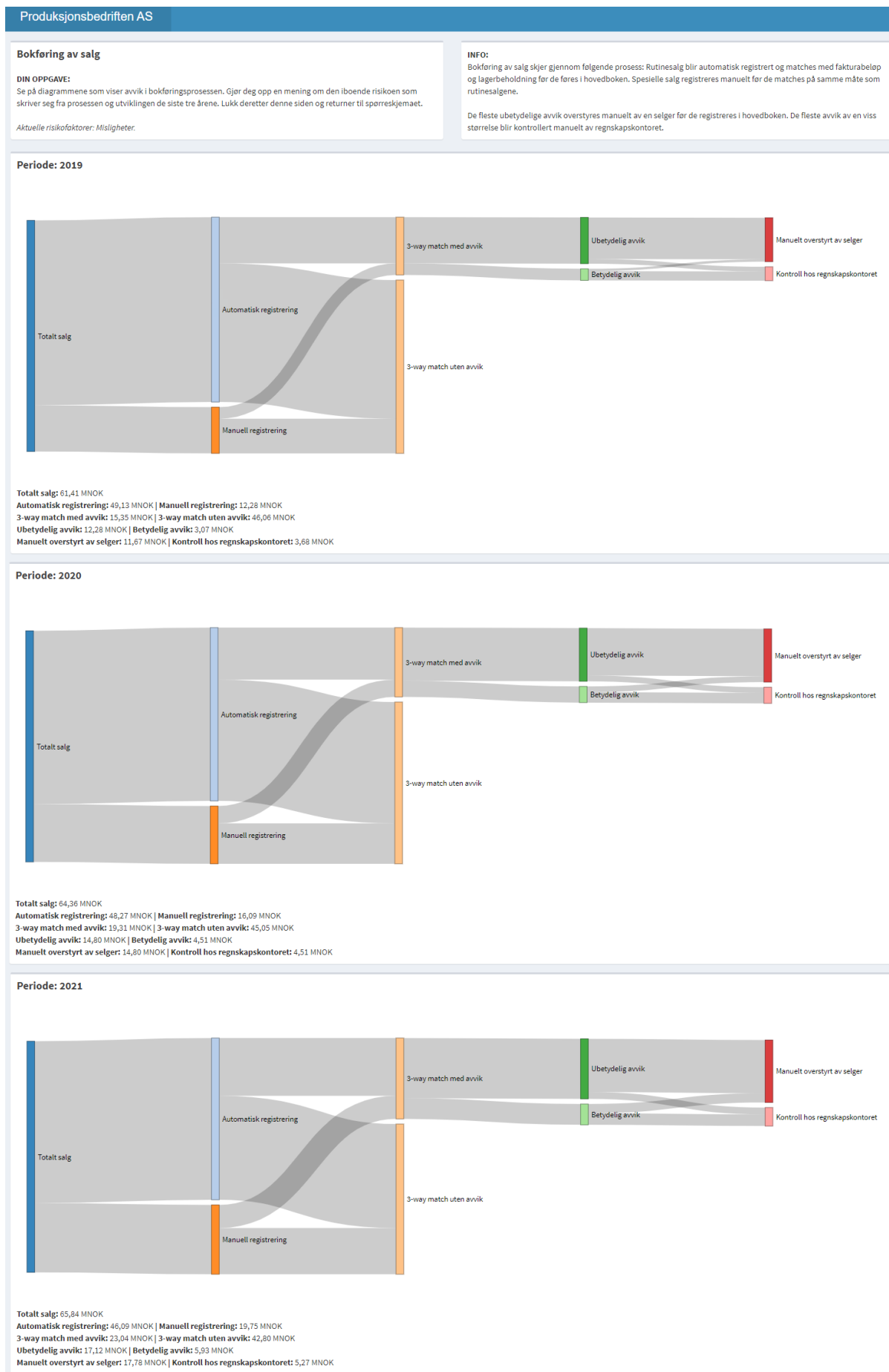
Figur A2.3: Skjermdump av webapplikasjon: Ordsky - (Positiv cue / Lavere risiko) - Statisk versjon



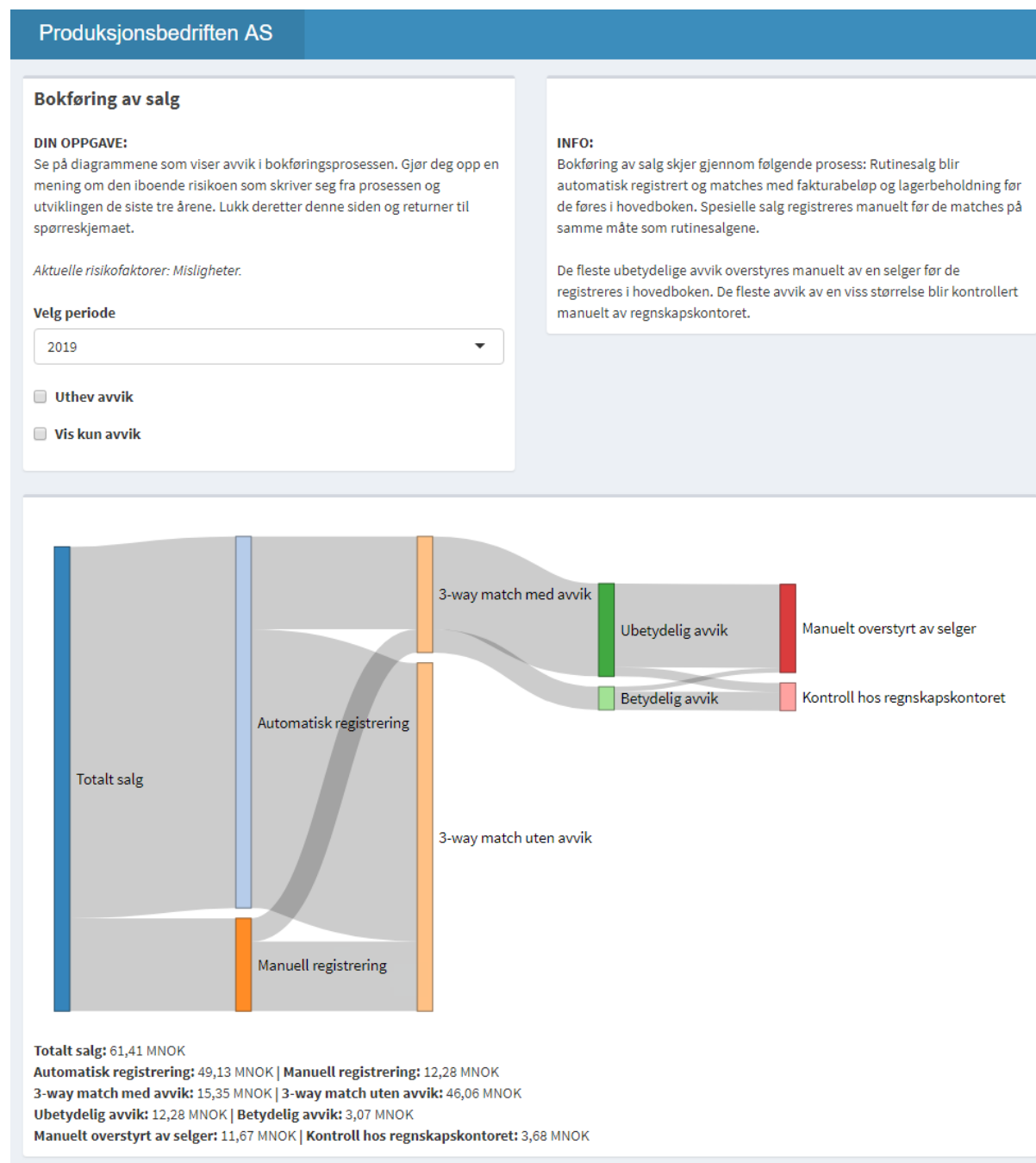
Figur A2.4: Skjermdump av webapplikasjon: Ordsky - (Positiv cue / Lavere risiko) - Interaktiv versjon



Figur A2.5: Skjermdump av webapplikasjon:
Flytskjema - (Negativ cue / Høyere risiko) - Statisk versjon



Figur A2.6: Skjermdump av webapplikasjon:
Flytskjema - (Negativ cue / Høyere risiko) - Interaktiv versjon




Figur A2.7: Skjermdump av webapplikasjon:
Kart - (Negativ cue / Høyere risiko) - Statisk versjon

Produksjonsbedriften AS

Leverandørendringer

DIN OPPGAVE:
Se raskt gjennom informasjonen du finner her. Gjør deg opp en mening om den iboende risikoen som skrives seg fra disse leverandørendringene. Lukk deretter denne siden og returner til spørreskjemaet.

Aktuelle risikofaktorer: Endringer, geografi, likviditet.



INFO:
Produksjonsbedriften AS har i løpet av de siste tre årene skiftet ut leverandørene som er knyttet til den aktuelle produktlinjen. Her kan du se en oversikt over finansiell og geografisk informasjon.

Land	By	CPI Indeks	AML Indeks
Norge	Hamar	7	3,35
Sverige	Gävle	3	3,36
Tyskland	Dortmund	9	4,43
Kina	Qingdao	78	6,70
India	Jetpur	86	5,15
Bangladesh	Gazipur	146	5,88

Forklaring av indekser:
AML Indeks rangerer land etter risiko for hvitvasking og terrorfinansiering. Høyere score betyr høyere risiko. I 2021 scoret Andorra best med 2,73 og Haiti dårligst med 8,49.
CPI Indeks rangerer land etter hvor korrupt deres offentlige sektor er antatt å være. Danmark er rangert som nummer 1 (minst korrupt) og Sør-Sudan er rangert som nummer 179 (mest korrupt).

Splitt for mobilsjerm

PERIODE: 2019

Leverandør-ID	Land	Leverandør siden	Transaksjoner i perioden	Gj.snitt per transaksjon (NOK)	Sum transaksjoner (NOK)	Gj.snitt dager kreditt
73EJD9947	Norge	2012-02-08	27	864 588	23 343 876	34,85
83GDE8843	Sverige	2012-03-14	14	668 554	9 359 756	36,73
92OLP7027	Tyskland	2014-04-18	19	443 997	8 435 943	31,20
					Tot: 41 139 575	

PERIODE: 2020

Leverandør-ID	Land	Leverandør siden	Transaksjoner i perioden	Gj.snitt per transaksjon (NOK)	Sum transaksjoner (NOK)	Gj.snitt dager kreditt
73EJD9947	Norge	2012-02-08	18	875 336	15 756 048	35,16
83GDE8843	Sverige	2012-03-14	17	612 597	10 414 149	36,24
92OLP7027	Tyskland	2014-04-18	23	387 149	8 904 427	30,80
42UUY0826	Kina	2020-01-09	14	377 468	5 284 552	26,80
					Tot: 40 359 176	

PERIODE: 2021

Leverandør-ID	Land	Leverandør siden	Transaksjoner i perioden	Gj.snitt per transaksjon (NOK)	Sum transaksjoner (NOK)	Gj.snitt dager kreditt
42UUY0826	Kina	2020-01-09	34	454 772	15 462 248	25,74
60YWB8273	India	2021-01-24	22	375 994	8 271 868	21,36
09QCT5416	Bangladesh	2021-02-12	29	476 394	13 815 426	18,79
					Tot: 37 549 542	

Figur A2.8: Skjermdump av webapplikasjon:
Kart - (Negativ cue / Høyere risiko) - Interaktiv versjon

Produksjonsbedriften AS

Leverandørendringer

DIN OPPGAVE:
Se raskt gjennom informasjonen du finner her. Gjør deg opp en mening om den iboende risikoen som skriver seg fra disse leverandørendringene. Lukk deretter denne siden og returner til spørreskjemaet.

Aktuelle risikofaktorer: Endringer, geografi, likviditet.

INFO:
Produksjonsbedriften AS har i løpet av de siste tre årene skiftet ut leverandørene som er knyttet til den aktuelle produktlinjen. Her kan du se en oversikt over finansiell og geografisk informasjon.

Forklaring av indekser:
AML Indeks rangerer land etter risiko for hvitvasking og terrorfinansiering. Høyere score betyr høyere risiko. I 2021 scoret Andorra best med 2,73 og Haiti dårligst med 8,49.
CPI Indeks rangerer land etter hvor korrupt deres offentlige sektor er antatt å være. Danmark er rangert som nummer 1 (minst korrupt) og Sør-Sudan er rangert som nummer 179 (mest korrupt).

Velg periode

2019

Kartet er interaktivt. Du kan zoome inn/ut og få mer informasjon om den enkelte leverandør ved å trykke på markørene.

Splitt for mobilsjerm

Leverandør-ID	Land	Leverandør siden	Transaksjoner i perioden	Gj.snitt per transaksjon (NOK)	Sum transaksjoner (NOK)	Gj.snitt dager kreditt
73EJD9947	Norge	2012-02-08	27	864 588	23 343 876	34,85
83GDE8843	Sverige	2012-03-14	14	668 554	9 359 756	36,73
92OLP7027	Tyskland	2014-04-18	19	443 997	8 435 943	31,20
					Tot: 41 139 575	

A3 Forutsetninger for variansanalyser

Tabell A3.1: Resultater fra Shapiro-Wilk -test for normalitet. Alle p-verdier er godt over 0.05 og består testen for normalitet.

Gruppe	Rekkefølge	Format	Vurdering	Test-score	P-verdi
Gruppe 1	PPNN	Statisk	Samlet	0.963	0.723
Gruppe 2	PPNN	Interaktiv	Samlet	0.948	0.458
Gruppe 3	NNPP	Statisk	Samlet	0.921	0.201
Gruppe 4	NNPP	Interaktiv	Samlet	0.949	0.434
Gruppe 1	PPNN	Statisk	P-cues	0.983	0.982
Gruppe 2	PPNN	Interaktiv	P-cues	0.914	0.133
Gruppe 3	NNPP	Statisk	P-cues	0.962	0.730
Gruppe 4	NNPP	Interaktiv	P-cues	0.951	0.479
Gruppe 1	PPNN	Statisk	N-cues	0.926	0.208
Gruppe 2	PPNN	Interaktiv	N-cues	0.979	0.952
Gruppe 3	NNPP	Statisk	N-cues	0.948	0.495
Gruppe 4	NNPP	Interaktiv	N-cues	0.972	0.856

Tabell A3.2: Resultater fra Levenes test, både med gjennomsnitt og median som sentrering. Gjennomsnittet er den originale Levenes test og median er en mer robust versjon. Alle p-verdier er godt over 0.05 og består testen for homogenitet.

Modell	Sentrering	Test-score	P-verdi
Samlet vurdering	Median	0.85	0.472
Samlet vurdering	Gjennomsnitt	0.92	0.439
Vurdering av P-cues	Median	1.27	0.292
Vurdering av P-cues	Gjennomsnitt	1.50	0.224
Vurdering av N-cues	Median	1.10	0.355
Vurdering av N-cues	Gjennomsnitt	1.18	0.327

A4 Alternativ tolkning av avhengig variabel

$$y'_{Pos} = \frac{1}{100} \left(\frac{T_{Sann}T_{Kons} + O_{Sann}O_{Kons}}{2} \right) = \mathbf{P-Cues}'_{Risk} \in [0, 100]$$

$$y'_{Neg} = \frac{1}{100} \left(\frac{F_{Sann}F_{Kons} + K_{Sann}K_{Kons}}{2} \right) = \mathbf{N-Cues}'_{Risk} \in [0, 100]$$

$$y'_{Sam} = \frac{1}{100} (S_{Sann}S_{Kons}) = \mathbf{Samlet}'_{Risk} \in [0, 100]$$

hvor

T_{Sann} og T_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med tidsserier.

O_{Sann} og O_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med ordsky.

F_{Sann} og F_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med flytskjema.

K_{Sann} og K_{Kons} er måling av sannsynlighet og konsekvens for nettsiden med kart.

S_{Sann} og S_{Kons} er måling av sannsynlighet og konsekvens for den samlede vurderingen.

Tabell A4.1: Anova-tabell (Type II tester) for den alternative tolkningen av samlet vurdering.

Avhengig variabel (alternativ): Samlet vurdering								
	<i>Toveis ANOVA</i>				<i>Toveis ANCOVA</i>			
	Sum Sq	Df	F value	Pr(>F)	Sum Sq	Df	F value	Pr(>F)
Rekkefølge	14.3	1	0.0472	0.8288	0.0	1	0.0001	0.9906
Format	304.9	1	1.0072	0.3196	424.2	1	1.2455	0.2699
Rekkefølge:Format	33.2	1	0.1096	0.7417	82.8	1	0.2430	0.6243
Alder					355.9	1	1.0450	0.3117
Kjønn					114.4	2	0.1679	0.8459
Utdanning					751.8	4	0.5519	0.6985
Erfaring					146.3	4	0.1074	0.9794
Residualer	18,165.0	60			16,689.0	49		

*p<0.1; **p<0.05; ***p<0.01

Tabell A4.2: Anova-tabell (Type II tester) for den alternative tolkningen av positive cues.

Avhengig variabel (alternativ): Positive cues									
	<i>Toveis ANOVA</i>				<i>Toveis ANCOVA</i>				
	Sum Sq	Df	F value	Pr(>F)	Sum Sq	Df	F value	Pr(>F)	
Rekkefølge	1.6	1	0.0113	0.9156	0.2	1	0.0013	0.9713	
Format	598.3	1	4.2034	0.0447**	758.6	1	4.9857	0.0302**	
Rekkefølge:Format	72.8	1	0.5118	0.4771	22.5	1	0.1479	0.7022	
Alder					71.2	1	0.4679	0.4972	
Kjønn					95.4	2	0.3135	0.7323	
Utdannelse					228.8	4	0.3760	0.8246	
Erfaring					611.8	4	1.0053	0.4139	
Residualer	8,539.7	60			7,455.2	49			

*p<0.1; **p<0.05; ***p<0.01

Tabell A4.3: Anova-tabell (Type II tester) for den alternative tolkningen av negative cues.

Avhengig variabel (alternativ): Negative cues									
	<i>Toveis ANOVA</i>				<i>Toveis ANCOVA</i>				
	Sum Sq	Df	F value	Pr(>F)	Sum Sq	Df	F value	Pr(>F)	
Rekkefølge	3.6	1	0.0116	0.9146	0.0	1	0.0001	0.9908	
Format	181.7	1	0.5846	0.4475	243.4	1	0.6740	0.4156	
Rekkefølge:Format	1.4	1	0.0046	0.9462	14.6	1	0.0405	0.8413	
Alder					50.7	1	0.1404	0.7095	
Kjønn					17.2	2	0.0238	0.9765	
Utdannelse					460.1	4	0.3185	0.8642	
Erfaring					325.6	4	0.2254	0.9229	
Residualer	18,648.0	60			17,696.1	49			

*p<0.1; **p<0.05; ***p<0.01