



Effekten av irrelevant informasjon og personlighetstrekk i en mislighetsrisikovurdering

En eksperimentell studie

Marie Andersen & Matthias Kjølbro

Veileder: Jonas Gaudernack

Masteroppgave, regnskap og revisjon

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i regnskap og revisjon 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.

Forord

Denne masterutredningen er en del av masterstudiet i regnskap og revisjon ved Norges Handelshøyskole (NHH) i Bergen. Utredningen er skrevet våren 2023, og utgjør 30 studiepoeng.

Arbeidet med masterutredningen har vært en lærerik og interessant prosess. Vi har fått muligheten til å fordype oss i et aktuelt og spennende tema. Studien benytter eksperimentell metode, noe som har vært helt nytt for oss. Dette har gitt oss verdifulle erfaringer innenfor revisjonsforskning, kvantitativ metode og bruk av SPSS.

Vi ønsker å takke vår veileder Jonas Gaudernack, for nyttige innspill og tilbakemeldinger gjennom vårt arbeid med denne utredningen. Videre vil vi takke Ulf Mohrmann for god hjelp i forbindelse med statistiske analyser. Vi vil også takke alle som deltok i spørreundersøkelsen vår, samt våre medstudenter som deltok i pilot-testingen. Til slutt vil vi rette en stor takk til nær familie, kjæreste og venner for støtte og oppmuntring.

Bergen, juni 2023

Marie Andersen & Matthias Kjølbro

Sammendrag

Revisjonsbransjen står overfor en teknologisk endring, hvor det investeres betydelig i avanserte digitale løsninger. Dette gir revisor tilgang til store mengder data, og evnen til å skille mellom relevant og irrelevant informasjon er dermed avgjørende. Tidligere forskning indikerer at revisor er utsatt for dilution bias, som vil si at revisors vurderinger blir mindre ekstreme når den blir presentert for både irrelevant og relevant informasjon, sammenlignet med kun relevant informasjon. Videre peker forskning på personlighet og kognitive bias i beslutningstaking, i retning av at det er forskjell mellom personlighetstyper i hvor disponert en er for kognitive bias. Dette ønsker vi å studere nærmere gjennom en mislighetsrisikovurdering, og har derfor utarbeidet følgende problemstilling, som er todelt:

1. Er revisor utsatt for dilution bias i en mislighetsrisikovurdering?
2. Har personlighetstrekk en modererende effekt på dilution bias?

Vi har utført et eksperiment med 42 respondenter. Deltakerne ble delt inn i to grupper, hvor begge skulle utføre en mislighetsrisikovurdering. Kontrollgruppen ble kun eksponert for relevant informasjon, mens eksperimentgruppen ble eksponert for både relevant- og irrelevant informasjon. På denne måten ønsket vi å undersøke om revisor er utsatt for dilution bias. I tillegg til dette ble deltakerne bedt om å gjennomføre en Big Five Inventory-2 personlighetstest. Dette for å undersøke personlighetstrekk som moderator for en eventuell dilution bias.

Resultatene våre finner ingen dilution bias. Interessant nok, ser vi en signifikant forskjell mellom gruppene, hvor de som ble utsatt for irrelevant informasjon tok mer ekstreme vurderinger enn de som kun fikk relevant informasjon. Videre observerer vi trender mellom deltakernes risikovurderinger og personlighetstrekk. Disse er imidlertid ikke signifikante, og vi finner dermed ikke personlighetstrekk som en signifikant moderator.

Innholdsfortegnelse

| | |
|--|------------|
| Forord | II |
| Sammendrag | III |
| Innholdsfortegnelse | IV |
| 1. Introduksjon | 1 |
| 1.1 <i>Motivasjon for oppgaven</i> | 1 |
| 1.2 <i>Disposisjon for resten av oppgaven</i> | 2 |
| 2. Teoretisk rammeverk | 3 |
| 2.1 « <i>Judgement and decision making research</i> »..... | 3 |
| 2.2 <i>Revisors vurdering av mislighetsrisiko</i> | 3 |
| 2.3 <i>Dilution bias</i> | 4 |
| 2.4 <i>Femfaktormodellen for personlighet</i> | 5 |
| 2.5 <i>Personlighet og kognitive bias</i> | 7 |
| 3. Konseptuell modell og hypoteser | 8 |
| 3.1 <i>Konseptuell modell</i> | 8 |
| 3.2 <i>Hypoteser</i> | 9 |
| 4. Metode | 12 |
| 4.1 <i>Forskningstilnærming</i> | 12 |
| 4.2 <i>Valg av forskningsdesign</i> | 12 |
| 4.2.1 <i>Faktorielt design</i> | 12 |
| 4.2.2 « <i>Within-</i> » og « <i>between-subjects</i> »..... | 13 |
| 4.3 <i>Datainnsamling</i> | 14 |
| 4.3.1 <i>Utvalg</i> | 14 |
| 4.3.2 <i>Utforming av personlighetstest</i> | 15 |
| 4.3.3 <i>Utforming av case</i> | 15 |
| 4.3.4 <i>Pilot-test</i> | 16 |
| 4.3.5 <i>Praktisk gjennomføring</i> | 17 |
| 4.4 <i>Forskningsetikk</i> | 17 |

| | | |
|-----------|---|-----------|
| 4.5 | <i>Analysemetode</i> | 18 |
| 4.6 | <i>Validitet og reliabilitet</i> | 19 |
| 4.6.1 | Validitet | 20 |
| 4.6.2 | Reliabilitet..... | 21 |
| 5. | Resultater | 22 |
| 5.1 | <i>Rensing av data</i> | 22 |
| 5.2 | <i>Manipulasjonssjekk</i> | 22 |
| 5.3 | <i>Deskriptiv statistikk</i> | 24 |
| 5.3.1 | Fordeling mellom gruppene | 24 |
| 5.3.2 | Demografiske variabler..... | 25 |
| 5.4 | <i>Test av hypoteser</i> | 27 |
| 5.4.1 | Hypotese 1..... | 27 |
| 5.4.2 | Hypotese 2..... | 29 |
| 5.4.3 | Hypotese 3..... | 30 |
| 5.4.4 | Hypotese 4..... | 31 |
| 5.5 | <i>Andre observasjoner</i> | 32 |
| 5.5.1 | Fordeling av personlighetstrekk | 32 |
| 6. | Diskusjon og konklusjon | 33 |
| 6.1 | <i>Diskusjon av hypoteser</i> | 33 |
| 6.1.1 | Diskusjon av hypotese 1 | 33 |
| 6.1.2 | Diskusjon av hypotese 2 | 34 |
| 6.1.3 | Diskusjon av hypotese 3 | 35 |
| 6.1.4 | Diskusjon av hypotese 4 | 36 |
| 6.2 | <i>Konklusjon</i> | 36 |
| 6.3 | <i>Begrensinger</i> | 37 |
| 6.4 | <i>Anbefalinger til videre forskningstema</i> | 38 |
| | Litteraturliste | 40 |
| | Vedlegg | 47 |

Figurliste

| | |
|--|----|
| FIGUR 1: KONSEPTUELL MODELL..... | 8 |
| FIGUR 2: FORDELING MELLOM GRUPPENE | 25 |
| FIGUR 3: GJENNOMSNITTSVERDIER HYPOTESE 1 | 28 |
| FIGUR 4: GJENNOMSNITTSVERDIER HYPOTESE 2 | 29 |
| FIGUR 5: GJENNOMSNITTSVERDIER HYPOTESE 3 | 30 |
| FIGUR 6: GJENNOMSNITTSVERDIER HYPOTESE 4 | 31 |

Tabelliste

| | |
|--|----|
| TABELL 1: MANIPULASJONSSJEKK | 23 |
| TABELL 2: DESKRIPTIV STATISTIKK - FREKVENSTABELL | 26 |
| TABELL 3: MANN-WHITNEY U TEST HYPOTESE 1 | 28 |
| TABELL 4: T-TEST HYPOTESE 1..... | 28 |
| TABELL 5: TOVEIS ANOVA HYPOTESE 2..... | 30 |
| TABELL 6: TOVEIS ANOVA HYPOTESE 3..... | 31 |
| TABELL 7: TOVEIS ANOVA HYPOTESE 4..... | 32 |
| TABELL 8: FORDELING AV PERSONLIGHETSTREKK | 32 |

1. Introduksjon

1.1 Motivasjon for oppgaven

På lik linje med resten av verden, står revisjonsbransjen nå overfor et teknologisk skifte. I løpet av de siste årene har de store internasjonale revisjonsselskapene foretatt betydelige investeringer i utvikling og implementering av avanserte digitale løsninger (Eilifsen et al., 2020). Dette medfører at revisor i dag har tilgang til nærmest ubegrensede mengder av data (IAASB, 2017). ISA 500 sier at revisor gjennom revisjonen skal innhente tilstrekkelige og hensiktsmessige revisjonsbevis (IAASB, 2009a). I sitt søk etter revisjonsbevis, skal revisor vurdere *relevansen* til informasjonen som skal brukes til revisjonsbevis (IAASB, 2009a). Revisors evne til å skille mellom relevant og irrelevant informasjon har derfor aldri vært viktigere.

Tidligere psykologisk forskning viser imidlertid at mennesker har en tendens til å redusere betydningen av diagnostisk informasjon, ved tilstedeværelse av ikke-diagnostisk informasjon (Koch & Wüstemann, 2009). Dette begrepet kalles også «dilution bias» eller «dilution effect». Revisjonsforskning viser at også revisor er utsatt for dilution bias (Hoffman and Patton, 1997; Glover, 1997; Shelton, 1999; Hackenbrack, 1992).

I denne oppgaven vil vi benevne begrepet «ikke diagnostisk informasjon» som irrelevant informasjon, og «diagnostisk informasjon» som relevant informasjon. Videre vil vi, på grunn av mangel på et bedre norsk begrep, benevne den overnevnte effekten som «dilution bias» og «dilution effect», og definere den slik:

Et individs vurderinger og beslutninger blir mindre ekstreme når den blir presentert for både irrelevant og relevant informasjon, sammenlignet med kun relevant informasjon

Er det imidlertid slik at alle er like utsatt for dilution bias eller er det forskjeller i hvor disponert man er? Forskning på personlighet og kognitive bias i beslutningstaking innen økonomisk forskning, peker i retning av at det er forskjell mellom personlighetstyper i hvor disponert en er for kognitive bias. Durand et al., (2008, 2013, 2019) og Lin (2011) argumenterte blant annet for at investorers personlighetstrekk påvirker sannsynligheten for kognitive bias i investeringsbeslutninger. Videre er det foreslått at risikovurdering er sterkt forankret i investorers personlighet (Bye & Lamvik, 2007; Engelberg & Sjöberg, 2007;

Nicholson et al., 2005). Dette er forskning som tyder på at et individs personlighet har en modererende effekt på kognitive bias.

Det finnes imidlertid ulike typer kognitive bias. Denne oppgaven tar for seg personlighet og dilution bias, hos populasjonen revisorer, i en mislighetsrisikovurdering. Så langt vi vet finnes det ikke noen forskning på nettopp dette.

På bakgrunn av dette har vi kommet frem til følgende forskningsspørsmål, som er todelt:

1. Er revisor utsatt for dilution bias i en mislighetsrisikovurdering?
2. Har personlighetstrekk en modererende effekt på dilution bias?

1.2 Disposisjon for resten av oppgaven

Masterutredningen består i alt av 6 kapitler. I dette kapitlet har vi tatt for oss valg av problemstilling, samt motivasjon for denne. I kapittel 2 vil vi presentere teori og tidligere forskning vi anser som relevant for å besvare vår problemstilling. Her vil vi først ta for oss oppgavens faglige plassering i forskningsuniverset, deretter redegjøres det for revisors vurdering av mislighetsrisiko. Videre tar vi for oss tidligere forskning om dilution bias, samt teori bak femfaktormodellen. Kapittel 3 tar for seg konseptuell modell, samt utarbeiding av hypoteser. Videre redegjøres og argumenteres det for den samfunnsvitenskapelige metoden som er anvendt i studien i kapittel 4. I kapittel 5 presenteres resultatene, og hypotesetesting blir utført. Hypotesene og andre observasjoner blir videre diskutert i kapittel 6. Avslutningsvis konkluderes det, samt oppgavens begrensinger og forslag til videre forskningstema diskuteres.

2. Teoretisk rammeverk

I kapittel 2 vil vi innledningsvis definere studiens faglige plassering i forskningsuniverset. Videre vil vi ta for oss relevant litteratur og tidligere forskning. Dette omfatter revisors vurdering av mislighetsrisiko, dilution bias, samt femfaktormodellen. Avslutningsvis tar vi for oss tidligere forskning som ser på sammenhengen mellom kognitive bias og personlighetstrekk.

2.1 «Judgement and decision making research»

Oppgavens forskningsspørsmål går under «judgement and decision-making research» innen revisjon, forkortet JDM. Formålet med JDM er å forbedre revisors vurderinger. Revisor står overfor en rekke skjønnsmessige vurderinger, hvor misligheter er en av dem. Betydningen av riktig bruk av skjønn har vært anerkjent i mange tiår (Trotman et al., 2011).

For å forbedre revisors bruk av skjønn, vil det først være nødvendig å forstå hvordan vurderinger gjøres, i vårt tilfelle mislighetsrisikovurderinger. Vi ønsker å undersøke om revisor er utsatt for dilution bias i en mislighetsrisikovurdering, samt sammenhengen mellom revisors personlighetstrekk og en eventuell dilution bias. På denne måten ønsker vi å bidra til å kartlegge og forbedre revisors vurderinger ved misligheter.

2.2 Revisors vurdering av mislighetsrisiko

Revisor er allmennhetens tillitsperson ved lovfestet revisjon, og har som oppgave å bidra til å forebygge og avdekke økonomisk kriminalitet (revisorloven, 2020). PwC (2022) sin globale økonomiske kriminalitets- og mislighetsundersøkelse for 2022, som dekker 53 land og 1296 virksomheter, stadfester at misligheter fortsatt utgjør en stor trussel. 46% av respondentene rapporterte å ha opplevd misligheter eller annen form for økonomisk kriminalitet i løpet av de to foregående årene.

ISA 240 definerer misligheter som: «En bevisst handling begått av en eller flere personer innen ledelsen, av personer som har overordnet ansvar for styring og kontroll, av ansatte eller av andre, som innebærer uredelighet for å oppnå en urettmessig eller ulovlig fordel» (IAASB, 2009b). Videre definerer ISA 240 to typer tilsiktet feilinformasjon som er relevant

for revisor: feilinformasjon som resultat av uredelig regnskapsrapportering, og feilinformasjon som resultat av underslag av eiendeler (IAASB, 2009b). For å redusere sannsynligheten for at misligheter begås, har kriminologen Donald R. Cressey utviklet en modell under betegnelsen mislighetstriangelet. Han hevder at tre faktorer må være til stede for at misligheter skal begås: muligheter, motivasjon og rettferdiggjøring (Moen og Havstein, 2017).

Når revisor skal identifisere og vurdere risikoen for misligheter, vil en vurdering av de tre overstående faktorene være hensiktsmessig. Revisorloven fastsetter hvilke oppgaver og plikter revisor har knyttet til misligheter. Etter §5-1 tredje ledd skal revisor «gjennom revisjonen bidra til å forebygge og avdekke misligheter og feil» (revisorloven, 2020). Videre skal revisor etter §5-2 første ledd «utføre revisjonen etter beste skjønn, herunder vurdere risikoen for at det kan foreligge feilinformasjon i årsregnskapet som følge av misligheter og feil» (revisorloven, 2020).

Hvilke revisjonshandlinger revisor skal foreta seg utdypes videre i revisjonsstandardene. Etter ISA 315 skal revisor utføre risikovurderingshandlinger for å opparbeide seg kunnskap om virksomheten og dens omgivelser, inkludert virksomhetens internkontroll (IAASB, 2019). Dersom internkontrollen i bedriften ikke er tilfredsstillende, kan dette være et eksempel på en mulighet i mislighetstriangelet. Standardens punkt 17 legger opp til at revisjonstemaet skal diskutere i hvilken grad regnskapet kan inneholde vesentlig feilinformasjon (IAASB, 2019). Sentralt i denne diskusjonen er misligheter, og hvordan disse kan oppstå (IAASB, 2009b).

2.3 Dilution bias

I 1974 foreslo Tversky og Kahneman bruken av en heuristikk- og bias tilnærming ved forskning innen JDM. Forståelsesrammen er at mennesker stoler på heuristikker som fungerer som simple tommelfingerregler i sin beslutningstaking. Disse tommelfingerreglene kan være nyttige i det at de tillater en energibesparende beslutningstakingsprosess som i de fleste tilfeller viser seg å være tilstrekkelig. Ulempen er derimot at slike heuristikker kan indusere systematiske feil på grunn av at de er basert på vurderinger som ikke tar høyde for alle variablene i en situasjon. Slike systematiske feil kalles også bias (Koch & Wüstemann, 2009).

Nisbett et al. (1981) og Zukier (1982) fant at mennesker har en tendens til å moderere betydningen av relevant informasjon, ved tilstedeværelse av irrelevant informasjon. Revisorer som presenteres for irrelevant informasjon har en tendens til å moderere sin risikovurdering knyttet til misligheter (Hackenbrack, 1992; Hoffman & Patton, 1997), vesentlige feil (Glover, 1997) og konkurser (Shelton, 1999). En studie av Labella & Koehler (2004) tyder imidlertid på at dilution effect kun har virkning under gitte forhold. De fant at når relevant informasjon vurderes i lys av irrelevant informasjon, blir effekten eliminert, eller til og med reversert. Med andre ord, så kan irrelevant informasjon medføre en såkalt bekreftelseeffekt, hvor den irrelevante informasjonen bekrefter vurderingen foretatt basert på den relevante informasjonen.

Mens tidligere forskning tydelig indikerer at revisor er utsatt for dilution bias, har revisjonsforskning med en preventiv tilnærming vist seg å være begrenset. Det eksisterer lite forskning rundt hvordan revisor kan eliminere eller minimere effekten av dilution bias, og eksisterende forskning har slitt med å finne effektive måter å håndtere effekten på (Eilifsen et al., 2019). Hoffman & Patton (1997), og Glover (1997) fant at det å holde revisor ansvarlig, ikke hadde noen innvirkning på dilution effekt. Videre fant Glover (1997) at tidspress reduserte effekten. Shelton fant i 1999 at revisjonserfaring reduserte effekten. Til slutt har også Eilifsen et al. (2019) funnet at bruken av en «frequency response mode», fremfor en «probability response mode» til å vurdere relevant og irrelevant informasjon, reduserte effekten i en mislighetsrisikovurdering.

2.4 Femfaktormodellen for personlighet

Personlighet kan defineres som et individs karakteristiske tankemønster, følelser og atferd (Funder, 1997; Triandis & Suh, 2002). Trekkteorien anses som en av de viktigste teoriene innen personlighets psykologi (Ewen 2003; Najm, 2019), og forklarer menneskers personlighet gjennom en rekke karakteristika eller trekk. Teorien hevder at sammensetningen av skåren på ulike trekk utgjør personlighetsfungering. Innen trekkteori er femfaktormodellen den mest utbredte og anerkjente (Rammstedt et al. 2010; Najm, 2019).

Teorien om femfaktormodellen ble opprinnelig utarbeidet av Lewis R. Goldberg (Najm, 2019; Goldberg, 1992). Modellen er siden sin opprinnelse blitt videreutviklet og eksisterer i dag i flere varianter og på flere ulike språk (Najm, 2019; Føllesdal & Soto, 2022).

Femfaktormodellen hevder at personlighet kan måles og forklares ut ifra fem stabile dimensjoner, eller trekk. (Najm, 2019; Føllesdal & Soto, 2022). Hver dimensjon består av en bipolar og adjektiv skala og omfatter en rekke subtrekk. Antall subtrekk for hver dimensjon varierer etter versjon (Najm 2019).

Vi vil i denne oppgaven ta for oss versjonen Big Five Inventory-2 (BFI-2), av Oliver P. John og Christopher J. Soto (Føllesdal & Soto, 2022). Her har de definert dimensjonene og tilhørende subtrekk slik:

Omgjengelighet. Dimensjonen måler subtrekkene «medfølelse», «respektfullhet» og «tillit». Dimensjonen blir også referert til som *medmenneskelighet* på norsk, og måler graden av tillit og medfølelse, samt grad av respekt for andre. Individuer som skårer lavt på omgjengelighet vil derfor utvise større grad av fiendtlighet, selvhevdelse og sosial konflikt (Kennair, 2022). Høy og lav skår må ikke forveksles med positivt og negativt. Hvilke trekk som anses som positive og negative vil i stor grad være kontekstavhengig.

Åpensinnhet. Dimensjonen måler subtrekkene «intellektuell nysgjerrighet», «estetisk sensitivitet» og «kreativ fantasi». Individuer som skårer høyt på denne dimensjonen vil typisk beskrives som åpne, med en levende fantasi og et rikt indre liv. Individet tiltrekkes estetikk, typisk uttrykt gjennom ulike kunstarter, og har interesse for betydningen av eget følelsesliv. Videre er slike individer åpne for nye erfaringer og utradisjonelle idéer og verdier (Kennair, 2022; Costa & McCrae, 1992). De som skårer lavt på åpensinnhet vil ofte foretrekke stabilitet og tradisjon fremfor endring, og vil i større grad utvise skepsis.

Samvittighetsfullhet, også ofte referert til som *planmessighet,* måler subtrekkene «ordenssans», «produktivitet» og «pålitelighet». Et individ som skårer høyt på samvittighetsfullhet vil karakteriseres som ryddige, ordentlige, organiserte, produktive og tillitsfulle med høye aspirasjoner (Kennair, 2022).

Ekstraversjon. Dimensjonen måler subtrekkene «sosiabilitet», «selvmarkering» og «energinivå». Med andre ord, så måles det i hvilken grad et individ trives i, får energi av og gjerne oppsøker sosiale sammenhenger. Videre måles det sosial dominans, eller selvmarkering.

Negativ emosjonalitet, også ofte referert til som *nevrotisisme,* måler subtrekkene «angst», «depresjon» og «følelsmessig instabilitet». Med andre ord, så måler dimensjonen graden

av emosjonell ubalanse og psykisk ubehag, men også graden av depressive og engstelige følelser. Individuer som skårer lavt på denne dimensjonen, vil altså fremstå som mer rolige med et stabilt følelsesliv.

2.5 Personlighet og kognitive bias

Forskning innen evolusjonspsykologi viser at kognitive bias ikke er å anse som feil ved det menneskelige sinn, men heller en egenskap, eller et trekk ved sinnet (Haselton et al., 2015). Slike heuristikker og kognitive bias har imidlertid en vesentlig påvirkning på beslutningstaking (Tversky & Kahneman, 1975).

En studie av Kumar et al (2021) tok for seg forholdet mellom kognitive bias og personlighetstrekk fra femfaktormodellen. De fant en signifikant årsakssammenheng mellom personlighetstrekket ekstraversjon og de kognitive bias flokkmentalitet og tapsaversjon. De fant også en signifikant årsakssammenheng mellom trekket åpensinnhet og den kognitive biasen overmodighet.

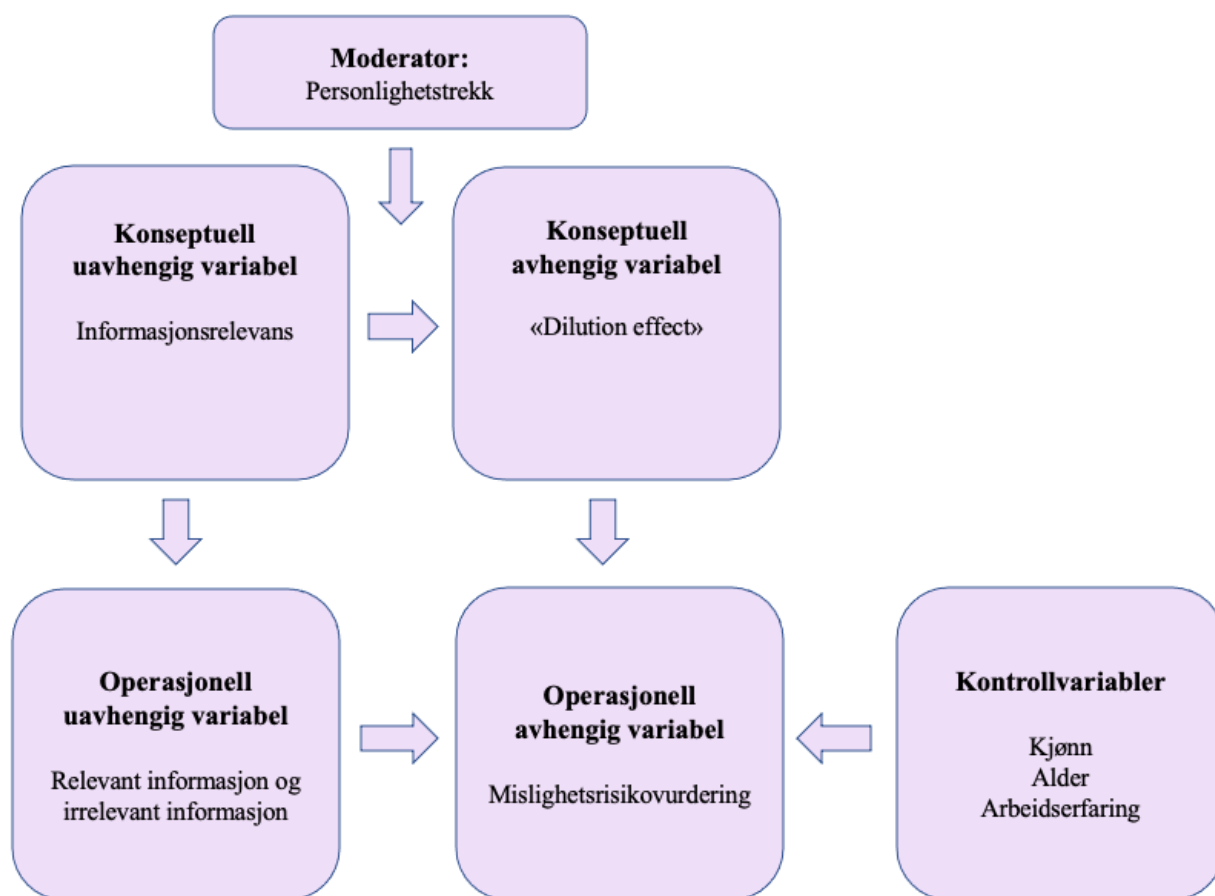
Videre rapporterte en lignende studie av Ahmand (2020) en signifikant årsakssammenheng mellom trekket negativ emosjonalitet og det kognitive bias flokkmentalitet. Til slutt fant de også en sammenheng mellom trekket samvittighetsfullhet, og de kognitive bias overmodighet og disposisjonseffekt.

Dette er eksempler på studier som tyder på at personlighetstrekk har en modererende effekt på kognitive bias. Under revisors arbeid utøves det betydelig grad av skjønn, hvor vurderingene kan være påvirket av kognitive bias. Vi ønsker derfor å bidra til litteraturen med en preventiv tilnærming, hvor vi undersøker om personlighetstrekk har en modererende effekt på dilution bias i en mislighetsrisikovurdering.

3. Konseptuell modell og hypoteser

I dette kapitlet vil vi redegjøre for valideringsrammeverket som benyttes i studien. Videre vil hypotesene fremlegges.

3.1 Konseptuell modell



Figur 1: Konseptuell modell

For å illustrere sammenhengene vi ønsker å undersøke, benytter vi Libby (1981) sitt valideringsrammeverk. Libby hevder at ingen teori kan testes direkte, men at den testes ved å vurdere forholdet mellom de operasjonelle definisjonene av sentrale begreper i teorien. For at testen skal være gyldig, må koblingene mellom de konseptuelle og de operasjonelle definisjonene være gyldige, og andre faktorer som kan påvirke den avhengige variabelen må kontrolleres eller ikke ha noen effekt (Libby, Bloomfield & Nelson, 2002).

Vår uavhengige variabel er informasjonsrelevans, som blir operasjonalisert til relevant- og irrelevant informasjon. Den avhengige variabelen er dilution effect, som er operasjonalisert til mislighetsrisikovurdering. Ved å utsette en gruppe for kun relevant informasjon, og en annen gruppe for både relevant- og irrelevant informasjon i en mislighetsrisikovurdering, ønsker vi å se om det foreligger grunnlag for dilution effect. Fordelingen av respondenter mellom de ulike gruppene er randomisert for å redusere systematiske skjevheter som kan påvirke resultatene. Videre ønsker vi å benytte personlighetstrekk som moderator for å undersøke om en eventuell dilution effect modereres med respondentens personlighetstrekk. Kontrollvariabler vil være kjønn, alder og arbeidserfaring. Dette for å sikre at disse variablene ikke skal påvirke resultatene i eksperimentet (Trotman, 2001).

3.2 Hypoteser

En hypotese kan beskrives som en påstand man ikke er sikker på er sann, og som man derfor ønsker å undersøke nærmere. Gjennom hypotesetesting ønsker man gjerne å undersøke om det er forskjell mellom ulike grupper. Nullhypotesen utgjør at det ikke forskjell mellom dem som undersøkes, og det man ønsker å undersøke. I motsetning til denne, vil alternativhypotesen hevde at det finnes forskjeller mellom gruppene man undersøker. Alternativhypotesen er den vi i utgangspunktet tror på, og ønsker å undersøke om det finnes støtte for (Grenness, 2012).

Innledningsvis ønsker vi å undersøke hvorvidt revisor er utsatt for dilution bias i en mislighetsrisikovurdering. Vi har derfor utarbeidet følgende hypotese:

H₁: Revisor vil foreta en mindre ekstrem mislighetsrisikovurdering når han blir presentert for både relevant- og irrelevant informasjon, sammenlignet med kun relevant informasjon

H₀: Revisor vil ikke foreta en mindre ekstrem mislighetsrisikovurdering når han blir presentert for både relevant- og irrelevant informasjon, sammenlignet med kun relevant informasjon

Videre vil vi undersøke sammenhengen mellom personlighetstrekk, og en eventuell dilution bias. Som påpekt tidligere utgjøres et individs personlighetsfungering av skårene på de ulike dimensjonene. I de fleste tilfeller er det derfor ikke slik at personlighet kan forklares av en enkelt dimensjon, men heller av samspillet mellom skåren på de ulike dimensjonene. Likevel kan de enkelte dimensjonene til en viss grad predikere handlingsmønstre i spesifikke situasjoner. Det ville imidlertid vært for omfattende for en masteroppgave å se på interaksjonseffekter mellom trekkene. I samråd med veileder har vi derfor sett det nødvendig å begrense oppgaven til enkelte dimensjoner. Vi vil altså se på enkelte dimensjoner som moderator for en eventuell dilution bias.

Individer som skårer høyt på trekket negativ emosjonalitet karakteriseres ofte som engstelige, og opplever høyere grad av emosjonell ubalanse og psykisk ubehag. Tidligere forskning indikerer at slike individer opplever et mer intenst stressnivå, samt er generelt mindre risikovillige (Bolger & Schilling, 1991; Bolger & Zuckerman, 1995; Suls et al., 1998; Gunthert et al. 1999; Nicholson et al., 2006). Vi tror derfor at slike individer vil oppleve informasjonsmengden som overveldende og ha vanskeligere med å skille mellom relevant- og irrelevant informasjon. Vi har derfor utarbeidet følgende hypotese:

H_{1,2}: Forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket negativ emosjonalitet

H₀: Forskjeller i risikovurderinger mellom gruppene kan ikke forklares av personlighetstrekket negativ emosjonalitet

Individer som skårer høyt på trekket åpensinnhet kjennetegnes typisk som mer åpne og mottakelige for nye inntrykk, samt mer villige til å utforske alternative idéer og forklaringer. Vi tenker at individer som skårer høyt på dette trekket vil være mer mottakelig for manipulasjonen vår, og derfor ha vanskeligere med å skille ut irrelevant informasjon. Vi har derfor utarbeidet følgende hypotese:

H_{1,3}: Forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket åpensinnhet

H₀: Forskjeller i risikovurderinger mellom gruppene kan ikke forklares av personlighetstrekket åpensinnhet

Dimensjonen samvittighetsfullhet, eller planmessighet brukes hyppig innen arbeids- og organisasjonspsykologi, ettersom dimensjonen antas å kunne predikere yrkesprestasjoner på tvers av mange områder (Kennair, 2022; Kennair, & Hagen, 2015). Slike individer karakteriseres typisk som produktive, ryddige, organiserte og ordentlige. Vi tror derfor at individer som skårer høyt på denne dimensjonen vil være bedre rustet til å skille mellom relevant- og irrelevant informasjon. Vi har derfor utarbeidet følgende hypotese:

H1₄: Forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket samvittighetsfullhet

H0: Forskjeller i risikovurderinger mellom gruppene kan ikke forklares av personlighetstrekket samvittighetsfullhet

4. Metode

I dette kapittelet gjør vi rede for vår forskningstilnærming, og diskuterer valg av forskningsdesign. Videre forklarer vi hvordan datainnsamlingen ble gjennomført, samt studiens forskningsetikk og analysemetode. Avslutningsvis diskuterer vi studiens reliabilitet og validitet.

4.1 Forskningstilnærming

Metode betegner hvilke teknikker som anvendes for å tilegne seg kunnskap om virkeligheten. Ved valg av metode skiller man mellom en deduktiv og induktiv tilnærming. En deduktiv tilnærming vil si at man går fra teori til empiri, mens en induktiv tilnærming vil si at man går fra empiri til teori. Vi har valgt å benytte en deduktiv tilnærming i vår studie, ettersom det foreligger en rekke tidligere forskning og teori innenfor emnet. Basert på dette har vi utviklet hypoteser, og ønsker å teste om disse stemmer overens med virkeligheten (Jacobsen, 2015).

4.2 Valg av forskningsdesign

Trochim et al. (2016) skiller mellom tre ulike typer forskningsdesign: eksplorativt, beskrivende og eksperimentelt. Studien vår benytter seg av et eksperimentelt forskningsdesign, som er den vanligste metoden innenfor JDM-forskning. Et eksperimentelt forskningsdesign vil si at man manipulerer minst en uavhengig variabel, mens andre faktorer holdes konstant. Basert på en deduktiv tilnærming, samt hypoteser vi ønsker å besvare, anser vi en eksperimentell studie som mest hensiktsmessig for å besvare vårt forskningsspørsmål. Vi vil da undersøke om det er en årsaks-virkning-sammenheng mellom informasjonsrelevans og personlighetstrekk i en mislighetsrisikovurdering.

4.2.1 Faktorielt design

Vi benytter oss av et faktorielt design ettersom vi har flere uavhengige variabler vi ønsker å undersøke: informasjonsrelevans og personlighetstrekk. Et faktorielt design muliggjør undersøkelse av hoved- eller interaksjonseffekter mellom de uavhengige variablene, også kalt faktorer (Trochim et al., 2014). En interaksjonsvariabel er en variabel som påvirker

styrken på sammenhengen mellom den avhengige og den uavhengige variabelen (Johannessen et. al., 2020). Som interaksjonsvariabel benytter vi personlighetstrekk. Gjennom BFI-2 personlighetstesten kartlegger vi ulike dimensjoner av personlighetstrekk, hvor vi videre ønsker å studere om disse har en interaksjon med en eventuell dilution bias. Vi har valgt å undersøke tre av dimensjonene nærmere.

4.2.2 «Within-» og «between-subjects»

En eksperimentell studie kan benytte seg av ulike metoder for å manipulere uavhengige variabler. Man skiller vanligvis mellom «within-subjects», «between-subjects», og en kombinasjon av «within-subjects» og «between-subjects». For å fastsette hva som er mest hensiktsmessig for vår studie er det nødvendig med en avveining av fordeler og ulemper mellom de ulike metodene (Trotman, 2001).

Ved bruk av «within-subjects» vil alle respondentene bli utsatt for alle nivåene av den uavhengige variabelen, slik at de blir sin egen kontrollgruppe. Dersom vi skulle benyttet oss av «within-subjects» ville alle respondenter fått tildelt både caset med relevant- og irrelevant informasjon. Fordelen med dette er at man vil trenge færre respondenter, noe som er nyttig i en masteroppgave med begrenset tid og ressurser. En ulempe er imidlertid at respondentens slutt svar og vurderinger kan bli påvirket av at deltakeren gjennomskuer hypotesene gjennom gjentakning av oppgavene. I tillegg vil man kunne få en såkalt rekkefølgeeffekt, som vil si at rekkefølgen på oppgavene i eksperimentet vil kunne påvirke deltakerens vurderinger (Trotman, 2001). Vi anser at ulempene med «within-subjects» vil overskygge fordelene, da en eventuell dilution effect vil kunne bli preget av gjentakelse og rekkefølgeeffekt i eksperimentet.

På bakgrunn av dette har vi valgt å benytte oss av «between-subjects». Respondentene vil da deles i to grupper, slik at man har en eksperimentgruppe og en kontrollgruppe (Trotman, 2001). Gjennom randomisering vil dermed eksperimentgruppen tildeles et case med både relevant- og irrelevant informasjon, mens kontrollgruppen vil tildeles et case bestående av kun relevant informasjon. Dette vil øke reliabilitet til designet, da en eventuell dilution effect ikke vil være påvirket av ulempene «within-subjects» medfører. En ulempe med «between-subjects» er imidlertid at det krever flere respondenter. Dette vil som nevnt være krevende i en masteroppgave, men vi anser likevel denne metoden som mest hensiktsmessig for vår studie.

4.3 Datainnsamling

Studien vår benytter seg av en kvantitativ tilnærming. Vi anser dette som hensiktsmessig, ettersom vi ønsker å teste hypoteser gjennom innhenting av målbare data. Fordelen med kvantitative data er at de er standardisert, slik at vi enklere kan analysere data gjennom statistiske analyser. Vi benytter oss av både primær- og sekundærdata. Primærdata innebærer opplysninger forskeren henter inn for første gang direkte fra kilden, mens sekundærdata er data som er innhentet av andre (Jacobsen, 2015). Vi innhenter primærdata gjennom eksperimentell spørreundersøkelse, mens sekundærdata omfatter data innhentet til teoretisk rammeverk.

4.3.1 Utvalg

Et utvalg er prosessen for å velge ut respondenter fra en forhåndsdefinert populasjon, med hensikt å generalisere resultatene fra utvalget til populasjonen (Trochim et al., 2016). Man kan overordnet skille mellom to typer utvalgsmetoder; sannsynlighet- og ikke-sannsynlighetsutvalg (Trochim et al., 2016).

Vårt forskningsdesign tar utgangspunkt i populasjonen revisorer, og målgruppen vår er revisjonsstudenter, i tillegg til individer med revisjonserfaring. Ashton & Kramer (1980) argumenterer for at bruken av studenter som surrogat innen JDM-revisjonsforskning er utbredt, og bør anses som rimelige alternativer i overordnede og generaliserende forskningsdesign.

Vi har altså et forhåndsdefinert utvalg med nærmere gitte egenskaper. Trochim et al. (2016) kaller denne type utvalg for ekspertutvalg, og er en form for ikke-sannsynlighetsutvalg. Ulempen med en slik utvalgsprosess er at ikke-sannsynlighetsutvalg kan være vanskelige å generalisere til populasjonen (Trochim, 2016). Bakgrunnen for utvalgsprosessen er imidlertid de naturlige begrensninger med å få tak i praktiserende revisorer. Utvalget er med andre ord blitt definert ut ifra et bekvemmelighetsvalg, noe som betyr at vi har valgt den del av populasjonen som var lettest å få tak i (Jacobsen, 2015).

4.3.2 Utforming av personlighetstest

For å kartlegge personlighetstrekk valgte vi å benytte en norsk versjon av typen *Big Five Inventory-2* (BFI-2). Testen er, som tidligere påpekt, basert på femfaktormodellen og videreutviklet av Oliver P. John og Christopher J. Soto (Føllesdal & Soto, 2022). Modellen er oversatt til norsk av Hallvard Føllesdal (Føllesdal & Soto, 2022). Vi har fått tilgang og tillatelse til bruk av testen gjennom personlig korrespondanse med Føllesdal.

Bakgrunnen for valget er at testen er forholdsvis kort og reduserer derfor sannsynligheten for å miste respondenter grunnet tidsbruk. Videre er testen tilpasset norsk kultur og språk, og validiteten er utprøvd (Føllesdal & Soto, 2022).

4.3.3 Utforming av case

Caset vårt er utformet med inspirasjon fra fem vignetter, utført i et eksperiment av Hansen & Kvamm (2012). Eksperimentet deres hadde som hensikt å kartlegge regnskaps- og revisjonsstudenters evne til å identifisere mislighetsrisikofaktorer, samt å evaluere risikoen for uredelig regnskapsrapportering.

Vedlagt artikkelen til Hansen og Kvamm var en liste av mislighetsrisikofaktorer som var benyttet i vignettene deres, fordelt på elementene i mislighetstriangelet. For å tilstrebe at caset vårt inneholdt faktiske mislighetsrisikofaktorer, valgte vi å oversette og innarbeide åtte av faktorene deres, se *vedlegg 3*. Av praktiske årsaker har vi utarbeidet et case med alle tre elementer tilstede.

Kontrollgruppen ble presentert for caset slik det følger av *vedlegg 1*. For eksperimentgruppen la vi til fire irrelevante påstander, se *vedlegg 4*. Disse påstandene har vi hentet og oversatt fra en studie til Hoffman & Patton (1997), som vedlegg til masteroppgaven til Lillefosse og Tran (2018). Hoffman & Patton har i studiet sitt brukt påstander fra et eldre eksperiment av Heiman-Hoffman et al. (1996), hvor de har rangert relevansen til en rekke påstander i en mislighetsrisikovurdering. Påstandene som vi har brukt, ble her vurdert som irrelevante i en mislighetsrisikovurdering. Artikkelen er imidlertid fra 1997, og det er dermed fare for at den er delvis utdatert. Påstandene er likevel benyttet i nyere eksperimenter innen samme tema (Eilifsen et al., 2017; Wood, 2012, Lillefosse & Tran, 2018). Etter en grundig gjennomgang vurderte vi at de påstandene vi ønsket å bruke, på tross av alder, fortsatt var relevante.

4.3.4 Pilot-test

Formålet med en pilot-test er å luke ut feil og uklarheter før spørreundersøkelsen publiseres. På denne måten finner vi ut om spørreskjemaet fungerer slik vi har tenkt, samt endringer som bør foretas (Jacobsen, 2015). Pilot-testen vår ble utført i tre omganger, fordelt på elleve personer.

Første pilottest ble utført på tre revisjonsstudenter. Tilbakemeldingene var at caset var forståelig og realistisk. Gruppen som mottok caset med irrelevant informasjon, opplevde informasjonsmengden som overveldende. På grunn av oppgavens hensikt, valgte vi å ikke endre på informasjonsmengden.

Etter pilot-testen merket vi oss at svarene var preget av subjektivitet vedrørende fastsetting av risiko for misligheter gjennom en 101-punktsskala. I oppfølgingsspørsmål med respondentene fremkom det at tolkningen av %-nivå var ulik. To av respondentene som mottok samme case, mente at det forelå en høy risiko for misligheter, men svarte henholdsvis 40% og 85%.

Etter en gjennomgang av caset kartla vi ikke noe som skulle tilsi en ulik tolkning av vurderingsskalaen. Subjektiviteten antas derfor å bli randomisert bort i store populasjoner, da ulikheten vil bli tilfeldig fordelt mellom gruppene. På grunn av naturlige begrensninger for vår populasjon, valgte vi imidlertid å endre vurderingsskalaen til en 5-punktsskala, med hensikt å redusere støy forårsaket av ulik tolkning. Begrunnelsen var at vi da kunne sette fornuftige merkelapper på de enkelte svaralternativene (Jacobsen, 2015).

Den andre pilot-testen ble utført på ytterligere fire respondenter. Pilot-testen avdekket nå at alle respondenter vurderte risikoen som «høy», uavhengig av tildelt case. I oppfølgingsspørsmål med respondentene fremkom det et behov for en ytterligere merkelapp, da respondentene ble tvunget til å runde av risikovurderingen til «høy». For å klare å fange opp mindre forskjeller i vurderinger, valgte vi derfor å gå for en syvpunktsskala med merkelapper.

Den tredje pilot-testen ble utført på ytterligere fire respondenter, og vi kartla nå differanser i risikovurderingene. Etter en helhetsvurdering vurderte vi at endringene fra første til tredje pilot-test samstemte med våre antagelser, og valgte derfor å avslutte pilot-testingen.

4.3.5 Praktisk gjennomføring

Eksperimentet ble utført som et anonymisert spørreskjema i dataprogrammet Qualtrics, hvor vi sikret oss at hverken IP-adresser eller annen personlig informasjon ble lagret.

Innledningsvis ble caset distribuert gjennom personlig korrespondanse til eget nettverk, i tillegg til at den ble publisert på facebook-gruppen «Platform for MRR students» d.

06.03.23. For å stimulere til ytterligere deltakelse valgte vi, fire dager senere, å lodde bort to gavekort på 600 kr fra Rema1000 som insentiv, og sende ut en påminnelse på samme facebook-gruppe. Gavekortene ble finansiert av oss.

På grunn av dårlig respons, tok vi som et siste forsøk kontakt med førsteamanuensis og programleder for MRR, Kjell Ove Røsok. Spørreskjemaet vårt ble da publisert på Canvas gruppen «Generell informasjon til MRR-studentene» for studenter med oppstart høsten 2022. Canvas er læringsplattformen som brukes på Norges Handelshøyskole.

Spørreskjemaet ble publisert d. 13.03.23. Videre ble studentene på kurs ACC410 i Bergen og Oslo oppfordret til deltakelse, henholdsvis d. 15.03.23 og 24.03.23.

Vi valgte å avslutte innhentingene d. 01.04.23 med totalt 83 registrerte svar.

4.4 Forskningsetikk

Samtykke til å delta i undersøkelsen ble gitt ved å svare ja på følgende spørsmål: «Jeg samtykker til å delta i denne undersøkelsen». Respondentene kunne også når som helst trekke seg fra undersøkelsen om ønskelig.

Alle deltakerne som gjennomførte spørreundersøkelsen var anonyme, og det utløses dermed ikke meldeplikt til Norsk senter for forskningsdata (NSD). For å sikre anonymitet var alle funksjonene som kunne identifisere personer deaktivert i Qualtrics. Spørreundersøkelsen inneholdt heller ingen spørsmål som gjorde det mulig å identifisere respondentene. På denne måten ivaretar studien vår hensynet til personvern.

4.5 Analysemetode

For å analysere datasettet vårt benytter vi IBM SPSS. Datasettet er direkte overført fra Qualtrics til SPSS slik at risikoen for menneskelige feil reduseres.

Gjennom våre statistiske analyser ønsker vi å finne sannsynligheten for om resultatene er realistiske eller tilfeldige (Trochim et al., 2016). Hypotesetestene utføres med et 95% konfidensintervall som er vanlig i denne type forskning. Det vil si at man med 95% sikkerhet kan påstå noe om populasjonen basert på resultatene fra utvalget. Vi vil da redusere sannsynligheten for type 1-feil, samtidig som sannsynligheten for type 2-feil ikke er for høy. Type 1-feil vil si at man forkaster en nullhypotese som er sann, mens type 2-feil forekommer når man beholder en nullhypotese til tross for at den er gal (Grenness, 2012).

For å teste hypotese 1 benyttet vi en T-test som tester om det er signifikante forskjeller i gjennomsnittet mellom to grupper (Jacobsen, 2015). Videre benyttet vi oss av en toveis-ANOVA for å teste hypotese 2 og 3. En toveis-ANOVA tester om det er en signifikant effekt av de to uavhengige variablene på den avhengige variabelen.

For å kunne gjennomføre en T-test og en toveis-ANOVA må følgende forutsetninger være oppfylt (Lærd Statistics, u.å.):

1. Den avhengige variabelen er kontinuerlig
2. De uavhengige variablene består av to kategoriske variabler
3. Observasjonene må være uavhengige
4. Det skal ikke foreligge noen signifikante utliggere (observasjoner som avviker signifikant fra resten av utvalget)
5. Den avhengige variabelen må være tilnærmet normalfordelt
6. Det må være homogenitet i variansen

Den avhengige variabelen vår måles gjennom en mislighetsrisikovurdering på en syvpunkts-skala, og den første forutsetningen er dermed oppfylt. Forutsetning to er også oppfylt da uavhengig variabel «informasjonsrelevans» består av tildelt case en eller to, mens uavhengig variabel «personlighetstrekk» består av om den enkelte respondent har skåret høyt eller ikke

på de ulike personlighetstrekkene. I denne oppgaven er «høyt» definert som en skår over medianverdien for utvalget. Videre er observasjonene uavhengige gjennom utarbeiding av spørreundersøkelsen.

For å vurdere forutsetning fire har vi foretatt en visuell vurdering av frekvenshistogram og boksplot for den avhengige variabelen, som altså ikke viser noen ekstreme uteliggere. Videre er verdien på 5% trimmet gjennomsnitt og gjennomsnitt svært nærme, noe som viser at uteliggere ikke har noen sterk innvirkning på gjennomsnittsverdien, se *vedlegg 5*.

Normalfordelingen har vi testet ved å benytte en Kolmogorov-Smirnov test og en Shapiro-Wilk test, i tillegg til visuell vurdering av frekvenshistogram, se *vedlegg 5*. Begge testene indikerer at svarene ikke er normalfordelt for både kontrollgruppen og eksperimentgruppen. Det skal likevel legges til at det kan være problematisk å vurdere normalfordeling for et så lite utvalg. I samråd med veileder har vi derfor valgt å benytte en Mann-Whitney U test som ikke forutsetter normalfordeling, supplert med en T-test.

En utført Levene's test viser at det foreligger homogenitet i variansen, og forutsetning 6 er derfor også oppfylt, se *vedlegg 7*.

4.6 Validitet og reliabilitet

Resultatet vi kommer frem til i oppgaven avhenger av datamaterialet som er innhentet. Det er derfor avgjørende at datamaterialet er av god kvalitet slik at resultatene er pålitelige og gyldige (Grønmo, 2016). For å vurdere dette er reliabilitet og validitet to viktige begrep. Reliabiliteten omfatter resultatenes pålitelighet eller nøyaktighet. Validiteten sier noe om studien måler det den faktisk er ment til å måle, slik at datamaterialet faktisk er gyldig for problemstillingen. For at det skal foreligge validitet, er det en forutsetning at vilkårene for reliabilitet er innfridd. En reliabel undersøkelse gir imidlertid ingen garanti for at denne også er valid, og begrepene må derfor vurderes både i sammenheng og hver for seg (Grenness, 2012).

4.6.1 Validitet

For å vurdere validiteten er tre begreper sentrale: begrepsvaliditet, intern validitet og ekstern validitet.

Begrepsvaliditet sier noe om måleinstrumentet har klart å fange opp begrepet man ønsker å måle (Grenness, 2012). Gjennom vår studie ønsker vi å måle begrepet dilution effect gjennom en mislighetsrisikovurdering, samt personlighetstrekk gjennom femfaktormodellen. For å opprettholde begrepsvaliditeten er det nødvendig at respondentene har kunnskap om vurdering av misligheter. Mislighetstriangelet er pensum i Revisjon I og er en del av ISA 240, og vi anser derfor MRR-studenter og personer med revisjonserfaring som kompetente til å besvare caset.

For å måle personlighetstrekk har vi benyttet personlighetstesten BFI-2, ettersom at denne er tilpasset norsk kultur og språk, i tillegg til at validiteten er utprøvd (Føllesdal & Soto, 2022).

Intern validitet refererer til hvorvidt studien kan påvise et årsaks-virkningsforhold (Johannesen et al., 2020). Som diskutert i 4.3.3, har vi vurdert de relevante og irrelevante påstandene brukt i caset til å være relevante for å undersøke om det foreligger dilution effect. Vår interne validitet er styrket gjennom bruk av randomisert kontrollgruppe og eksperimentgruppe i Qualtrics. Dette gjør at alle respondentene har samme mulighet til å havne i de ulike gruppene, og vil øke muligheten for årsaksforklaring (Grenness, 2012). En ulempe er at Qualtrics også registrerer ufullstendige svar, slik at gruppene kan bli noe skjevfordelt etter filtreringen av svarene er gjennomført. Vi ser imidlertid i kapittel 5.3.1 at vi har fått til en tilnærmet lik fordeling mellom gruppene.

Bruk av manipulasjonssjekk er også en styrke for den interne validiteten. Vi sikrer da at respondentene har oppfattet informasjonen riktig i caset. I tillegg styrkes den interne validiteten ved at vi benytter kontrollvariabler som eksempelvis arbeidserfaring. Da vil sannsynligheten for at korrelasjonen mellom uavhengig- og avhengig variabel skyldes andre variabler som ikke er tatt hensyn til, være redusert (Cozby, 2012). Bruk av internett kan imidlertid være en ulempe for den interne validiteten da respondentene kan samarbeide og bli distraheret (Trotman, 2001). For å redusere sannsynligheten for at svarene kan være preget av dette inneholder spørreundersøkelsen et kontrollspørsmål, hvor respondenter som samarbeider under spørreundersøkelsen blir luket ut av analysen.

Ekstern validitet sier noe om i hvilken grad man kan generalisere resultatene til å gjelde hele populasjonen, samt er gjeldende for andre situasjoner enn den man studerer (Jacobsen, 2015). En ulempe med en masteroppgave er som nevnt at mangel på tid og ressurser medfører et begrenset utvalg. Utvalget vårt er et bekvemmelighetsvalg, og består i hovedsak av MRR-studenter. Det kan tenkes at resultatene ville blitt annerledes dersom utvalget bestod av eksempelvis revisorer med lang arbeidserfaring. Det er blitt hevdet at bruk av studenter hvor det kreves inngående forståelse om et fagfelt kan svekke den eksterne validiteten (Friedman, 1994). På den annen side er det som nevnt i 4.3.1 argumentert for at bruk av studenter innen JDM-revisjonsforskning bør anses som rimelige alternativer (Ashton & Kramer, 1980). Det er vanskelig å maksimere både intern og ekstern validitet, og forholdet mellom disse blir dermed på mange måter en «trade-off» (Grenness, 2012). Vi har derfor valgt å fokusere på at studien vår har rimelig god intern validitet. Samtidig vil det være muligheter for å samle inn gode data for å hele populasjonen, men vi vil ikke kunne generalisere resultatene (Grenness, 2012).

4.6.2 Relabilitet

Relabilitet vurderes gjennom undersøkelsens data, hvilke data som brukes, måten den samles inn på, og hvordan data bearbeides (Johannesen, 2015). Dersom undersøkelsesopplegget og datainnsamlingen gir pålitelige svar, er reliabiliteten høy. Endringer i datamaterialet skal skyldes endringer i faktiske forhold, ikke et ustabil forskningsopplegg eller en annen forsker (Grønmo, 2016). En styrke i vår studie er at vi er to studenter som sammen innhenter og analyserer datamaterialet, slik at sannsynligheten for feiltolkninger er redusert. Før publisering av spørreundersøkelsen utførte vi som nevnt flere pilot-tester på til sammen elleve personer. På denne måten luket vi ut feil, og rettet opp i elementer som skapte ulike tolkninger. I tillegg er det en styrke at undersøkelsen er anonym, ettersom respondentene da kan svare uten fare for å bli gjenkjent. Dette anser vi som spesielt viktig for å få pålitelige svar gjennom bruk av BFI-2. Videre er reliabiliteten også styrket som følge av at datasettet ble renset for ufullstendige og ugyldige svar før analysene. En fordel ved bruk av Qualtrics er som nevnt at data kan overføres til SPSS, noe som reduserer risikoen for menneskelige feil, og dermed også er med på å styrke reliabiliteten.

5. Resultater

I det følgende presenteres resultatene fra eksperimentet vårt. Kapittelet tar først for seg rensing av data og resultatene av manipulasjonssjekken. Deretter presenteres deskriptiv statistikk. Videre utføres hypotesetesting og tolkning av resultatene. Avslutningsvis presenteres andre observasjoner.

5.1 Rensing av data

Det var totalt 83 personer som åpnet spørreundersøkelsen vår. Av disse var 35 ufullstendige, og ble dermed fjernet. Dette gir en frafallsprosent på 42.2%. Dette er ikke uvanlig da vi gjennomførte et ikke-kontrollert eksperiment over internett (Dandurand et al., 2008). Videre var det tre respondenter som ikke gjennomførte personlighetstesten. De hadde imidlertid gjennomført caset og manipulasjonstesten. Vi valgte derfor å beholde disse for å teste hypotese 1.

I tillegg valgte vi også å fjerne respondenter som hadde mer enn fire av ni feil på manipulasjonssjekken, som omfattet seks antall svar. Dette er begrunnet i at vi anser det essensielt for analysen at respondentene kan skille mellom relevant- og irrelevant informasjon ved gjennomføring av caset. Vi har valgt å benytte en grense på fire ettersom vi ønsker valide resultater, samtidig som antall respondenter skal være på et akseptabelt nivå. Kapittel 5.2 tar for seg ytterligere diskusjon av manipulasjonssjekk. Etter rensing av data satt vi dermed igjen med totalt 42 respondenter som vil benyttes til videre analyse. Vi merker oss at dette er et relativt lavt antall respondenter fra et forskningsståsted. Vi ser imidlertid at lav deltakelse i eksperimentelle masteroppgaver er vanlig, og historisk sett er en akseptert risiko.

5.2 Manipulasjonssjekk

Vår manipulasjonssjekk bestod av ni påstander som respondentene skulle klassifisere som enten «relevant» eller «irrelevant». Påstandene ble hentet fra caset som skulle besvares i forbindelse med mislighetsrisikovurderingen. Respondentene som ble tatt med videre i analysen besvarte manipulasjonssjekken som vist i tabell 1 under. Svarene markert i grønn indikerer riktig svar.

| Nr. | Påstander | N | Relevant | Irrelevant |
|-----|---|----|----------|------------|
| 1 | Ledelsen er lite samarbeidsvillig. De setter ikke av tilstrekkelig tid til å møte deg, og er treg med å oversende dokumentasjon | 42 | 100% | |
| 2 | En stor andel av ledelsens lønn er resultatbetinget | 42 | 100% | |
| 3 | Godene som er tilgjengelige for de ansatte er mer sjenerøse enn bransjegjennomsnittet | 42 | 60% | 40% |
| 4 | Ledelsens og arbeidstakernes representanter indikerer at det er mulighet for streik i det kommende året | 42 | 37,5% | 62,5% |
| 5 | Ikke tilstrekkelig bemanning i bedriftens økonomiavdeling | 42 | 97,5% | 2,5% |
| 6 | Fjorårets revisor gjennomførte revisjonen litt under budsjettet | 42 | 22,5% | 77,5% |
| 7 | Økt konkurranse | 42 | 90% | 10% |
| 8 | En endring i kommunens skattestruktur vil føre til en reduksjon i selskapets eiendomsskatt | 42 | 20% | 80% |
| 9 | Behov for betydelig lånefinansiering for å holde seg konkurransedyktig. Banken vil da kreve mer egenkapitalfinansiering og personlige garantier for bedrifters gjeld. | 42 | 90% | 10% |

Tabell 1: Manipulasjonssjekk

Ut fra tabell 1 ser man at respondentene treffer bra på de fleste påstandene. Fem av ni påstander har mellom 90-100% klassifisert riktig. Videre har påstand fire og seks riktig

svarprosent på henholdsvis 62,5% og 77,5%. Dette ser vi på som akseptable verdier, da dette er godt over halvparten av respondentene.

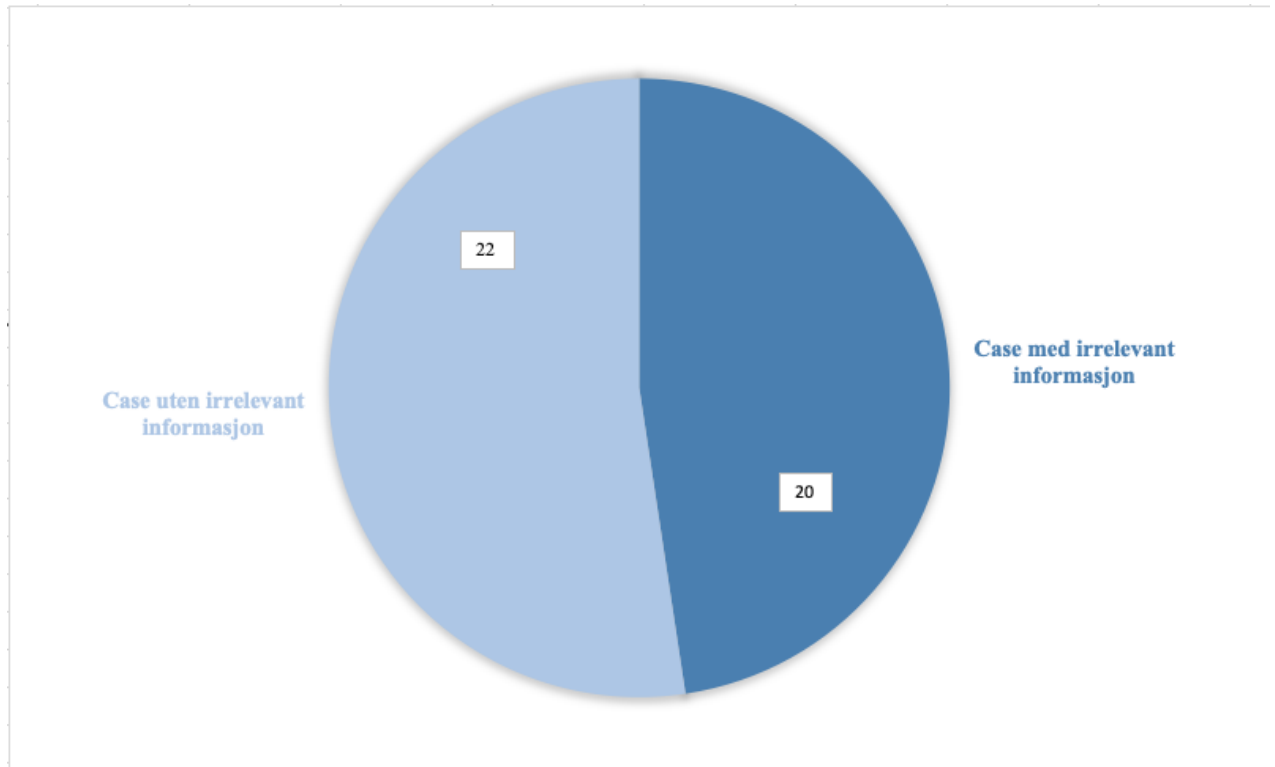
Påstanden «godene som er tilgjengelige for de ansatte er mer sjenerøse enn bransjegjennomsnittet» skiller seg ut hvor kun 40% av respondentene har klassifisert den som irrelevant. Vi velger å vurdere manipulasjonssjekken som helhet, og aksepterer derfor denne verdien da resterende påstander i det alt vesentlige er besvart riktig. Gjennom vår manipulasjonssjekk er vi dermed trygge på at respondentene evner å skille mellom relevant- og irrelevant informasjon på en god måte, slik at svarene kan brukes i videre analyser.

5.3 Deskriptiv statistikk

Deskriptiv statistikk vil si å beskrive dataene ved å se på hvordan verdiene i materialene fordeler seg, og kan vise særtrekk ved utvalget (Grenness, 2012). I det følgende vil vi ved hjelp av deskriptiv statistikk illustrere fordeling mellom gruppene som fikk henholdsvis case uten irrelevant informasjon og case med irrelevant informasjon, samt demografiske variabler. Dette er for å sikre at de to gruppene vi studerer er sammenlignbare, slik at forskjeller mellom gruppene ikke kommer av ulik fordeling eller ulike karakteristika. På denne måten styrkes den interne validiteten i studien.

5.3.1 Fordeling mellom gruppene

Våre 42 respondenter fordeler seg på de to ulike gruppene som vist i figur 2. Gjennom randomisering i Qualtrics var formålet å fordele gruppene likt. Man ser imidlertid en minimal skjevfordeling. Dette skyldes at ufullstendige svar ble fjernet som forklart i kapittel 5.1. Vi har oppnådd en fordeling tilnærmet slik vi ønsket, og den interne validitet er dermed ikke svekket som følge av dette forholdet.



Figur 2: Fordeling mellom gruppene

5.3.2 Demografiske variabler

Tabell 2 viser en oversikt over demografiske variabler for våre 42 respondenter i en frekvenstabell, som er en av de vanligste metodene for beskrivende statistikk (Trochim et al., 2016). Denne informasjonen ble innhentet gjennom spørreundersøkelsen fordi vi ønsket et bedre bilde av utvalget vårt. I tillegg vil de ulike faktorene brukes til kontrollvariabler i analysen.

| | Case uten irrelevant informasjon | Case med irrelevant informasjon | Totalt |
|--------------------------|---|--|---------------|
| <i>Antall per gruppe</i> | 22 | 20 | 42 |
| Kjønn | | | |
| <i>Mann</i> | 12 (55%) | 10 (50%) | 22 (52%) |
| <i>Kvinne</i> | 10 (45%) | 10 (50%) | 20 (48%) |
| Alder | | | |
| <i>22-25</i> | 13 (59%) | 11 (55%) | 24 (57%) |
| <i>26-30</i> | 7 (32%) | 7 (35%) | 14 (33%) |
| <i>31 eller mer</i> | 2 (9%) | 2 (10%) | 4 (10%) |
| Utdanning | | | |
| <i>MRR</i> | 22 (100%) | 19 (95%) | 41 (98%) |
| <i>Ikke tatt MRR</i> | 0 (0%) | 1 (5%) | 1 (2%) |
| Arbeidserfaring | | | |
| <i>Ja</i> | 6 (27%) | 4 (20%) | 10 (24%) |
| <i>Nei</i> | 12 (55%) | 10 (50%) | 22 (52%) |
| <i>Kun internship</i> | 4 (18%) | 6 (30%) | 10 (24%) |

Tabell 2: Deskriptiv statistikk - Frekvenstabell

Fordelingen mellom menn og kvinner er omtrent lik i begge gruppene, og svarene representerer dermed begge kjønn på en god måte. Videre ser vi at alder i begge gruppene har en overvekt av personer mellom 22-25 år. Dette skyldes at flertallet av personer i vårt nettverk ligger i denne aldersgruppen. Alder er imidlertid ganske likt fordelt mellom de to gruppene.

Alle respondenter tar eller har tatt MRR, med unntak av en person. Denne personen har imidlertid arbeidserfaring, og vi velger dermed å beholde respondenten i utvalget vårt. Videre observerer vi at begge gruppene har overvekt av personer uten arbeidserfaring. Dette er forventet da vi i hovedsak har sendt ut spørreundersøkelsen til personer som fortsatt er studenter. Antall personer som har arbeidserfaring er relativt likt fordelt mellom gruppene, og vi anser derfor ikke dette som en faktor som kan skape forskjell mellom de to ulike gruppene.

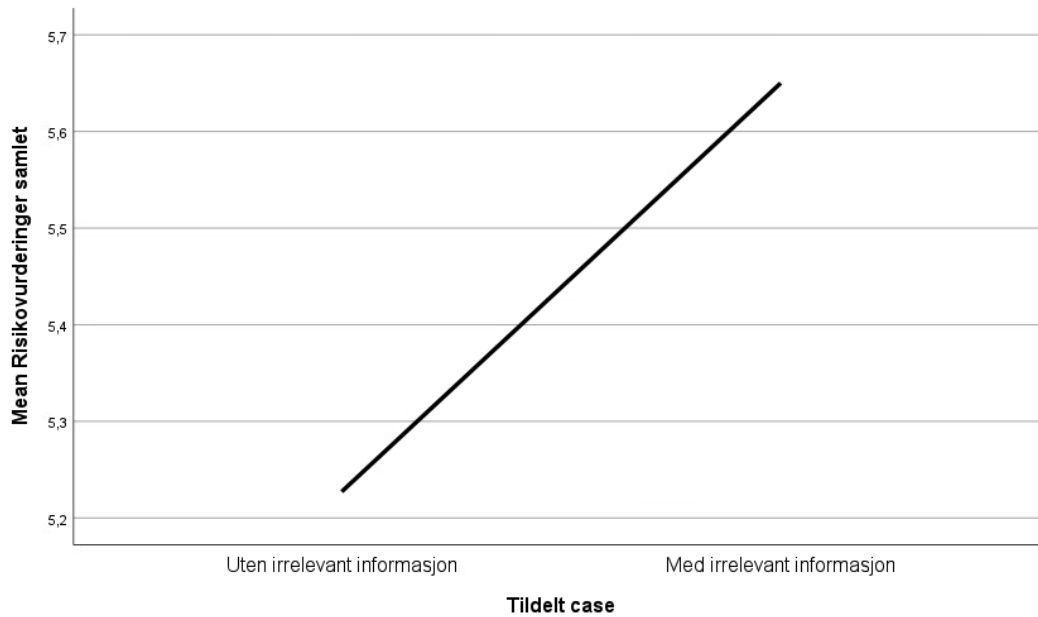
Samlet sett er de to gruppene relativt likt fordelt, og randomiseringen i Qualtrics har dermed vært vellykket. Vi finner ingen skjevfordelinger i demografiske variabler som kan svekke sammenlignbarheten mellom dem. Gjennomgang av deskriptiv statistikk viser oss dermed at datamaterialet er tilfredsstillende for videre analyseformål.

5.4 Test av hypoteser

I dette delkapittelet vil vi teste hypotesene våre. I den forbindelse utførte vi også enveis og toveis ANCOVA («Analysis of covariance») for å undersøke om noen av kontrollvariablene påvirket den avhengige variabelen. Resultatene viser at ingen av kontrollvariablene har en signifikant betydning for resultatene, se *vedlegg 6*.

5.4.1 Hypotese 1

I hypotese 1 vil vi undersøke om eksperimentgruppen vil foreta en mindre ekstrem risikovurdering, sammenlignet med kontrollgruppen. For å teste hypotese 1 ble både en Mann-Whitney U-test og en T-test anvendt.



Figur 3: Gjennomsnittsverdier hypotese 1

Figur 3 viser gjennomsnittsverdien på risikovurderingene. Her ser vi en forskjell mellom gruppene, hvor kontrollgruppen har en gjennomsnittsverdi på 5.23 mens eksperimentgruppen har en gjennomsnittsverdi på 5.62.

| | Mann-Whitney U | Z | Sig (2-tailed) |
|-----------------|----------------|--------|----------------|
| Risikovurdering | 142,5 | -2,078 | 0,038 |

Tabell 3: Mann-Whitney U test hypotese 1

Resultatene i tabell 3 fra Mann-Whitney U testen viser en signifikant forskjell mellom kontrollgruppen og eksperimentgruppen med $p = 0.038 < 0.05$.

| | Levene's Test for equality of Variances | | | | | | 95% Confidence Interval of the Difference | |
|-----------------|---|-------|--------|----|----------------|----------------|---|--------|
| | F | Sig | t | df | Sig (1.tailed) | Sig (2.tailed) | Lower | Upper |
| Risikovurdering | 3,857 | 0,057 | -1,396 | 40 | 0,85 | 0,17 | -1,035 | -0,189 |

Tabell 4: T-test hypotese 1

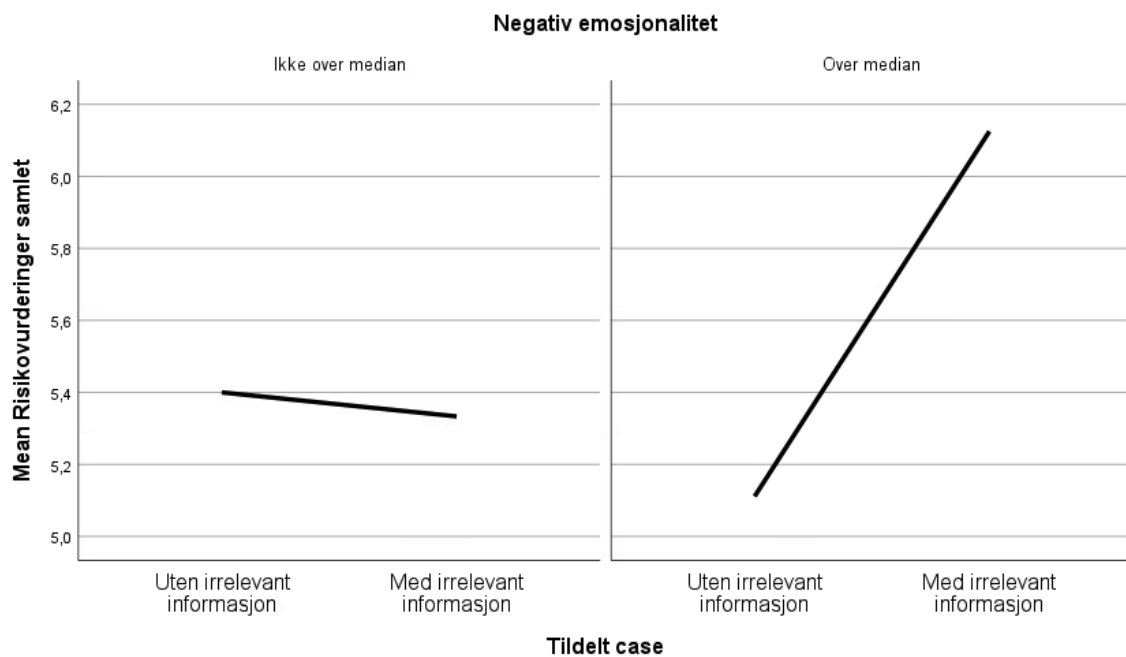
Resultatene i tabell 4 fra T-testen viser imidlertid ingen signifikante forskjeller med enveis $p = 0.085 > 0.05$ og toveis $p = 0.17 > 0.05$.

Interessant nok, går resultatene i motsatt retning av den opprinnelige hypotesen. Funnene indikerer altså at revisor foretar en mer ekstrem risikovurdering når den blir presentert for både relevant- og irrelevant informasjon, sammenlignet med kun relevant informasjon.

På bakgrunn av dette beholder vi H0.

5.4.2 Hypotese 2

I hypotese 2 vil vi undersøke om forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket negativ emosjonalitet. Dette gjør vi ved hjelp av en toveis ANOVA analyse.



Figur 4: Gjennomsnittsverdier hypotese 2

Fra figur 4 ser vi en tydelig trend som indikerer at individer som skårer høyt på trekket negativ emosjonalitet øker sin risikovurdering betraktelig når den blir presentert for irrelevant informasjon, mens resterende i utvalget ser ut til å redusere sin risikovurdering litt, basert på gjennomsnittsverdiene.

| | Type III Sum of Squares | df | Mean Square | F | Sig. |
|-----------------------------------|-------------------------|----|-------------|-------|-------|
| Gruppe * Negativ Emosjonalitet | 2,784 | 1 | 2,784 | 2,968 | 0,094 |

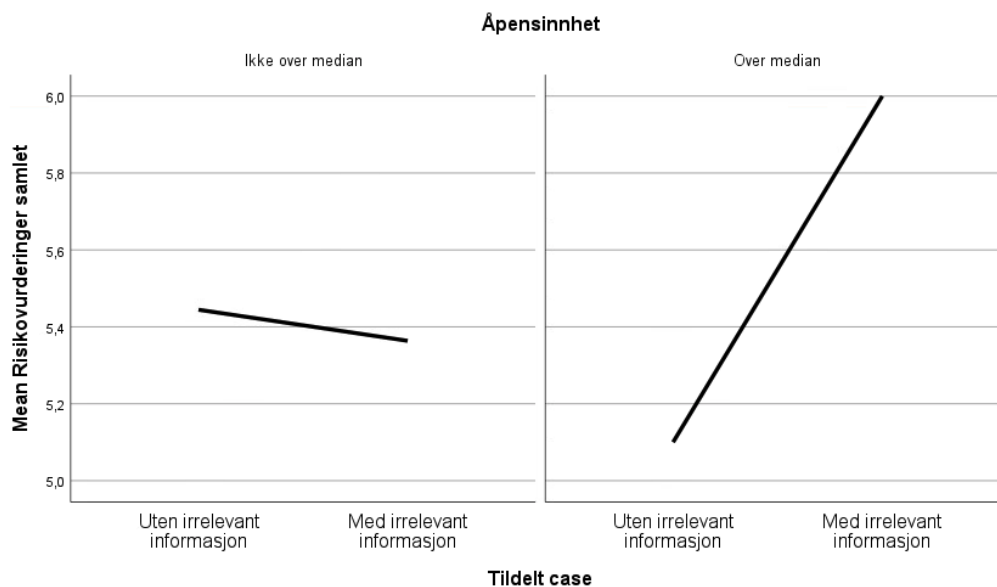
Tabell 5: Toveis ANOVA hypotese 2

Resultatene i tabell 5 fra toveis ANOVA analysen viser imidlertid ingen signifikante forskjeller, med $p = 0.094 > 0.05$.

På bakgrunn av dette beholder vi H_0

5.4.3 Hypotese 3

I hypotese 3 vil vi undersøke om forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket åpensinnhet. Dette gjør vi også ved hjelp av en toveis ANOVA analyse.



Figur 5: Gjennomsnittsverdier hypotese 3

Også her ser vi fra figur 5, en tydelig trend som viser at individer som skårer høyt på trekket åpensinnhet i gjennomsnitt øker sin risikovurdering betraktelig når den blir presentert for

irrelevant informasjon. Resterende i utvalget ser ut til å redusere sin risikovurdering basert på gjennomsnittsverdier.

| | Type III Sum of Squares | df | Mean Square | F | Sig. |
|----------------------|-------------------------|----|-------------|-------|-------|
| Gruppe * Åpensinnhet | 2,329 | 1 | 2,329 | 2,421 | 0,129 |

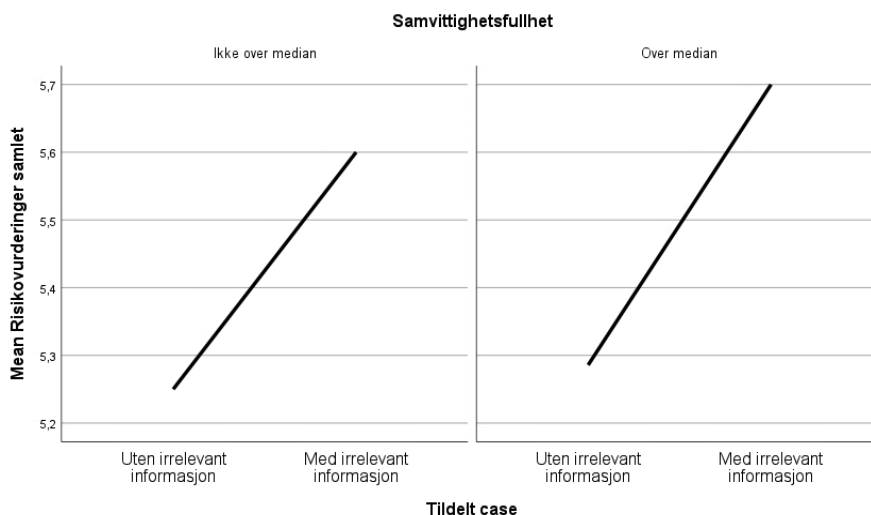
Tabell 6: Toveis ANOVA hypotese 3

Resultatene i tabell 6 fra toveis ANOVA analysen viser imidlertid heller ingen signifikante forskjeller, med $p = 0.129 > 0.05$.

På bakgrunn av dette beholder vi H_0 .

5.4.4 Hypotese 4

I hypotese 4 vil vi undersøke om forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket samvittighetsfullhet. Dette gjør vi også ved hjelp av en toveis ANOVA analyse.



Figur 6: Gjennomsnittsverdier hypotese 4

Figur 6 indikerer at individer som skårer høyt på trekket øker sin risikovurdering tilnærmet likt som resterende i utvalget, når den blir presentert for irrelevant informasjon.

| | Type III Sum of Squares | df | Mean Square | F | Sig. |
|--|-------------------------------|----|-------------|-------|-------|
| Gruppe * Samvittighetsfullhet | 0,01 | 1 | 0,01 | 0,009 | 0,923 |

Tabell 7: Toveis ANOVA hypotese 4

Resultatene i tabell 7 fra toveis ANOVA analysen viser ingen signifikante forskjeller, med $p = 0.923 > 0.05$.

På bakgrunn av dette beholder vi H_0 .

5.5 Andre observasjoner

I det følgende vil vi presentere andre observasjoner, som omfatter fordelingen mellom skåren på de ulike personlighetstrekkene.

5.5.1 Fordeling av personlighetstrekk

Under omregningen av svarene fra personlighetstesten til en endelig skår, la vi merke til at trekkene hadde en uforventet fordeling. Vi la merke til at utvalget kjennetegnes av to trekk.

| Tildelt case | | Ekstraversjon | Omgjengelighet | Samvittighetsfullhet | Negativ Emosjonalitet | Åpensinnhet |
|-----------------------------------|--------|---------------|----------------|----------------------|--------------------------|--------------|
| Uten irrelevant informasjon | Mean | 38,11 | 41,95 | 45,42 | 28,84 | 40,05 |
| | Median | 37,00 | 41,00 | 47,00 | 28,00 | 41,00 |
| Med irrelevant informasjon | Mean | 38,70 | 47,15 | 47,05 | 29,70 | 40,55 |
| | Median | 40,50 | 48,00 | 47,50 | 27,00 | 39,50 |
| Total | Mean | 38,41 | 44,62 | 46,26 | 29,28 | 40,31 |
| | Median | 39,00 | 45,00 | 47,00 | 28,00 | 40,00 |

Tabell 8: Fordeling av personlighetstrekk

Fra tabell 8 ser vi at utvalget vårt skårer høyt på trekket samvittighetsfullhet, vurdert ut fra median- og gjennomsnittsverdi. Videre ser vi at utvalget skårer lavt på trekket negativ emosjonalitet, vurdert ut fra median- og gjennomsnittsverdi.

6. Diskusjon og konklusjon

I dette kapittelet vil vi diskutere hypotesene våre, før vi videre konkluderer. Avslutningsvis diskuteres studiens begrensinger, samt forslag til videre forskning.

6.1 Diskusjon av hypoteser

6.1.1 Diskusjon av hypotese 1

Hypotese 1: Revisor vil foreta en mindre ekstrem mislighetsrisikovurdering når han blir presentert for både relevant- og irrelevant informasjon, sammenlignet med kun relevant informasjon

Figur 3 viser at gjennomsnittsverdien til respondentenes risikovurdering øker med irrelevant informasjon. Resultatene fra Mann-Whitney U testen viser at forskjellen mellom gruppene er statistisk signifikant med p-verdi på $0,038 < 0,05$. Videre fant T-testen en enveis p-verdi på 0,085 og toveis p-verdi på 0,17. Resultatene fra Mann-Whitney U testen indikerer altså at det er en signifikant forskjell, mens T-testen ikke finner noen signifikante forskjeller. P-verdien fra t-testen er imidlertid nær signifikansnivået, noe som indikerer at det kan foreligge en forskjell mellom gruppene.

Vi finner dermed det motsatte av hypotese 1, at eksperimentgruppen foretar en mer ekstrem risikovurdering ved tillegg av irrelevant informasjon. Dette er i strid med tidligere forskning fra Hackenbrack (1992) og Hoffman & Patton (1997), som fant at revisor som presenteres for irrelevant informasjon har en tendens til å redusere sin risikovurdering knyttet til misligheter. Labella & Koehler (2004) argumenterer imidlertid for at dilution effect kun har virkning under gitte forhold. De trekker frem at dersom ny irrelevant informasjon presenteres i etterkant av relevant informasjon, benyttes den nye informasjonen til å bekrefte den vurdering som allerede ble gjort basert på den relevante informasjon. Dette kaller de en bekreftelseeffekt, som resulterer i at dilution effect reduseres eller til og med reverseres. Av praktiske årsaker ble caset vårt utarbeidet med en høy risiko for uredelig regnskapsrapportering. Vi ser også at kontrollgruppen i gjennomsnitt har vurdert risikoen et sted mellom «moderat til høy» og «høy». Sett i lys av funnene til Labella & Koehler kan det tenkes at opplevd irrelevant informasjon utilsiktet har skapt en bekreftelseeffekt for vurderingen deres, da respondentene allerede hadde vurdert risikoen som høy.

En annen mulig forklaring kan være at tilleggsinformasjonen har skapt usikkerhet, og som respons har deltakerne økt sin risikovurdering.

En tredje forklaring kan være at deltakerne har opplevd en såkalt «framing effect». En framing effect kan kort forklares som en kognitiv bias hvor ens vurdering påvirkes av hvordan informasjonen er presentert, eller pakket inn (Plous, 1997; Beratšová et al., 2016). Tidligere revisjonsforskning viser at revisor, under visse forutsetninger, også er utsatt for denne effekt (Emby & Finley, 1997). Plasseringen av påstandene våre kan utilsiktet ha påvirket fremstillingen og presentasjonen av informasjonen. For eksempel avsluttes caset med påstand fire om at ledelsens og arbeidstakernes representanter indikerer at det er mulighet for streik i det kommende året, og caset avsluttes dermed med en negativ, irrelevant påstand.

Til slutt er en mulig forklaring at respondentene i eksperimentgruppen faktisk opplevde å få flere relevante mislighetsrisikofaktorer, enn kontrollgruppen. Vi ser fra manipulasjonstesten at 60% av respondentene vurderte påstand tre som relevant, på tross av at denne var irrelevant. Videre ser vi samme trend for påstand fire, seks og åtte, hvor henholdsvis 37.5%, 22,5% og 20% vurderte disse som relevante på tross av at de var klassifisert som irrelevante.

Ettersom vi ikke finner dilution bias, har vi heller ikke grunnlag for å vurdere en eventuell sammenheng med personlighet. Hypotese to, tre og fire søker å forklare forskjellen i risikovurderinger mellom gruppene, noe vi antok ville være forårsaket av dilution bias. Mann-Whitney U testen fant imidlertid en statistisk signifikant forskjell mellom gruppene, noe som indikerer at eksperimentgruppen er blitt manipulert. På grunn av hypotesenes utforming vil vi nå ta for oss sammenhengen mellom manipulasjonen og personlighet.

6.1.2 Diskusjon av hypotese 2

Hypotese 2: Forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket negativ emosjonalitet

Resultatene gir ikke statistisk grunnlag for å påstå at forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket negativ emosjonalitet, med et signifikansnivå på 5%, da p-verdien ligger over 0,05 (0,094).

Vi ser imidlertid av figur 4 en trend hvor personer som skårer høyt på negativ emosjonalitet øker sin risikovurdering ved irrelevant informasjon, i sammenligning med resten av utvalget som nå svakt reduserer sin risikovurdering. Dette kan tyde på at personlighetstrekket kan være med å forklare forskjellen mellom de to ulike gruppene. Til tross for at forskjellen ikke er statistisk signifikant, vil en p-verdi på 0,094 kunne sies å være marginal signifikant. Hadde vi benyttet et konfidensintervall på 90% ville vi hatt grunnlag for å hevde at resultatene var signifikante.

Trenden som vi observerer, er i tråd med forventningene våre; at individer som skårer høyt på trekket vil være mer utsatt for manipulasjonen vår. Dette på tross av at manipulasjonen går i motsatt retning av hva vi hadde forventet.

6.1.3 Diskusjon av hypotese 3

Hypotese 3: Forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket åpensinnhet

Resultatene gir ikke statistisk grunnlag for å påstå at forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket åpensinnhet, med et signifikansnivå på 5%, da p-verdien ligger over 0,05 (0,129).

Fra figur 5 ser vi også her en tydelig trend mellom gruppene. Personer som skårer høyt på personlighetstrekket øker sin risikovurdering ved irrelevant informasjon.

Når vi ser på gruppene justert for individer som skårer høyt på trekket åpensinnhet, reduserer de nå sin risikovurdering når de blir presentert for irrelevant informasjon. Denne trenden observeres også for trekket negativ emosjonalitet. Det kan derfor tenkes at trekkene, sett under ett, kan være med på å forklare hvorfor resultatene fra hypotese 1 ble slik de ble.

Selv om vi ikke finner signifikante resultater, betyr ikke nødvendigvis en p-verdi på 0.129 at hypotesen er feil. Dersom utvalget hadde vært større, ville det vært interessant å se om forskjellen mellom gruppene hadde blitt statistisk signifikant.

6.1.4 Diskusjon av hypotese 4

Hypotese 4: Forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket samvittighetsfullhet

Resultatene gir ikke statistisk grunnlag for å påstå at forskjeller i risikovurderinger mellom gruppene kan forklares av personlighetstrekket samvittighetsfullhet, med et signifikansnivå på 5%, da p-verdien ligger over 0,05 (0,923).

Fra tabell 8 observerer vi imidlertid at utvalget vårt som helhet har skåret høyt på dette trekket. Gjennomsnittsverdien ligger på 46,26 og medianverdien på 47. Dette stemmer overens med tidligere forskning innen personlighetstrekk hos revisorer, som indikerer at revisorer generelt skårer høyere enn befolkningen på trekket samvittighetsfullhet (Asare et al., 2023). I oppgaven vår definerte vi en høy skår som en skår over median-verdien for utvalget. Det faktiske antallet av individer som skårer høyt på trekket sammenlignet med befolkningen generelt, er derfor sannsynligvis høyere. Det foreligger derfor en mulighet for at et skjevt utvalg kan ha påvirket resultatet, da det vil være vanskelig å oppdage noen forskjell når store deler av utvalget i realiteten har skåret høyt på trekket. Det hadde derfor vært interessant å sett om resultatene hadde vært annerledes med et mer balansert utvalg, eller en annen definisjon av høy skår.

6.2 Konklusjon

Formålet med vår masterutredning var å undersøke om revisor er utsatt for dilution bias i en mislighetsrisikovurdering, samt om personlighet har en modererende effekt på en eventuell dilution bias. Dette ble gjort gjennom et case hvor mislighetsrisiko skulle vurderes, samt tilhørende kartlegging av personlighetstrekk gjennom personlighetstesten BFI-2.

Resultatene våre finner ingen dilution bias. Interessant nok, ser vi en signifikant forskjell mellom gruppene, hvor de som ble utsatt for irrelevant informasjon tok mer ekstreme vurderinger enn de som kun fikk relevant informasjon. Videre ser vi trender mellom gjennomsnittlige risikovurderinger og personlighetstrekk. Trekket som utpeker seg er negativ emosjonalitet, hvor personer som skårer høyt på trekket i større grad påvirkes av irrelevant informasjon i sine vurderinger. Funnet er imidlertid ikke signifikant, og våre resultater finner dermed ikke personlighetstrekk som en signifikant moderator.

6.3 Begrensinger

Studien vår inneholder flere begrensinger. Grunnet masteroppgavens omfang og naturlige tids- og ressursbegrensinger har vi et utvalg på 42 respondenter. Dette er et lite utvalg, som øker sannsynligheten for at resultatene er tilfeldige, og dermed ikke kan generaliseres til hele populasjonen revisorer. Som diskutert i 4.3.1 ble det foretatt et bekvemmelighetsutvalg.

Dette medførte at store deler av våre respondenter er MRR-studenter ved NHH, og studenter ved andre skoler i landet er dermed ikke representert. Videre er studenter brukt som surrogat for praktiserende revisorer. Flertallet av våre respondenter har dermed ikke praksis erfaring, i kombinasjon med at mislighetsrisikovurderinger ofte er komplekse. En kan dermed ikke utelukke at resultatene ville blitt annerledes dersom utvalget besto av revisorer med flere års arbeidserfaring.

Selve caset inneholder også begrensinger. Studien vår består av et eksperiment, og er dermed ikke en reell situasjon. For det første skal caset besvares individuelt. I virkeligheten ville man gjerne diskutert dette i et team, samt diskutert videre med klienten. Respondenten vil også kunne være påvirket av at den ikke bærer ansvar for sin beslutning. Svarene kan derfor være preget av mindre motivasjon til å gjøre gode vurderinger, i kombinasjon med mindre konsekvenstenkning enn man ville hatt i en jobbsammenheng. I tillegg er caset relativt forenklet i forhold til virkeligheten. Dette ble gjort slik at caset ikke skulle ta for lang tid å gjennomføre, for å unngå for stort frafall. Dette gjør at caset ikke er helt realistisk i forhold til den virkelige verden.

Det er også en fare for at svarene i personlighetstesten kan inneholde mangler. Til tross for at spørreundersøkelsen var anonym, kan personlige spørsmål være vanskelig å svare på, i tillegg til at de krever refleksjon og selvinnsikt. Videre kan vår definisjon av en høy skår være mangelfull, da vi kun ser på skåren i forhold til de andre respondentene i utvalget og ikke befolkningen generelt. Til slutt har vi valgt å begrense oppgaven til å se på skåren på hver enkelt dimensjon, fremfor samspillet mellom skåren på de ulike dimensjonene. Dette kan være problematisk, ettersom at et individs personlighetsfungering utgjøres av samspillet mellom skårene på de ulike dimensjonene.

6.4 Anbefalinger til videre forskningstema

Revisor blir stadig utsatt for større mengder informasjon. Å kunne skille mellom relevant og irrelevant informasjon er derfor avgjørende for å utføre en effektiv revisjon, og videre forskning på dilution bias vil dermed være viktig. Utførelse av vår studie uten de overnevnte begrensningene ville vært interessant. Flere av våre resultater er ikke signifikante, men vi observerer likevel flere trender. Det ville vært spennende å se hvordan trendene utviklet seg dersom man undersøkte et større utvalg.

Videre ville det vært interessant å gjennomføre samme studie på nytt hvor utvalget utelukkende består av revisorer med lengre erfaring, da de gjerne har mer kunnskap om mislighetsrisikovurderinger.

Funnene våre tyder på at dilution bias kun foreligger under visse rammer. En videre forskning på hva disse rammene innebærer ville vært givende. I lys av diskusjonen under 6.1.1 kan en ide være å gjennomføre studien med enten utelukkende negativ eller positiv irrelevant informasjon, for å se hvordan dette vil påvirke resultatene.

Anbefaling til videre forskningstema vil også innebære å utføre en studie hvor gjennomføring av caset er mer realistisk. Dette vil kunne øke validiteten og relevansen av studien, da resultatene vil være mer overførbare og gyldige for den virkelige verdien. Dersom man utfører studien med økt tid og ressurser, kan et forslag være å utføre en faktisk mislighetsrisikovurdering i et revisjonsselskap, samt tilhørende personlighetstest. Da vil utvalget også kunne kjenne på ansvar for sine beslutninger, da dette utføres i en jobbsammenheng. I tillegg får man et mer klart bilde på hvordan en mislighetsrisikovurdering faktisk utføres i praksis.

Videre kunne det vært interessant å gjennomføre samme studie på nytt, men uten de praktiske begrensningene som vi har tatt for oss vedrørende personlighetstesten. For eksempel kan fremtidig forskning se på samspillet mellom skåren på de ulike personlighetstrekkene, fremfor å se på hver enkelt trekk isolert sett. Et annet forslag kan være å sammenligne skårene fra personlighetstesten med befolkningen generelt, fremfor å benytte median verdien til utvalget.

Avslutningsvis vil det være nyttig å studere flere typer faktorer som kan påvirke dilution bias. Ved økt kunnskap om hva som forårsaker effekten, vil man også kunne forske videre på løsninger som kan begrense den. Et eksempel, som også vil gjøre studien mer realistisk, er å studere dilution bias ved gjennomføring av mislighetsrisikovurderinger i team.

Litteraturliste

- Ahmad, F. (2020). Personality traits as predictor of cognitive biases: moderating role of risk-attitude. *Qualitative Research in Financial Markets* 12(4), 465-484.
<https://doi.org/10.1108/QRFM-10-2019-0123>
- Asare, S. K., Brenk, H. V., Demek, K. C. (2023). Evidence on the homogeneity of personality traits within the auditing profession. *Critical Perspectives on Accounting*.
<https://doi.org/10.1016/j.cpa.2023.102584>
- Ashton, R. H. & Kramer, S. S. (1980). Students As Surrogates in Behavioral Accounting Research: some Evidence. *Journal of Accounting Research*, 18(1), 1-15.
<https://doi.org/10.2307/2490389>
- Beratšová, A., Krchová, K., Gažová, N. & Jirásek, M. (2016). Framing and Bias: A literature Review of Recent Findings. *Central European Journal of Management*, 3(2).
<https://doi.org/10.5817/CEJM2016-2-2>
- Bolger, N., & Schilling, E. A. (1991). Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality*, 59, 355-386. <https://doi.org/10.1111/j.1467-6494.1991.tb00253.x>
- Bolger, N., & Zuckerman, A. (1995). A framework for studying personality in the stress process. *Journal of Personality and Social Psychology*, 69, 890-902.
<https://doi.org/10.1037/0022-3514.69.5.890>
- Costa, P. T. & McCrae, R. R. (1992). *Neo PI-R professional manual*. Odessa, FL: Psychological Assessment Resources
- Cozby, P. (2012). *Methods in behavioral research* (11th ed., international ed., pp. XV, 416). McGraw-Hill.
- Dandurand, F., Shultz, T. R. & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods* 40(2), 428-434.
<https://doi.org/10.3758/BRM.40.2.428>

-
- Durand, R. B., Fung, L. & Limkriangkrai, M. (2019). Myopic Loss Aversion, Personality, and Gender. *Journal of Behavioral Finance*, 20(3), 1-15.
<https://doi.org/10.1080/15427560.2018.1511562>
- Durand, R. B., Newby, R. & Sanghani, J. (2008). An Intimate Portrait of the Individual Investor. *Journal of Behavioral Finance*, 9(4), 193-208.
<https://doi.org/10.1080/15427560802341020>
- Durand, R., Newby, R., Tant, K. & Trepongkaruna, S. (2013). Overconfidence, overreaction, and personality. *Review of Behavioural Finance*, 20(3), 104-133.
<https://doi.org/10.1108/RBF-07-2012-0011>
- Eilifsen, A., Kinserdal, F., Messier, W. F. Jr. & McKee, T. E. (2020). An Exploratory Study into the Use of Audit Data Analytics on Audit Engagement. *American Accounting Association*, 30(4), 75-103. <https://doi.org/10.2308/HORIZONS-19-121>
- Eilifsen, A., Kochetova, N. & Messier Jr., W. F. (2017). Mitigating the Dilutions Effect in Auditors' Judgment Using a Frequency Response Mode. Working Paper.
<https://www.nhh.no/globalassets/departments/accounting-auditing-and-law/working-papers/wp-rrr-2017-eilifsen-et-al.-mitigating-the-dilution-effect-in-auditors-judgments-using-a-frequency-response-mode.pdf>
- Eilifsen, A., Kochetova, N. & Messier Jr., W. F. (2019). Mitigating the Dilutions Effect in Auditors' Judgment Using a Frequency Response Mode. *Behavioral Research in Accounting*. <http://dx.doi.org/10.2139/ssrn.2609392>
- Emby, C. & Finley, D. (1997). Debiasing Framing Effects in Auditors' Internal Control Judgments and Testing Decisions*. *Contemporary Accounting Research*, 14(2), 55-77. <https://doi.org/10.1111/j.1911-3846.1997.tb00527.x>
- Engelberg, E., & Sjöberg, L. (2007). Money obsession, social adjustment, and economic risk perception. *The Journal of Socio-Economics*, 36(5), 686-697.
<https://doi.org/10.1016/j.socec.2007.01.005>
- Ewen, R. B. (2003). *An Introduction to Theories of Personality*. New Jersey: Lawrence Erlbaum Associates, Inc.

-
- Friedman, D. (1994). *Experimental methods: a primer for economists* (pp. XIV, 229). Cambridge University Press
- Funder, D. (1997). *The personality puzzle*. WW Norton & Co.
- Føllesdal, H. & Soto, C. J. (2022). The Norwegian Adaption of the Big Five Inventory-2. *Frontiers in Psychology 13 – 2022*. <https://doi.org/10.3389/fpsyg.2022.858920>
- Glover, S. M. (1997). The Influence of Time Pressure and Accountability on Auditors' Processing of Nondiagnostic Information. *Journal of Accounting Research*, 35(2) 213-226. <https://doi.org/10.2307/2491361>
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26-42. <https://doi.org/10.1037/1040-3590.4.1.26>
- Grønmo, S. (2016). *Samfunnsvitenskapelige metoder* (2.utg). Fagbokforlaget
- Gunthert, K. C., Cohen, H. L. & Armeli, S. (1999). The role of Neuroticism in Daily Stress and Coping. *Journal of Personality and Social Psychology*, 77(5), 1087-1110. <https://web.s.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=b8dee975-d676-4808-b020-eb16d8f75780%40redis>
- Hansen, J. & Klamm, B. K. (2012). A STUDY OF ACCOUNTING STUDENTS' ABILITY TO RECOGNIZE AND EVALUATE FRAUD RISK. *Journal of Theoretical Accounting Research*, 8(1), 1-23.
- Haselton, M. G., Nettle, D. & Murray, D. R. (2015). *The evolution of cognitive bias*. The handbook of evolutionary psychology.
- Heiman-Hoffman, V., Morgan, K. P., & Patton, J. M. (1996). The warning signs of fraudulent financial reporting. *Journal of Accountancy*, 182(4), 75. <https://www.proquest.com/trade-journals/warning-signs-fraudulent-financial-reporting/docview/206772468/se-2>
- Hoffman, V. B. & Patton, J. M. (1997). Accountability, the Dilution Effect, and Conservatism in Auditors' Fraud Judgments. *Journal of Accounting Research*, 35(2) 227-237. <https://doi.org/10.2307/2491362>

-
- IAASB (2009a). *ISA 500: Revisjonsbevis*. Oversatt av revisorforeningen i 2010.
<https://www.revisorforeningen.no/globalassets/fag/standarder-og-veiledninger/revisjonsstandardene/pr-20032023/isa-500-0922-20032023.pdf>
- IAASB (2009b). *ISA 240: Revisors oppgaver med og plikter til å vurdere misligheter ved revisjon av regnskaper*. Oversatt av revisorforeningen 2010.
<https://www.revisorforeningen.no/globalassets/fag/standarder-og-veiledninger/revisjonsstandardene/pr-09022021-regnskapsar-fra-15122021/isa-240-0221-fra-regnskapsar-som-begynner-15.12.21-eller-senere.pdf>
- IAASB (2019). *ISA 315 (revidert 2019): Identifisering og vurdering av risikoene for vesentlig feilinformasjon*. Oversatt av revisorforeningen i 2019).
<https://www.revisorforeningen.no/globalassets/fag/standarder-og-veiledninger/revisjonsstandardene/pr-20032023/isa-315-0922-20032023.pdf>
- IAASB. (2017). *Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics*. International Auditing and Assurance Standards Board.
<https://www.iaasb.org/publications/exploring-growing-use-technology-audit-focus-data-analytics>
- Jacobsen, D. I. (2015). *Hvordan gjennomføre undersøkelser?* (3 utg.). Cappelen Damm.
- Johannessen, A., Christoffersen, L., & Tufte, P. A. (2020). *Forskningsmetode for økonomisk-administrative fag*. Abstrakt forlag.
- Kennair, L. E. O. (2022). *Femfaktormodellen*. Store Norske Leksikon.
<https://snl.no/femfaktormodellen>
- Kennair, L. E. O. & Hagen, R. (2015). *Personlighetspsykologi*. Fagbokforlaget
- Koch, C. & Wüstemann, J. (2009). A Review of Bias Research in Auditing: Opportunities for Integrating Experimental Psychology and Experimental Economics. *Social Science Research Network*. <http://dx.doi.org/10.2139/ssrn.1032961>

-
- Kowert, P. A. & Hermann, M. G. (1997). Who Takes Risks? Daring and Caution in Foreign Policy Making. *The Journal of Conflict Resolution*, 41(5), 611-638.
<https://www.jstor.org/stable/174466>
- Kumar, V., Dudani, R. K., L. (2021). The big five personality traits and psychological biases: an exploratory study. *Curr Psychol* 42, 6587-6597.
<https://doi.org/10.1007/s12144-021-01999-8>
- Libby, R. (1981). Accounting and Human Information Processing: Theory and Applications. Prentice-Hall
- Lillefosse, M., L., R. & Tran, M. L. T. (2018). Effekten av irrelevant informasjon og erfaring i en mislighetsrisikovurderinger. *NHH Brage*. <https://openaccess.nhh.no/nhh-xmlui/handle/11250/2560264>
- Lin, H. W. (2011). Elucidating rational investment decisions and behavioral biases: evidence from the Taiwanese stock market. *African Journal of Business Management*, 5(5), 1630-1641. <https://academicjournals.org/journal/AJBM/article-full-text-pdf/AC702DF18509.pdf>
- Lærd Statistics (u.å.). Independent T-test using SPSS Statistics.
<https://statistics.laerd.com/spss-tutorials/independent-t-test-using-spss-statistics.php>
- Moen, T.G. og Havstein, B. (2017) *Regnskapsorganisering*. 7. utg. Oslo, Cappelen Damm
- Najm, Najm. (2019). Big Five Traits: A Critical Review. *Gadjah Mada International Journal of Business*, 21(2), 159-186.
https://www.researchgate.net/publication/335834113_Big_Five_Traits_A_Critical_Review
- Nicholson, N., Soane, E., Fenton-O'Creevy, M. & Willman, P. (2006). Personality and domain-specific risk taking. *Journal of Risk Research*, 8(2), 157-176.
<https://doi.org/10.1080/1366987032000123856>
- Nicholson, N., Soane, E., Fenton-O'Creevy, M. & Willman, P. (2005).

-
- Nisbett, R. E., Zukier, H. & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, 13(2), 248-277. [https://doi.org/10.1016/0010-0285\(81\)90010-4](https://doi.org/10.1016/0010-0285(81)90010-4)
- Personality and domain-specific risk taking. *Journal of Risk Research*, 8(2), 157-176. <https://doi.org/10.1080/1366987032000123856>
- Plous, S. (1993). *The Psychology of Judgment and Decision Making*. New York: McGraw-Hill Education
- PwC. (2022). *Global Economic Crime and Fraud Survey 2022*. PwC. <https://www.pwc.com/gx/en/services/forensics/economic-crime-survey.html>
- Rammstedt, B., Goldberg, L. R. & Borg, I. (2010). The measurement equivalent of big-five factor markers for persons with different levels of education. *Journal of Research in Personality* 44(1), 53-61. <https://doi.org/10.1016/j.jrp.2009.10.005>
- Revisorloven. (2020). *Lov om revisjon og revisorer* (LOV-2020-11-20-128). Lovdata. <https://lovdata.no/dokument/NL/lov/2020-11-20-128>
- Samagaio, A. & Felício, T. (2022). The influence of the auditor's personality in audit quality. *Journal of Business Research*, 141(1), 794-807. <https://doi.org/10.1016/j.jbusres.2021.11.082>
- Shelton, S. W. (1999). The Effect of Experience on the Use of Irrelevant Evidence in Auditor Judgment. *The Accounting Review*, 74(2), 217-224. <https://doi.org/10.2308/accr.1999.74.2.217>
- Suls, J., Martin, R., & David, J. P. (1998). Person-environment fit and its limits: Agreeableness, neuroticism, and emotional reactivity to interpersonal conflict. *Personality and Social Psychology Bulletin*, 24, 88-98. <https://doi.org/10.1177/0146167298241007>
- Triandis, H: & Suh E. M. (2002). Cultural Influences on Personality. *Annual Review of Psychology*, 53(1), 133-160. <https://doi.org/10.1146/annurev.psych.53.100901.135200>

-
- Trochim, W., Donnelly, J., & Arora, K. (2016). *Research methods: The essential knowledge base* (2nd ed.). Wadsworth Publishing
- Trotman, K. T. (2001). Design Issues in Audit JDM Experiments. *International Journal of Auditing*, 5, 181-192.
- Tversky, A., & Kahneman, D. (1975). *Judgment under uncertainty: Heuristic and biases. Utility, probability and human decision making*. Springer
- Wood, L. I. (2012). The impact of decision aid use on the dilution effect when assessing fraud. *Journal of Finance and Accountability*, 23-42.
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3bda120d852c042280d27d6b21f31d1f95053b43>
- 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*, 43(6), 1163-1174. <https://doi.org/10.1037/0022-3514.43.6.1163>

Vedlegg

Vedlegg 1 – Fullstendig eksperiment uten irrelevant informasjon

Du er revisor for Aktivt Liv AS som er lokalisert på Vestlandet. Dette selskapet driver med salg av treningsutstyr, og har en visjon om at alle skal ha råd til et aktivt liv. Forretningsmodellen går ut på å kjøpe varelagre fra avviklede bedrifter, for deretter å videreselge varene til enkeltpersoner. Aktivt Liv AS får kjøpt inn varene til sterkt redusert pris, og har dermed mulighet til å videreselge sportsutstyret til under markedspris. Selskapets totale eiendeler er på 20 millioner kroner, og resultatet er på 3 millioner. Gjeldsgraden er på 40%.

Dette er ditt første år som revisor for Aktivt Liv AS. Selskapet har akseptert tilbudet ditt, og det vil være tilstrekkelig med tid til å gjennomføre revisjonen. Ved påbegynt revisjon legger du raskt merke til at ledelsen er lite samarbeidsvillig. Ledelsen setter ikke av tilstrekkelig tid til å møte deg, og er treg med å oversende dokumentasjon.

I løpet av de siste årene har flere konkurrenter åpnet lignende butikker. For å møte den økende konkurransen ønsket eierne av Aktivt Liv AS å styrke insentivene til lederne ved å gi en bonus basert på årsresultatet. I fjor utgjorde bonusen 50% av ledernes totale lønn. Til tross for økt konkurranse er selskapet overbevist om at de vil dominere bransjen, og planlegger derfor å utvide lagerkapasiteten. De anslår at utvidelsen fører til at totale forpliktelser vil doble seg. Banken vil dermed kreve mer egenkapitalfinansiering og personlige garantier for bedriftens gjeld.

Aktivt Liv AS har tre avdelinger: markedsføring, økonomi og drift. Som følge av økt salg har bemanningen i markedsføring og drift doblet seg. Økonomiavdelingen er imidlertid begrenset, og driften av denne beskrives som uformell. Dette har ført til at eksempelvis kundefordringene ikke følges opp i tide.

Vennligst vurder den samlede risikoen for at det foreligger uredelig regnskapsrapportering som følge av eventuelle misligheter (fraud):

Svært lav risiko Lav risiko Lav til moderat risiko Moderat risiko Moderat til høy risiko Høy risiko

Kontrollspørsmål:

Kjønn

Hvor gammel er du?

Har du tatt eller tar du MRR?

Har du arbeidserfaring fra revisjon?

Diskuterte du caset med andre under gjennomføring av spørreundersøkelsen?

Vedlegg 2 – Fullstendig eksperiment med irrelevant informasjon

Du er revisor for Aktivt Liv AS som er lokalisert på Vestlandet. Dette selskapet driver med salg av treningsutstyr, og har en visjon om at alle skal ha råd til et aktivt liv. Forretningsmodellen går ut på å kjøpe varelagre fra avviklede bedrifter, for deretter å videreselge varene til enkeltpersoner. Aktivt Liv AS får kjøpt inn varene til sterkt redusert pris, og har dermed mulighet til å videreselge sportsutstyret til under markedspris.

Selskapets totale eiendeler er på 20 millioner kroner, og resultatet er på 3 millioner. Gjeldsgraden er på 40%.

Dette er ditt første år som revisor for Aktivt Liv AS. Selskapet har akseptert tilbudet ditt, og det vil være tilstrekkelig med tid til å gjennomføre revisjonen. Du får opplyst at avtroppende revisor klarte å gjennomføre revisjonen litt under budsjett året før. Ved påbegynt revisjon legger du raskt merke til at ledelsen er lite samarbeidsvillig. Ledelsen setter ikke av tilstrekkelig tid til å møte deg, og er treg med å oversende dokumentasjon. De ansatte er imidlertid godt fornøyd med ledelsen, da godene som er tilgjengelig for de ansatte er mer sjenerøse enn bransjegjennomsnittet.

I løpet av de siste årene har flere konkurrenter åpnet lignende butikker. For å møte den økende konkurransen ønsket eierne av Aktivt Liv AS å styrke insentivene til lederne ved å gi en bonus basert på årsresultatet. I fjor utgjorde bonusen 50% av ledernes totale lønn. Til tross for økt konkurranse er selskapet overbevist om at de vil dominere bransjen, og planlegger derfor å utvide lagerkapasiteten. De anslår at utvidelsen fører til at totale forpliktelser vil doble seg. Banken vil dermed kreve mer egenkapitalfinansiering og personlige garantier for bedriftens gjeld. Du får videre opplyst at en endring i kommunens skattesatsstruktur vil føre til en reduksjon i selskapets eiendomsskatt.

Aktivt Liv AS har tre avdelinger: markedsføring, økonomi og drift. Som følge av økt salg har bemanningen i markedsføring og drift doblet seg. Økonomiavdelingen er imidlertid begrenset, og driften av denne beskrives som uformell. Dette har ført til at eksempelvis kundefordringene ikke følges opp i tide. I tillegg får du opplyst at ledelsens og arbeidstakernes representanter indikerer at det er mulighet for streik i det kommende året.

Vennligst vurder den samlede risikoen for at det foreligger uredlige regnskapsrapportering som følge av eventuelle misligheter (fraud):

Svært lav risiko Lav risiko Lav til moderat risiko Moderat risiko Moderat til høy risiko Høy risiko

Vedlegg 3 – Relevante påstander benyttet fra Hansen og Klamm (2012)

- Several Competitors
- Growth...sales increasing from \$400,000 to \$8.3 M
- Debt ratio expected to increase...bank will no doubt require more equity financing and personal guarantees
- More competition...target of doubling sales from \$8.3 M to \$16 M during the upcoming year
- Manager bonus plans based on yearly sales; bonuses totaled nearly 50% of managers' total salaries
- Senior officers uncooperative during the financial audit
- Accounting staff is limited and operations are described as informal
- Policies not followed, e.g. credit manager does not age accounts receivable monthly

Vedlegg 4 – Irrelevante påstander benyttet fra Hoffman og Patton (1997)

The irrelevant unfavorable cues (irrelevant to fraud, but unfavorable in nature) were:

- Management and labor representatives indicate that there is a possibility of a strike in the coming year

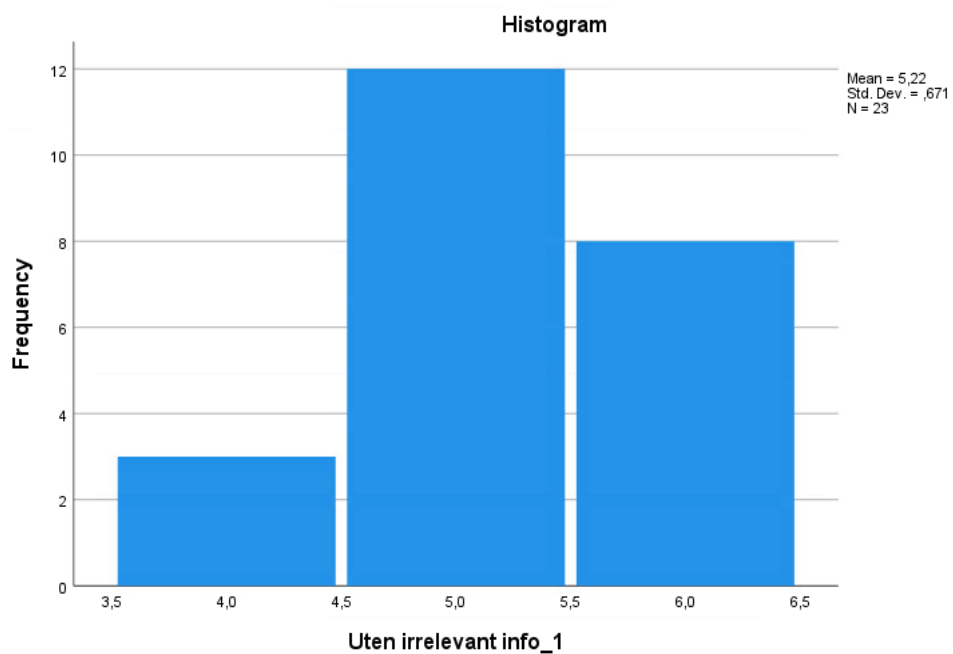
The irrelevant favorable cues (irrelevant to fraud, but favorable in nature) were:

- Last year the auditors were able to complete the engagement slightly under budget
- The benefit plans made available to client employees are more generous than the industry average.
- A change in the local tax rate structure caused a decrease in the client's property taxes

Vedlegg 5 – Signifikante uteliggere og normalfordeling

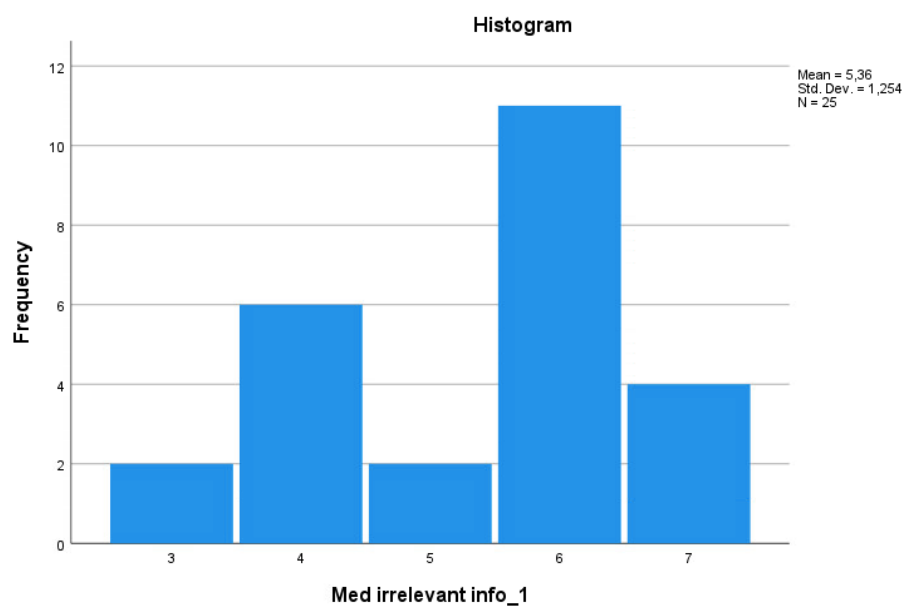
Histogram – uten irrelevant informasjon:

Uten irrelevant info_1

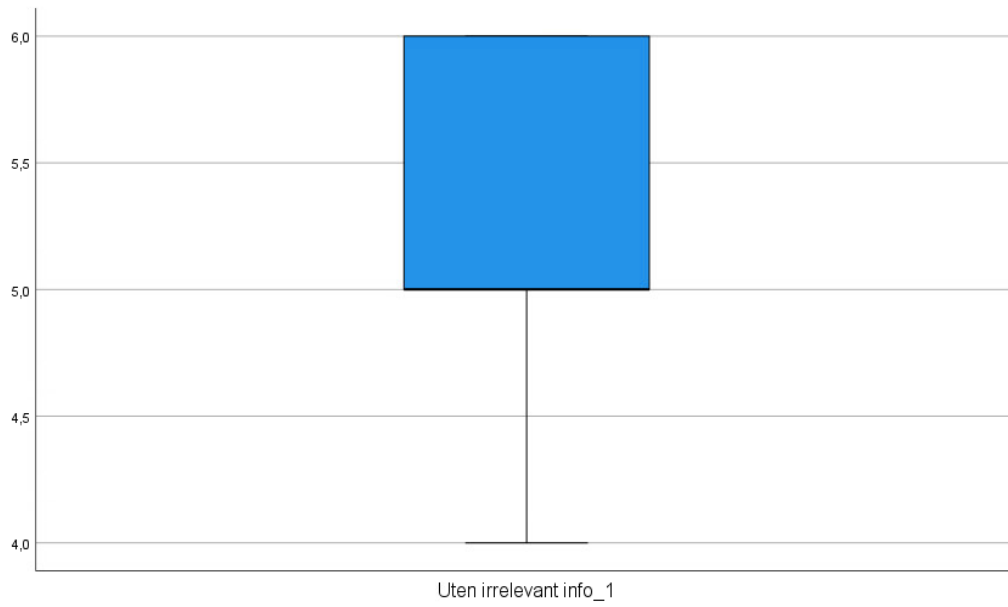


Histogram – med irrelevant informasjon:

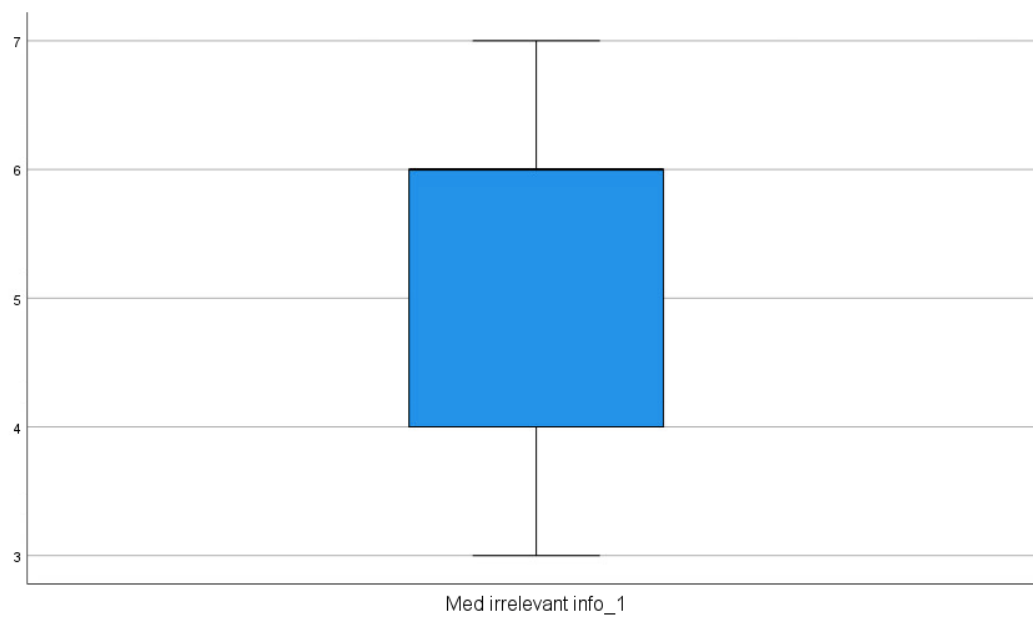
Med irrelevant info_1



BoksploTT – uten irrelevant informasjon:



BoksploTT – med irrelevant informasjon:



Descriptives:**Descriptives**

| | | Statistic | Std. Error | |
|------------------------|----------------------------------|-------------|------------|--|
| Uten irrelevant info_1 | Mean | 5,22 | ,140 | |
| | 95% Confidence Interval for Mean | Lower Bound | 4,93 | |
| | | Upper Bound | 5,51 | |
| | 5% Trimmed Mean | 5,24 | | |
| | Median | 5,00 | | |
| | Variance | ,451 | | |
| | Std. Deviation | ,671 | | |
| | Minimum | 4 | | |
| | Maximum | 6 | | |
| | Range | 2 | | |
| | Interquartile Range | 1 | | |
| | Skewness | -,280 | ,481 | |
| | Kurtosis | -,627 | ,935 | |
| Med irrelevant info_1 | Mean | 5,36 | ,251 | |
| | 95% Confidence Interval for Mean | Lower Bound | 4,84 | |
| | | Upper Bound | 5,88 | |
| | 5% Trimmed Mean | 5,40 | | |
| | Median | 6,00 | | |
| | Variance | 1,573 | | |
| | Std. Deviation | 1,254 | | |
| | Minimum | 3 | | |
| | Maximum | 7 | | |
| | Range | 4 | | |
| | Interquartile Range | 2 | | |
| | Skewness | -,481 | ,464 | |
| | Kurtosis | -,963 | ,902 | |

Tests of Normality

| | Kolmogorov-Smirnov ^a | | | Shapiro-Wilk | | |
|------------------------|---------------------------------|----|------|--------------|----|------|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Uten irrelevant info_1 | ,279 | 23 | ,000 | ,792 | 23 | ,000 |
| Med irrelevant info_1 | ,295 | 25 | ,000 | ,861 | 25 | ,003 |

a. Lilliefors Significance Correction

Vedlegg 6 – ANCOVA

Hypotese 1 – Enveis ANCOVA:

Tests of Between-Subjects Effects

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|--------|-------|---------------------|
| Corrected Model | 3,462 ^a | 4 | ,865 | ,870 | ,491 | ,086 |
| Intercept | 20,349 | 1 | 20,349 | 20,446 | <,001 | ,356 |
| Kjønn | ,542 | 1 | ,542 | ,544 | ,465 | ,014 |
| Alder | ,373 | 1 | ,373 | ,375 | ,544 | ,010 |
| Erfaring | 1,249 | 1 | 1,249 | 1,255 | ,270 | ,033 |
| Gruppe | 1,747 | 1 | 1,747 | 1,756 | ,193 | ,045 |
| Error | 36,824 | 37 | ,995 | | | |
| Total | 1278,000 | 42 | | | | |
| Corrected Total | 40,286 | 41 | | | | |

a. R Squared = ,086 (Adjusted R Squared = -,013)

Hypotese 2 – Toveis ANCOVA:

Tests of Between-Subjects Effects

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|--------|-------|---------------------|
| Corrected Model | 7,078 ^a | 6 | 1,180 | 1,233 | ,316 | ,188 |
| Intercept | 20,820 | 1 | 20,820 | 21,763 | <,001 | ,405 |
| Kjønn | 1,286 | 1 | 1,286 | 1,345 | ,255 | ,040 |
| Alder | ,918 | 1 | ,918 | ,960 | ,335 | ,029 |
| Erfaring | ,918 | 1 | ,918 | ,959 | ,335 | ,029 |
| Gruppe | 1,390 | 1 | 1,390 | 1,453 | ,237 | ,043 |
| N | ,746 | 1 | ,746 | ,780 | ,384 | ,024 |
| Gruppe * N | 2,388 | 1 | 2,388 | 2,496 | ,124 | ,072 |
| Error | 30,614 | 32 | ,957 | | | |
| Total | 1201,000 | 39 | | | | |
| Corrected Total | 37,692 | 38 | | | | |

a. R Squared = ,188 (Adjusted R Squared = ,035)

Hypotese 3 – Toveis ANCOVA:**Tests of Between-Subjects Effects**

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|--------|-------|---------------------|
| Corrected Model | 6,205 ^a | 6 | 1,034 | 1,051 | ,412 | ,165 |
| Intercept | 21,331 | 1 | 21,331 | 21,678 | <,001 | ,404 |
| Kjønn | ,766 | 1 | ,766 | ,778 | ,384 | ,024 |
| Alder | 1,100 | 1 | 1,100 | 1,118 | ,298 | ,034 |
| Erfaring | 1,470 | 1 | 1,470 | 1,494 | ,230 | ,045 |
| Gruppe | 1,120 | 1 | 1,120 | 1,138 | ,294 | ,034 |
| Å | ,860 | 1 | ,860 | ,874 | ,357 | ,027 |
| Gruppe * Å | 1,171 | 1 | 1,171 | 1,190 | ,283 | ,036 |
| Error | 31,487 | 32 | ,984 | | | |
| Total | 1201,000 | 39 | | | | |
| Corrected Total | 37,692 | 38 | | | | |

a. R Squared = ,165 (Adjusted R Squared = ,008)

Hypotese 4 – Toveis ANCOVA:**Tests of Between-Subjects Effects**

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|--------|-------|---------------------|
| Corrected Model | 4,347 ^a | 6 | ,725 | ,695 | ,655 | ,115 |
| Intercept | 20,650 | 1 | 20,650 | 19,817 | <,001 | ,382 |
| Kjønn | 1,670 | 1 | 1,670 | 1,603 | ,215 | ,048 |
| Alder | ,883 | 1 | ,883 | ,848 | ,364 | ,026 |
| Erfaring | 1,874 | 1 | 1,874 | 1,799 | ,189 | ,053 |
| Gruppe | 1,164 | 1 | 1,164 | 1,117 | ,299 | ,034 |
| S | ,266 | 1 | ,266 | ,255 | ,617 | ,008 |
| Gruppe * S | ,161 | 1 | ,161 | ,154 | ,697 | ,005 |
| Error | 33,345 | 32 | 1,042 | | | |
| Total | 1201,000 | 39 | | | | |
| Corrected Total | 37,692 | 38 | | | | |

a. R Squared = ,115 (Adjusted R Squared = -,051)

Vedlegg 7 – T-test og Mann-Whitney U test

T-test:

Group Statistics

| | Tildelt case | N | Mean | Std. Deviation | Std. Error Mean |
|--------------------------|-----------------------------|----|------|----------------|-----------------|
| Risikovurderinger samlet | Uten irrelevant informasjon | 22 | 5,23 | ,685 | ,146 |
| | Med irrelevant informasjon | 20 | 5,65 | 1,226 | ,274 |

Independent Samples Test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | |
|--------------------------|-----------------------------|---|------|------------------------------|--------|--------------|-------------|-----------------|-----------------------|---|-------|
| | | F | Sig. | t | df | Significance | | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | One-Sided p | Two-Sided p | | | Lower | Upper |
| Risikovurderinger samlet | Equal variances assumed | 3,847 | ,057 | -1,396 | 40 | ,085 | ,170 | -,423 | ,303 | -1,035 | ,189 |
| | Equal variances not assumed | | | -1,361 | 29,199 | ,092 | ,184 | -,423 | ,311 | -1,058 | ,212 |

Independent Samples Effect Sizes

| | | Standardizer ^a | Point Estimate | 95% Confidence Interval | |
|--------------------------|--------------------|---------------------------|----------------|-------------------------|-------|
| | | | | Lower | Upper |
| Risikovurderinger samlet | Cohen's d | ,980 | -,431 | -1,042 | ,184 |
| | Hedges' correction | ,999 | -,423 | -1,022 | ,181 |
| | Glass's delta | 1,226 | -,345 | -,956 | ,275 |

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control group.

Mann-Whitney U test:

| | | Ranks | | |
|--------------------------|-----------------------------|--------------|-----------|--------------|
| Tildelt case | | N | Mean Rank | Sum of Ranks |
| Risikovurderinger samlet | Uten irrelevant informasjon | 22 | 17,98 | 395,50 |
| | Med irrelevant informasjon | 20 | 25,38 | 507,50 |
| | Total | 42 | | |

Test Statistics^a

| Risikovurderinger samlet | |
|--------------------------|---------|
| Mann-Whitney U | 142,500 |
| Wilcoxon W | 395,500 |
| Z | -2,078 |
| Asymp. Sig. (2-tailed) | ,038 |

a. Grouping Variable: Tildelt case

Vedlegg 8 – Toveis ANOVA

Between-Subjects Factors

| | | Value Label | N |
|-----------------------|---|-----------------------------|----|
| Tildelt case | 1 | Uten irrelevant informasjon | 19 |
| | 2 | Med irrelevant informasjon | 20 |
| Negativ emosjonalitet | 0 | Ikke over median | 22 |
| | 1 | Over median | 17 |

Descriptive Statistics

Dependent Variable: Risikovurderinger samlet

| Tildelt case | Negativ emosjonalitet | Mean | Std. Deviation | N |
|-----------------------------|-----------------------|------|----------------|----|
| Uten irrelevant informasjon | Ikke over median | 5,40 | ,516 | 10 |
| | Over median | 5,11 | ,782 | 9 |
| | Total | 5,26 | ,653 | 19 |
| Med irrelevant informasjon | Ikke over median | 5,33 | 1,303 | 12 |
| | Over median | 6,13 | ,991 | 8 |
| | Total | 5,65 | 1,226 | 20 |
| Total | Ikke over median | 5,36 | 1,002 | 22 |
| | Over median | 5,59 | 1,004 | 17 |
| | Total | 5,46 | ,996 | 39 |

Tests of Between-Subjects Effects

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|----------|-------|---------------------|
| Corrected Model | 4,862 ^a | 3 | 1,621 | 1,728 | ,179 | ,129 |
| Intercept | 1150,704 | 1 | 1150,704 | 1226,743 | <,001 | ,972 |
| Gruppe | 2,139 | 1 | 2,139 | 2,280 | ,140 | ,061 |
| N | ,603 | 1 | ,603 | ,642 | ,428 | ,018 |
| Gruppe * N | 2,784 | 1 | 2,784 | 2,968 | ,094 | ,078 |
| Error | 32,831 | 35 | ,938 | | | |
| Total | 1201,000 | 39 | | | | |
| Corrected Total | 37,692 | 38 | | | | |

a. R Squared = ,129 (Adjusted R Squared = ,054)

Hypotese 3:**Between-Subjects Factors**

| | | Value Label | N |
|--------------|---|-----------------------------|----|
| Tildelt case | 1 | Uten irrelevant informasjon | 19 |
| | 2 | Med irrelevant informasjon | 20 |
| Åpensinnhet | 0 | Ikke over median | 20 |
| | 1 | Over median | 19 |

Descriptive Statistics

Dependent Variable: Risikovurderinger samlet

| Tildelt case | Åpensinnhet | Mean | Std. Deviation | N |
|-----------------------------|------------------|------|----------------|----|
| Uten irrelevant informasjon | Ikke over median | 5,44 | ,726 | 9 |
| | Over median | 5,10 | ,568 | 10 |
| | Total | 5,26 | ,653 | 19 |
| Med irrelevant informasjon | Ikke over median | 5,36 | 1,362 | 11 |
| | Over median | 6,00 | 1,000 | 9 |
| | Total | 5,65 | 1,226 | 20 |
| Total | Ikke over median | 5,40 | 1,095 | 20 |
| | Over median | 5,53 | ,905 | 19 |
| | Total | 5,46 | ,996 | 39 |

Tests of Between-Subjects Effects

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|----------|-------|---------------------|
| Corrected Model | 4,025 ^a | 3 | 1,342 | 1,395 | ,261 | ,107 |
| Intercept | 1161,771 | 1 | 1161,771 | 1207,746 | <,001 | ,972 |
| Gruppe | 1,624 | 1 | 1,624 | 1,689 | ,202 | ,046 |
| Å | ,206 | 1 | ,206 | ,214 | ,646 | ,006 |
| Gruppe * Å | 2,329 | 1 | 2,329 | 2,421 | ,129 | ,065 |
| Error | 33,668 | 35 | ,962 | | | |
| Total | 1201,000 | 39 | | | | |
| Corrected Total | 37,692 | 38 | | | | |

a. R Squared = ,107 (Adjusted R Squared = ,030)

Hypotese 4:**Between-Subjects Factors**

| | | Value Label | N |
|----------------------|---|-----------------------------|----|
| Tildelt case | 1 | Uten irrelevant informasjon | 19 |
| | 2 | Med irrelevant informasjon | 20 |
| Samvittighetsfullhet | 0 | Ikke over median | 22 |
| | 1 | Over median | 17 |

Tests of Between-Subjects Effects

Dependent Variable: Risikovurderinger samlet

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|-----------------|-------------------------|----|-------------|----------|-------|---------------------|
| Corrected Model | 1,514 ^a | 3 | ,505 | ,488 | ,693 | ,040 |
| Intercept | 1118,745 | 1 | 1118,745 | 1082,300 | <,001 | ,969 |
| Gruppe | 1,371 | 1 | 1,371 | 1,326 | ,257 | ,037 |
| S | ,043 | 1 | ,043 | ,042 | ,839 | ,001 |
| Gruppe * S | ,010 | 1 | ,010 | ,009 | ,923 | ,000 |
| Error | 36,179 | 35 | 1,034 | | | |
| Total | 1201,000 | 39 | | | | |
| Corrected Total | 37,692 | 38 | | | | |

a. R Squared = ,040 (Adjusted R Squared = -,042)

Descriptive Statistics

Dependent Variable: Risikovurderinger samlet

| Tildelt case | Samvittighetsfullhet | Mean | Std. Deviation | N |
|-----------------------------|----------------------|------|----------------|----|
| Uten irrelevant informasjon | Ikke over median | 5,25 | ,622 | 12 |
| | Over median | 5,29 | ,756 | 7 |
| | Total | 5,26 | ,653 | 19 |
| Med irrelevant informasjon | Ikke over median | 5,60 | 1,265 | 10 |
| | Over median | 5,70 | 1,252 | 10 |
| | Total | 5,65 | 1,226 | 20 |
| Total | Ikke over median | 5,41 | ,959 | 22 |
| | Over median | 5,53 | 1,068 | 17 |
| | Total | 5,46 | ,996 | 39 |