



Effekten av irrelevant informasjon og tidspress på revisjonskvalitet

En eksperimentell studie

Jon Marius Hagerud og Kristine Eide Haugen

Veileder: Jonas Gaudernack og Cardamine Carmen Olsen

Masteroppgave i økonomi og administrasjon, økonomisk styring

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i 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.

Innhold

Tabelliste	3
Forord	4
Sammendrag	5
1. Introduksjon	6
2. Tidligere teori og forskning.....	7
2.1 Judgement and Decision making	7
2.2 Dilution effect.....	8
2.3 Tidspress.....	10
2.4 Fortsatt drift vurderingen.....	11
2.5 Kontrollvariabler	12
3. Metode.....	14
3.1 Eksperimentelt design.....	14
3.2 Within- og between subjects.....	15
3.3 Design	18
3.3.1. Hypoteser	18
3.3.2 Operasjonalisering av kvalitet.....	19
3.3.3 Casedesign - operasjonalisering av relevant og irrelevant informasjon	20
3.3.4 Operasjonalisering av tidspress.....	21
3.3.5 Operasjonalisering av fortsatt drift	22
3.3.6 Manipulasjonsjekk	23
3.4 Praktisk gjennomføring	23
3.5 Analysemetoder	23
3.5.1 ANOVA repeated measures.....	24
3.5.2 ANOVA between subjects	25
3.5.3 SPANOVA.....	26
4. Resultater og tolkning	26

4.1 Deskriptiv statistikk.....	27
4.1.1 Hypotese 1	28
4.1.2 Hypotese 2	28
4.1.3 Hypotese 3	29
4.2 Manipulasjonssjekk	30
4.2.1 Tidspress	30
4.2.2 Informasjon	30
4.3 Presentasjon og tolkning.....	33
4.3.1 Hypotese 1	33
4.3.2 Hypotese 2	36
4.3.3 Hypotese 3	38
5. Konklusjon	41
5.1 Resultater	41
5.2 Begrensninger med oppgaven	42
5.3 Forslag til videre forskning.....	43
5.4 Betydning for praksis.....	44
6. Referanser.....	45
Vedlegg	48
Vedlegg 1: Eksperimentet	48
Innledende spørsmål	48
Oppgave 1: Tidspress gruppen får 3 min og 20 sek.....	48
Oppgave 2: Tidspressgruppen får 1 min og 10 sek.....	50
Manipulasjonssjekk.....	51
Vedlegg 2: Definisjon av variabler.....	52
Vedlegg 3: Original case - Fra Shelton 1999	53

Tabelliste

1. Deskriptiv statistikk for hypotese 1
2. Deskriptiv statistikk for hypotese 2
3. Deskriptiv statistikk for hypotese 3
4. Resultater av manipulasjonsjekk for informasjon
5. Resultater av hypotese 1, med bruk av Within-ANOVA
6. Resultater av hypotese 1, med bruk av sammenkoblede T-test (Paired T-test)
7. Resultater av hypotese 2, med bruk av Between-ANOVA
8. Resultater av hypotese 2, med bruk av uavhengig T-test (Independent T-test)
9. Resultater av hypotese 3, med bruk av SPANOVA

Forord

Denne masterutredningen er den avsluttende delen av vår Mastergrad i økonomi og administrasjon (MØA) ved Norges Handelshøyskole (NHH). Dette har vært et spennende, utfordrende og ikke minst lærerik prosess for oss begge. Utredningen var spennende ved at vi kunne kombinere revisjon med psykologiske faktorer. Temaet for oppgaven ble mer og mer spennende jo mer vi jobbet med det.

Vi vil rette en stor takk til våre veiledere, Jonas Gaudernack og Cardamine Carmen Olsen. Deres hjelp langs veien har vært helt avgjørende for oppgaven. Vi har hatt stort utbytte av at Carmen alltid har hatt tid til besøk på kontoret. I tillegg har vi fått delta på flere spennende seminarer og fått et innblikk i en forskernes verden. Takk for gode råd og innspill!

Vi har vært så heldige å skrive samme type masteroppgave som en gruppe andre studenter. Vi vil takke dere for diskusjoner og innspill. Det er godt å vite at vi er flere i samme båt når vi støter på problemer. Takk for samarbeidet!

Takk til studievenner, uten dere hadde ikke de to siste årene vært like givende. Til slutt vil vi takke familie og kjærester for støtte og råd. Vi setter pris på at dere er våre personlige heia-gjenger!

Bergen, juni 2016

Kristine Eide Haugen

Jon Marius Hagerud

Sammendrag

Tidligere forskning har avdekket fenomenet “dilution effect”¹ som går ut på at revisorer gjør mindre ekstreme beslutninger når de får både relevant og irrelevant informasjon, enn hvis de kun får relevant informasjon. Flere studier har undersøkt faktorer som kan moderere denne effekten, et av forslagene er tidspress. Vi har brukt fortsatt drift-vurderingen som revisjonshandling i denne studien.

Vi har i denne studien utført et eksperiment, der vi ønsker å finne en “dilution effect”, undersøke om revisorer som jobber under tidspress er mer effektive enn andre revisorer og om tidspress kan være en moderator for “dilution effect”. Eksperimentet ble utført på 30 studenter som studerer revisjon eller har tatt ett revisjonskurs. De ble delt i to grupper, der en gruppe ble utsatt for tidspress og en gruppe fikk bruke så lang tid de ønsket. Begge gruppene besvarte de samme oppgavene. I første oppgave var det bare relevant informasjon som var tilgjengelig. De ble da bedt om å vurdere sannsynligheten for fortsatt drift. Deretter fikk de beskjed om at informasjonen fra forrige oppgave fortsatt gjaldt, men de fikk litt ytterligere informasjon. Den nye informasjonen var irrelevant, og de ble bedt om å gjøre en ny vurdering av sannsynligheten for fortsatt drift.

I resultatene fant vi ingen “dilution effect”. Vi fant at de som ble utsatt for tidspress svarte raskere på oppgavene enn de uten tidspress. Siden vi ikke fant en “dilution effect” er det vanskelig å finne en løsning på “dilution effect”. Dermed fant vi heller ikke at tidspress var løsningen på “dilution effect”.

¹ Vi bruker det engelske uttrykket «dilution effect» fordi vi ikke har funnet noen egnet norsk oversettelse som inneholder det samme som det engelske uttrykket.

1. Introduksjon

I denne masterutredningen skal vi se på “dilution effect” og hvordan denne kan modereres. Under revisjonen vil revisor få mye informasjon. Hvilken type informasjon revisor får vil ha betydning for vurderingen og beslutningen hans/hennes. Vi kan dele informasjonen inn i to grupper. Den ene gruppen er relevant informasjon². Dette er informasjon som er nyttig og behøves for å gjøre vurderinger og ta beslutninger i revisjonen. Den andre gruppen er irrelevant informasjon³. Disse typene informasjon er grunnlaget for vår problemstilling, som er to-delt;

1. Hvordan påvirkes vurderinger av hvilken type informasjon revisor får og hvordan påvirkes denne sammenhengen av tidspress?
2. Hvordan påvirkes vurderinger av hvor mye tid revisor får?

Det er flere forhold som kan føre til at vurderingene og beslutningene i revisjonen ikke blir utført med høyest mulig kvalitet. Et av disse forholdene er hvor lang tid revisoren får på å utføre vurderingene og beslutningene. Et annet forhold er “dilution effect”. Dette fenomenet består i at vurderinger og beslutninger blir mindre ekstreme når revisor får både relevant og irrelevant informasjon, enn hvis han/hun bare får relevant informasjon. Vi ønsker først å se om vi kan påvise en “dilution effect”. Forskning og praksis søker å finne løsninger eller forhold som modererer systematiske problemer i revisjonen, slik som “dilution effect”. En tidligere studie har foreslått tidspress som en løsning på “dilution effect” (Glover 1997).

Tidspress tolkes av Glover (1997) som at deltakerne har nok tid til å skimme igjennom informasjonen, men ikke har tid til å fullt ut prosessere informasjonen. Vi legger til grunn samme tolkning for tidspress i vårt eksperiment. Eksperimentet undersøker hvordan tidspress direkte påvirker vurderinger, samt om tidspress modererer “dilution effect”.

Vi finner ingen sammenheng mellom hvilken informasjon deltakerne får og vurderingen som gjøres. Dette betyr at vi ikke finner noen “dilution effect” i dette utvalget. Når vi ikke finner en “dilution effect” er det vanskelig å finne en faktor som modererer effekten. Vi finner dermed ikke at tidspress endrer styrken på “dilution effect” i vårt utvalg. Dette betyr at vi ikke finner “dilution effect” i verken gruppen som er utsatt for tidspress eller gruppen som ikke er

² Relevant informasjon er også kalt diagnostic information i tidligere forskning på engelsk.

³ Irrelevant informasjon er også kalt nondiagnostic information i tidligere forskning på engelsk.

utsatt for tidspress. Vi finner at hvor mye tid deltakerne får på oppgaven påvirker vurderingen. Deltakerne tar en vurdering raskere når de blir utsatt for tidspress enn hvis de ikke blir utsatt for tidspress.

Videre skal vi i kapittel 2 plassere vår studie blant den forskningen som finnes fra før. Deretter vil vi gå gjennom tidligere teori og forskning innen “dilution effect” og tidspress, samt se på teori knyttet til fortsatt drift-vurderingen. I kapittel 3 ser vi på metoden vi bruker. Her vil vi først begrunne valget av eksperimentelt design, før vi forklarer grundig designet på vår studie. Vi går igjennom hypotesene og begrunner valgene som er tatt og vurderingene som er gjort videre i undersøkelsen. I kapittel 4 går vi igjennom resultatene av analysene våre og tolker disse. Konklusjonen i kapittel 5 ser på begrensninger ved studien og forslag til videre forskning.

2. Tidligere teori og forskning

2.1 Judgement and Decision making⁴

Vår studie kan plasseres innen forskning på “Judgement and Decision making” (heretter kalt for JDM). Dette er altså forskning innen vurderinger og beslutninger. Vi fokuserer på JDM innen revisjon. Innen JDM-forskning ønsker man å finne de personene eller omgivelsene som gir høy JDM-kvalitet. Det er to hovedsynspunkter på hvordan man skal definere JDM-kvalitet; prosess- og resultat-perspektivet (Bonner 2008). I denne studien tar vi utgangspunkt i resultat-perspektivet. I våres studie vil høy JDM-kvalitet vises ved at hvert individ er enig med seg selv over tid, så lenge forutsetningene tilsier at svaret burde være likt.

Vi har plassert vår studie innen JDM-forskningen, dette er et stort felt. Bonner (2008) har laget et rammeverk for studier innen JDM-forskning. Her har hun listet opp tre faktorer som kan påvirke eller forårsake forskjeller i JDM-kvaliteten. De tre kategoriene er person, oppgave og omgivelser. I vår undersøkelse ser vi på hvilken type informasjon man får når man utfører oppgaven, samt tidspress, som er foreslått som en løsning på “dilution effect”. Vi

⁴ Vi har valgt å beholde det engelske uttrykket «judgement and decision making» fordi vi ikke har funnet noen god norsk oversettelse til dette uttrykket. I kursopplegget ble dette uttrykket brukt både skriftlig og muntlig, noe som gir oss inntrykk av at dette er allment innen denne grenen av forskning.

plasserer oss dermed inn i både oppgave- og omgivelse-kategorien. Type informasjon knyttes til oppgaven og tidspress til omgivelsene.

2.2 Dilution effect

Flere studier har vist at revisorer blir påvirket av hvilken type informasjon man får (Hackenbrack 1992; Waller og Zimbelman 2003; Shelton 1999; Glover 1997; Eilifsen, Kochetova-Kozolske og Messier 2015; Hoffman og Patton 1997). Disse studiene har vist at vurderingene og beslutningene blir mindre ekstreme når revisor får både relevant og irrelevant informasjon enn hvis de bare får relevant informasjon. Dette har blitt kalt “dilution effect”. Waller og Zimbelman (2003) forklarer “dilution effect” slik: “the presence of nondiagnostic cues, when processed with diagnostic cues, causes judges to under-weigh the diagnostic cues”. Det blir altså tillagt for lite vekt på den relevante informasjonen. Eilifsen, Kochetova-Kozloski og Messier (2015) definerer “dilution effect” når irrelevant informasjon er tilgjengelig som at “...individuals make less extreme (more regressive) judgements than those in the presence of diagnostic evidence only.”. I denne studien har vi definert “dilution effect” som

Vurderinger og beslutninger blir mindre ekstreme når revisor får både relevant og irrelevant informasjon enn hvis han/hun bare får relevant informasjon.

En av de første studiene som fant en “dilution effect” var Nisbett, Zukier og Lemley i 1981. Denne studien hører til i psykologi-litteraturen og har ingen tilknytning til revisjon. Grunnen til at “dilution effect” oppstår mener Nisbett, Zukier og Lemley (1981) er fordi mennesker gjør vurderinger på bakgrunn av sammenligning. Vi kan se dette i vår studie. Vi ser på fortsatt drift-vurderingen. Revisor har erfaringer med både bedrifter som fortsetter i drift og bedrifter som går konkurs. Jo mer informasjon revisor får som gjør at bedriften ligner på en konkursbedrift, jo større sannsynlighet vil han sette for konkurs. Hvis revisor også får irrelevant informasjon ligner bedriften litt mindre på en konkursbedrift og revisor vil sette sannsynligheten for konkurs lavere. Generelt kan vi si at hvis informasjonen ligner på det vi vet at et av de alternative utfallene innebærer, velger vi dette utfallet. Får man informasjon som ikke er relevant, blir utfallet og den totale informasjonen man har fått, mindre like. Dermed blir sannsynligheten for dette utfallet også mindre. Zukier (1982) uttaler følgende i en

studie om “dilution effect”: “The dilution effect rests on an apparent paradox: subjects consider information non diagnostics for the prediction of a criterion and nevertheless use it in their judgments”.

En av de første studiene som påviste en “dilution effect” innen revisjon var Hackenbrack (1992). Resultatet her viser at vi får en “dilution effect” av både fordelaktig og ufordelaktig informasjon. Hackenbrack (1992) påpeker at “dilution effect” kan forventes å påvirke mange beslutninger innen revisjon fordi revisorer som oftest får betydelige mengder irrelevant informasjon.

Når man finner fenomener som “dilution effect” i forskning ønsker man å finne ut hvorfor dette fenomenet oppstår og hvordan man kan minske effekten av fenomenet (hvis det anses som skadelig). Vi har ikke funnet noe forskning innen revisjon som forklare fullt ut hvorfor vi finner en “dilution effect”. Det finnes noe forskning på hvordan man kan minske effekten av “dilution effect”. Glover (1997) undersøkte hvordan styrken på “dilution effect” endret seg når man påfører tidspress og når man holder revisor ansvarlig for vurderingen som blir gjort. Han fant at tidspress reduserte styrken på “dilution effect”, men eliminerer den ikke. Det var ingen effekt av å holde revisor ansvarlig for vurderingen. Hoffman og Patton (1997) undersøkte også hvordan ansvarlighet påvirket “dilution effect”. Resultatene deres viser at “dilution effect” ikke blir påvirket av at revisor blir holdt ansvarlig for vurderingen. Shelton (1999) undersøkte om revisjonserfaring påvirket styrken på “dilution effect”. Her fikk han meget gode resultater, der “dilution effect” var helt borte hos de erfarne revisorene. En studie har også vist at “dilution effect” blir mindre i vurdering av sannsynlighet for misligheter når man oppgir sannsynligheten med frekvens enn med prosent (Eilifsen, Kochetova-Kozloski og Messier 2015). “Dilution effect” ble mindre når revisorene fikk irrelevant, *fordelaktig* informasjon og bruker frekvens-metoden. Dette er et viktig funn fordi bedriftsledelse som har noe å skjule, slik som misligheter, som regel vil peke på irrelevant informasjon som er fordelaktig. (Eilifsen, Kochetova-Kozloski og Messier 2015).

Vår studie baseres i stor grad på Glover (1997) og Shelton (1999). Det vi adopterer fra Shelton (1999) er hvilken vurdering som gjøres, altså fortsatt drift-vurderingen og hvilke konkrete påstander som inkluderes i caseteksten. I Glover (1997) gjorde revisorene en vurdering på risiko for misligheter på kundefordringskontoen. Hypotesene til Glover (1997) går ut på at tidspress reduserer “dilution effect” og ansvar øker “dilution effect”. Man tenker

at tidspress skal minske effekten fordi revisor ikke har tid til å gå nøye gjennom all informasjonen. Dermed bruker man den informasjonen som ved første øyekast virker som den mest relevante. Studien til Glover ble gjennomført i to deler der del 1 inneholdt åtte korte case med bare relevant informasjon og del 2 tok opp igjen en av de korte casene og supplerte med irrelevant informasjon. Vår studie er bygd opp på samme måte, men fordi dette er en liten studie har vi kun en case med bare relevant informasjon. I Glover (1997) fikk noen i utgangspunktet informasjon om lav risiko, mens andre fikk informasjon om høy risiko, dermed kunne man undersøke om “dilution effect” virket i begge ekstremene. Resultatene til Glover (1997) viser at tidspress reduserte “dilution effect”, men eliminerte den ikke. I vår studie ønsker vi å finne en “dilution effect” og se om tidspress løser dette problemet hos oss også.

2.3 Tidspress

I revisjonen vil det alltid finnes en viss grad av tidspress. Det er mange oppdrag som skal være ferdig samtidig, derfor blir det ofte satt opp et budsjettert antall timer til hver revisjon. Dette kan hjelpe bedriften til å holde kontroll på kostnadene. Tidspress er en ytre faktor som pålegges av andre enn revisor. Asare, Trompeter og Wright (2000) påpeker at det finnes to måter å påføre tidspress på. Man kan få en tidsbegrensning der man ikke kan bruke lengre tid enn de tildelte minuttene, timene eller dagene. Alternativt kan man få et budsjettert tidsbruk, da vil det være mulig å overskride budsjettet. I forskning er det mest vanlig å bruke en tidsbegrensning, noe som er uvanlig i praksis (Asare, Trompeter og Wright 2000). Vi har i vår studie valgt en tidsbegrensning på oppgavene som skal utføres.

Det finnes ingen absolutt definisjon av tidspress. Glover (1997) har i sin studie valgt å gi deltakerne i undersøkelsen så lite tid at de ikke rekker å fullt ut vurdere all informasjonen. De skal derimot rekker å skimme igjennom, slik at informasjonen filtreres. Vi har lagt samme syn til grunn for hva tidspress skal være i vår studie. Rent intuitivt forstår vi at personer som blir påført tidspress, slik at de ikke rekker å ta hensyn til all informasjon på samme måte som uten tidspress, vil føle en viss grad av stress. De fleste har også opplevd dette på ett eller annet tidspunkt i livet. Choo (1995) mener at det er personens egne oppfatning av stressnivået i situasjonen og bedømmelsen av ens egen evne til å takle stress som til slutt definerer stress.

Stress blir altså bedømt subjektivt og personens tillit til at man selv håndterer stresset vil påvirke hvor stort tidspress man føler.

Choo (1995) undersøker hvilke teorier som forklarer sammenhengen mellom stress og vurderinger best. Han finner at teorien som gir en invers U-form forklarer sammenhengen best. Dette betyr at ved lavt nivå av stress vil vurderingene ha lav kvalitet. Når stressnivået øker vil kvaliteten på vurderingene også øke, man blir altså mer effektiv. Blir stressnivået for høyt vil man ikke lengre gi høy kvalitet på vurderingene. Teorien som Choo (1995) hentet dette fra kalles “Cue Utilization Theory” og foreslår at når stressnivået øker vil man ignorere den minst relevante informasjonen. Dette når et optimalt nivå når all irrelevant informasjon blir ignorert. Hvis stressnivået øker ytterligere utover dette nivået vil informasjon som er relevant også bli ignorert. Noe som fører til en dårligere vurdering. Tidspress kan dermed være en fordel opp til et visst punkt der man er mest effektiv i vurderingene.

I tillegg til Choo (1995) viser McDaniel (1990) og Glover (1997) at tidspress fører til en filtreringsstrategi. All informasjon blir ikke vurdert som å være like viktig lengre. Wrights studie fra 1974 forklarer at en person under tidspress ekskluderer mindre relevant informasjon, selv om han vurderer denne informasjonen som relevant nok til å bli med i vurderingen under mindre stressende omstendigheter. Tidspress blir dermed lagt fram som en løsning på “dilution effect” fordi man ikke lengre vurderer den irrelevante informasjonen (Glover 1997). I studien til Glover (1997) viser han at tidspress reduserer men ikke eliminerer “dilution effect”. Vi skal delvis replikere dette i vår studie. Hvis våre resultater viser at tidspress reduserer “dilution effect” vil dette styrke tidspress som løsningen på “dilution effect”.

2.4 Fortsatt drift vurderingen

Alle revisjonsoppdrag inneholder en fortsatt drift vurdering. At en bedrift skal fortsette driften er en forutsetning når man vurderer verdien av eiendeler og gjeld (Eilifsen et al 2014). ISA 570 sier at “Under fortsatt drift-vurderingen antas det at enheten vil fortsette driften i overskuelig fremtid.” (ISA 570). “Overskuelig fremtid” ifølge Eilifsen et al (2014) er minst ett år fra avslutningen av regnskapsåret. Revisor skal gå gjennom tre trinn i vurderingen om fortsatt drift (Eilifsen et al 2014; ISA 570). Først går man gjennom revisjonshandlingene som

er utført i revisjonen og ser etter bevis på betydelig tvil om fortsatt drift. Deretter må man innhente informasjon fra ledelsen i bedriften om hvilke planer har de for å unngå et problem i forhold til fortsatt drift. Hvis det var betydelig tvil om fortsatt drift etter trinn en, men ledelsens planer er gode og realistiske kan dette føre til at det ikke lengre er betydelig tvil om fortsatt drift. Har ikke ledelsen gode strategier knyttet til fortsatt drift kan det fortsatt være betydelig tvil knyttet til fortsatt drit. Trinn tre i vurderingen er altså å vurdere ledelsens planer.

Forhold som kan skape tvil om fortsatt drift kan knytte seg til tre hovedområder; finansielle forhold, driftsrelaterte forhold eller andre forhold. (ISA 570). Negative nøkkeltall og manglende evne til å betale kreditorer ved forfall er eksempler i ISA 570 på finansielle forhold som kan skape tvil i fortsatt drift vurderingen. Driftsrelaterte forhold kan være tap av nøkkelpersonell uten at de blir erstattet eller problemer med tilgang på viktige innsatsfaktorer. Andre forhold som påvirker bedriftens evne til fortsatt drift kan være politiske forhold. ISA 570 påpeker at forhold som i utgangspunktet skaper tvil om fortsatt drift kan kompenseres av andre faktorer, f.eks.; "...tap av en hovedleverandør oppveies av at en alternativ leverandør er tilgjengelig." (ISA 570).

I vår studie er fortsatt drift-vurderingen hovedspørsmålet deltakerne våre skal besvare. Vi kommer til å gå mer inn på hvilke faktorer vi har valgt som del av informasjonen for å innføre betydelig tvil om fortsatt drift. Fortsatt drift-vurderingen ble valgt fordi den gjøres i alle revisjonsoppdrag og er den underliggende antagelsen i de fleste regnskap. Dette er med andre ord en viktig revisjonshandling. I forskningen på fortsatt drift-vurderingen kommer det fram at langt fra alle de bedriftene som gikk konkurs hadde merknader i revisjonsberetningen til fortsatt drift i året før de gikk konkurs (Rau og Moser 1999).

2.5 Kontrollvariabler

I forskning vil man gjerne finne hvor mye en uavhengig variabel påvirker en avhengig variabel. Hvis man gjennomfører analysene med bare disse to variablene kan den uavhengige variabelen få for mye eller for lite forklaringskraft fordi noe av variasjonen avhenger av andre utelatte variabler. Disse variablene kan man ta med i analysen og vi vil få et bedre estimat for den uavhengige variabelen vi er interessert i. Kontrollvariablene kan og gi oss svar på om den avhengige variabelen varierer med disse. Vi har valgt tre kontrollvariabler som vi innhenter

informasjon om, i tillegg til det som er diskutert til nå. Disse spørsmålene kommer først i eksperimentet.

Den første kontrollvariabelen vi har valgt er kjønn. Bonner (2008) har funnet en rekke faktorer som påvirker JDM kvaliteten og kan være interessante for vår studie, der man har funnet forskjeller mellom kjønnene. Man har funnet bevis for at menn og kvinner er forskjellige med hensyn til kognitive evner. Det er derimot ingen forskjell i generell intelligens. Det handler om at kvinner scorer bedre når man tester verbale evner, mens menn scorer bedre når man tester resonneringsevner. Det er bevist at kvinner har lavere mestringstro enn menn i tradisjonelt maskuline jobber. Tradisjonelt sett har det vært mange menn i revisjonsyrket. For ett år siden var det rundt 59% menn som var revisor/finansrådgiver (Olsson 2015). Man kan også finne bevis for at menn er mer selvsikre og har overdreven selvtillit, spesielt i maskuline jobber. Det er stor forskjell mellom kjønnene når det kommer til holdninger til risiko. Kvinner er klart mer risikoaverse enn menn. Studier har funnet at kvinner har bedre korttidshukommelse. Til slutt har en studie antydnet at kvinner har større evne til å prosessere informasjon. Vi ser derfor på kjønn som en variabel som kan forklare noe av variasjonen i den uavhengige variabelen.

Den andre kontrollvariabelen vi har valgt er karaktersnitt. Evner vil kunne ha en positiv effekt på JDM-kvalitet. Helt intuitivt forstår vi at evner ikke kan ha positive effekter på JDM-kvalitet hvis revisorene mangler kunnskap. I tillegg finnes det en studie som foreslår at hvor fort mennesker henter fram og prosesserer informasjon varierer mellom forskjellige individer. (Bonner 2008). Vi bruker karaktersnitt som en proxy for evner og kunnskap. Siden alle deltakerne i eksperimentet vårt må gå Master i Revisjon og Regnskap (MRR) eller tatt masterkurset i Revisjon (BUS426/MRR411) har vi sikret et visst kunnskapsnivå innen revisjon. Dette er viktig for at deltakerne skal forstå hva vurderingen de skal gjøre går ut på og forstå alle begreper i oppgaveteksten. Etter vår erfaring er de som har høye karaktersnitt gode til å skille ut relevant og irrelevant informasjon. Dette er helt nødvendig når man skal gjennom et stort pensum og finne hovedpunktene. Karaktersnitt kan derfor påvirke variasjonen i den uavhengige variabelen.

Den siste kontrollvariabelen vi har valgt er tidligere erfaring. En stor del av de som går MRR har jobbet innen revisjon tidligere. Intuitivt kan man tenke seg at jo mer erfaring man har med fortsatt drift-vurderingen jo flinkere er man til å skille mellom relevant og irrelevant

informasjon. Dette er bevist i Shelton (1999). Her ble det vist at erfarne revisorer ikke viser noen “dilution effect”, mens mindre erfarne revisorer viser en “dilution effect”. Shelton (1999) mener dette henger sammen med måten en revisor lærer seg å organisere det de finner i revisjonsprosessen. Vi mener derfor tidligere erfaring kan forklare noe av variasjonen i den uavhengige variabelen.

3. Metode

3.1 Eksperimentelt design

Den mest vanlige metoden for JDM-forskning innen revisjon er eksperimentelt design (Trotman 2001). Hvis et eksperimentelt design utføres riktig er det antagelig et av de sterkeste designene med hensyn på intern validitet (Trochim og Donnelly 2008). Dette skyldes at vi har gått inn og manipulert en av variablene. Dermed vet vi at endringen i den avhengige variablene er et resultat av manipulasjonen. Den høye interne validiteten kan være en av grunnene til at eksperiment er en så populær metode. I et eksperiment vil den eksterne validiteten være lav. Dette skyldes at man ofte har en urealistisk situasjon. I vårt eksperiment er det urealistisk at en revisor kun har den informasjonen vi gjør tilgjengelig og kun gjør fortsatt drift-vurderingen. Et eksperimentelt design kjennetegnes ved at en eller flere av de uavhengige variablene manipuleres av forskeren (Trotman 2001). Læreboken til Trochim og Donnelly (2008) påpeker at vi både vil påvise (1) at hvis vi manipulerer X (uavhengig variabel) får vi Y (avhengig variabel) og (2) at hvis vi ikke manipulerer X (uavhengig variabel) får vi ikke Y (avhengig variabel). I vår studie er det fortsatt drift-vurderingen som er avhengig variabel, Y. Denne variabelen blir oppgitt i sannsynlighet for fortsatt drift på en 101-poengsskala, der 0 er garantert konkurs og 100 er garantert fortsatt drift. Vi skal manipulere både hvilken informasjon subjektene får og om hvilken tid de får på å gjennomføre oppgaven. X-variablene består dermed av informasjon og tidspress.

3.2 Within- og between subjects⁵

Videre må man bestemme om uavhengige variabler skal manipuleres “within subjects” eller “between subjects”. “Within subjects” betyr at alle subjektene får alle nivåene av de uavhengige variablene. (Trotman 2001). Vi har i vår studie to uavhengige variabler. Derfor ville et fullstendig “within subjects”-design betydd at alle subjektene gikk gjennom de fire rutene i figuren under. Det vil si at alle subjektene gikk gjennom oppgaven med bare relevant informasjon og både relevant og irrelevant informasjon, uten tidspress og i tillegg fikk oppgaven med bare relevant informasjon og både relevant og irrelevant informasjon, med tidspress. Altså rute 1, 2, 3 og 4 under.

Figur 1: 2x2 matrise av design.

	Ikke tidspress	Tidspress
Bare relevant info	1	2
Begge typer info	3	4

“Between subjects” betyr at hvert subjekt bare får en “treatment”⁶. En “treatment” vil i et eksperiment være å manipulere en av variablene. I vår studie kunne et subjekt kun fått en oppgave og hadde havnet i en av rutene i figuren over. Altså f.eks. kun relevant informasjon og ikke tidspress (rute 1). Vi har to typer “treatment”; manipulering av informasjon og manipulering av tid. I figuren over blir rute nr 3 og 4 utsatt for “treatment” innen informasjon, altså at her får man også irrelevant informasjon. Rute nr 2 og 4 blir utsatt for manipulasjon innen tid. De får begrenset med tid, mens i rute 1 og 3 kan man bruke så mye tid man ønsker.

⁵ Vi har valgt å bruke de engelske uttrykkene «within subjects» og «between subjects» fordi det ikke finnes noen gode norske oversettelser for disse uttrykkene.

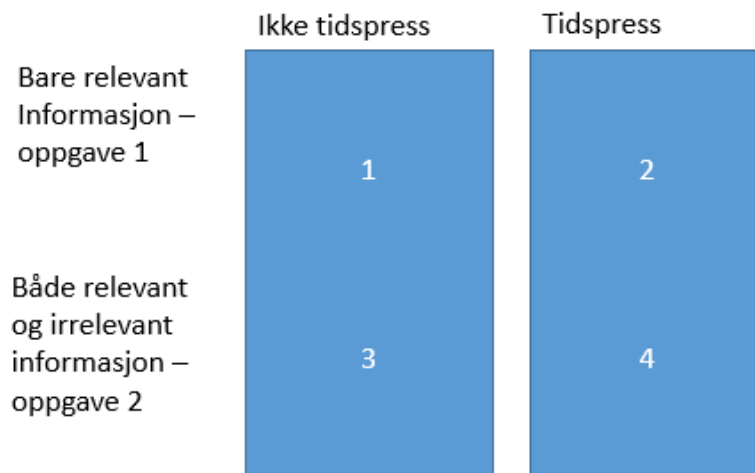
⁶ Vi har valgt å bruke det engelske uttrykket «treatment». Den nærmeste norske oversettelsen vil være manipulasjon. Vi har inntrykk av at «treatment» er et allment uttrykk innenfor eksperimentelt design og brukes gjerne på norsk både skriftlig og muntlig.

Begge typene design har både fordeler og ulemper. I stor grad vil fordelene med den ene være ulempe ved den andre. Trotman (2001) har valgt ut tre fordeler og tre vesentlige ulemper man bør ta hensyn til når man velger design. Fordelene med “within subjects” består i at man trenger færre subjekter. Dette vil være en stor fordel i en masterutredning som må samle inn alle deltakerne blant en travel studentmasse. Dette vil typisk være en ulempe ved “between subjects”. I “within subjects” vil man få større statistisk kraft fordi subjektene blir sin egen kontrollgruppe. I forskningen er det kun ved bruk av en kontrollgruppe man kan sjekke om resultatene ble annerledes enn hvis man ikke hadde innført “treatment”. Ulempen med en kontrollgruppe er at det krever flere deltakere, men siden subjektene kan virke som sin egen kontrollgruppe unngår vi dette problemet. I vår studie gir samme subjekt et svar uten manipulasjon av informasjonsvariabelen og et svar med manipulasjon. Disse svarene kan sammenlignes direkte mot hverandre. Da vil også alle andre variabler som kjønn, erfaring og kunnskap holdes like mellom de to svarene. Den tredje fordelene innebærer at “within subjects” er effektiv når man undersøker læringseffekter eller bruk av hint. Denne fordelene er ikke relevant for oss, siden vi ikke ønsker å undersøke læringseffekter. I vårt case bør læringseffekten være minimal da begge oppgavene omhandler samme bedrift med begrenset informasjon.

En av ulempene med “within subjects” er at når subjektene blir utsatt for flere “treatments” vil de ha større sjans for å finne ut hva hypotesene er. Hvis de klarer å gjette hypotesene kan de gi vurderinger og svar de tror forskeren ønsker eller forventer. Den andre ulempen ved “within subjects” er at man “øver” på oppgavene fordi de gjentas. Dette kan være positivt fordi subjektet utfører oppgaven bedre enn estimert på forhånd eller negativt fordi subjektet kjeder seg. Den tredje ulempen ved “within subjects” er effekten av en “treatment” kan smitte over på neste oppgave. (Trotman 2001).

Vi har i vår oppgave valgt å ikke bruke verken “within subjects” eller “between subjects” fullt ut, men en kombinasjon av begge. Vi kan illustrere dette med en figur.

Figur 2: Kombinasjon av within- og between subjects design



Denne figuren viser at vi har valgt et “within subjects” design på informasjonsvariabelen. Dette gjennomføres ved at hvert enkelt subjekt svarer på to oppgaver. Vi har valgt “between subjects” på tidspress-variabelen. Et subjekt kan bare bli plassert i “tidspress”-gruppen eller “ikke tidspress”-gruppen, ikke begge. På denne måten kan vi utnytte fordelene og unngå noen av ulempene med de to designtypene. Når informasjonsvariabelen er “within subjects” vil dette føre til at vi trenger færre subjekter og at hvert subjekt fungerer som en kontrollgruppe for seg selv. Når vi sammenligner 3 og 1 vil alle kontrollvariablene være like. Når vi bruker “within subjects” kombinert med “between subjects” mener vi sannsynligheten for at subjektene forstår hva hypotesene er, blir mindre. Siden det kun er to oppgaver som gjennomføres av samme subjekt og de får forskjellig informasjon vil man ikke kunne øve seg i noen særlig grad. Oppgavene er korte, slik at det ikke skal bli for kjedelig, som kunne ført til at subjektene ikke gjennomfører. Effekten av en “treatment” kan ikke smitte over på en annen oppgave fordi vi gir oppgaven uten “treatment” (bare relevant informasjon) først og deretter oppgaven med treatment (både relevant og irrelevant). Rekkefølgen er viktig her fordi den forhindrer at “treatment” på informasjonsvariabelen smitter over på oppgaven uten “treatment”. Effekten av tidspress kan ikke smitte over på en annen oppgave da man enten blir utsatt for tidspress eller ikke.

“Eksperimenter er internt valide når variasjonen i avhengig variabel entydig kan tilskrives manipulasjonen i den uavhengige variabelen.” (Fritt oversatt fra Trotman 2001). Dette betyr at man må kontrollere at variasjonen i den avhengige variabelen ikke blir påvirket av eksterne variabler. Den vanligste løsningen på dette er tilfeldig fordeling av subjekter til gruppene

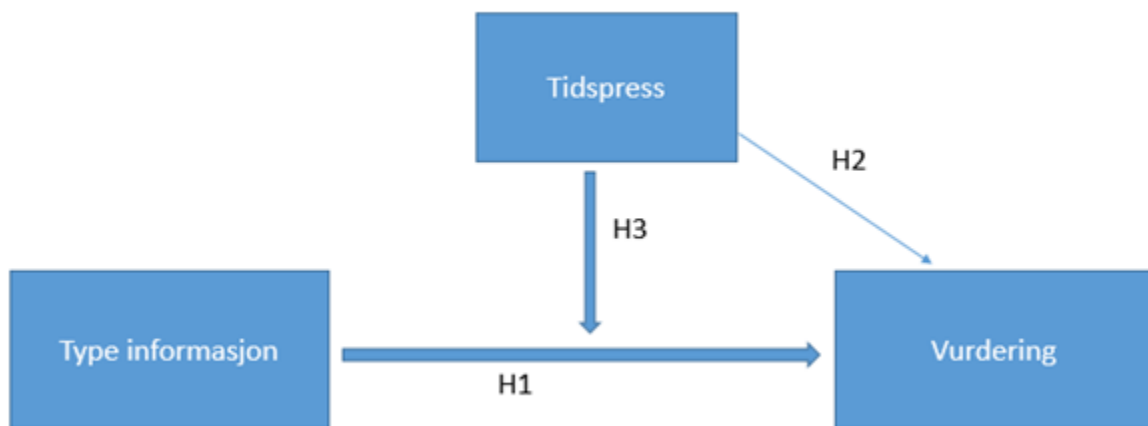
(Trotman 2001). I vår oppgave blir subjektene tilfeldig fordelt mellom tidspress og ikke tidspress gruppen. Dette ble i praksis utført ved at dataprogrammet eksperimentet ble utført i fordelte annet hvert subjekt til hver av de to gruppene.

3.3 Design

3.3.1. Hypoteser

Vi ønsket å replikere en studie av Glover (1997) der han ser på hvordan tidspress øker eller minker graden av “dilution effect”. Tidspress er en modererende variabel fordi den regulerer styrken på sammenhengen, den kan ikke forklare sammenhengen. I tillegg har vi hentet mye inspirasjon fra Shelton (1999). For å styrke tidspress som en løsning må vi i studien påvise “dilution effect” i “base case”. Vi har laget en enkel modell for å illustrere de sammenhengene vi vil undersøke.

Figur 3: Sammenhengen mellom informasjon, vurdering og tidspress.



Den første boksen gir hvilken type informasjon subjektene i eksperimentet får. Her er det to alternativer, bare relevant eller både relevant og irrelevant informasjon. Der relevant informasjon er informasjon som er nyttig for vurderingen eller beslutningen, mens irrelevant informasjon er informasjon som ikke er nyttig for vurderingen eller beslutningen. Vurdering i boksen til høyre representerer den vurderingen eller beslutningen som blir gjort. I vår studie er dette en fortsatt drift-vurdering der deltakerne skal vurdere sannsynligheten for at dette selskapet fortsatt er i drift om 12 måneder, gitt informasjonen de har fått. Hypotese 1 gir sammenhengen mellom type informasjon og vurderingen. H1a er alternativhypotesen og H1b

er nullhypotesen. I analysen vil vi teste H1b og vi ønsker å forkaste denne. Essensen av H1 er “dilution effect”.

H1a: Når revisor får både relevant og irrelevant informasjon er vurderingene mindre ekstreme enn hvis revisor kun får relevant informasjon

H1b: Når revisor får både relevant og irrelevant informasjon er vurderingen lik som hvis revisor kun får relevant informasjon.

Den siste boksen i modellen er tidspress. Denne mener vi vil påvirke både vurderingen i seg selv, men også forholdet mellom type informasjon og vurderingen (altså “dilution effect”). Hypotese 2 forteller sammenhengen mellom tidspress og vurderinger. H2a er alternativhypotesen og H2b er nullhypotesen. Vi ønsker å forkaste nullhypotesen og stå igjen med alternativhypotesen. Det er ikke mulig å måle om vurderingene er dårligere under tidspress i vår studie, fordi det ikke er noe rett eller galt svar på hvor sannsynlig det er at bedriften fortsetter driften. Dette er en vurderingssak. Dermed handler hypotesen om hvor raskt subjektene gir et svar, altså hvor effektive de er.

H2a: Ved tidspress tas beslutningene mer effektivt enn uten tidspress.

H2b: Ved tidspress tas beslutningene like raskt som uten tidspress.

Hypotese 3 bygger på studien til Glover (1997) som fant at tidspress modererer “dilution effect”. H3a er alternativhypotesen og H3b er nullhypotesen.

H3a: Ved tidspress vil “dilution effect” bli mindre

H3b: “Dilution effect” er like stor med og uten tidspress.

3.3.2 Operasjonalisering av kvalitet

Vi har tidligere sagt at kvalitet i vårt eksperiment betyr at hvert individ er enig med seg selv over tid. Eksperimentet er lagt opp slik at i oppgave 1 får man all relevant informasjon og i oppgave 2 tilfører man bare irrelevant informasjon. Den relevante informasjonen som ligger til grunn for begge vurderingene er dermed den samme. Høy JDM-kvalitet finner vi hvis svaret på oppgave 2 minus svaret på oppgave 1 blir 0, altså at svaret på de to oppgavene er det samme. I figur 2 vil dette tilsvare rute 3 minus rute 1 og rute 4 minus rute 2.

3.3.3 Casedesign - operasjonalisering av relevant og irrelevant informasjon

Når vi skulle utforme caseteksten måtte vi velge ut noen momenter som er relevante for fortsatt drift-vurderingen og noen som er irrelevante. I tillegg måtte vi finne opp en bedrift å bygge casen rundt. Bedriften i casen vår er derfor satt sammen av litt fantasi og en del personlig erfaring fra en faktisk bedrift. Hva som er relevant og irrelevant informasjon kan ikke vi bestemme. Vi har derfor tatt utgangspunkt i påstandene i Shelton (1999). Hun har brukt samme vurdering, altså fortsatt drift, i sin studie. Det er listet opp 11 elementer som er pretestet og brukt i hennes studie. Av disse er seks relevante og fire irrelevante. Vi oversatte disse til norsk og inkludert dem i caseteksten. Dette kunne vi gjøre fordi de var godt testet fra før og tilgjengelige i artikkelen til Shelton (1999). Noen av påstandene måtte modifiseres litt for å passe inn i caseteksten. Shelton (1999) har påstanden; "The company has begun a management rotation program.". Dette måtte modifiseres til "Selskapet har innført et internt rotasjonsprogram for de ansatte" fordi selskapet vi lagde var så lite at det ikke gir mening å innføre et rotasjonsprogram for ledelsen.

I tillegg har vi brukt ISA 570 om fortsatt drift. Her er det listet opp en rekke "hendelser eller forhold som vil skape tvil om enhetens evne til fortsatt drift" (ISA 570). Vi valgte ut punktet "manglende evne til å betale kreditorer ved forfall" (ISA 570). Dette passet inn i bedriften vi skapte og stemte for bedriften vi ble inspirert av. Flere av punktene til Shelton (1999) finner vi igjen i ISA 570. Til slutt ønsket vi å ha en irrelevant påstand til. Vi valgte "Ledelsen består av 50 % kvinner". Dette mente vi intuitivt var irrelevant.

Siden vi innførte noen nye påstander og modifiserte noen påstander ønsket vi å gjennomføre en liten pretest. Vi kontaktet derfor personer vi kjenner og som jobber med revisjon eller regnskap. Pretesten har kun to deltakere. En av deltakerne er statsautorisert revisor med utdanning fra NHH. Han har både undervist innen faget revisjon og regnskap og jobbet med regnskap i rundt 25-30 år. Den andre deltakeren har en bachelorgrad i revisjon og har arbeidet med regnskap i 2-3 år. Vi setter mest tillit til de svarene vi fikk fra deltaker nr 1, siden han har en høyere utdanning og har vesentlig mer erfaring. Oppgaven de fikk var å preteste påstandene på samme måte som i Shelton (1999). Her blir påstandene rangert på en skala fra -5 til 5, der 0 er irrelevant informasjon, -5 er relevant informasjon som påvirker bedriftens

evne til fortsatt drift negativt og 5 er relevant informasjon som påvirker bedriftens evne til fortsatt drift positivt.

De viktigste påstandene vi fikk pretestet er de som er nye eller modifisert, dette handler kun om tre påstander. “Ledelsen består av 50% kvinner” mente begge deltakerne var irrelevant. Vi brukte derfor dette videre inn i caseteksten vår. “Manglende evne til å betale kreditorer ved forfall” ble rangert til -5 og -4. Dette stemmer med ISA 570 og vi har dermed brukt dette som en relevant, negativ påstand videre. “Selskapet har innført et rotasjonsprogram for de ansatte” ble vurdert som irrelevant av den deltakeren med mest erfaring. Den andre deltakeren rangerte denne påstanden som svakt positivt. Vi har valgt å bruke dette som en irrelevant påstand i caseteksten.

Man kan finne alle påstandene i caseteksten hvis man ser etter. Noen står eksplisitt i teksten, mens andre har vi inkludert i tallmateriale, slik som resultatregnskap eller nøkkeltall. Dette betyr at man må prosessere informasjonen litt for å finne alle påstandene. De relevante påstandene er valgt for å antyde at det er betydelig tvil om bedriftens evne til fortsatt drift.

3.3.4 Operasjonalisering av tidspress

Vi gjennomførte flere pilottester for å bestemme hvor mye tid tidspress-gruppen skulle få til rådighet. McDaniels (1990) sin studie bestemte tidsbegrensningene ved å først gjennomføre oppgaven med 22 deltakere. Deretter valgte hun fire nivåer av tidspress, f.eks. tiden 10% av de raskeste deltakerne brukte ble tiden som brukes som “høyt tidspress”. Vi har gjennomført dette på en lignende måte.

I pilottest 1 hadde vi åtte deltakere og ingen tidsbegrensning. Denne ble utført på studenter som har samme veileder som oss på masterutredningen. I denne testen lå gjennomsnittet på oppgave 1 på 307,5 sekunder og oppgave 2 på 109 sekunder. Vi ønsker at deltakerne skal kunne lese fort igjennom oppgaven, men ikke ha tid til å prosessere informasjonen like nøye som de ellers ville gjort. Vi valgte derfor de 25% raskeste og satt dette som maks tidsbruk på pilottest 2. Vi fikk i tillegg tilbakemeldinger på innholdet og oppsettet av casen som førte til noen endringer.

Pilottest 2 hadde en tidsbegrensning på 230 sekunder på oppgave 1 (3 min og 50 sek) og 72 sekunder på oppgave 2 (1 min og 21 sek). Denne pilottesten ble utført på 4 masterstudenter fra Norges Handelshøyskole (NHH) innen økonomi og administrasjon. De hadde ikke tatt faget Revisjon (BUS426/MRR411). Dette mente vi ikke er viktig i pilottest 2 fordi vi ikke var opptatt av hvilken vurdering de gjorde, men hvor lang tid de brukte på å lese teksten. Disse studentene brukte i gjennomsnitt 167 sekunder på oppgave 1 og 52 sekunder på oppgave 2. Særskilt på oppgave 1 er det et stort avvik mellom tiden brukt og tid til rådighet. Deltakerne besvarte også spørsmålet “Følte du tidspress?”. Alle fire følte et tidspress. I tillegg kom de med skriftlige tilbakemeldinger, slik at vi kunne evaluere nivået av tidspress. Deltakerne kommenterte at de ble stresset og følte de hadde lite tid til å sette seg inn i informasjonen. På bakgrunn av resultatene ønsket vi å justere tidsbegrensningen noe mer.

I pilottest 3 fikk deltakerne en tidsbegrensning på 200 sekunder (3 min og 20 sek) på oppgave 1 og 70 sekunder (1 min og 10 sek) på oppgave 2. Her hadde vi ytterligere 2 deltakere. En av disse har tatt faget Revisjon på masternivå på NHH. Den andre har ikke tatt fag i Revisjon, men går på NHH på masternivå. Disse brukte i gjennomsnitt 138 sekunder på oppgave 1 og 54 sekunder på oppgave 2. En av deltakerne følte stort tidspress, mens den andre følte hun hadde tid til å ta en liten vurdering etter hun hadde lest informasjonen. Vi valgte å beholde tidsbegrensningene på 200 sekunder og 70 sekunder.

3.3.5 Operasjonalisering av fortsatt drift

Deltakerne i vårt eksperiment har gjort en fortsatt drift-vurdering. Denne ble gjort på en 101-poengsskala fra 0 til 100, der 0 er ingen sannsynlighet for fortsatt drift og 100 er garantert fortsatt drift. Vi ønsker å finne ut om vurderingen blir mer ekstrem med ulik type informasjon. Det er to måter å tolke hva «ekstrem» er. Det første alternativet er å anse 50% som midtpunktet. Da vil en vurdering på 50% være nøytral. Deretter vil man bli mer ekstrem, jo lengre mot 100 eller 0 man nærmer seg. På den andre siden er det mest vanlige resultatet av en fortsatt drift-vurderingen at bedriften har evne til fortsatt drift, altså ikke kommer til å gå konkurs. Dermed kan vi tolke ekstrem som mindre enn 100% sannsynlighet for fortsatt drift. Vi kan si at jo nærmere 0 man kommer, jo mer ekstrem er vurderingen. Vi har valgt den siste måten å tolke begrepet ekstrem.

3.3.6 Manipulasjonsjekk

Når man driver forskning er det vanlig å sjekke om man faktisk har manipulert variablene. En slik sjekk gjøres f.eks. i Glover (1997) og Zukier (1982). Vi har manipulasjon på to variabler, informasjon og tid. For å sjekke om deltakerne ble utsatt for tidspress har vi bedt deltakerne selvrapportere om de følte tidspress. På informasjonsvariabelen har vi skilt ut 14 påstander som er hovedpunktene i oppgaveteksten. Dette er de samme variablene som vi bygde opp oppgaveteksten av. Det siste spørsmålet i eksperimentet bestod i å rangere disse påstandene som enten relevant og positivt for bedriftens evne til fortsatt drift, irrelevant eller relevant og negativt for bedriftens evne til fortsatt drift. Denne testen vil undersøke studentenes kunnskap på hva som er relevant og irrelevant informasjon i forhold til fortsatt drift-vurderingen.

3.4 Praktisk gjennomføring

Når vi har fastlagt alle begrepene og hvordan de skal operasjonaliseres, blir eksperimentet gjennomført i praksis. Vi lagde først casen. Dette ble deretter testet gjennom tre pilottester. Pilottest 1 ble gjennomført på studenter som skriver masteroppgave for samme veileder som oss. De to neste pilottestene ble gjennomført med studenter vi kjenner fra NHH. Resultatene fra pilottestene formet casen vår. Den ferdige casen ble distribuert til så mange aktuelle studenter på NHH som mulig. Undersøkelsen ble først distribuert gjennom "It's learning", som er læringsplattformen som brukes på NHH. Vi sendte melding gjennom "It's learning" til alle som er meldt opp i utvalgte fag. Deretter klarte vi å melde oss inn i en gruppe for de som går MRR på Facebook. Her la vi ut innlegg med link til undersøkelsen vår og oppfordringer til å ta undersøkelsen. Begge disse strategiene ble gjennomført flere ganger for å oppnå et tilfredsstillende antall deltakere. I tillegg fikk vi maillister til de som har tatt faget BUS 426 (Revisjon) etter 2013 og fortsatt er studenter på NHH fra studieadministrasjonen. Disse fikk dermed og en henvendelse om å ta undersøkelsen vår på mail.

3.5 Analysemetoder

I analysen har vi valgt å bruke IBM SPSS da vi synes dette er mer brukervennlig enn for eksempel Microsoft Excel og er ellers mye brukt innenfor forskning. For å unngå uteliggere er det vanlig å fjerne topp og nedre 5% av dataene. Dette kalles det trimmede gjennomsnittet. Det er ikke stor forskjell mellom det vanlige gjennomsnittet og trimmede gjennomsnittet for

noen av oppgavene eller noen av gruppene. Før vi kan teste hypotesene undersøker vi om dataene i eksperimentet er pålitelige. Dette testes med “Cronbach alfa”⁷. Dataene er pålitelige hvis denne alfaen er høyere enn 0.7, men helst 0.8 (Pallant 2010).

Vi har valgt tre forskjellige analysemetoder til de tre hypotesene våre. Dette skyldes forskjellig design og forskjellige data som ligger til grunn for de tre hypotesene. Vi presenterer først de tre analysemetodene før vi presenterer resultatene fra dem. Vi har valgt å bruke et 95% konfidensintervall i analysen.

3.5.1 ANOVA repeated measures

I hypotese 1 har vi som tidligere nevnt et “within subjects”-design, der vi ønsker å finne en “dilution effect”. Vi valgte derfor å bruke “ANOVA repeated measures”⁸ (“within”) (heretter kalt “within-ANOVA”⁹) som analysemetode. Denne analysemetoden gjør at vi kan fjerne noe av feilen som kan forekomme ved å bruke samme deltaker flere ganger. Analysemetoden er altså spesielt tilpasset et “within”-design og er derfor godt egnet i hypotese 1.

For å bruke “within-ANOVA” som analysemetode så krever dette at en del forutsetningene blir oppfylt. Den første forutsetningen er uavhengighet i observasjonene av deltakerne. Dette kan være et problem ved at deltakere snakker med hverandre under testen. Eventuelt at noen deltakere forteller andre om oppgaven før/under de tar undersøkelsen. Vi forutsetter at dette ikke er et problem i denne utredningen. Videre forutsettes det at dataene må være normalfordelte. Dette sjekker vi med Shapiro-Wilks testen (Lund Research Ltd, 2013), siden vi har et utvalg som er lavere enn 50 stykker (Pallant 2010). For at variablene skal være normalfordelt må de måles på en kontinuerlig skala.

En tredje forutsetning er at det burde være homogenitet i variansen. Vi ønsker med andre ord at de forskjellige oppgavene har relativt like varianser. Dette sjekker vi ved å se på F-max forholdstallet. Man regner ut F-max ved å ta den gruppen med høyest varians og dividerer på den gruppen med lavest variansen. Dette forholdstallet skal ikke være høyere enn 3 (Coakes

⁷ Vi har valgt å bruke det engelske uttrykket «Cronbach alfa» fordi det ikke finnes noen norsk oversettelse.

⁸ Vi bruker her det engelske uttrykket fordi det ikke finnes noen god norsk oversettelse. Dette er navnet på en analysemetode og vil være mest presis på sitt originalspråk.

⁹ «Within-ANOVA» bygger på «within subjects» og vi har derfor valgt å fortsette og bruke det engelske uttrykket «within».

2013). Grunnen til at vi bruker denne metoden i denne hypotesen er at vi ser på samme gruppe og må derfor bruke en alternativ metode til for eksempel Levenes test, som vi bruker i hypotese 2. Fjerde og siste forutsetning er "Sphericity"¹⁰. Denne tester vi med Mauchlys Test for "Sphericity". Forutsetningen holder hvis testen ikke er signifikant ($p > 0.05$) (Coakes 2013).

I "within-ANOVA" kan man ikke legge inn kontrollvariabler som selvstendige uavhengige variabler. Dette kommer av at det er de samme deltakerne som har svart på oppgave 1 og 2, og kontrollvariablene vil dermed være like for begge oppgavene. "Within-ANOVA" integrerer kontrollvariablene ved å lage interaksjonsledd mellom den uavhengige variabelen som skal forklare forskjellen mellom de to gruppene og hver av kontrollvariablene. Dette vil si at analysen multipliserer "oppgave" med f.eks. karakter. Den uavhengige variabelen "oppgave" sier at deltakerne enten utfører oppgave 1 med bare relevant informasjon eller oppgave 2 med både relevant og irrelevant informasjon. Dette er dermed en dummy. Hvis man lager interaksjonsledd med informasjon og f.eks. karakter betyr det at hvilken type informasjon man får avhenger av karaktersnittet man har. Dette gir ingen mening i vår oppgave og vi har valgt å utelate kontrollvariablene helt fra hypotese 1.

3.5.2 ANOVA between subjects

I hypotese 2 ønsker vi å finne ut om tidspress-gruppen er mer effektiv enn ikke tidspress-gruppen. Vi sammenligner derfor tiden de to gruppene brukte i oppgave 1. Dette blir dermed et "between subjects"-design. Vi har derfor valgt å bruke "ANOVA one-way between subject"¹¹, da dette er mer hensiktsmessig å bruke enn for eksempel ANCOVA. Fordelen med ANCOVA er at man kan ta hensyn til kontrollvariabler. Problemet med ANCOVA er at en av forutsetningene for denne metoden krever linearitet mellom kontrollvariablene og den uavhengige variabelen (Coakes 2013). Dette kravet blir ikke oppfylt i vår utredning. Kjønn er en binær variabel og vil aldri være lineær. Dette gir da en systematisk "bias" i analysen, som igjen kan påvirke signifikansen i den uavhengige variabelen. Vi mener da at vi kan få problemer av såkalt type-II feil. Det vi vil si at vi aksepterer hypotesen selv om hypotesen er falsk. Vi fant ut at kun to av kontrollvariablene, erfaring og karakter, er tilnærmet lineær med

¹⁰ Vi har valgt å bruke dette engelske uttrykket fordi vi ikke har funnet noen god norsk oversettelse.

¹¹ Dette engelske uttrykket er mer presist enn et oversatt norsk uttrykk.

den avhengige variabelen, som her er tiden. Vi mener derfor at det er mer hensiktsmessig å bruke “ANOVA one-way between subject” (heretter kalt “between-ANOVA”¹²), selv om vi da ikke tar med kontrollvariablene.

Forutsetningene for “between-ANOVA” består av at utvalget er normalfordelt og at det er homogenitet i variansen. Dette er de samme forutsetningene som for “within-ANOVA”. Vi bruker samme test i hypotese 2 for sjekke normaliteten som i hypotese 1, altså Shapiro-Wilks. For sjekke om det er homogenitet i variansen må vi bruke en annen test enn i hypotese 1. Vi har valgt å bruke Levenes test for denne hypotesen. Denne forutsetningen holder hvis Levenes test ikke er signifikant ($p > 0.05$). (Coakes 2013).

3.5.3 SPANOVA

I hypotese 3 har vi et blandet design. Som tidligere nevnt har vi her både “within”- og “between subjects”-design, kombinert. Når vi ser på oppgaveløsningen internt i gruppen så har vi et “within subjects”-design og mellom gruppene har vi et “between subjects”-design. Vi ønsker å finne ut om tidspress påvirker “dilution effect”. Vi har derfor valgt å analysere med SPANOVA. Styrken til denne testen er at vi kan kombinere “within”- og “between subjects”-design i en analyse. Dessuten kan vi ta høyde for kontrollvariablene.

SPANOVA bygger samme forutsetninger som “within-ANOVA”. I tillegg har vi en forutsetning om homogenitet i interkorrelasjonen (Coakes 2013). Denne blir testet med bruk av Box M test. Denne forutsetningen holder også hvis testen ikke er signifikant ($p > 0.05$).

4. Resultater og tolkning

I dette kapitlet ser vi først på den deskriptiv statistikken for å se hva slags utvalg vi hadde. Deretter ser vi nærmere på manipulasjonsjekken. Her har vi undersøkt om deltakerne “fikk pillen”, altså om vi klare å manipulere variabelen. Til slutt ser vi på hvilke resultater vi har fått og tolke disse.

¹² Vi bruker et engelsk uttrykk fordi denne analysen bygger på «between subjects»-design, der vi også har brukt det engelske uttrykket.

4.1 Deskriptiv statistikk

Det var totalt 52 stykker som svarte på undersøkelsen. Deltakerne i undersøkelsen var studenter på masternivå fra NHH. Det første spørsmålet var om deltakeren har hatt faget Revisjon (BUS426) eller gikk på MRR. Syv stykker svarte nei og ble fjernet fra eksperimentet. Videre svarte 13 stykker så ufullstendig at vi måtte utelate dem fra eksperimentet. 2 stykker svarte delvis og blir tatt med videre der disse dataene kunne brukes. Datagrunnlaget vårt består derfor av svarene til 32 deltakere. Vi har relativt få deltakere i eksperimentet. Dette skyldes at vi utfører en masterutredning med begrenset tilgang til deltakere. Dette er en del av kursopplegget og dermed godkjent av veiledere. Analysene og eksperimentet har blitt gjennomført som om vi har mange nok deltakere. Videre kan dette påvirke resultatene fordi det er vanskelig å få signifikante resultater når man har få deltakere.

De to deltakerne som ikke svarte fullstendig på alle spørsmål blir kun tatt med i analysen av hypotese 2. De resterende 30 svarte fullstendig på alle spørsmål. Dette utvalget består av 15 kvinner og 15 menn. 50% av hvert kjønn er en god blanding. Karaktersnitt ble rapportert på en skala fra F (0) til A (5). Deltakerne kunne legge inn karaktersnitt med en desimal. Gjennomsnittskaraktter for deltakerne ligger på 4.04 av 5. Dette tilsvarer et snitt på litt over B. Laveste karakter er C (3) og høyeste er midt mellom B og A (4.5). Den siste kontrollvariabelen er tidligere revisjonserfaring. Dette ble rapportert på intervallnivå; ingen erfaring, 1-2 års erfaring og 3 år eller mer. De fleste deltakerne hadde ikke erfaring innen revisjon. Dette utgjorde 21 av 30 deltakere. 7 deltakere hadde 1-2 års erfaring. Kun 2 deltakere hadde 3 eller flere år med revisjonserfaring.

Vi valgte å gjennomføre eksperimentet i Qualtrics. Her kunne eksperimentet gjøres helt anonymt og vi kunne randomisere om deltakerne ble utsatt for tidspress eller ikke. Dette foregikk ved at annenhver deltaker som startet undersøkelsen ble utsatt for tidspress. Siden antallet deltakere som svarte ufullstendig ikke var likt for begge gruppene endte vi med 14 deltakere i gruppen uten tidspress og 18 deltakere i gruppen med tidspress. Kjønn ble ikke likt fordelt mellom de to gruppene. Gruppen uten tidspress består av 8 menn og 4 kvinner. Karaktersnittet i denne gruppen er på 4.14. 8 stykker i denne gruppen har ikke tidligere revisjonserfaring, mens 5 stykker har 1-2 års revisjonserfaring. Gruppen med tidspress består

av 8 menn og 10 kvinner. Her var karaktersnittet på 3.97. Vi ser altså at karaktersnittet er litt lavere i gruppen uten tidspress enn gruppen med tidspress. I tidspressgruppen var det 2 deltakere med 3 eller flere års erfaring innenfor revisjon, 3 stykker hadde 1-2 års erfaring og 13 deltakere har ingen revisjonserfaring.

4.1.1 Hypotese 1

Tabell 1: Deskriptiv statistikk for hypotese 1

Deskriptiv statistikk			
	Gjennomsnitt	Standardavvik	N
Oppgave 1	54,000	21,123	12
Oppgave 2	56,083	18,976	12

I hypotese 1 er det deltakerne som ikke ble utsatt for tidspress som analyseres. Tabellen over viser deskriptiv statistikk for denne gruppen. Det er 12 deltakere som ikke ble utsatt for tidspress. De vurderte i oppgave 1 og 2 sannsynligheten for fortsatt drift gitt den informasjonen de hadde fått. I oppgave 1 vurderte denne gruppen i gjennomsnitt at det var 54% sannsynlighet for fortsatt drift. I oppgave 2 vurderte de gjennomsnittlig 56% sannsynlighet for fortsatt drift. Vi ser at både gjennomsnittet og standardavviket ligger nærme hverandre på oppgave 1 og 2.

4.1.2 Hypotese 2

Tabell 2: Deskriptiv statistikk for hypotese 2

Deskriptiv statistikk			
Tid	N	Gjennomsnitt	Standardavvik
Uten tidspress	14	199,189	66,746
Tidspress	18	123,007	56,922
Total	32	156,337	71,550

I hypotese 2 ser vi både på gruppen uten tidspress og gruppen med tidspress. Vi ser kun på oppgave 1. Det er 14 deltakere i gruppen uten tidspress og 18 deltakere i gruppen med tidspress. Her ser vi at de uten tidspress bruker i gjennomsnitt 199 sekunder på oppgave 1 og de med tidspress bruker i gjennomsnitt 123 sekunder på oppgave 1. De som blir utsatt for tidspress bruker altså vesentlig mindre tid på oppgaven. Gjennomsnittet for begge gruppene er 156 sekunder på oppgave 1.

4.1.3 Hypotese 3

Tabell 3: Deskriptiv statistikk for hypotese 3

Deskriptiv statistikk				
Tidspress		Gjennomsnitt	Standardavvik	N
Oppgave 1	Uten tidspress	54,00	21,12	12
	Tidspress	57,00	16,58	18
	Total	55,80	18,24	30
Oppgave 2	Uten tidspress	56,08	18,98	12
	Tidspress	56,50	21,78	18
	Total	56,33	20,36	30

I hypotese 3 bruker vi data fra begge gruppene og begge oppgavene. Det er totalt 30 deltakere. På oppgave 1 er gjennomsnittlig vurdering for begge gruppene på 55.8% sannsynlighet for fortsatt drift. Gruppen uten tidspress har et gjennomsnitt på 54%, mens gruppen med tidspress har et gjennomsnitt på 57%. Standardavviket ligger mellom 16 og 21. På oppgave 2 ligger gjennomsnittet for begge gruppene på 56% sannsynlighet for fortsatt drift. Gjennomsnittet for gruppen uten tidspress er 56.08% og for gruppen med tidspress er den 56.5%. Vi ser at gruppene ligger veldig tett opp til hverandre på oppgave 2. Standardavviket ligger mellom 18 og 21. Dette er noe høyere enn på oppgave 1, men likevel relativt like hverandre.

4.2 Manipulasjonssjekk

4.2.1 Tidspress

I etterkant av eksperimentet testet vi om vi har påført de manipulasjonene vi ønsket. Tidspress er en meget subjektiv variabel. Det som oppleves som kort tid på en oppgave for noen kan være mer enn nok tid for andre. I etterkant av eksperimentet spurte vi om deltakerne hadde følt tidspress. Dette ble bare gjennomført for de som fikk denne “treatmenten”, altså tidspressgruppen. 17 av 18 deltakere svarte på dette spørsmålet. Av de som svarte var det 88.2% (15 av 17) som følte tidspress og 11.8% (2 av 17) følte ikke et tidspress. Vi mener dette er et godt resultat som viser at tidsbegrensingen og klokken som teller ned i eksperimentet førte til at de fleste følte et tidspress. Her mener vi deltakerne “fikk pillen”.

4.2.2 Informasjon

Vi testet i etterkant av eksperimentet hva deltakerne rangerte de ulike påstandene i casetaksten som. De kunne velge mellom (1) relevant og positivt for bedriftens evne til fortsatt drift, (2) irrelevant og (3) relevant og negativt for bedriftens evne til fortsatt drift. I tabellen under har i listet opp alle påstandene i alfabetisk rekkefølge. N gir hvor mange som svarte på spørsmålet. Gjennomsnittet blir regnet ut ved å kode relevant og negativt som -1, irrelevant som 0 og relevant og positivt som 1. Vi har lagt inn hva som er rett svar og hvor mange, i antall deltakere, som svarte de forskjellige alternativene. Det er lett å se om de fleste klarte å skille ut hva som er relevant og ikke. For de påstandene som er irrelevante vil vi ha et gjennomsnitt så nærme 0 som mulig.

Tabell 4: Resultater av manipulasjonsjekk for informasjon.

	N	Gj.snitt	Std.	Rett svar	Relevant & negativt	Irrelevant	Relevant & positivt
Alle revisjonsbevis som har blir forespurt av revisorteamet, har blitt forberedt og forelagt av klienten innen rimelig tid	27	0,44	0,506	Irrelevant	0	15	12
Høyere innkjøpskostnader har ført til tap av viktig kunde	27	-0,93	0,385	Relevant, negativt	26	0	1
I inneværende år har selskapet pådratt seg negativ kontantstrøm fra driften	27	-0,85	0,456	Relevant, negativt	24	2	1
Klienten har periodiske opptelling av varelageret for å korrigere feil	27	0,41	0,501	Irrelevant	0	16	11
Ledelsen består av 50% kvinner	27	0,00	0,277	Irrelevant	1	25	1
Ledelsens kompensasjon har blitt endret i tråd med bedriftens økte fokus på å treffe budsjetterte mål	27	0,33	0,734	Irrelevant	4	10	13
Lederen for økonomiavdelingen har gått av med pensjon, og har blitt erstattet av en med høyere utdannelse	27	0,22	0,641	Irrelevant	3	15	9
Manglende evne til å betale kreditorer ved forfall	27	-0,93	0,385	Relevant, negativt	26	0	1
Selskapet har innført et rotasjonsprogram internt for de ansatte	27	0,48	0,509	Irrelevant	0	14	13
Selskapet har pådratt et tap som har ført til et underskudd i inneværende- og foregående år	27	-0,93	0,385	Relevant, negativt	26	0	1
Selskapet har restrukturert lånet til en lavere lånerente	26	0,92	0,272	Relevant, positivt	0	2	24
Selskapets gjeldsgrad er høyere enn gjennomsnittet for bransjen	27	-0,93	0,267	Relevant, negativt	25	2	0
Sterk konkurranse har ført til økte markedsføringskostnader og lavere marginer	27	-0,96	0,192	Relevant, negativt	26	1	0

Alle påstandene som er relevante og negative for bedriftens evne til fortsatt drift har fått gjennomsnitt svært nær -1. Gjennomsnittene ligger mellom -0.85 og -0.96. Dette er et høyt nok til at vi kan slå fast at denne informasjonen ble ansett som relevant og negativ av deltakerne. Den eneste påstanden som er relevant og positiv; “Selskapet har restrukturert lånet til en lavere lånerente”, har fått et gjennomsnitt på 0.92. Dette er svært nærme 1, og vi kan fastslå at de aller fleste deltakerne anser denne informasjonen som relevant og positiv.

De irrelevante påstandene er deltakerne mer usikre på. Det eneste de fleste er enige om er at “Ledelsen består av 50% kvinner” er irrelevant. Her var det én deltaker som mente dette var relevant og positivt og én deltaker som mente det var relevant og negativt. Påstanden fikk dermed et gjennomsnitt på 0. De resterende irrelevante påstandene har et gjennomsnitt svakt positivt. “Selskapet har innført et rotasjonsprogram internt for de ansatte” har et gjennomsnitt på 0.48 (std 0.509) og like under halvparten av deltakerne mente det var irrelevant. Denne påstanden ble modifisert fra Shelton (1999) for at den skulle gi mening i vår case og i pretesten fikk den en svak positiv rangering. Den påstanden som færrest mente var irrelevant i manipulasjonssjekken var “Ledelsens kompensasjon har blitt endret i tråd med bedriftens økte fokus på å treffe budsjetterte mål. Kun 10 deltakere mente dette var irrelevant. Gjennomsnittet har likevel blitt 0.33. Dette kommer av at fire deltakere mente dette var relevant og negativt og 13 mente det var relevant og positivt. De positive og negative svarene vil veie opp for hverandre når vi regner ut gjennomsnittet. Derfor er det verdt å merke seg at flest deltakere mente dette var relevant og positivt. Den samme effekten finner vi på påstanden “Lederen for økonomiavdelingen har gått av med pensjon, og har blitt erstattet av en med høyere utdanning”. Her mener flest deltakere at dette er irrelevant (15). 3 deltakere mener dette er relevant og negativt og 9 deltakere mener det er relevant og positivt.

I de fleste tilfeller ser vi at gjennomsnittet ligger nærme det “riktige” gjennomsnittet og mer enn halvparten av deltakerne har svart rett. Kun påstanden om at “Ledelsens kompensasjon har blitt endret i tråd med bedriftens økte fokus på å treffe budsjetterte mål” har flere ment var relevant enn de som mente at det var irrelevant, noe som er motsatt av vår oppfatning på forhånd. Generelt mener vi at deltakerne har forstått hva som var relevant informasjon. Det ser derimot ut til at de fleste ikke forstod hva som var irrelevant. Dermed kan det se ut til at deltakerne ikke hadde nok kunnskap til å “få pillen” på informasjonsvariabelen.

4.3 Presentasjon og tolkning

En viktig begrensning som gjelder for alle hypotesene er at vi kun har et utvalg på 30 deltakere (32 i hypotese 2) og gruppene er relativt små. Dette medfører at dataene ikke alltid er normalfordelt. Dette bryter med normalitetsforutsetningen for ANOVA, både “within” og “between” og for SPANOVA. En annen forutsetningen er at variablene må måles på en kontinuerlig skala. I hypotese 1 og 3 måles den avhengige variabelen, altså sannsynligheten for fortsatt drift, på en 101-poengs skala. I hypotese 2 måler vi tiden deltakerne bruker før de har gjort en vurdering, i sekunder. Denne forutsetningen holder derfor for alle variablene.

4.3.1 Hypotese 1

De videre forutsetningene for “within-ANOVA” som vi bruker i analysen av hypotese 1 er homogenitet i variansen og “Sphercity”. Det første vi tester er likevel om dataene vi bruker i analysene er pålitelige gjennom “Cronbach alfa”-testen. Vi fant her at dataene er pålitelige. Vi fant også at forutsetningen om “Sphercity” holder. For å se om forutsetningen for at det er homogenitet i variansen holder, regner vi ut F-max. Dette gir da F-max på 1.24, noe som er lavere enn 3. Dette tilsier da at forutsetningen om homogenitet i variansen holder. Vi sjekket om variablene vi bruker i analysene av hypotese 1 var normalfordelte. Her fant vi at de uavhengige variablene er normalfordelt, mens ingen av kontrollvariablene er normalfordelt. Shapiro-Wilks signifikansnivået for de uavhengige variablene er på 0.906 og 0.499 og dermed høyere enn 0.05. Disse er dermed normalfordelte. Vi går dermed videre med analysen. I tabellen under ser vi resultatene av analysen. Den er kjørt med et konfidensnivå på 95%.

Tabell 5: Resultater av hypotese 1, med bruk av «Within-ANOVA».

Tests of Within-Subjects Effects								
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power ^a
Oppgave	Sphericity Assumed	26,042	1	26,042	0,488	0,499	0,043	0,098
	Greenhouse-Geisser	26,042	1	26,042	0,488	0,499	0,043	0,098
	Huynh-Feldt	26,042	1	26,042	0,488	0,499	0,043	0,098
	Lower-bound	26,042	1	26,042	0,488	0,499	0,043	0,098
Error(info)	Sphericity Assumed	586,458	11	53,314				
	Greenhouse-Geisser	586,458	11	53,314				
	Huynh-Feldt	586,458	11	53,314				
	Lower-bound	586,458	11	53,314				
a. computing using alpha =.05								

For at vi skal forkaste nullhypotesen $H_0 = \text{gjennomsnitt oppgave 1} = \text{gjennomsnitt oppgave 2}$ må signifikansnivået være over 0.05 siden vi kjører analysen med et konfidensnivå på 95%. Her har vi fått et signifikansnivå på 0.499 og vi blir nødt til å beholde nullhypotesen. Vi kan ikke si det er en signifikant forskjell mellom gruppene. Vi finner med andre ord ikke en “dilution effect” i dette eksperimentet. “Partial Eta Squared”¹³ forteller hvor mye av variansen som har blitt forklart. Her ligger den på 0.043. Dette er relativt lavt og har sammenheng med at det ikke finnes noen signifikant forskjell mellom oppgavene.

“Observed Power”¹⁴ er oppgitt som 1-b. b= sannsynligheten for en type II feil. En type II feil er at man feilaktig godtar en nullhypotese. Man gjør en type II feil hvis man konkluderer med

¹³ Dette uttrykket er på engelsk fordi det ikke finnes en god norsk oversettelse. Vi ønsker å beholde de uttrykkene som brukes i analyseprogrammet.

¹⁴ Dette uttrykket er på engelsk fordi det ikke finnes en god norsk oversettelse. Vi ønsker å beholde de uttrykkene som brukes i analyseprogrammet.

at nullhypotesen er sann, selv om den egentlig er falsk. (University of Idaho, no date). Vi ønsker derfor en høy “Observed Power” fordi det betyr at sannsynligheten for å godta en falsk nullhypotese er liten. “Observed Power” er i dette tilfellet 0.098, noe som er relativt lavt. Dette betyr at det er stor sannsynlighet for at vi godtar en falsk nullhypotese. Hadde forskjellen vært signifikant og vi hadde hatt en lav “Observed Power” ville dette gitt grunn til videre analyse. Vi har både en ikke signifikant forskjell og en lav “Observed Power”. Dette betyr at det er stor sannsynlighet for at vi har gjort en type II feil. Denne typen feil innebærer at man har forkastet en alternativ forklaring, selv om den egentlig stemmer. (Trochim & Donnelly 2008). I statistikken er type I og type II feil gjensidig utelukkende, har man liten sannsynlighet for type I feil vil man ha høy sannsynlighet for type II feil og motsatt. I analyser vil man heller ha stor sannsynlighet for type II feil enn type I feil (Trochim & Donnelly 2008). Det er mer alvorlig å akseptere en alternativ forklaring som ikke stemmer enn å forkaste en alternativ forklaring, selv om den stemmer.

Vi har valgt å dobbeltsjekke dataene med bruk av T-test, og dette gir samme signifikansnivå som i analysen ovenfor, med et signifikansnivå på 0.499. I tabellen over er det oppgitt F-verdi. Dette vil ikke være relevant å bruke i våre analyser fordi vi ikke har flere enn én hypotese. Derfor har vi utført en t-test i stedet for.

Tabell 6: Resultater av hypotese 1, med bruk av sammenkoblede T-test (Paired T-test).

T-test for Sammenkoblede utvalg									
		Sammenkoblede forskjeller							
		Gjennomsnitt	Standardavvik	Standardfeil	95% Konfidensintervall				
					Nedre	Øvre			
Par 1	Oppgave 1, uten tidspress - Oppgave 2, uten tidspress	-2,083	10,326	2,981	-8,644	4,478	-0,699	11	0,499

Her ser vi at gjennomsnittlig forskjell mellom oppgave 1 og 2 er -2 prosentpoeng. Standardavviket er 10.3 og standardfeilen er 2.98. Konfidensintervallet ligger mellom -8.64 og 4.48. Det vil si at 95% sannsynlighet for at den sanne forskjellen mellom oppgavene ligger

mellom -8.6 og 4.5. Den kritiske verdien for t i denne testen (to-halet og 95% konfidensnivå med 11 frihetsgrader) er 2.2. Analysen viser at vi har en t-verdi på -0.699. Denne er altså for lav til å være signifikant. Dette støtter konklusjonen over og vi må beholde nullhypotesen.

Generelt ser vi at vi ikke får støtte for hypotese 1. Dette kan skyldes at vi har relativt få deltakere. Disse resultatene kan vi se i sammenheng med manipulasjonssjekken. I resultatene har vi funnet at deltakerne mener oppgave 1 og 2 gir den samme relevante informasjonen, noe den også gjør. I manipulasjonssjekken fant vi at deltakerne ikke mente den irrelevante informasjonen var irrelevant, de mener altså at denne informasjonen er relevant. Vi har dermed ikke funnet “dilution effect” i den forstand at deltakerne lar beslutningen bli påvirket av irrelevant informasjon. *Vi har funnet at deltakerne ikke lar beslutningen bli påvirket av det de mener er relevant informasjon.* Dette er meget interessant. Vi tror dette kommer av at deltakerne ikke hadde nok kunnskap og erfaring rundt fortsatt drift-vurderingen. Dermed er de ikke sikre på hva som er relevant og irrelevant i en slik sammenheng.

4.3.2 Hypotese 2

I hypotese 2 har vi valgt å se på om de med tidspress bruker mindre tid på å løse oppgavene enn de uten tidspress. Dette gjør vi ved å sammenligne første oppgave mot hverandre mellom gruppene.

Først undersøker vi om alle forutsetningene er overholdt. Dataene er pålitelige i forhold til “Cronbach alfaen” og “Sphericity”-forutsetningen holder. Shapiro-Wilks normalitetstest viser at tiden brukt av gruppen uten tidspress er normalfordelt, mens tiden brukt av gruppen med tidspress ikke er normalfordelt. Dette gir en begrensning. Vi vet ikke hvor stor påvirkning dette har på påliteligheten til resultatet, men dette kan være en relativ stor begrensning.

Tabell 7: Resultater av hypotese 2, med bruk av «Between-ANOVA».

ANOVA					
Tid	Sum of Squares	df	Mean Square	F	Sig.
Between	45703,610	1	45703,610	12,134	0,002
Within	112997,181	30	3766,573		
Total	158700,791	31			

ANOVA-analysen viser at det er en signifikant forskjell mellom gruppene for oppgave 1. Dette ga et signifikansnivå på 0.002 ($p < 0.05$). Vi legger også merke til at denne hypotesen også er signifikant på et 99% konfidensnivå. Analysen sier ikke noe om hvilken gruppe som er raskest. Utfra den deskriptive statistikken vi har sett på tidligere vet vi at det er gruppen som er utsatt for tidspress som svarte raskest. Tidspress-gruppen ligger i gjennomsnitt på 123 sekunder og uten tidspress-gruppen ligger i gjennomsnitt på 199 sekunder.

Tidspress-gruppen fikk 200 sekunder på å gjennomføre oppgaven og kunne dermed i teorien brukt like lang tid på vurderingen som de uten tidspress endte opp med å bruke. Dette tyder på at tidspress-gruppen følte seg presset og antagelig stresset. Noe som førte til at de ikke brukte all tiden de fikk tildelt. Analysene støtter hypotese 2, de som ble påført tidspress gjennomførte oppgave 1 raskere enn de som ikke fikk påført tidspress.

Vi velger også her å dobbeltsjekke resultatet med bruk av t-test, og dette gir samme signifikansnivå som i «Between-ANOVA» analysen ovenfor.

Tabell 8: Resultater av hypotese 2, med bruk av uavhengig T-test (Independent T-test).

		Levens test for like varianser		T-test for like gjennomsnitt						
		F	Sig.	t	df	Sig. (2-halet)	Gj.sn. forskjeller	Std. feil forskjeller	95% Konfidensintervall	
									Nedre	Øvre
Tid	Lik varians forutsatt	1,073	0,309	3,483	30	0,002	76,182	21,870	31,517	120,846
	Lik varians ikke forutsatt			3,413	25,6	0,002	76,182	22,321	30,266	122,098

Vi forutsetter som sagt at homogeniteten i variansen holder. Gjennomsnittlig løste gruppen med tidspress oppgaven 76 sekunder raskere enn gruppen uten tidspress. Standardfeilen ligger på 21.87. Konfidensintervallet ligger mellom 31.517 og 120.846, noe som forteller oss at det er 95% sannsynlighet at den sanne verdien ligger i dette intervallet. Den kritiske verdien for t i denne testen (to-halet og 95% konfidensnivå med 30 frihetsgrader) er 2.04. Vi fikk en t-verdi på 3.483, noe som er høyere enn den kritiske verdien og er derfor signifikant, med et signifikansnivå på 0.002. Signifikansnivået er det samme som vi fikk i ANOVA-testen.

Disse resultatene kan sees i sammenheng med manipulasjonsjekken. Her har deltakerne “fått pillen” og vi har fått signifikante resultater. Her har vi fått bekreftet hypotesen og denne stemmer emd tidligere teori.

4.3.3 Hypotese 3

I analysen av hypotese 3 har vi valgt å bruke SPANOVA. Denne analysemetoden har samme forutsetningene som “within-ANOVA”, men har i tillegg en forutsetning om homogenitet i interkorrelasjonen. Forutsetningene som er oppfylt i hypotese 1, vil også være oppfylt her, fordi vi bruker det samme datagrunnlaget, men legger til en gruppe deltakere til. Vi ser derfor på om forutsetningene også er oppfylt i gruppen med tidspress.

Normalitetstesten viser at alle de uavhengige variablene er normalfordelt, unntatt oppgave 2 for gruppen med tidspress. Normalitet testes med Shapiro-Wilks testen og hvis denne testen er

signifikant er variablene ikke normalfordelt. Oppgave 2 i tidspress-gruppen har et signifikansnivå på 0.025, noe som er mindre enn 0.05 og dermed signifikant og ikke normalfordelt. Kontrollvariablene var ikke normalfordelt i hypotese 1 og dette vil også gjelde her. Forutsetningen om “Sphericity” og pålitelighet gjennom “Cronbach alfa” er oppfylt. Når det gjelder forutsetningen om homogenitet i variansen så holder denne forutsetningen for alle fire oppgavene. Ingen av disse var signifikant med Levenes test.

Forskjellig fra hypotese 1 må vi også ta hensyn til homogenitet i interkorrelasjonen i SPANOVA. Dette testes med Boxs M. Denne forutsetningen holder, da den ikke er signifikant med et signifikansnivå på 0.364 ($p > 0.05$).

Vi er i analysen hovedsakelig opptatt av om “dilution effect” påvirkes av tidspress. “Dilution effect” måles som forskjellen mellom oppgave 2 og oppgave 1. Analysen gjennomføres med et konfidensnivå på 95%. SPANOVA-analysen gir følgende resultater:

Tabell 9: Resultater av hypotese 3, med bruk av SPANOVA

Tests of Between-Subjects Effects							
Transformed variable average							
Source	Type III Sum of squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power ^a
Intercept	80,241	1	80,241	0,115	0,737	0,115	0,062
Kjønn	6,569	1	6,569	0,009	0,924	0,009	0,051
Karakter	433,332	1	433,332	0,621	0,438	0,621	0,118
Erfaring	1799,284	1	1799,284	2,577	0,121	2,577	0,339
Tidspress	71,923	1	71,925	0,103	0,751	0,103	0,061
Error	17455,204	25	698,208				

a. computing using alpha =.05

I resultatene kan vi fokusere på radene kjønn, karakter, erfaring og tidspress. Her er kontrollvariablene inkludert i analysen. Første raden omhandler kontrollvariablene kjønn. Her vil nullhypotesen være at kvinner og menn i gjennomsnitt svarer det samme. Kjønn har et

signifikansnivå på 0.924 noe som er dramatisk større enn 0.05. Denne er dermed ikke signifikant og vi må beholde nullhypotesen. “Partial Eta Squared” er her 0.000. Denne variabelen forklarer altså ikke noe av variansen i forskjellen mellom gruppene. “Observed Power” ligger på 0.051 noe som betyr at sannsynligheten for at vi godtar en falsk hypotese er relativt stor. Dette er likevel bra fordi vi ønsker en lav sannsynlighet for at man aksepterer en alternativ forklaring som ikke stemmer.

Videre ser vi på karaktersnitt. Nullhypotesen vil her være at karaktersnitt ikke påvirker forskjellen mellom tidspress- og ikke tidspress-gruppen. Signifikansnivået på karaktersnitt ligger på 0.438. Dette er høyere enn 0.05 og denne variabelen er ikke signifikant. Vi må dermed beholde nullhypotesen. “Partial Eta Squared” er 0.024. Dette betyr at denne variabelen forklarer 2.4% av variansen i den avhengige variabelen. “Observed Power” er 0.118, noe som også er relativt lavt. Dette betyr at sannsynligheten for en type II feil er stor, mens sannsynligheten for en type I feil er lav.

Den siste kontrollvariabelen er erfaring. Nullhypotesen vil her være at erfaring ikke påvirker forskjellen mellom tidspress- og ikke tidspress-gruppen. Signifikansnivået ligger på 0.121 og vi må derfor beholde nullhypotesen. Dette er likevel den mest signifikante variabelen vi har. Vi mener at denne variabelen viser en tendens til at erfaring påvirker forskjellen mellom gruppene, men vi må beholde nullhypotesen. “Partial Eta Squared” ligger på 0.093, det vil si at 9.3% av variansen mellom gruppene forklares av erfaring. Dette er den variabelen som forklarer mest av variansen. “Observed Power” ligger på 0.339. Det er dermed fortsatt relativt stor sjans for å feilaktig godta en falsk nullhypotese. I dette eksperimentet er 0.339 relativt høyt.

Til slutt har vi den viktigste uavhengige variabelen, tidspress. Vi ønsker å undersøke om tidspress påvirker “dilution effect”. Nullhypotesen vil være at tidspress ikke gir noen forskjell mellom gruppene. Signifikansnivået ligger på 0.751 og dermed dramatisk større enn 0.05. Vi blir dermed nødt til å beholde nullhypotesen. “Partial Eta Squared” er 0.004, noe som betyr at 0.4% av variansen mellom gruppene forklares av tidspress. Dette er lavt og henger sammen med den lave signifikansen. “Observed Power” er 0.061 og vi har dermed lav sannsynlighet for type I feil, men desto større sannsynlighet for type II feil.

Generelt viser disse resultatene at vi ikke har støtte for hypotese 3. Denne hypotesen skal undersøke om tidspress er en løsning på “dilution effect” og bygger mye på hypotese 1. Det vil uansett være vanskelig å finne en løsning på et problem når man ikke har funnet problemet. Vi fant ingen “dilution effect” i hypotese 1 og vil da aldri kunne få signifikante resultater i den retningen vi ønsker i hypotese 3. Når vi ser denne hypotesen i sammenheng med manipulasjonssjekken får vi samme resultater som i hypotese 1. Deltakerne mener den irrelevante informasjonen er relevant, men lar ikke vurderingen bli påvirket av denne informasjonen likevel.

5. Konklusjon

5.1 Resultater

Gjennom denne masterutredningen har vi gjennomført et eksperiment som ser på “dilution effect” og tidspress. Problemstillingen vår var;

1. Hvordan påvirkes vurderinger av hvilken type informasjon revisor får og hvordan påvirkes denne sammenhengen av tidspress?
2. Hvordan påvirkes vurderinger av hvor mye tid revisor får?

Vi undersøkte første del av problemstillingen gjennom hypotese 1 og hypotese 3. I resultatene får vi ikke støtte for hypotese 1. Dette betyr at vi ikke fant noen “dilution effect”. Dette betyr imidlertid ikke at denne effekten ikke finnes. Fenomenet er godt bevist i litteraturen. Det kan være flere grunner til at vi ikke finner det samme fenomenet her. Eksperimentet vårt har relativt få deltakere. De fleste deltakerne har ingen revisjonserfaring og de som har erfaring har lite erfaring. Vårt inntrykk er at det gjerne er ansvarlig revisor på revisjonsteamet som gjør fortsatt drift-vurderingen. Derfor kan det være at studentene som deltar i vårt eksperiment ikke har noe særlig erfaring med fortsatt drift-vurderingen. I selve eksperimentet ble det først oppgitt den relevante informasjonen i oppgave 1 og deretter den irrelevante informasjonen i oppgave 2. Det kan være at deltakerne gjennomskuet at informasjonen i oppgave 2 var irrelevant fordi all den nye informasjonen de fikk oppgitt her var irrelevant. Gjennom manipulasjonssjekken oppdaget vi at deltakerne ikke var enige med oss om hva som var irrelevant informasjon. Dette burde dra i retning av at sannsynligheten for oppgave 1 og 2

burde være forskjellig. Likevel finner vi ingen signifikant forskjell mellom sannsynligheten i de to oppgavene.

Hypotese 2 undersøker sammenhengen mellom tidspress og vurderingen som ble gjort. Helt konkret så vi på om de som ble utsatt for tidspress utførte oppgaven raskere enn de som ikke ble utsatt for tidspress. Denne hypotesen fikk vi støtte for i resultatene. De som blir utsatt for tidspress utfører oppgaven raskere enn de andre. Vi mener disse deltakerne antagelig føler et visst press og blir stresset til å prosessere informasjonen raskere enn de andre. Vi ser likevel at de kunne brukt like lang tid i denne gruppen som uten tidspress-gruppen brukte i gjennomsnitt. Det var derfor ikke nødvendig å bli stresset, men det skjedde likevel. Til slutt ønsket vi å finne ut om tidspress kunne være en løsning på “dilution effect”. Denne hypotesen er dermed basert på at vi kan forkaste nullhypotesen i H1. Siden vi ikke fant en “dilution effect” er det vanskelig å kunne finne en løsning. Med andre ord er det vanskelig å teste ut en løsning når man ikke har funnet problemet. Denne hypotesen finner vi ikke støtte for. Man kan likevel ikke avskrive dette som en løsning på “dilution effect”.

5.2 Begrensninger med oppgaven

Vi ser at det finnes flere begrensninger ved denne masterutredningen. Den første er at det er NHH studenter innen revisjon og ikke utøvende revisorer som er deltakere i eksperimentet. Mange av deltakerne mangler dermed den praktiske erfaringen innen revisjon. Alle deltakerne har grunnleggende kunnskaper om revisjon og studerer på et masternivå. Dette vil dermed kunne begrense oppgaven, men det er relativt vanlig å utføre eksperimenter på studenter. Noe av grunnen til dette kan være at de er mer tilgjengelige enn utøvende revisorer.

I eksperimentet hadde vi 30 deltakere. Dette er relativt få deltakere, men i en masterutredning er det tilstrekkelig. Grunnen til at vi hadde så få deltakere er at vi har begrenset tilgang til deltakere som oppfyller vårt krav om grunnleggende revisjonskunnskap. En av forutsetningene for analysene er normalfordeling. Dette er enklere å overholde hvis man har mange deltakere. Vi har likevel kjørt de samme analysene som vi ville brukt dersom vi hadde hatt normalitet. En av kontrollvariablene, kjønn, er en binær variabel og kan aldri bli normalfordelt. Deltakerne ble delt i to grupper, noe som førte til at gruppene ble svært små.

Datagrunnlaget som brukes i hypotese 1 består dermed av svaret til kun 12 deltakere og hva hver enkelt av dem svarer vil da ha forholdsvis mye å si på resultatene.

Tidspress ble i denne undersøkelsen innført som en tidsbegrensning. Dette er en vanlig måte å innføre tidspress på i forskningen. Asare, Trompeter og Wright (2000) påpeker likevel at dette ikke er vanlig i praksis. I praksis er det vanligere med tidsbudsjetter der det vil være mulig å overskride et budsjett. Vår oppgave tillater ikke å bruke mer tid enn det man får tildelt i tidspress-gruppen. I tillegg ser vi etter eksperimentet er gjennomført, at de uten tidspress i gjennomsnitt brukte like lang tid på oppgave 1 som tidspress-gruppen fikk lov til å bruke. Vi mener at vi kunne satt tidsbegrensningen til tidspress-gruppen enda lavere. Likevel ser det ut til at deltakerne følte seg stresset og de aller fleste rapporterer i manipulasjonssjekken at de følte et tidspress. Hovedformålet med tidspress er dermed nådd.

Generelt i eksperimentelt design utsetter vi deltakerne for urealistiske settinger. Vår oppgave er begrenset til å handle om fortsatt drift-vurderingen. I tillegg får deltakerne betraktelig mindre informasjon enn man vil få i en ekte revisjon. Dermed kan ikke resultatene våre generaliseres til en populasjon.

5.3 Forslag til videre forskning

Vi mener JDM-forskning er en viktig gren av forskningen og det kommer den til å være videre framover også. Revisjon som fag har fått økt fokus etter finanskrisen. Det er tydelig at revisjonen må ha en meget høy kvalitet for å avdekke hvilke selskaper som har vesentlige feil eller mangler i regnskapet. Revisjon er gjennom dette en viktig samfunnsoppgave.

Hovedtema for oppgaven vår er “dilution effect”. Dette fenomenet har man ikke fullt ut klart å forklare eller løse. Det er flere alternative løsninger som er lagt på bordet, men disse er gjerne kun undersøkt i én studie. Vi ser et stort rom for flere studier som undersøker erfaring og tidspress som løsninger. Dette bør gjerne gjøres under andre omstendigheter eller gjennom andre revisjonshandlinger enn de som er valgt til nå. Da kan man styrke disse løsningene vesentlig. Det har blitt foreslått noen forklaringer på hvorfor en “dilution effect” oppstår, men ingen av de studiene retter seg mot revisjon. Dette er et felt som er utforsket. Videre forskning kan gjerne bruke utøvende revisorer i forskningen, da det er vist at erfaring kan ha påvirkning på “dilution effect”.

Et annet videre forskningsområde er hvordan tidspress påvirker vurderingene. I større studier vil det være mulig å dele deltakerne i forskjellige tidsbegrensninger, slik at noen nesten ikke rekker å lese gjennom mens andre får mindre press. I tillegg kan man gjøre studier der man bruker tidsbudsjetter, ikke tidsbegrensninger. Dette ville vært mer likt praksis i revisjonsbedrifter i dag.

5.4 Betydning for praksis

Siden vi ikke får signifikante resultater i hypotese 1 og 3 er det vanskelig å anbefale endringer i praksis ut fra denne studien. Hypotese 2 fikk signifikante resultater. Dette tyder på at revisorer arbeider raskere under tidspress enn uten. Vi har inntrykk av at de fleste revisjonsselskaper utøver et visst tidspress på sine ansatte. Teorien om tidspress som er presentert i denne studien gir revisjonsselskaper enda en grunn til å gjennomføre revisjoner med et tidsbudsjett. I tillegg til at tidsbudsjetter er et godt verktøy for kontroll av tiden som brukes på et oppdrag. Humankapital er den viktigste ressursen i revisjonsselskaper og hvor mye som brukes av denne ressursen kan måles i tid. Teorien vi har brukt og resultatene i denne studien peker i retning av at revisorer er mer effektive når de jobber under tidspress.

Generelt kan teorien om “dilution effect” som vi har gått gjennom her gi flere retningslinjer for praksis. “Dilution effect” vil antagelig fortsette å være et problem. Teorien trekker fram at hos mer erfarne revisorer finner man ikke foreløpig en “dilution effect”. Dette fremmer at mer erfarne revisorer skal gå gjennom eller godkjenne arbeidet til de mindre erfarne revisorene. Tidspress bør man fortsette å påføre revisorer. Man vil få mer effektive revisorer og en lavere grad av “dilution effect”. Langtidsvirkningene av å jobbe under tidspress hele tiden bør imidlertid tas med i vurderingen når man pålegger sine ansatte slike arbeidsforhold.

6. Referanser

Asare, S.K., Trompeter, G.M. & Wright, A. (2000). The Effect of Accountability and Time Budgets on Auditors' Testing Strategies. *Contemporary Accounting Research*, vol 17 (nr. 4), 539-560.

Becker, L. A. (1999). General Linear Model (GLM): 2-way, Between-Subjects Designs. Hentet fra: http://www.uccs.edu/lbecker/glm_2way.html

Bonner, S. (2008). *Judgment and Decision Making in Accounting*. (1. Utg). New Jersey: Pearson Prentice Hall.

Choo, F. (1995). Auditors' Judgment Performance under Stress: A Test of Predicted Relationship by Three Theoretical Models. *Journal of Accounting, Auditing & Finance*, ??, 611-641.

Coakes, S. (2013). *SPSS version 20.0 for Windows: Analysis without Anguish*. (1. Utg). Milton: Wiley.

Eilifsen, A., Messier Jr., W. F., Glover, S. M. & Prawitt, D. F. (2014). *Auditing & Assurance services*. (3. Utg). Berkshire: McGraw-Hill Education

Eilifsen, A., Kochetova-Kozloski, N. & Messier Jr., W. F. 2015. Mitigation the Dilution Effect of Non-Diagnostic Information on Auditors' Judgment Using a Frequency Response Mode. Ikke utgitt.

Glover, S. M. (1997). The Influence of Time Pressure and Accountability on Auditors' Processing of Nondiagnostic Information. *Journal Accounting Research*, vol 35 (nr 2), 213-226.

Hackenbrack, K. (1992). Implications of Seemingly Irrelevant Evidence in Audit Judgment. *Journal of Accounting Research*, vol 30 (nr 1), 126-136.

Hoffman, V. B. & Patton, J. M. (1997). Accountability, the Dilution Effect, and Conservatism in Auditors' Fraud Judgments. *Journal of Accounting Research*, vol 35 (nr. 2), 227-237.

International Auditing and Assurance Standards Board (IAASB). (2009). *ISA 570: Fortsatt drift*. Oversatt av Revisorforeningen i 2010. Hentet fra:
<https://www.revisorforeningen.no/globalassets/fag/standarder-og-veiledninger/revisjonsstandardene/isa-570-fortsatt-drift.pdf>

Lund Research Ltd. (2013). Testing for Normality Using SPSS Statistics. Hentet fra:
<https://statistics.laerd.com/spss-tutorials/testing-for-normality-using-spss-statistics.php>

McDaniel, L. (1990). The Effects of Time Pressure and Audit Programme Structure on Audit Performance. *Journal of Accounting Research*, vol 28 (nr 2), 267-285.

Nisbett, R.E., Zukier, H. & Lemley, R. E. (1981). The Dilution Effect: Nondiagnostic Information Weakens the Implications of Diagnostic Information. *Cognitive Psychology*, (nr. 13), 248-277.

Olsson, S.V. (2015, 28. Mai). Så kjønnsdelt er arbeidslivet. Hentet 12.05.2016 fra
<https://www.nrk.no/norge/sa-kjonnsdelt-er-arbeidslivet-1.12382396>

Pallant, J. (2010). *SPSS: Survival Manual*. (4. Utg). Maidenhead: McGraw-Hill Open University Press.

Rau, S. E. & Moser, D. V. (1999). Does Performing Other Audit Tasks Affect Going-Concern Judgments?. *Accounting Review*, vol 74 (nr. 4), 493-508.

Shelton, S. W. (1999). The effects of Experience on the Use of Irrelevant Evidence in Auditor Judgment. *The Accounting Review*, vol 74 (nr 2), 217-224.

Trochim, W. M. K., & Donnelly, J. P. (2008). *The research methods knowledge base*. (3. Utg). Mason: Cengage Learning.

Trotman, K. T. (2001). Design Issues in Audit JDM Experiments. *International Journal of Auditing*, (nr. 5), 181-192.

University of Idaho. (No date). Homework Set 4: Comparing conditions using repeated measures in SPSS. Hentet fra:

<http://www.webpages.uidaho.edu/~bdyre/psyc512/Homework/HW4%20comparing%20repeated%20measures%20in%20SPSS.doc>

Waller, W. S., & Zimbelman, M. F. (2003). A cognitive footprint in archival data: Generalizing the dilution effect from laboratory to field settings. *Organizational Behavior and Human Decision Processes*, (91), 254-568.

Wright, P. (1974). The Harassed Decision Maker: Time Pressures, Distractions and the Use of Evidence. *Journal of Applied Psychology*, vol 59 (nr 5), 555-561.

Zukier, H. (1982). The dilution Effect: The Role of the Correlation and the Dispersion of Predictor Variables in the Use of Nondiagnostic Information. *Journal of Personality and Social Psychology*, vol 43 (nr 6), 1163-1174.

Vedlegg

Vedlegg 1: Eksperimentet

Innledende spørsmål

Tar du MRR eller faget BUS426 (Revisjon) ?

- Ja
- Nei

Kjønn?

- Mann
- Kvinne

Gjennomsnittskarakter fra høyere utdanning? (Der F=0 og A=5)

- Svarer på en kontinuerlig skala med et desimal fra 0.0 til 5.0.

Har du erfaring som revisor?

- Nei
- Ja, 1-2 år
- Ja, 3 år eller mer

Oppgave 1: Tidspress gruppen får 3 min og 20 sek

Du er ansvarlig revisor for Trendsetteren AS og skal foreta en fortsatt-drift-vurdering.

Trendsetteren AS er en liten butikkjede med fem butikker spredt rundt på Sørlandet. De selger interiør, glass og serviser. Sommeren 2014 gikk de ut av Tilbords. Ledelsen i Trendsetteren AS forklarer at de valgte å gå ut av Tilbords for å få mer handlefrihet, både i forhold til vareutvalg og i forhold til utforming av butikken. Alle varene kom tidligere fra Tilbords og etter de gikk ut måtte de ta kontakt med leverandører og få nye avtaler. I tillegg måtte de finne en del nye leverandører. Dette arbeidet viste seg vanskelig og førte til et tap i 2014 og 2015. Enkelte varer har Tilbords enerett på å selge i Norge og de ble dermed ukurante for Trendsetteren AS og måtte selges med tap.

I 2015 etablerer Tilbords nye butikker i Sørlandsområdet, noen av disse er på samme kjøpesenter som Trendsetterens butikker. Dette fører til svært sterk konkurranse.

Trendsetteren har derfor sett seg nødt til å senke prisene og bruke mer på markedsføringen for å konkurrere med de store interiørkjedene i området.

En annen bivirkning av at kjeden gikk ut av Tilbords var at de ikke kunne kjøpe like store kvantum av de enkelte produktene lengre. Dette førte til høyere innkjøpskostnader fra de leverandørene de fikk avtale med.

En av Trendsetterens største kunder er en liten restaurantkjede som også holder til på Sørlandet. De kjøper jevnlig store kvanta av vinglass og tallerkener når de har behov. Når innkjøpskostnadene til Trendsetteren øker kan de ikke lengre tilby samme gode pris til denne restaurantkjeden, de har derfor ikke kjøpt noe siden begynnelsen av 2015 og anses som tapt til konkurrentene.

Trendsetteren AS opplever av og til at de ikke har likviditet til å betale leverandørene sine. Dette fører til at de ikke får levert nye ordrer før de har betalt forrige faktura. I 2015 hadde Trendsetteren AS en negativ kontantstrøm fra driften.

Trendsetteren har samme arbeidsstyrken som før de gikk ut av Tilbords. Dette er lojale mennesker med lang erfaring innen produktgruppene til Trendsetteren. Ledelsen har god tro på framtiden og jobber med restrukturering av bedriften for å tilpasse seg den nye situasjonen. Tidlig i 2015 fikk Trendsetteren refinansiert banklånet. De fikk dermed en lavere rente enn tidligere. Ledelsen har vært i kontakt med nye mulige investorer for å få tak i frisk kapital.

Resultatregnskap			2015	2014
Inntekter			71 235	73 126
Varekostnader			61 150	60 000
Lønn			5700	5700
Avskrivning på butikkinnredning			250	250
Nedskrivning av varelager			1800	900
Markedsføring			1000	800
Husleie			2500	2500
Andre driftskostnader			2730	2700
Sum driftskostnader			75 130	72 850
Driftsresultat			-3 895	276
Gjeldsrenter banklån			425	656
Gjeldsrenter kassekreditt			180	120
Sum totalekostnader			75 735	73 626
Resultat			-4 500	-500
Nøkkeltall			2015	2014
Avanse i %			21.88 %	16.49 %
Bankrente			5 %	8 %
Kassakredittrente			12 %	12 %
Gjeldsgrad			2.91	2.34
Gjennomsnittlig gjeldsgrad			2.04	2.04

Hvilken sannsynlighet mener du det er for fortsatt drift om ett år på en skala fra 0 til 100, der 0 gir ingen drift om ett år og 100 gir garantert drift om ett år?

- Svarer på en kontinuerlig skala fra 0 til 100, ingen desimaler

Oppgave 2: Tidspressgruppen får 1 min og 10 sek

I tillegg til informasjonen i forrige oppgave finner du flere opplysninger når du snakker med ledelsen. Ledelsen i Trendsetteren består av 50% menn. De forteller deg at økonomidirektøren gikk av med pensjon i 2015. De har ansatt en ny økonomidirektør som har bakgrunn som regnskapsfører og utdanning fra NHH.

Revisjonsteamet ditt har hatt et godt samarbeid med ledelsen og andre ansatte i Trendsetteren AS. Alle revisjonsbevis som er forespurt har blitt forberedt og forelagt av klienten innen rimelig tid.

Ledelsens kompensasjonspakke er endret for å øke fokuset på bedriftens budsjetterte mål. Bedriften har laget et rotasjonsprogram for de ansatte, slik at de ansatte kan jobbe i den butikken det er mest behov for dem.

Trendsetteren har kvartalsvise varetellinger for å korrigere feil i varelageret. Det forekommer feil i varelageret hvis varene blir solgt med feil strekkode eller hvis det ikke blir tatt varemottak.

Hvilken sannsynlighet mener du det er for fortsatt drift om ett år på en skala fra 0 til 100, der 0 gir ingen drift om ett år og 100 gir garantert drift om ett år?

- Svarer på en kontinuerlig skala fra 0 til 100, ingen desimaler

Manipulasjonssjekk

Hvordan mener du disse elementene påvirker bedriftens evne til fortsatt drift?

- På alle elementene under kunne deltakerne svare en av tre følgende svar;
 - Relevant og positivt
 - Irrelevant
 - Relevant og negativt
- Alle revisjonsbevis som har blitt forespurt av revisorteamet, har blitt forberedt og forelagt av klienten innen rimelig tid
- Høyere innkjøpskostnader har ført til tap av viktig kunde
- I inneværende år har selskapet pådratt seg negativ kontantstrøm fra driften
- Klienten har periodiske opptellinger av varelageret for å korrigere feil
- Ledelsen består av 50% kvinner
- Ledelsens kompensasjon har blitt endret i tråd med bedriftens økte fokus på å treffe budsjetterte mål
- Lederen for økonomiavdelingen har gått av med pensjon, og har blitt erstattet av en med høyere utdannelse
- Manglende evne til å betale kreditorer ved forfall
- Selskapet har innført et rotasjonsprogram internt for de ansatte
- Selskapet har pådratt et tap som har ført til et underskudd i inneværende- og foregående år
- Selskapet har restrukturert lånet til en lavere lånerente
- Selskapets gjeldsgrad er høyere enn gjennomsnittet for bransjen
- Sterk konkurranse har ført til økte markedsføringskostnader og lavere marginer

Vedlegg 2: Definisjon av variabler

Relevant informasjon - informasjon som er nyttig og behøves for å gjøre vurderinger og ta beslutninger i revisjonen.

Irrelevant informasjon - informasjon som ikke er nyttig eller behøves for å gjøre vurderinger og ta beslutninger i revisjonen.

Dilution effect - Vurderinger og beslutninger blir mindre ekstreme når revisor får både relevant og irrelevant informasjon enn hvis de bare får relevant informasjon.

Tidspress - deltakeren føler stress og vil ha nok tid til å skimme igjennom informasjonen, men ikke prosessere den.

Kjønn - mann eller kvinne.

Karaktersnitt - gjennomsnittlig karakter fra høyere utdanning (høyskole eller universitet, både bachelor- og masternivå).

Erfaring - antall år deltakeren har arbeidet med revisjon.

Vedlegg 3: Original case - Fra Shelton 1999

Relevante elementer

- The company incurred a net loss in the current and immediately preceding year.
- During the current year the company experienced negative cash flows from operating activities.
- The company's debt-to-equity ratio is higher than the industry average.
- Strong competition has led to increasing marketing costs and lower margins.
- Production costs have led to a loss of a principal customer.
- The company was successful in restructuring its loan commitment at a lower borrowing rate.

Irrelevante elementer

- The company's chief internal auditor retired and was replaced by a better-trained individual.
- All supporting schedules requested by the auditing team were prepared and submitted by the client in a timely manner.
- The management compensation package was revised to reflect the company's increased emphasis on achieving budgeted targets.
- The company has begun a management rotation program.
- The client makes periodic counts of physical inventory to correct errors in the individual perpetual inventory records.

Disse elementene er brukt i studien til Shelton.

Kilde: Shelton, S. W. (1999). The effects of Experience on the Use of Irrelevant Evidence in Auditor Judgment. *The Accounting Review*, vol 74 (nr 2), s. 220..