



Profesjonell skepsis når revisor møter kunstig intelligens og drone-teknologi: Nye metoder for inspeksjon og tilsyn

En eksperimentell studie

Eystein Sylte & Tobias Even Bull

Veileder: Jonas Gaudernack

Masteroppgave i Regnskap & Revisjon

NORGES HANDELSHØYSKOLE

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

Forord

Denne masterutredningen er skrevet som en del av mastergradstudiet i regnskap og revisjon (MRR) ved Norges Handelshøyskole (NHH). Masterutredningen ble skrevet våren 2023, og utgjør 30 studiepoeng.

Masteroppgaven er skrevet innenfor revisjonsforskning, noe som har vært en spennende og lærerik prosess. Vi har fått en god innsikt i hvordan eksperimentell metode fungerer som et forskningsverktøy, samt god innføring i programvaren SPSS. Videre har vi tilegnet oss kunnskap om et høyaktuelt tema som omhandler digitale verktøy i revisjonen, og hvordan dette vil forme fremtidens revisjon. Dette har vært svært relevant og nyttig for vårt fremtidige arbeid i revisjonsbransjen.

Vi ønsker å takke vår veileder Jonas Gaudernack for gode innspill og tilbakemeldinger underveis i prosessen. Videre vil vi også rette takk til Ulf Mohrmann for verdifulle innspill vedrørende gjennomføring og tolkning av statistiske analyser. Til slutt ønsker vi å takke våre medstudenter som har gitt oss påfyll av motivasjon, samt alle som tok seg tid til å gjennomføre undersøkelsen vår.

Norges Handelshøyskole

Bergen, mai 2023

Eystein Amund Sylte

Tobias Even Bull

Sammendrag

Revisjonsbransjen er avhengig av å investere betydelige summer i ny teknologi for å opprettholde kvalitet og effektivitet, noe som er en konsekvens av den raske teknologiske utviklingen i de siste årene. Imidlertid henger reguleringsmyndighetene og revisjonsstandardene etter når det gjelder å tilpasse seg denne utviklingen (Kinserdal, 2017). Dette gapet mellom teknologi og regulering fører til økt usikkerhet for sanksjoner og erstatningsansvar for revisorer som tar i bruk avanserte analytiske metoder i revisjonsprosessen (Eilifsen et al., 2020). Som et resultat øker revisors profesjonelle skepsis når de bruker teknologisk avanserte verktøy i revisjonsarbeidet (Nearon, 2005).

Vi gjennomførte et eksperiment med 70 deltakere som ble tilfeldig fordelt i tre grupper. Alle gruppene ble presentert med en case om revisjon av et stort varelager, og deretter ble de gitt forskjellig informasjon om hvordan de skulle utføre revisjonen. Gruppe 1 skulle revidere varelageret ved hjelp av en tradisjonell metode, mens gruppe 2 skulle bruke en analytisk metode med drone og kunstig intelligens. Gruppe 3 skulle også bruke samme metode som gruppe 2, men fikk tilleggsinformasjon fra reguleringsmyndigheter om risikoen for erstatningsansvar ved bruk av metoden. For å måle gruppenes profesjonelle skepsis, ble alle gruppene presentert med fem påstander om revisjonsbeviset som de skulle vurdere.

Resultatet fra eksperimentet viser at profesjonell skepsis var høyere ved bruk av analytisk metode i revisjonen sammenlignet med tradisjonell metode. Imidlertid var resultatet ikke signifikant nok til å konkludere med sikkerhet. Resultatene antyder også at profesjonell skepsis var høyere når respondentene ble gjort oppmerksom på risikoen ved bruk av analytisk teknologi i revisjonen av reguleringsmyndigheter, men resultatet var ikke signifikant nok til å trekke en endelig konklusjon.

Innholdsfortegnelse

Innhold

FORORD	II
SAMMENDRAG	III
INNHALDSFORTEGNELSE	IV
FIGURLISTE	VII
TABELLISTE	VII
1. INNLEDNING	1
1.1 MOTIVASJON FOR OPPGAVEN	1
1.2 FORSKNINGSMETODE.....	2
1.2.1 <i>Judgement and decision making</i>	2
1.3 AVGRENSNINGER.....	3
1.4 DISPOSISJON OVER RESTEN AV STUDIEN	3
2. LITTERATUR OG FORSKNING	4
2.1 PROFESJONELL SKEPSIS	4
2.2 REVISJONSBEVIS.....	6
2.3 TRADISJONELL REVISJONSMETODIKK AV VARELAGER	7
2.4 ANALYTISKE REVISJONSMETODER.....	8
2.4.1 <i>Audit Data Analytics</i>	8
2.4.2 <i>Dronerteknologi</i>	10
2.4.3 <i>Kunstig intelligens</i>	12
2.4.4 <i>Tidligere forskning av droner og kunstig intelligens i revisjon</i>	14
2.4.5 <i>First mover</i>	15
2.5 BEVISSTGJØRING	16

3.	KONSEPTUELL MODELL OG HYPOTESER.....	18
3.1	FORSKNINGSDESIGN	18
3.1.1	<i>Uavhengig Variabel</i>	19
3.1.2	<i>Avhengig variabel</i>	19
3.1.3	<i>Kontrollvariabel</i>	20
3.2	HYPOTESER	20
3.2.1	<i>Hypotese 1</i>	20
3.2.2	<i>Hypotese 2</i>	21
4.	METODE.....	22
4.1	DEDUKTIV TILNÆRMING.....	22
4.2	FORSKNINGSDESIGN: EKSPERIMENTELL METODE.....	23
4.2.1	<i>«Between-subjects»-design</i>	24
4.3	DATAINNSAMLING.....	25
4.3.1	<i>Pilottest</i>	25
4.3.2	<i>Utvalg</i>	25
4.4	PRAKTISK GJENNOMFØRING AV CASE.....	26
4.4.1	<i>Utforming av case</i>	26
4.4.2	<i>Vurderingsspmål</i>	27
4.5	RELIABILITET OG VALIDITET	29
4.6	ANALYSEMETODE.....	30
4.6.1	<i>Dataanalyse</i>	30
4.6.2	<i>Forutsetninger for analysen</i>	31
4.7	ETIKK	32

5. RESULTAT	33
5.1 RENSING AV DATA	33
5.2 DESKRIPTIV STATISTIKK	33
5.2.1 <i>Demografiske variabler</i>	34
5.3 HYPOTESETESTING	36
5.3.1 <i>Forutsetninger</i>	36
5.3.2 <i>Resultater hypotese 1:</i>	36
5.3.3 <i>Resultater hypotese 2:</i>	38
5.4 ANDRE FUNN	40
5.4.1 <i>Profesjonell skepsis som en tilstand</i>	40
5.4.2 <i>Positiv til bruk av drone i revisjon og vektleggelse av bevisstgjøring</i>	41
6. KONKLUSJON OG DISKUSJON.....	43
6.1 DISKUSJON AV HYPOTESER.....	43
6.1.1 <i>Hypotese 1</i>	43
6.1.2 <i>Hypotese 2</i>	44
6.1.3 <i>Andre funn</i>	46
6.2 KONKLUSJON.....	47
6.3 BEGRENSNINGER	48
6.4 VIDERE FORSKNING	49
LITTERATURLISTE.....	50
VEDLEGG.....	56
VEDLEGG 1: SPØRREUNDERSØKELSEN	57
VEDLEGG 2: TRIMMET GJENNOMSNIITT	67
VEDLEGG 3: CRONBACH'S ALPHA:.....	67

VEDLEGG 4: NORMALFORDELING – SHAPIRO-WILK TEST:.....	68
VEDLEGG 5: HYPOTESE 1.....	68
VEDLEGG 6: HYPOTESE 2.....	69
VEDLEGG 7: ANDRE FUNN	70
<i>Positiv til bruk av drone/Finanstilsynet.....</i>	<i>71</i>
<i>Gjennomsnitt profesjonell skepsis mot arbeidserfaring</i>	<i>72</i>

Figurliste

Figur 1: Dataoverføring fra ERP-systemer (Kinserdal,2017)	10
Figur 2: Sammenligning av Revisjonsmetodikk (Nehmer & Appelbaum, 2017)	12
Figur 3: Konseptuell modell: Libby bokser (Libby, 1981)	18
Figur 4: Deduktiv tilnærming (Trochim, William M. et al., 2014)	22
Figur 5: Faktorielt design	24
Figur 6: Hypotese 1 - Skeptisk holdning.....	37
Figur 7: Hypotese 2 – Skeptisk holdning	38

Tabelliste

Tabell 1: Fordeling av gruppene.....	34
Tabell 2: Fordeling av kjønn i gruppene	34
Tabell 3: Aldersfordeling i gruppene	35
Tabell 4: Arbeidserfaring i gruppene	35
Tabell 5: Hypotese 1 (Mann-Whitney-U-Test)	37
Tabell 6: Hypotese 2 (T-Test)	39
Tabell 7: Deskriptiv statistikk for Robinson score	40
Tabell 8: Sammenheng mellom skeptisk tilstand og skeptisk holdning.....	41
Tabell 9: Sammenligning positivitet og grad av vektlegging av kommentar fra FT.....	42

1. Innledning

1.1 Motivasjon for oppgaven

De siste års teknologiutvikling har vært så omfattende at det har vært vanskelig å holde følge. Kunstig intelligens har blitt et begrep man hører daglig, og programvare som blant annet ChatGPT¹ har muliggjort at dette blir mer og mer implementert i hverdagslige settinger og gjøremål. Man hører stadig at en stor del av arbeidsstokken vil bli automatisert, og revisjonsbransjen er intet unntak. I tillegg så vi under Covid-19 pandemien at det er mulig å gjennomføre revisjon uten å faktisk være fysisk til stede hos klienten (Sjanifer, 2020). Dette ved bruk av direktesendt videooverføring og bruk av droner til å inspisere bedriftene. Sistnevnte er noe som blant annet revisjonsselskapet PWC har vært offensive på til bruk av opptelling av større varebeholdninger, og som nå blir sett på som et viktig verktøy i den fremtidige revisjonen (PricewaterhouseCoopers, 2019). Ettersom droneteknologien også har utviklet seg til å bli mer avansert, har denne teknologien vist seg å både være tidsbesparende og presis ved innhenting av revisjonsbevis (Christ et al., 2021). Dette har ført til at den har blitt med i miksen av det som omtales som Audit Data Analytics, forkortet «ADA».

Det er likevel ikke slik at man kan benytte disse verktøyene «fritt frem» ved innhenting av revisjonsbevis. Standardsettere, offentlige reguleringsmyndigheter og revisorer har havnet i en «first-mover»-situasjon om hvem som skal ta steget om å benytte verktøyene først. Dette fordi nåværende standarder ikke gir tilstrekkelig veiledning om hva som er innenfor bruk, eller ikke (Kinserdal, 2017). De fremtidsrettede analytiske revisjonsmetodene har enda ikke blitt testet i en mulig erstatningssak mot revisor, noe som gjør at det heller ikke er noe rettspraksis på temaet. Dette betyr at mange revisorer ønsker å benytte disse metodene, men verger seg å ta det ut i full bruk grunnet usikkerheten rundt retningslinjene og risikoen dette medfører for revisor (Eilifsen et al., 2020).

En av grunnsteinene i revisjon er profesjonell skepsis. Dette vil si at revisor skal være spørrende og kritisk ved innhenting av revisjonsbevis. Det store spørsmålet i revisjonen i

¹ Chat-robot utviklet av OpenAI

forbindelsen med teknologiutviklingen er dermed hvordan skal revisor stole på revisjonsbevisene som er innhentet ved bruk av automatiserte verktøy? Og hvordan påvirker de lite tilrettelagte retningslinjene valgene som revisor tar? Dette vil kunne være med på å påvirke revisors profesjonelle skepsis. Vi har dermed utviklet to problemstillinger vi ønsker å se nærmere på:

- *Vil revisors profesjonelle skepsis være annerledes ved forskjellige revisjonsmetoder?*
- *Vil bevisstgjøring av risikoene knyttet til analytiske metoder påvirke revisors profesjonelle skepsis?*

Innføringen i disse problemstillingene og hvordan de digitale verktøyene vil påvirke fremtidens revisjon i kurset «Digital Revisjon» ved NHH, gav oss interesse for å utforske temaet nærmere. Dette i tillegg til den stadige aktualiteten med både droner og kunstig intelligens både i forskjellige virksomheter, og i hverdagen.

1.2 Forskningsmetode

1.2.1 Judgement and decision making

Beskrivelse og bedømmelse av revisors vurderings- og beslutningsevne er viktige temaer innen revisjonsforskning, der Judgment and decision making (JDM) utgjør en sentral del. JDM-forskningen har som mål å forklare hvordan vurderinger gjøres, samt å gi innsikt i de kognitive prosessene som påvirker revisors evne til å fatte beslutninger på en kvalitativ måte (Trotman et al., 2011).

Vårt bidrag til JDM-forskning handler om hvordan bruk av analytiske revisjonsmetoder påvirker revisors skeptiske holdning kontra det å benytte de tradisjonelle metodene. De internasjonale revisjonsstandardene krever at revisor utøver profesjonell skepsis gjennom sin virksomhet ved vurdering om regnskapet inneholder vesentlig feil (PCAOB, 2002). Ved å også undersøke effekten av bevisstgjøring på risikoene tilknyttet revisjonsbevisene, kan studien bidra til å forbedre revisors profesjonelle skepsis generelt sett, og se om påminnelse av erstatningsrisiko fungerer.

1.3 Avgrensninger

I denne masterutredningen har vi gjort noen avgrensninger for at oppgaven ikke skal bli for lang og kompleks. Vi har i oppgaven valgt å se bort fra regulatoriske myndigheter sine generelle regler for å fly drone. I tillegg har vi også valgt å avgrense ved å ikke se på ISA 520, som omhandler analytiske handlinger ved evaluering av finansiell informasjon i revisjonen. Vi anser standarden som ikke relevant ettersom vi ønsker å se på innhenting av revisjonsbevis isolert sett, mens standarden dekker mer å se på forskjeller i avvik av finansiell informasjon når revisjonsbevisene allerede er innhentet.

Videre er denne studien en masteroppgave, noe som gjør at det er en akseptert risiko for at antallet respondenter kan bli lavere enn det som er påkrevd fra et forskningsperspektiv. Dermed kan de statistiske analysene ofte gi resultater uten signifikans, men dette vil ikke hindre det overordnede læringsformålet med å gjennomføre et fullverdig forskningseksperiment.

1.4 Disposisjon over resten av studien

Kapittel to vil vi gi en oversikt over relevant litteratur og tidligere forskning som vi mener er sentral for forskningsområdet. Videre så vil kapittel tre gi en oversikt over den konseptuelle modellen og hvordan modellen er oppbygd. I tillegg vil den gi en oversikt over de to hypotesene vi har utarbeidet. Kapittel fire har som hensikt å gi en oversikt over hvilken metode som blir brukt, og hvordan eksperimentet blir utført. Resultatet vil bli presentert i kapittel fem. I kapittel seks vil vi kommentere funnene våre, og komme med en konklusjon. Det vil også bli diskutert begrensninger i studiet og videre forskning.

2. Litteratur og forskning

I dette kapitlet vil vi presentere relevant litteratur og tidligere forskning for våre problemstillinger.

2.1 Profesjonell skepsis

Profesjonell skepsis er et svært sentralt begrep innenfor revisjonslitteraturen og i praksis. Til tross for at begrepet er så sentralt i revisjonen, har man ikke klart å definere hva det faktisk er. ISA 200 definerer profesjonell skepsis som følgende: «En holdning som innebærer at revisor stiller spørsmål og er oppmerksom på forhold som kan indikere mulige feilinformasjon som følge av feil eller misligheter, og foretar en kritisk vurdering av revisjonsbevisene (IAASB, 2009, Pkt. 13 bokstav l). I dette ligger at det skal ligge i revisors natur å ha skepsis holdninger når han reviderer. PCAOB² definerer profesjonell skepsis som en holdning der revisor stiller kritiske spørsmål og er kritisk til vurderingen av revisjonsbevisene. Revisor skal likevel være objektiv og utføre revisjonen med god tro og integritet ved innsamling og vurdering av revisjonsbevisene (PCAOB, 2002).

En økende bruk av teknologi i revisjonen kan bidra til at revisjonen blir mer effektiv og øke kvaliteten på hele revisjonen. En høyere kvalitet på revisjonen kan bidra til at profesjonelle skepsisen forbedres. Likevel så stiller Arthur A. Hayes seg kritisk til om at profesjonell skepsis holdningen kan reduseres, fordi revisor stoler for mye på de digitale virkemidlene. Han utdyper det ved at de teknologiske hjelpemidlene ikke skal være et substitutt for den menneskelige vurderingen, og at menneskelig vurdering er den vurderingen som må ligge til grunn (Hayes Jr., Arthur A., 2014).

For å kunne måle profesjonell skepsis så har det blitt utviklet ulike måleverktøy. Kathy Hurtt har utviklet en modell som forklarer profesjonell skepsis på individ nivå. Modellen deler opp profesjonell skepsis som en tilstand og som et personlighetstrekk. Der personlighetstrekket er varig, mens tilstanden er kortvarig og oppstår på grunn av en gitt situasjon. For å kunne måle

² PCAOB – Public Company Accounting Oversight Board

profesjonell skepsis så har Hurtt utarbeidet en skala på 30 spørsmål, kalt *Hurtt's Professional Scepticism Scale* (HPSS). Hurtt's skala baserer seg på seks sentrale områder: et spørrende sinn, evne til overveielse, søken etter kunnskap, mellommenneskelig forståelse, selvtillit og autonomi. De tre første karakteristikene dreier seg om revisors innhenting av revisjonsbevis og undersøkelse av bevisene. Den mellommenneskelige forståelsen handler om at revisor må forstå det mellommenneskelige aspekter som ligger bak revisjonsbeviset. De to resterende karakteristikene dreier seg om revisors evne til å handle på den innhentede informasjonen (Hurtt, 2010).

Robinson, Curtis og Robertson har videre utviklet en skala basert på Hurtt sitt 30 spørsmåls skala for å kunne måle profesjonell skepsis. Den videreutviklede modellen ser profesjonell skepsis som en tilstand hos revisor, og tar utgangspunkt i at personligtrekk er stabil over tid. Ved å ta utgangspunkt i at personligtrekket til revisor er stabilt over tid, så tar den nye skalaen i bruk tre karakteristikker og ikke fem som Hurtt bruker. De tre karakteristikene som Robinson, Curtis og Robertson tar i bruk er: et spørrende sinn, evne til overveielse og søken etter kunnskap. Det gjør at den videreutviklede skalaen tar i bruk 12 spørsmål for å kunne måle profesjonell skepsis (Robinson et al. , 2018).

Nelson har et litt annet syn på profesjonell skepsis. Han definerer profesjonell skepsis som risikoen for at regnskapet inneholder vesentlig feilinformasjon. Skal revisor utøve profesjonell skepsis i revisjonssammenheng må både en skeptisk vurdering og en skeptisk handling være til stede. En skeptisk handling er basert på en skeptisk vurdering. Det vil si at skeptisk vurdering er revisors evne til å oppdage vesentlig feil i regnskapet, og at regnskapet dermed må undersøkes grundigere. Skeptisk handlingen skjer deretter på bakgrunn av revisors vurdering (Nelson, 2009).

2.2 Revisjonsbevis

Målet for revisor er å utforme og utføre revisjonshandlinger på en slik måte at revisor kan innhente *tilstrekkelig* og *hensiktsmessig* revisjonsbevis for å kunne trekke rimelige konklusjoner som grunnlag for meningen i revisjonsberetningen (IAASB, 2009b). IAASB definerer revisjonsbevis som «informasjon revisor bruker for å komme frem til konklusjonene som revisors mening bygger på». Revisjonsbevis er både informasjonen som finnes i regnskapsmaterialet, og informasjon som er hentet fra andre kilder. ISA 500 er revisjonsstandarden i det internasjonale rammeverket som forklarer hva som utgjør revisjonsbevis i en revisjon, og hva som er hensiktsmessig og tilstrekkelig revisjonsbevis.

Hensiktsmessighet og tilstrekkelighet av revisjonsbevis henger sammen med hverandre (IAASB, 2009b). Hensiktsmessighet er et mål på revisjonsbevis, som forklares med bevisets relevans og pålitelighet for å gi støtte for konklusjoner revisors uttalelser er bygd på. Hvor relevant beviset avhenger av retningen for testing, altså hvor relevant det er for påstanden som skal undersøkes. Pålitelighet til beviset påvirkes av dens kilde og natur, og er avhengig av de individuelle omstendighetene som det er oppnådd under.

Tilstrekkelighet er et mål av *mengden* revisjonsbevis, hvor mengden revisjonsbevis påvirkes av revisors vurdering av risikoen for feilinformasjon og bevisets kvalitet (IAASB, 2009b). Jo høyere risikoen for feilinformasjon er, jo mer revisjonsbevis er påkrevd. Om man har revisjonsbevis av høy kvalitet trenger man sannsynligvis færre revisjonsbevis for å konkludere påstanden. ISA 500 presiserer også at det er viktig at man ikke kompenserer mangel på kvalitetsbevis med å benytte seg av flere bevis av samme kvalitet.

Det har lenge blitt diskutert om digitale revisjonsbevis alene ikke er nok for at man kan konkludere at påstanden er hensiktsmessig og tilstrekkelig. Bruce H. Nearon mente allerede i 2005 at man i tillegg til å benytte digitale revisjonsbevis må man gjennomføre videre revisjonshandlinger for å skaffe betryggende bevis. Dette grunnet risikoen for at de digitale bevisene enklere kan endres og manipuleres. Nearon mener derfor at om man benytter digitale revisjonsbevis, må man ha en større grad av profesjonell skepsis, i forhold til bruk av de allerede etablerte metodene (Nearon, 2005).

2.3 Tradisjonell revisjonsmetodikk av varelager

Når revisor skal revidere et selskap er varelageret ofte en sentral del av revisjonen. Hos en handels eller produksjonsbedrift, vil ofte 20 til 60 prosent av balansen bestå av varelageret (Mjøsund & Stenvold, 2020). ISA 501 er den regulerende standarden for varelager, og har som mål at varelageret må oppfylle regnskapspåstandene eksistens og tilstand. For å kunne oppnå målet krever ISA 501 pkt. 4 at revisor skal være tilstede på varetelling (IAASB, 2009b). Dette for å kunne gjennomføre substanshandlinger som observasjon og stikkprøver under varetelling.

Revisor har frem til i dag ikke hatt mulighet til å teste alle transaksjonene og postene i regnskapet. Revisor benytter derfor stikkprøver for å kunne utale seg med betryggende sikkerhet for at regnskapet ikke inneholder vesentlig feilinformasjon. En stikkprøve er at revisor kontrollerer et utvalg av en regnskapslinje og ikke 100 prosent av dataen (Eilifsen et al., 2014). I praksis vil en stikkprøve være at revisor gjennomfører et utvalg³ av hele populasjonen. Revisor bruker deretter resultatet fra utvalget til å vurdere og trekke konklusjon om hele populasjonen inneholder vesentlig feil. Risikoen for at revisor konkluderer annerledes ved stikkprøve enn når revisor tester hele populasjonen, kalles utvalgsrisiko. Dette er risiko som vil være til stede, ettersom revisor kan trekke ut et utvalg som ikke representerer hele populasjonen (IAASB, 2009c).

Tradisjonell revisjonsmetodikk per dags dato består av svært mange manuelle operasjoner. Revisor oppsøker kunden fysisk for å inspisere varene på varelageret, for å så sjekke manuelt om fakturaene faktisk stemmer med reskontrosystemet til kunden (Kinserdal, 2017). Ved at dagens revisjon inneholder svært mange manuelle operasjoner, så fører det til at revisor bruker

³ Det finnes to forskjellige utvalgsmetoder. (1) statistisk metode – benytter sannsynlighetsteori for å finne utvalgsstørrelse og evaluering av utvalgsresultater. (2) ikke-statisk metode - benytter ingen sannsynlighetsteorier og innehar derfor ikke noen av egenskapene som en statistikk metode. Det er revisors skjønn og vurderinger som ligger til grunn ved bruk av metoden (IAASB, 2009c, s. 530)

80 prosent av tiden på å innhente nødvendig data, og kun 20 prosent av tiden går til å analysere dataen. Ved å kunne endre fordelingen, vil revisjonen kunne gi mer merverdier til kunden enn den gjør i dag (Pedersen, Jo Sigurd, 2016).

2.4 Analytiske revisjonsmetoder

2.4.1 Audit Data Analytics

Tiden hvor revisor sitter med tjukke billags-permer er for lengst over. Digitalisering er på full fart inn i alle deler av næringslivet, og revisjonsbransjen er ikke et unntak. Endringene vil føre til store endringer i hvordan revisjonsarbeidet utføres, og hvilken verdi klienten mottar fra revisjonsselskapene (Deloitte, 2019). Bare i de siste årene har de 4 store revisjonsselskapene investert over 90 milliarder dollar på kunstig dataanalyse produkter og opplæring av ansatte for å implementere dette både i konsulenttjenester og revisjon (Bloomberg, 2020). To nøkkelbegreper innenfor digitalisering er automatisering⁴ og Big Data⁵. Både nåværende og fremtidige revisorer vil derfor jobbe i større grad mer analytisk og dynamisk i forhold til den tradisjonelle revisjonsmetodikken (Kinserdal, 2017). Bruken av digitale verktøy i revisjonssammenheng er riktignok ikke noe som har kommet de siste årene. Den første generasjonen av digitale verktøy er den som omtales som CAATs (Computer Assisted Audit Techniques). CAATs er et samlebegrep for enklere dataprogrammer som har forenklet deler av revisjonsprosessen, blant annet ved den tradisjonelle bilagskontrollen hvor revisor tar et utvalg av alle de inngående fakturaene som en bedrift har arkivert for å identifisere avvik. Med hjelp av disse verktøyene kan man i større grad benytte seg av revisjonshandlingen «scanning», hvor revisor gjennomgår regnskapsdata for å finne «signifikante eller unormale

⁴ Automatisering er teknikken å få systemer til å fungere uten, eller med liten grad av menneskelig medvirkning. Automatisering benyttes på alle områder der man vil erstatte menneskelig arbeidskraft med selvvirkende systemer, for eksempel i industri og transport. (Redaksjonen, u.å.)

⁵ Big Data er et begrep som brukes om datasett som er for store eller komplekse å håndtere ved hjelp av tradisjonelle teknologier og metoder for databehandling. (PricewaterhouseCoopers, u.å.-a)

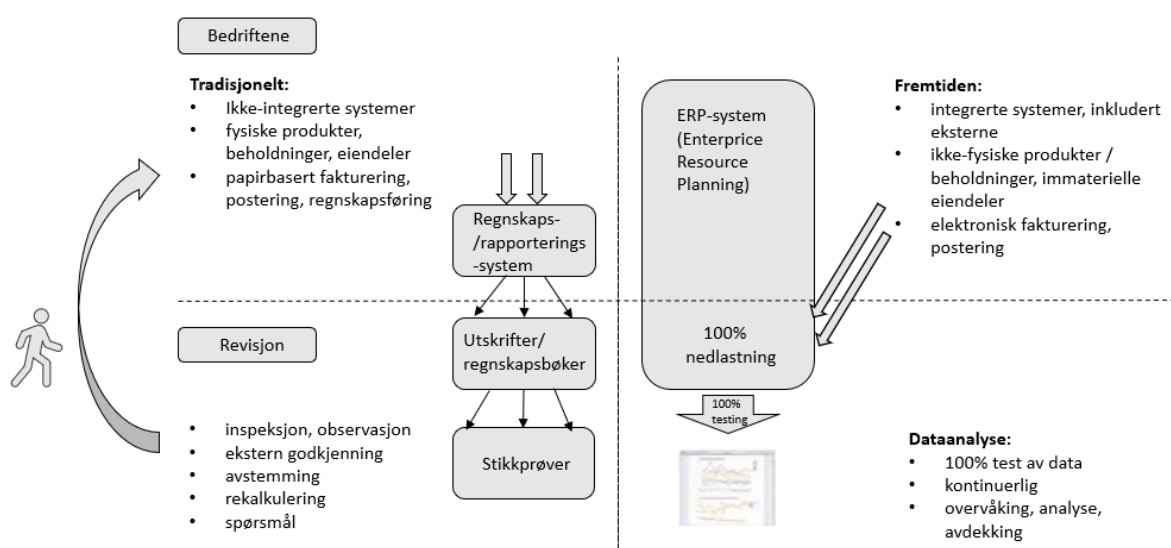
enheter» (Kulset, 2020). Noen eksempler på CAATs er programvare som ACL (Audit Command Language), Excel og SQL server (DataConsulting, 2009).

Audit Data Analytics (ADA) bygger videre på de analytiske verktøyene som kan benyttes i revisjonen. Audit Data Analytics defineres som: Vitenskapen og kunsten å oppdage og analysere mønstre, identifisere anomalier og utvinne andre nyttige informasjon i data som ligger til grunn for eller relatert til gjenstanden for revisjon, gjennom analyse, modellering og visualisering for formålet med å planlegge eller utføre revisjonen (Eilifsen et al., 2020). Enklere forklart er ADA et samlebegrep for bruk av dataanalyser, automatisering, kunstig intelligens og maskinlæring⁶, for å trekke konklusjoner om revisjonspåstander. Målet for ADA er å hjelpe revisorer å skaffe en dypere forståelse av de finansielle og operasjonelle aktivitetene til organisasjonen som revideres, identifisere mulige risikoer, og forbedre effektiviteten i revisjonsprosessen. Ved å kunne analysere store mengder av data, kan man i revisjonen finne trender og avvik (anomalier) som kan indikere misligheter, regnskapsfeil eller andre avvik som kan lede til at man gjør flere undersøkelser. Eksempler på programvare som går under ADA er Excel, Tableau, PowerBI, eller mer revisjonsorienterte som ACI eller CaseWare (Brataas, 2022). De store revisjonsselskapene har også utviklet sine egne programvarer som de benytter internt i sine organisasjoner. Blant annet har PWC utviklet dataanalyseverktøyet Halo, hvor man enkelt kan analysere hvordan 100% av transaksjonene flyter gjennom en viss forretningsprosess, og dernest identifisere unormale transaksjoner som bør undersøkes nøyere (PricewaterhouseCoopers, u.å.-c).

ADA har potensiale til å benyttes i alle stegene i revisjonen, fra klientaksept, til man skal konkludere og avlegge revisjonsberetningen. En stor fordel med ADA er at man med hjelp av programvare kan teste 100% av populasjonen, snarere enn den vanlige revisjonsmetoden hvor man tester et lite utvalg av dataen, blant annet ved hjelp av substanshandlingen stikkprøver. For at dette skal kunne brukes i revisjonen, avhenger dette også at klienten har datamaterialet tilgjengelig. Når kundene i liten grad er automatisert og digitalisert innen økonomiområdet, blir også revisjonen manuell. I fremtiden kommer kundenes ERP-systemer (Enterprise Resource Planning) i større grad til å være mer integrerte, også for økonomifunksjonene i

⁶ Maskinlæring er en spesialisering innen kunstig intelligens hvor man bruker statistiske metoder for å la datamaskinen finne mønstre i store datamengder. Man sier at maskinen «lærer», i stedet for å bli programmert. (Tidemann & Elster, 2023)

foretaket. Ved å kombinere dette med felles kontoplaner, integrerte konsolideringsløsninger og kontrollsystemer, vil man få mulighet til å revidere på en helt ny måte (Kinserdal, 2017). Dette kan så opplastes direkte til revisjonsforetaket, noe som også åpner for at revisjonen blir kontinuerlig gjennom hele året, såkalt «continius auditing».⁷ Man kan derfor ved et integrert ERP-system bygge revisjonen i større grad på kontrollhandlinger og «overvåkning», og minimere bruken av substanshandlinger.



Figur 1: Dataoverføring fra ERP-systemer (Kinserdal, 2017)

2.4.2 Dronerteknologi

Covid-19 pandemien gjorde at man måtte starte å tenke nytt også i revisjonsbransjen. Grunnet nedstengingene var det ikke slik at revisorene alltid hadde mulighet til å gjøre de planlagte varetellingene fysisk slik som det er krevd at man skal gjøre. Dette gjorde at man begynte å benytte seg av allerede eksisterende teknologi som droner og videooverføring for å gjøre varetellingene (Sjanifer, 2020). Droner i seg selv har ikke blitt kategorisert som ADA, men

⁷ Continius auditing omhandler at revisjonen foregår året rundt, og ikke bare i små perioder i løpet av året. Dette inkluderer overvåking av viktige KPI'er, internkontroller og transaksjoner i sanntid. Misligheter og avvik er noe som ofte oppdages etter at det har skjedd, og dette vil hjelpe til å forhindre dette. (PricewaterhouseCoopers, u.å.-b)

med den raske og avanserte utviklingen av droner har gjort slik at dette er noe som kan innhente viktig informasjon, og dermed være med i verktøy-miksen i den digitale revolusjonen i revisjonsbransjen. Droner kan gjøre seg nyttig i mange av forskjellige sektorer i fremtiden. PWC estimerer at det fremtidige globale markedet for droneløsninger i forskjellige sektorer som infrastruktur, transport, landbruk og forsikring vil være på 127 milliarder dollar. (PricewaterhouseCoopers, 2016). Ettersom hver industri har forskjellige behov, så vil dronene kunne være spesifikt tilpasset for hvilken bruk den skal gjøre. Noen av dem vil kunne fly raskt eller ha stor lastekapasitet, hvor noen kan levere data i sanntid gjennom sensorer i høy kvalitet, på en kostnadseffektiv måte. Sistnevntes kvaliteter er det som trengs i revisjonen for å gjennomføre substanshandlinger. Videre er det forskjell på bemannede droner og ubemannede. De bemannede er de man er mest kjent med som har egen pilot, hvor droner som er ubemannet kjører av seg selv og følger et mønster etter hvordan de er innstilt.

Dronene som er benyttet i mange industrier er typisk utstyrt med kameraer, videokameraer, RFID-brikker⁸, sensorer og GEO-lokaliseringsutstyr (Nehmer & Appelbaum, 2017). Videooverføring med RFID/sensor sporing blir sett på som en automatisk form for datainnhenting, avhengig av hvilken data revisor ønsker å anskaffe. Et eksempel på dette i praksis har blitt gjort med bilprodusenten Daimler hvor de festet RFID-merker på bilene de hadde på varelageret. Ettersom bilene stod tett i tett var det vanskelig å skanne dette manuelt, men ved bruk av drone med RFID-sensor kunne man skanne bilene ovenfra, noe som fungerte utmerket (Banker, 2016). Dette åpner for at man kan feste tilsvarende merker på andre typer varebeholdninger. Videre åpner det også for at kundens egne kontrollører også kan benytte denne teknologien ved at man for eksempel måler og observerer beholdningen en gang i timen. Dermed får man mulighet til å innhente sanntidsdata fra beholdningen som videre kan sendes til ekstern revisor, som gjennom sine programvarer kan observere og analysere dataen. Dette blir da en form for dynamisk revisjon eller continuous auditing, og som kan danne et godt grunnlag når den endelige revisjonen skal gjøres (Nehmer & Appelbaum, 2017).

⁸ Radiofrekvensidentifikasjon: en teknologi for automatisk identifikasjon av objekter ved hjelp av radiobølger. Systemet består av en leser og en «tag», som er små enheter som kan festes til objekter som skal spores (Techtarget.com, u.å.)

Bruken av droner vil i revisjonsprosessen være mest nyttig på to områder når revisjonsbevisene skal innhentes: fysisk inspeksjon og observasjon. Nehmer og Appelbaum har laget et rammeverk hvor og hvordan bemannede droner og ubemannede droner kan benyttes i revisjonen:

	Nåværende tradisjonell revisjonsmetode	Revisjonsmetodikk med bemannet drone	Revisjonsmetodikk med ubemannet drone
<u>Fysisk inventar:</u>			
1. Evaluere	Verifisere at prosedyrer er i samsvar med det som er rapportert	Dronen kan innhente bilder av flytskjemaer og lese / analysere resultater	Dronen kan innhente bilder av flytskjemaer og lese / analysere resultater
2. Observasjon	Observere/følge prosedyrer	Kan observere og følge prosedyrer som innstilt / styrt av pilot	Dronen kan observere og følge prosedyrer basert på video og sensor-sporing
3. Inspeksjon	Visuelt og/eller fysisk inspeksjon av varelageret	Dronen kan observere / vise varelageret selv eller observere at det blir gjort av mennesker	Dronen kan observere og følge inspeksjon med video og sensor-sporing
4. Utførelse	Må kanskje gjøre gjentakelse av opptelling av varelageret	Dronen kan gjøre gjentakelse av opptelling -> Data lastes opp i applikasjon	Dronen gjør konstant gjentakelse av opptelling -> lastes direkte opp i applikasjon som gjentar prosessen
<u>Hendelser:</u>			
1. Observasjon	Følge med på prosesser og aktiviteter	Dronen kan følge med på prosesser og kontrollaktiviteter	Dronen kan se og observere basert på video-input og analyser
<u>Verdsettelse:</u>			
1. Inspeksjon	Visuelt inspiserer eiendelen for avskrivninger / nedskrivninger	Dronen kan visuelt inspiserer eiendelen for avskrivninger/nedskrivninger. Sikkert for ansatte	Dronen kan visuelt gjøre inspeksjon basert på video-overføring og GPS-sensor

Figur 2: Sammenligning av Revisjonsmetodikk (Nehmer & Appelbaum, 2017)

Det vil naturligvis ikke alltid passe å benytte droneteknologi i revisjonen av varelager og fysiske eiendeler. Et spørsmål man kan stille er hvordan dronen skal inspiserer varer som ligger bak andre varer i en varehyll. Dette kan være problematisk fordi det kan være varer som ikke blir plukket opp av scanningen til dronen, noe som gjør at man vil få avvik. Dette gjør da at revisor likevel må flytte på gjenstandene, og gjøre manuelle inspeksjoner av eiendelene.

2.4.3 Kunstig intelligens

Etter flere tiår med relativt lite fremgang på kunstig intelligens-feltet, ser flere forretningsområder nå en ny «vår» med massive og overbevisende fordeler. Dette gjelder også i revisjonen, hvor teknologien vil kunne utvikle og effektivisere revisjonen videre i fremtiden (Issa et al., 2016). Kunstig intelligens er programvare som er utviklet for å etterligne

mennesket sin dømmekraft og kognitive evne, som kan ta hensyn til miljøet rundt seg. Dette gjør at kunstig intelligens kan vurdere risiko for å ta beslutninger, spådommer eller utføre handlinger. Denne type programvare er utformet slik at den lærer og utvikler seg over tid. Jo mer informasjon den får, jo smartere blir den uten å bli programmert av mennesker (Munoko et al., 2020).

Begrepet kunstig intelligens kan deles inn i to områder, svak og sterk KI (AI). Svak kunstig intelligens er laget for å utføre en spesiell oppgave og utfører den særdeles godt. Den vil derfor ha begrensninger, siden dens oppgave er veldig smale (PricewaterhouseCoopers, u.å.-d). Et eksempel på svak kunstig intelligens vil typisk være teknologien bak førerløse biler eller teknologien bak google sin søkemotor. Sterk kunstig intelligens er teknolog som kan løse alle type problemer, generalisere læring, og høste erfaring fra et problem til et annet. Dermed klare å tenke helhetlig rundt komplekse spørsmål. Sterk KI er per dags dato kun et akademisk begrep og brukes ikke i praksis (PricewaterhouseCoopers, u.å.-d).

Det er mulig å benytte seg av kunstig intelligens i ulike deler av revisjonsprosessen. En av bruksområdene er å identifisere ulike trender i hovedboken. PWC har utviklet en teknologi kalt GL.AI som spesifikt setter søkelys på å oppdage unormale transaksjoner og aktiviteter i hovedboken (PricewaterhouseCoopers, 2018). Ved å fange opp såkalte "røde flagg" vil teknologien gi revisor en indikasjon på hvilke transaksjoner som bør undersøkes nærmere. Dette vil være et nyttig verktøy for å øke revisjonens effektivitet.

En annen form for kunstig intelligens som kan benyttes i revisjonsprosessen, er automatisert opptellings-teknologi, for eksempel tjenesten *Countthings.com*. Revisoren kan ta i bruk denne teknologien under varetelling ved å ta bilder med en drone og laste dem inn i programvaren, som deretter kan telle antall objekter i bildet. Dette verktøyet vil trolig øke effektiviteten i varetellingen og frigjøre tid til andre oppgaver.

En annen teknologi som kan bidra til økt effektivitet og redusert risiko for vesentlige feil i revisjonsprosessen er process-mining. Process-mining innebærer automatisert kontroll og sjekk av alle kontrollene i en prosess for å sikre at alle transaksjoner går gjennom de ulike stegene. Med andre ord, process-mining fungerer som en automatisk "Walk-Trough". Teknologien varsler revisoren om transaksjoner som unngår enkelte av kontrollstegene i transaksjonsprosessen, slik at revisoren kan identifisere svake kontroller og iverksette tiltak.

Ved å bruke process-mining kan revisoren teste alle transaksjonene i selskapet og ikke bare et utvalg, slik det gjøres i dagens revisjonsprosess (Jans et al., 2013).

Anvendelse av kunstig intelligens i fremtiden vil føre til betydelige endringer i dagens tradisjonelle revisjonsmetoder. En stor utfordring for bruken av kunstig intelligens i dag er revisjonsstandardene. For å implementere kunstig intelligens i revisjonsprosessen, må revisjonsstandardene oppdateres og oppmuntre revisjonselskapene til å ta i bruk denne teknologien (Issa et al., 2016).

2.4.4 Tidligere forskning av droner og kunstig intelligens i revisjon

Spørsmålet er om denne teknologien kan forbedre revisjonskvaliteten. (Christ et al., 2021) gjennomførte et eksperiment hvor man benyttet droner og kunstig intelligens for å innhente revisjonsbevis fra varelager. Varelageret i dette eksperimentet bestod av 33 000 levende kveg og sau fra en gård, som var spredt utover et større inngjerdet område. I stedet for å gjøre manuell opptelling av dyrene, fløy man en operert drone over området, og fikk tatt bilder av hele populasjonen. Disse bildene ble så overført til programvare (Countthings.com) som benytter kunstig intelligens for å gjøre opptelling av populasjonen. For å finne ut om bruken av disse verktøyene ville forbedre revisjonskvaliteten benyttet de tre dimensjoner av revisjonskvalitet; *efficacy*, *effectiveness of quality of audit*. Effektivitet måles av hvor lang tid man benytter, og hvor stor feilmargin det er. Resultatene fra undersøkelsen viste at ved hjelp av drone og kunstig intelligens gikk fra 681 timer benyttet til 19 timer. I tillegg fikk man redusert feilmarginen på tellingen fra 0,15% til 0,03%. Dette viser at det er potensiale å spare mye tid, og at software som teller automatisk kan være veldig presise om man benytter teknologien riktig. Videre for å måle kvaliteten av revisjonen intervjuet de revisjonspartnere og medlemmer av standardsetterne AICPA og PCAOB. Intervjuobjektene var enige i at den analytiske revisjonsmetodikken var bedre egnet enn den tradisjonelle metoden å gjøre det på, men at man trenger klarere veiledning for å implementere dette i revisjonen. I tillegg var revisjonspartnerne usikker om de nåværende standardene tillater bruken av denne type verktøyer (Christ et al., 2021).

2.4.5 First mover

De nye digitale verktøy gjør at man kan fornye og effektivisere revisjonsprosessen. Imidlertid er reguleringsmyndigheter og de internasjonale revisjonsstandardene i liten grad tilpasset den nye digitale, nye revisjonsverden (Kinserdal, 2017). Etersom de fleste av standardene kom før nye digitale verktøy, er disse utformet for at man skal gjennomføre revisjonen på den tradisjonelle metoden. Dette betyr at det ikke er definert hva som er lov og hva som ikke er lov å benytte av dataanalyse-verktøy for å innhente revisjonsbevis. Revisjonsstandardene er ikke noe som oppdateres fortløpende, og det kan ta mange år fra det kommer et forslag til det blir implementert. Dette gjør at det har oppstått en «first mover» situasjon, altså om hvem som skal sette inn det første støtet om å virkelig benytte denne type teknologi i en revisjonsprosess.

I Norge er det Finanstilsynet som fører tilsyn over at revisor følger de internasjonale reglene som er satt av standardsetterne. Finanstilsynet fører jevnlig kontroller over at revisjonen er utført i samsvar med lover, forskrifter og god revisjonsskikk, og revisor er pliktig til å utgi informasjon og dokumentasjon om hvordan arbeidet er utført (Finanstilsynet, 2017). Tilsynet betyr at revisor må klare å forsvare bruken av nye revisjonsmetoder som ikke er definert i revisjonsstandardene i en mulig rettsak.⁹ I en tradisjonell revisjon er planlagt bra og man stoler på kontrollene, er en utvalgsstørrelse på 25 godt nok i en rettsak om man skal forsvare seg. Ved bruk av digitale verktøy har ingen vært i retten enda, og risikoen blir derfor stor selv om man har testet 100% av populasjonen. Etersom bruken av denne teknologien ikke er definert i standardene og aldri har vært forsøkt testet i en rettsak, hindrer dette revisorer fra å innlemme teknologien. En tidligere undersøkelse gjort av (Eilifsen et al., 2020) viser til at partnere i revisjonsselskapene ikke benytter tilgjengelig teknologi grunnet usikkerheten om at tilsyn vil akseptere det som revisjonsbevis eller ikke. Likevel er bruken av ADA høyt oppe på agendaen, selv om det ikke brukes i oppdrag der det er forventet at man burde benyttet det.

Et annet eksempel er om man benytter drone og kunstig intelligens under inspeksjon og observasjon av varelager. Når revisor har et krav om å være fysisk til stede under varetellingen, vil man da være fysisk til stede ved bruk av en drone eller videooverføring? Og er det

⁹ Revisor kan bli erstatningsansvarlig ved påført skade som han/hun forsettlig eller uaktsomt har voldt under utførelsen av sitt oppdrag. (*Revisors erstatningsansvar for vesentlige feil i årsregnskapet | Revisjon og Regnskap, u.å.*)

revisjonsbevis? Diskusjonen er om dette er godt nok eller ikke, og det har ikke Finanstilsynet svart på. De internasjonale standardsetterne er klar over problemstillingen om at nåværende standarder ikke oppmuntrer til bruk av avansert digital teknologi, indikerer når man kan benytte det eller utlyser risikoen tilknyttet (PCAOB, 2020). De ønsker dog ikke å gjøre store endringer på standardene ettersom man er i tvil om at revisjonsbevisene er gode nok, og at man ikke har sett nok i praksis. Dette fører til at man har endt i en spiral hvor det er kamp mellom revisjonshusene (partnerne), standardsetterne, og de offentlige tilsynene om å ta det første steget.

2.5 Bevisstjøring

Innenfor revisjonsfeltet oppstår JDM når man evaluerer revisjonsbevis, ved sannsynlighetsvurderinger og når man velger mellom forskjellige alternativer (Eilifsen et al., 2014). Ved å ha en økt bevissthet rundt risiko vil man kunne redusere muligheten til å gjøre kognitive feil (KPMG, 2011), og dermed minimere sannsynligheten for å gjøre feil hvor man kan bli erstatningsansvarlig. Bevissthet er ikke et lett begrep å definere, men det kan tolkes som «vår evne til å oppleve og registrere hva som påvirker oss og hva som fortløpende hender i samspillet mellom oss og våre omgivelser (Svartdal, 2023). En modifisert definisjon av begrepet er at det er subjektivt, selektivt og blir påvirket av både omgivelser, planer, forventinger, minner, følelser og annet (Svartdal, 2023). En økt grad av bevissthet kan føre til en større oppmerksomhet rundt våre egne tankemønstre og atferd, noe som igjen kan gi oss muligheten til å reagere på en annen måte enn vår automatiske respons (LeadersharpGroup, 2019).

En velkjent psykologisk teori er prospektteorien fra Kahnemann og Tversky (Kahneman & Tversky, 1979). De utarbeidet en modell for å se hvordan folk tenker og handler, når det skal ta avgjørelser som involverer finansiell risiko. Prospektteorien vurderer menneskers mulige gevinster og tap forskjellig, og at man er i større grad påvirket av tap enn gevinst. Man vil også reagere annerledes på samme informasjon, avhengig av hvordan informasjonen blir presentert. I og med at man vektlegger tapet større enn den mulige gevinsten, vil et tap av en bestemt

størrelse påvirke mer negativt hvis man tidligere har hatt en større gevinst, enn hvis det skjer etter en mindre gevinst.

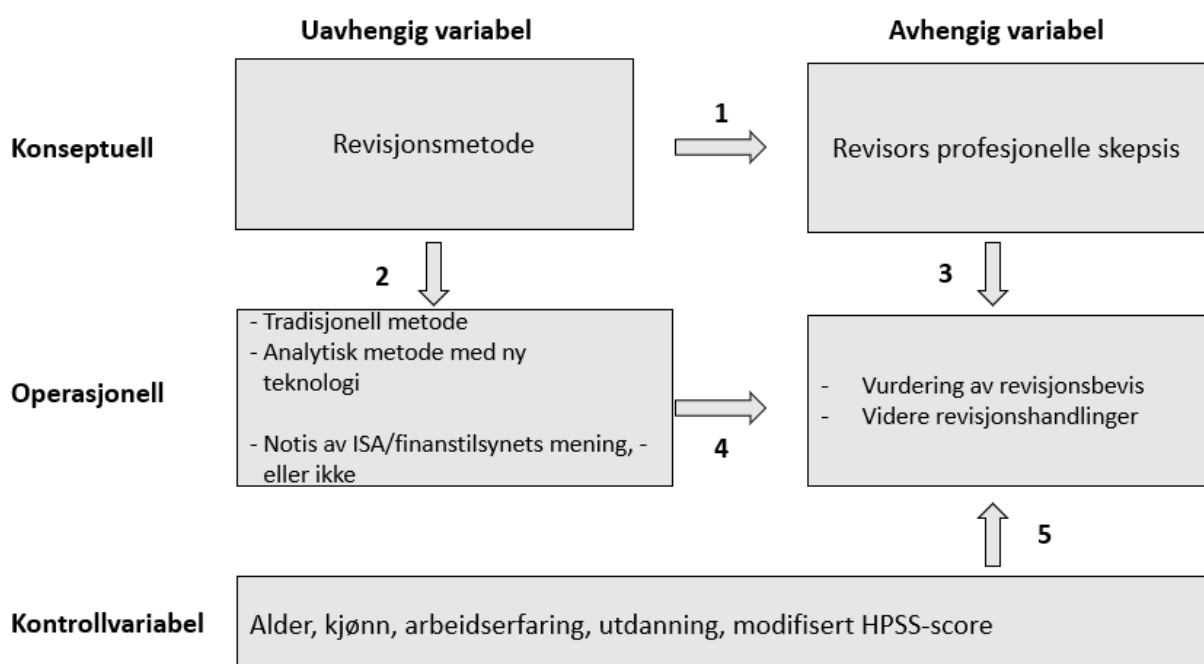
Selv om prospektteorien i stor grad handler om risiko knyttet til penger, kan man benytte den i andre situasjoner hvor man står overfor et valg. Man kan dermed benytte dette overfor bruken av den trygge tradisjonelle revisjonsmetoden eller bruke mer avansert teknologi hvor gevinsten er en mer effektiv og potensielt mer nøyaktig revisjon, hvor risikoen er at man kan bli erstatningsansvarlig. Ved å være bevisst over at det finnes en risiko for bruken av teknologien er det derfor sannsynlig at man vil foretrekke tradisjonell metode, og ha en høyere profesjonell skepsis ved bruk av analytisk metode for å innhente revisjonsbevis.

3. Konseptuell modell og hypoteser

Basert på litteraturen og tidligere forskning har vi utformet en konseptuell modell og to hypoteser, som vi i dette kapittelet vil presentere.

3.1 Forskningsdesign

For å besvare våre problemstillinger har vi utviklet en modell ved hjelp av Libby sitt valideringsrammeverk (Libby, 1981). Modellen består av uavhengige variabler, avhengig variabel og kontrollvariabler, hvor formålet er å sette lys på de viktigste faktorene for forskningsdesignets interne og eksterne validitet (Libby et al., 2002). I undersøkelsen er det de uavhengige variablene som skal manipuleres, og den avhengige variabelen fungerer som et mål på variasjonen av manipuleringen. Kontrollvariabelen er andre eksterne variabler som potensielt kan påvirke resultatet fra undersøkelsen (Trotman, 2001).



Figur 3: Konseptuell modell: Libby bokser (Libby, 1981)

3.1.1 Uavhengig Variabel

For å kunne måle de ulike variablene må de først operasjonaliseres. I undersøkelsen er det to uavhengige variabler som vi ønsker å teste. Den første uavhengige variabelen er “revisjonsmetode”. For å operasjonalisere bruken av analytisk metode i revisjonen har vi delt respondentene av undersøkelsen inn i to grupper; en gruppe som benytter drone og kunstig intelligens, og den andre gruppen som benytter tradisjonelle revisjonsmetoder av varelager. Ved å dele denne variabelen i to grupper kan man sammenligne resultatene og analysere om det er en sammenheng mellom bruk av drone og kunstig intelligens og økt profesjonell skepsis hos revisor.

Den andre uavhengige variabelen er “bevisstgjøring”. Operasjonaliseringen av denne variabelen gjør vi på samme måte ved å dele inn i to grupper, bare her at det er kun gruppene som har drone og kunstig intelligens som revisjonsmetode som blir utsatt. Dette er fordi vi ikke skal teste effekten av bevisstgjøring av risikoer ved bruk av ADA på tradisjonelle revisjonsmetoder. Operasjonaliseringen av bevisstgjøring i undersøkelsen er gjort ved å legge inn et avsnitt om standardsetternes og offentlige tilsyn sin stilling av bruk av ADA i revisjon, og risiko for revisor i caseoppgaven for den ene gruppen. Den andre gruppen vil ikke få dette avsnittet, og da ikke få en bevisstgjøring av risikoene tilknyttet ADA.

3.1.2 Avhengig variabel

Modellens avhengige variabel er revisors profesjonelle skepsis, hvor vi skal måle variasjonen av revisors profesjonelle skepsis som følge av bruk av tradisjonell metode og analytisk metode, i tillegg til variasjonen ved bevisstgjøring og ikke bevisstgjøring. Revisors profesjonelle skepsis gjøres målbar ved å operasjonalisere i to variabler: risiko for feilinformasjon og videre revisjonshandlinger. De to variablene er hentet fra (Nelson, 2009) sitt rammeverk for å måle profesjonelle skepsis. Grunnen til at vi bruker dette rammeverket er for at man her måler revisors skeptiske holdning for selve revisjonsbeviset og caseoppgaven, og ikke respondentens individuelle skeptiske holdning generelt sett.

3.1.3 Kontrollvariabel

For at eksperimentet skal være gyldig internt, så er det nødvendig å sikre at den observerte blir kun påvirket av uavhengige variabler og ikke blir påvirket av andre eksterne variabler (Trotman, 2001). For at ikke sammenhengen av uavhengige variabler og avhengig variabler skal bli påvirket av fremmed variabler, så tar vi i bruk kontrollvariabel. De kontrollvariabelen som vi mener er hensiktsmessig å ha med i eksperimenter er: alder, arbeidserfaring, utdanning og kjønn. I tillegg har vi lagt til modifisert HPSS-score etter rammeverket til (Robinson et al., 2018). HPSS-scoren måler respondentens profesjonelle skepsis gjennom undersøkelsen, og vil gi et bilde av revisors kontekstavhengige profesjonelle skepsis. Disse kontrollvariablene ble inkludert, for å kunne skille ut respondenter (variabler) som laget støy for resultatet.

3.2 Hypoteser

I dette delkapittelet vil vi presentere hypotesene som er basert på den konseptuelle modellen. En hypotese har to mulige utfall, en null hypotese og en alternativ hypotese. En alternativ hypotese representerer det man ønsker å test og finne ut av. En null hypotese representerer alle andre mulige utfall (*Trochim, William M. et al., 2014*).

3.2.1 Hypotese 1

Den første hypotesen tar utgangspunkt i valg av revisjonsmetode og grad av profesjonell skepsis. Bruken av droner og dataanalyse er noe som har blitt vurdert av flere revisjonsselskaper, og var noe som ble aktuelt gjennom Covid-19 pandemien da man ikke fikk tilgang til å inspisere varelager og andre eiendeler som man vanligvis ville gjort (Sjanifer, 2020). Det foreligger derimot en del usikkerhet om disse verktøyene vil innhente revisjonsbevis som er gode nok til at man kunne forsvare de i en potensiell erstatnings sak (Eilifsen et al., 2014; Kinserdal, 2017). Vi tar utgangspunkt i teori om at ved bruk av digitale revisjonsverktøy til å innhente revisjonsbevis, vil revisors grad av profesjonell skepsis blir høyere enn ved bruk av de tradisjonelle metodene (Nearon, 2005). Videre når risikoen blir

høyere vil revisor måtte gjennomføre flere revisjonshandlinger for å innhente revisjonsbevis som er tilstrekkelige og hensiktsmessige (Nelson, 2009).

Bruk av ADA og drone i revidering av varelager er relativt nytt, i tillegg til at standardene ikke har gitt en klar veiledning om hvordan man skal bruke det, og vi antar derfor at revisorer er mer skeptiske til at revisjonsbevisene er tilstrekkelig og hensiktsmessige kontra den tradisjonelle metodikken.

H1: Ved bruk av ADA i revisjonen av varelageret vil revisors profesjonelle skepsis bli høyere enn ved bruk av tradisjonelle revisjonsmetoder

H0: Ved bruk av ADA i revisjonen av varelageret vil revisors profesjonelle skepsis ikke bli høyere enn ved bruk av tradisjonelle revisjonsmetoder

3.2.2 Hypotese 2

Den andre hypotesen vi ønsker å teste er om det er en effekt ved å bevisstgjøre revisor for risikoene ved å bruke avansert teknologi i revisjonen, og om den profesjonelle skepsisen vil øke om man har et bevisst forhold til risikoene dette innebærer. Vi tar her utgangspunkt i teorien om at ved bevisstgjøring blir man mer oppmerksom på egne tankemønster og adferd, og at man da vil ta andre valg enn man vanligvis ville gjort (LeadersharpGroup, 2019). Dette støttes videre av prospektteorien fra Kahnemann og Tversky som tar utgangspunkt i at om det er større risiko innblandet en den potensiell gevinst, vil man velge det tryggere valget, altså den tradisjonelle revisjonsmetoden. Det kan derfor tenkes at ved bevisstgjøringen av risikoene tilknyttet de digitale verktøy, vil man ha en høyere grad av profesjonell skepsis, enn hvor man ikke fikk denne bevisstgjøringen.

H1: Ved bruk av ADA og bevisstgjøring i revisjonen av varelageret vil revisors profesjonelle skepsis bli høyere enn når man ikke har bevisstgjøring

H0: Ved bruk av ADA og bevisstgjøring i revisjonen av varelageret vil revisors profesjonelle skepsis ikke bli høyere enn når man ikke har bevisstgjøring

4. Metode

Vi vil i dette kapittelet redegjøre forskningsprosessen som er gjennomført og den metodiske tilnærmingen benyttet for å kunne svare på våre problemstillinger:

- *Vil revisors profesjonelle skepsis være annerledes ved forskjellige revisjonsmetoder?*
- *Vil bevisstgjøring av risikoene knyttet til analytiske metoder påvirke revisors profesjonelle skepsis?*

Metode er en fremgangsmåte i forskningsprosjekter for å få svar på spørsmål og innhente ny kunnskap og viten innen et felt (Larsen, 2017). Vi skiller mellom to hovedtyper av metode, kvalitative metoder og kvantitative metoder. Av undersøkelsens natur synes vi det var hensiktsmessig å gjennomføre en kvantitativ metode, herunder eksperimentell metode. Dette fordi vi ønsket å teste hypotesene med kvantitative data for å se om det var forskjell på gruppene vi testet.

4.1 Deduktiv tilnærming

For å svare på forskningsrelaterte problemstillinger kan man bruke forskjellige tilnærminger og strategier (Grønmo, 2016). Det foreligger mye teori om både revisors profesjonelle skepsis ved vurdering av revisjonsbevis og bruk av nye teknologiske verktøy i revisjonen, og har dermed benyttet en deduktiv forskningstilnærming i denne oppgaven. Basert på dette har vi kunnet gjøre våre egne forventninger og utformet både hypoteser og forskningsspørsmål av grunnlag for teorien som er funnet. Deduktiv tilnærming er en «ovenfra-og-ned»-tilnærming, hvor man går fra teori til data, altså at man har forventninger som man vil teste ut om det stemmer med virkeligheten (Trochim, William M. et al., 2014).



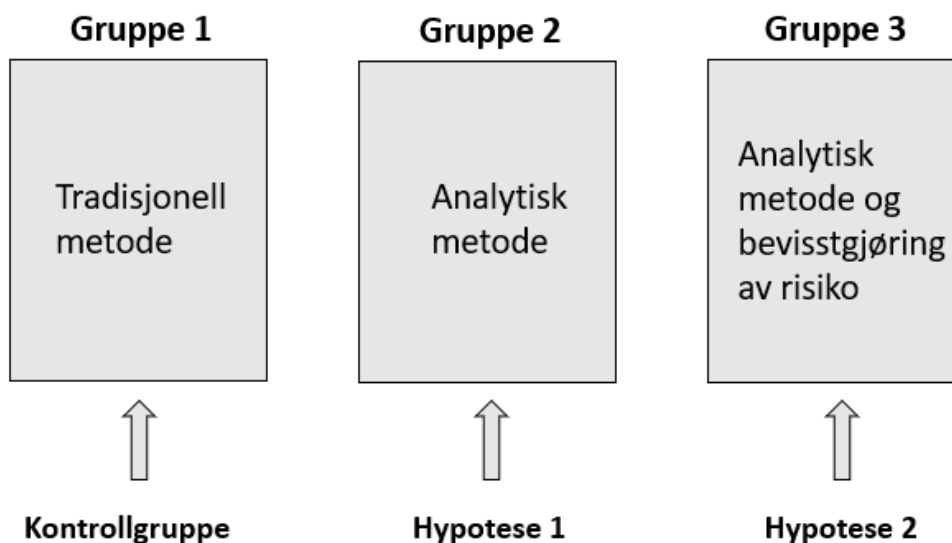
Figur 4: Deduktiv tilnærming (Trochim, William M. et al., 2014)

4.2 Forskningsdesign: Eksperimentell metode

Et forskningsdesign har som målsetting å utrette et svar på forskningsspørsmålet, samt å kontrollere for feilvarians (Trotman, 2001). Man deler gjerne inn i tre typer forskningsdesign; beskrivende, relasjonelle og eksperimentelle (Trochim, William M. et al., 2014). Da vi har en deduktiv tilnærming til vår oppgave, og har hypoteser vi ønsker å teste, passer det best at vi har et eksperimentelt design for å kunne svare problemstillingene våre på best mulig måte.

Den mest benyttede metoden for å undersøke revisorers beslutningstaking er å bruke eksperimentell metode. Når man gjennomfører en eksperimentell metode gjør man en undersøkelse hvor etterforskeren manipulerer og kontrollerer en eller flere uavhengige variabler, for å så undersøke hvordan den avhengige variabelen varierer på bakgrunn av manipulasjonen av de uavhengige variablene. Ved å benytte en eksperimentell metode kan man dermed se på effekten av revisjonsmetode og bevisstgjøring av risikoene, på revisors profesjonelle skepsis (Trotman, 2001).

For å se på om valg av revisjonsmetode og om bevisstgjøring har en effekt på revisors profesjonelle skepsis, velger vi å dele respondentene inn i tre grupper. Med dette får vi testet den uavhengige variabelen revisjonsmetode mellom gruppe 1 som fikk tildelt tradisjonell metode, og gruppe 2 som fikk analytisk revisjonsmetode. Videre får vi testet den andre uavhengige variabelen bevisstgjøring mellom gruppe 2 som ikke fikk bevisstgjøringseffekt, mot gruppe 3 som fikk dette. I og med at vi ikke skal teste den uavhengige variabelen bevisstgjøring på gruppe 1 som benytter tradisjonell metode, er det mest passende å benytte et 1x3 faktorielt design. Dette ville blitt gjort om man benyttet et mer vanlig 2x3 design, men dette er ikke hensiktsmessig ut fra teori og det vi ønsker å undersøke. Hvordan de uavhengige variablene og hypotesene er organisert ut fra gruppene kan man se i figuren under.



Figur 5: Faktorielt design

4.2.1 «Between-subjects»-design

I eksperimentet må det avgjøres om man skal gjennomføre manipulasjon «between-subjects» eller «within-subjects» av de uavhengige variablene. Det er også mulig å gjøre en kombinasjon av disse metodene (Trotman, 2001). Ved å benytte et «between-subjects»-design blir deltakerne av eksperimentet fordelt tilfeldig hvor den ene gruppen mottar manipulasjon, og den andre gruppen ikke. Den andre gruppen blir da en kontrollgruppe. Ved å gjennomføre et design «within-subjects» blir alle deltakerne i eksperimentet utsatt for manipulasjon, det vil si at alle deltakerne får alle nivåene av de uavhengige variablene. Ved å benytte en kombinasjon av metodene vil deltakerne bli tilfeldig delt inn i gruppene, og begge gruppene vil motta en grad for manipulering.

Ved å benytte et «between-subjects»-design har man mulighet til å gjøre tilfeldig hvem som skal få manipulasjon av de uavhengige variablene (Trotman, 2001). Dette betyr at man kan fordele de uavhengige variablene ut fra grupper, hvor hver gruppe får kun en manipulasjon hver (Charness et al., 2011). Man får da mulighet til å teste tradisjonell revisjonsmetodes profesjonelle skepsis mot analytisk metodes profesjonelle skepsis. Videre får man mulighet til å se om bevisstgjøringseffekten har effekt mot gruppen som ikke får dette. Det er dog både fordeler og ulemper ved å benytte et «between-subjects»-design. Sammenlignet med et

«within-subjects»-design hvor man kan teste flere uavhengige variabler på hver gruppe, trenger man ofte flere respondenter i et «between-subjects»-design, ettersom hver gruppe kun får en avhengig variabel hver og den statistiske kraften ofte vil være dårligere (Charness et al., 2011). Ulempen er at deltakere ofte vil fatte mistanke om baktanken av studien ved et «within-subjects»-design (Trotman, 2001). Dette kunne vært tilfelle om respondentene i studien først skulle gjennomføre tradisjonell metode, for å så gjøre en analytisk metode. I tillegg ville det da vært vanskeligere å måle effekten av bevissthet, noe som kunne dog vært løst med å benytte en slags kombinasjon av metodene.

Vi beslutter derfor å benytte et «between-subjects»-design på bakgrunn av at vi tror at vi kan anskaffe nok respondenter og at det blir mest hensiktsmessig å tilføre en uavhengig variabel på hver gruppe i undersøkelsen. Videre tror vi også det vil være mulig å gjennomskue hensikten med studiet om man skulle gjort begge de tradisjonelle metodene, noe som kunne ført til at målingen av profesjonell skepsis hadde blitt partisk mot en av metodene.

4.3 Datainnsamling

4.3.1 Pilottest

Før vi publiserte spørreundersøkelsen, så gjennomførte vi en pilottest på 4 revisjonsstudenter. Hensikten med en pilottest er å avdekke om det er noen uklarheter med caset eller spørsmålene. Samtidig gav det oss en indikator på hvor lang tid respondentene ville bruke på spørreundersøkelsen. Tilbakemeldingene fra pilottesten gjorde at vi fikk justert noen uklarheter med caset, endret på noen spørsmål, og fikk oppklart tidsbruken til respondentene.

4.3.2 Utvalg

Populasjonen er antall enheter som brukes til å besvare forskningsrelaterte problemstillinger. Populasjonen vi hovedsakelig ønsker å benytte er revisorer, men ettersom dette er en masterutredning, hadde vi ikke tilgang eller ressurser til å kontakte en stor del av populasjonen, og ble derfor nødt til å bruke studenter som en erstatning for revisorer. Ved å trekke ut det

letteste utvalget man får tak i, gjør man et bekvemmelighetsutvalg (Jacobsen, Dag Ingvar, 2018). Bruk av studenter som en erstatning for revisorer har blitt forsket på tidligere. Forskning viser at det er korrelasjon mellom praktiserende revisorer- og studenters meninger og holdninger (Abdolmohammadi & Wright, 1987). Ifølge nyere forskning som har blitt gjort, så finnes det ingen signifikante forskjeller mellom revisorer og revisjonsstudenter (Mortensen et al., 2012).

Utvalget vårt består av masterstudenter som tar master i regnskap og revisjon. Et minstekrav til deltakerne er at de hadde tatt faget ACC402/MRR411, eller lignede fag. Dette for å kunne sikre at respondentene hadde tilstrekkelig med grunnleggende kunnskap om revisjon for å delta i undersøkelsen.

For å klare å nå ut til utvalget så delte vi spørreundersøkelsen på facebook-gruppene *Linjeforeningen MRR NHH* og *Platform for MRR Students*, som til sammen har 180 medlemmer. I tillegg delte vi undersøkelsen ut på mail til førsteårs MRR-studenter, hvor vi nådde ut til ytterligere 137 studenter. Etter en uke sendte vi ut en påminnelse på mail for å fange opp studentene som ikke såg den første mailen.

4.4 Praktisk gjennomføring av case

4.4.1 Utforming av case

Vi ønsket å utforme en case og spørsmål på en måte slik at det skulle være enkelt og forståelig for respondentene, i tillegg til at det ikke skulle ta for lang tid å gjennomføre undersøkelsen. Samtidig var det viktig for oss at respondentene skulle få nok informasjon til at caset virket realistisk, og at man fikk en følelse av at man måtte handle slik man ville gjort i en ekte situasjon. Caset gikk ut på at man først fikk informasjon om selskapet Techzone, en revisjonskunde hvor varelageret skulle revideres. Varelageret til det fiktive elektronikk-selskapet Techzone var en stor del av selskapets balanse, og dermed vesentlig for revisjonen. Deltakerne fikk her opplyst om noen nøkkeltall fra selskapets regnskap, blant annet varelagerstørrelse og omløpshastighet, for at man skulle få en følelse av størrelsen. Vi ønsket

at respondentene skulle ha følelsen av at varelageret var stort slik at det var mulig å benytte drone og kunstig intelligens som verktøy.

Videre fikk respondentene tilfeldig utdelt revisjonsmetode, dette operasjonalisert ved å få forskjellige arbeidspapirer. Ved hjelp av spørreundersøkelse programmet Qualtrics fikk man plassert respondentene inn i «blokker», noe som gjorde at man fikk tildelt revisjonsmetode tilfeldig. Dette gjorde at noen fikk den tradisjonelle revisjonsmetode med stikkprøver og inspeksjon, mens noen fikk analytisk metode med drone og AI. Respondentene som fikk tradisjonell revisjonsmetode, hadde undersøkelsen basert på utvalgtesting. De fikk en visuell oversikt over et stikkprøveutvalg på 10% av varelageret, som sammenlignet kostprisen på varene og prisen oppgitt i faktura. De fikk også oversikt over antall opptelte varer under revisjonen mot det som lå på varelagerlista, noe som gjøres i praksis med telleliste. Dette ble gjort for å teste påstandene eksistens og verdsettelse. Noe avvik ble lagt til i listen for å gjøre respondenten mer kritisk.

Respondentene som skulle gjennomføre revisjonen av varelageret ved bruk av ADA, fikk først informasjon om hvordan verktøyet «DroneAudit.AI» (DAA), fungerer. Dette er et fiktivt analyseprogram som innhenter informasjon ved bruk av drone, som lastes opp i en programvare for å gjøre en 100% analyse. DAA ble visualisert og gir respondenten oversikt over avvik og anomalier i en graf, i tillegg til «flagg» som revisor skal sette oppmerksomhet til. Videre ble det også lagt inn et bilde som forsterker inntrykket at drone ble tatt i bruk, inspirert fra eksperimentet til (Christ et al., 2021). Den siste gruppen hvor vi skulle teste effekten av bevisstgjøring fikk samme case som de som gjennomførte på analytisk metode, men at de fikk inn et avsnitt i casen om dagens status i revisjonsstandardene om bruk av avanserte digitale verktøy, finanstilsynets stilling, og risikoen for revisor ved å ta i bruk dette. Vi ønsket ikke at dette skulle ta for mye plass i casen, men nok til at respondentene tenkte over problemsstillingen.

4.4.2 Vurderingsspmå

Den andre delen av spørreundersøkelsen bestod av spørsmål med pre-definerte svaralternativ, også kalt lukkede spørsmål (Larsen, 2017). For å måle deltakernes profesjonelle skepsis

tilknyttet revisjonsbevisene tok vi utgangspunkt i rammeverket til (Nolder & Kadous, 2018) og (Nelson, 2009), og utarbeidet 5 hovedspørsmål. Disse spørsmålene tok utgangspunkt i om respondentene ønsket å innhente flere revisjonsbevis og respondentenes mening om risiko for feilinformasjon. Det er flere metoder å måle profesjonell skepsis på, men vi mente disse 5 spørsmålene var hensiktsmessige nok for å måle profesjonell skepsis ut fra vår case. Videre utformet vi to spørsmål som omhandler feilinformasjon basert på eksistens og verdi i varelageret, ettersom dette er to sentrale påstander i revisjon av et varelager, og mulig den største utfordringen ved bruk av drone i revisjonen. Vi inkluderte i tillegg to ekstra spørsmål til respondentene som fikk ADA, og de som fikk ADA samt bevisstgjøring. Dette fordi det var interessant å se deres mening om verktøyene, og i hvor stor grad de la vekt på ISA/Finanstilsynets meninger i vurderingen.

Til slutt innhentet vi demografiske informasjon, som omhandler alder, kjønn, tidligere kurs og erfaring, i tillegg til spørsmål tilknyttet modifisert HPSS. Modifisert HPSS er hentet fra (Robinson et al., 2018) sitt rammeverk på 12 spørsmål, som måler profesjonell skepsis ut fra deltakernes kontekstavhengige skepsis i tilknytning til selve undersøkelsen, altså som en tilstand.

Vi benyttet Likert-skala¹⁰ (1-7) i utforming av alle spørsmålene i undersøkelsen, utenom de demografiske, hvor det ble benyttet nominal måling.¹¹ Ettersom modifisert HPSS-er utformet etter Likert-skala mente vi det var hensiktsmessig for å benytte dette i resten av undersøkelsen for sammenligningsgrunnlag. Skalaen på syv nivåer er også nyttig ettersom det påtvinger respondentene å gi en nyansert vurdering. Dette fordi profesjonell skepsis ikke har et fasitsvar, men det er mange ulike grader og faktorer som spiller inn. Se fullstendig case og spørreundersøkelse i vedlegg 1.

¹⁰ Likert skala benyttes i spørreundersøkelser for at respondenten skal ta stilling til et bestemt fenomen. Vanligvis er den fra 1 til 5. For eksempel (1) svært uenig og (5) svært enig.

¹¹ Nominal måling har respondentene like eller ulike verdier, for eksempel kjønn med verdiene mann og kvinne.

4.5 Reliabilitet og validitet

De to kriteriene for å måle kvalitet på en undersøkelse er reliabilitet og validitet. (Trochim, William M. et al., 2014). Reliabilitet refererer til konsistensen av målingene, som vil si at undersøkelsen gir de samme resultatene under identiske forhold. Reliabiliteten av undersøkelsen trues av målefeil, tilfeldige- og systematiske feil. (Trochim, William M. et al., 2014).

Før undersøkelsen ble sendt ut til utvalget vårt, gjennomførte vi en pilottest blant studenter som også gjennomfører eksperimentelle metode for å styrke den eksterne reliabiliteten til vår undersøkelse. En pilottest kan avdekke uklarheter eller feil både i selve caseundersøkelsen og i spørsmålene etter selve casen. I vårt tilfelle var det noen uklarheter når det gjaldt informasjon i casen, og noen av spørsmålene som ble stilt, noe som vi rettet opp i ettertid. Systematiske feil i resultatet kan komme av at det ikke er mange som ikke gjennomfører undersøkelsen, noe som kan føre til at det oppstår skjevfordeling i svarene (Trochim, William M. et al., 2014). For å unngå dette ønsket vi å gjøre casen så enkel og kort som vi kunne, uten at informasjon ble holdt utenfor. En annen type feil som kan påvirke reliabiliteten til undersøkelsen er tilfeldige feil. Respondentens holdning under undersøkelsen er noe som kan skape tilfeldige feil (Trochim, William M. et al., 2014). Dette er noe som ikke påvirker hele populasjonen, men kan skape «støy» i dataen. Denne typen feil har en tendens til å både være positive og negative for resultatet i undersøkelsen, og vil gjerne «jevne seg ut», og vil derfor ikke ha en stor betydning for resultatet.

Vi hentet dataen ut fra spørreundersøkelsesskjemaet Qualtrics, som ble direkte overført til dataanalyse-programmet SPSS. Dette styrker reliabiliteten til dataen fordi det reduserer risikoen for menneskelige feil. Etter dette luket vi ut ugyldige og ufullstendige svar, og svar som kan kategoriseres som useriøse. All ekstern reliabilitet kan ikke dekkes. Selv om undersøkelsen var anonym og man hadde den tiden man trengte til å svare på undersøkelsen, ble eksperimentet gjennomført i et tidsrom hvor både medstudenter og praktiserende revisorer hadde hektiske perioder med henholdsvis egen masterskriving/eksamensforberedelser og ferdigstilling av årsoppgjøret. Dette kan ha påvirket hvor lang tid man prioriterte på å gjennomføre undersøkelsen.

Validitet er den andre komponenten for å sikre kvalitet på undersøkelsen. Validitet står for gyldighet og relevans, noe dataen som innhentes må være for temaet som undersøkes (Trochim, William M. et al., 2014). Videre deles validitet opp i intern validitet og ekstern validitet. Studiens interne validitet vil si at man kan tilordne endringen i den avhengige variabelen til den avhengige variabelen (Peecher & Solomon, 2001). Vi har sikret den interne validiteten ved å vise til et kausalitetsforhold mellom variablene vi undersøker i eksperimentet, dette vist i den konseptuelle modellen.

Videre er den eksterne validiteten risikoen for at de sammenhengene man finner i eksperimentet er generalisert på bakgrunn av populasjonen man tar undersøkt (Trochim, William M. et al., 2014). Den eksterne validiteten sikres ved at man får et representativt utvalg, noe som avhenger av at alle i populasjonen har lik sannsynlighet for å bli valgt. Ved bruk av programmet Qualtrics ble utvalget delt tilfeldig inn i hvilken av revisjonsmetodene de skulle gjennomføre. Videre begrenset vi utvalget til at kun studenter som har tatt / tar mastergrad innen regnskap og revisjon, eller har tatt revisjonsfag, med i studien, slik at man har nødvendig kunnskap til å gjøre vurderingene. Vi har tidligere nevnt at vi i undersøkelsen tok sikte på å bruke studenter fra Norges Handelshøyskole som utvalg, noe som neppe kan sies å representere hele populasjonen av revisorer, men i forhold til ressurser og tid ville det vært vanskelig å anskaffe et utvalg som gjør nettopp dette.

4.6 Analysemetode

4.6.1 Dataanalyse

For å analysere datasettet vårt, benyttet vi analyseprogrammet IBM SPSS. Undersøkelsen ble gjennomført i programvaren Qualtrics, hvor dataen ble overført direkte fra Qualtrics til IBM SPSS for å eliminere risikoen for menneskelige feil.

Når man gjennomfører statistiske analyser, må man vurdere sannsynligheten for om resultatene er representative eller tilfeldige. Dette kan gjøres ved å sammenligne sannsynligheten med et forhåndsdefinert signifikansnivå for forskningen. Innen samfunnsvitenskapelige eksperimenter er det vanlig å benytte en kritisk verdi på 5 % for å

vurdere om resultatene er signifikante eller ikke (Trochim, William M. et al., 2014). Hvis eksperimentets resultater har en p-verdi som overstiger det forhåndsdefinerte signifikansnivået, vil resultatene ikke være statistisk signifikante, og nullhypotesen kan ikke forkastes. Dette innebærer at det er en 5 % sannsynlighet for å begå en type I-feil, det vil si å feilaktig forkaste nullhypotesen.

I eksperimentet vårt tester vi ikke mer en to populasjoner for hver hypotese, noe som gjør det hensiktsmessig å benytte en uavhengig t-test. Testen sammenligner gjennomsnittet mellom to ikke-relaterte grupper for den samme kontinuerlige avhengige variabelen (Laerd Statistics, u.å.-a). Denne metoden vil gjøre det mulig å teste om det er statistisk signifikante forskjeller i gjennomsnittet til gruppene (Trochim, William M. et al., 2014).

Den avhengige variabelen profesjonell skepsis blir operasjonalisert ved hjelp av 5 spørsmål tilknyttet revisjonsbevisene hvor man må benytte profesjonell skepsis. Vi måler signifikansen på spørsmålene både separat, og totalt i en indeks. Når man indekserer resultatene fra spørsmålene er det viktig at spørsmålene er konsistente og måler de samme tingene. For å sikre at de gjør dette bruker vi Crohnbachs Alpha-test. Denne strekker seg fra 0-1, hvor ved en høyere alfa, desto større er sannsynligheten for at spørsmålene vi stiller måler det samme underliggende konseptet. For å være sikker ønsker man generelt en alfa som er større en 0,7 (StatisticsHowTo, u.å.).

4.6.2 Forutsetninger for analysen

Når man benytter en T-test er det seks forutsetninger som må være tilfredsstilt for at testen skal kunne gjennomføres (Laerd Statistics, u.å.-a). Forutsetningene som ligger til grunn er (1) den avhengige variabelen måles kontinuerlig, (2) de uavhengige variablene skal bestå av to kategoriske variabler, (3) det må være uavhengige observasjoner, (4) det skal ikke være noen signifikante uteliggere¹², (5) den avhengige variabelen må være tilnærmet normalfordelt og (6), det må være homogenitet i variansen.

¹² Uteliggere er observasjoner som avviker signifikant fra resten av utvalget

Ved å gjennomføre de metodiske valgene som er gjort er de tre første forutsetningene allerede til stede. Forutsetning nummer fire kontrolleres gjennom et boksplokk av de observerte verdiene hvor mulige uteliggere blir fjernet. Videre testes forutsetning fem om normalfordeling av avhengig variabel ved hjelp av en Shapiro-Wilk test. Om det ikke er normalfordeling tilstede vil vi på de aktuelle hypotesene gjennomføre en Mann-Whitney U test i stedet, som tar hensyn til at dataen ikke er normalfordelt (Laerd Statistics, u.å.-b). Denne testen har de samme forutsetningene som en avhengig T-test, men at den ikke forutsetter normalfordeling. Hvilken av disse to metodene som blir gjennomført blir redegjort i delkapittelet forutsetninger før test av hypoteser. Ved den sjette og siste forutsetningen om homogenitet i variansen benyttes Levene's test.

4.7 Etikk

Spørreundersøkelsen ble gjennomført ved bruk av Qualtrics. Ved å gjennomføre spørreundersøkelsen i Qualtrics ble alle funksjoner som kan identifisere respondentene deaktivert. Det utløste dermed ingen meldeplikt til Sikt¹³, ettersom alle respondentene var anonyme.

¹³ Kunnskapssektorens tjenesteleverandør

5. Resultat

I dette kapittelet vil vi presentere resultatene fra eksperimentet. Vi tar for først for oss hvordan dataen ble rensset før vi gjennomførte analysen. Videre tar vi for oss deskriptiv statistikk som gir en oversikt over datagrunnlaget og utvalget. Vi vil så gå videre på testing av hypotesene og tolkning av resultat. Til slutt tar vi for oss andre funn fra eksperimentet.

5.1 Rensing av data

Totalt var det 91 respondenter som åpnet spørreundersøkelsen vår. Av disse var det 74 personer som fullførte undersøkelsen i sin helhet, noe som gjorde at vi ekskluderte 17 fra datagrunnlaget vårt. Videre ønsket vi at utvalget vårt skulle bestå av nåværende eller tidligere revisjonsstudenter, noe som gjorde at vi ekskluderte respondenter som ikke tar, eller har tatt, MRR eller relevante revisjonsfag, som var totalt 2 stykker. I tillegg ekskluderte vi 2 respondenter som brukte såpass lang tid på å gjennomføre at vi valgte å fjerne de. Dette gjorde at vi totalt endte opp med 70 respondenter i vårt datagrunnlag. Svarprosenten på undersøkelsen resulterte i 77%, noe som blir kategorisert som en god svarprosent på spørreundersøkelser (Jacobsen, Dag Ingvar, 2018).

5.2 Deskriptiv statistikk

Den deskriptive statistikken forklarer dataens grunnleggende egenskaper, og benyttes for gi et nyttig overblikk over utvalget og potensielle mønster (Trochim, William M. et al., 2014). Etter at vi rensset datagrunnlaget satt vi igjen med totalt 70 respondenter. Ved hjelp av programmet Qualtrics ble respondentene tilfeldig delt inn i hvilken metode de skulle gjennomføre, og dermed i hvilken gruppe de havnet i. Dette gjorde at vi fikk en jevn fordeling av respondenter i de forskjellige gruppene, noe som øker sannsynligheten for at endringen i den avhengige variabelen kommer fra manipulasjonen i de uavhengige variablene (Trochim, William M. et al., 2014).

	Antall	Prosent
Analytisk metode med bevisstgjøring	22	31.4
Analytisk metode	24	34.3
Tradisjonell metode	24	34.3
Totalt	70	100.0

Tabell 1: Fordeling av gruppene

5.2.1 Demografiske variabler

Tabellen viser at på kjønn så er det relativt skjevfordelt. Det totale utvalget består av 64,28 prosent menn mot 34,28 prosent kvinner. Gruppene tradisjonell metode og analytisk metode er relativt jevne, mens den store skjevheten er i gruppen analytisk metode ved bevisstgjøring hvor menn utgjør 75 prosent av utvalget. Fordelingen mellom kvinner og menn ved MRR på NHH er nesten helt jevnt (*nhh-arsrapport-2021.pdf*, u.å.), noe som tilsier at fordelingen burde vært mer lik. En sannsynlig forklaring er at nettverket vårt av revisjonsstudenter på høyskolen består av en større andel menn.

		Tradisjonell metode	Analytisk metode	Analytisk metode med bevisstgjøring	Total
Kjønn	Mann	15	13	17	45
	Dame	9	11	4	24
	Annet	0	0	1	1
Totalt		24	24	22	70

Tabell 2: Fordeling av kjønn i gruppene

Aldersfordeling blant det totale utvalget viser at 54,42 prosent er mellom 26-30 år, mens 40 prosent er mellom 20-25 og resterende 5,71 prosent er over 30 år. Dette forklares med at utvalget vi ønsket å teste var studenter, og de fleste studentene er i aldersgruppen 20-25 og 26-30 år. Aldersgruppen 26-30 år var den største gruppen. I gruppene tradisjonell metode og analytisk metode med bevisstgjøring så kan vi se at det er flest respondenter i aldersgruppen 26-30, mens i gruppen analytisk metode er det flest i aldersgruppen 20-25 år.

		Tradisjonell metode	Analytisk metode	Analytisk metode med bevisstgjøring	Total
Alder	20 - 25	6	14	8	28
	26 - 30	16	10	12	38
	Over 30	2	0	2	4
Totalt		24	24	22	70

Tabell 3: Aldersfordeling i gruppene

Fordelingen mellom arbeidserfaring av det totale utvalget viste at 64,28 prosent hadde ingen arbeidserfaring, 22,85 prosent hadde under 2 års arbeidserfaring, 11,42 prosent hadde arbeidserfaring mellom 2-5 år og kun 1,42 prosent hadde mer enn 5 års erfaring. Fordelingen viser en trend hvor antall respondenter minsker når arbeidserfaringen øker. Dette skyldes nok at vi sendte ut spørreundersøkelsen til studenter, og de har som oftest ikke mye arbeidserfaring. Gruppene analytisk metode og analytisk metode med bevisstgjøring er ganske like med omtrent 70 prosent respondenter som ikke har arbeidserfaring. I gruppen tradisjonell metode innehar 50 prosent av respondenten ingen arbeidserfaring.

		Tradisjonell metode	Analytisk metode	Analytisk metode med bevisstgjøring	Total
Arbeidserfaring innen revisjon	Ingen arbeidserfaring	12	17	16	45
	Under 2 år	8	5	3	16
	2 - 5 år	4	2	2	8
	Mer enn 5 år	0	0	1	1
Totalt		24	24	22	70

Tabell 4: Arbeidserfaring i gruppene

5.3 Hypotesetesting

5.3.1 Forutsetninger

Før vi analyserer resultatene, kontrollerer vi først om forutsetningene i delkapittel 4.6.2 er oppfylt. Boksplottet¹⁴ viser et 5% trimmet gjennomsnitt som indikerer at det ikke foreligger noen signifikante uteliggere, noe som gjør at vi ikke fjerner noe fra datasettet vårt.

For å kunne teste hypotesene med en T-test så er en forutsetning at den avhengige variabelen i hver gruppe er normalfordelte. For å teste dette så utførte vi en Shapiro-Wilk-Test¹⁵. Testen viste at gruppene analytisk metode og analytisk metode med bevisstgjøring oppfyller kravene for normalfordeling. Gruppen tradisjonell metode har en signifikantverdi på 0,003, og dermed avviker dataen i denne gruppen signifikant fra en normalfordeling. Vi vil dermed ved test av hypotese 1 benytte *Mann-Whitney U-test* som redegjort i *delkapittel 4.6.2*, hensyntar at dataen ikke er normalfordelt.

Vi gjennomførte Crohnbachs-Alfa¹⁶ test for å vurdere reliabiliteten til profesjonell skepsis indeksen vi har konstruert. Denne viser et resultat på 0,901, og er dermed vurdert som pålitelig nok til å bruke i analysen.

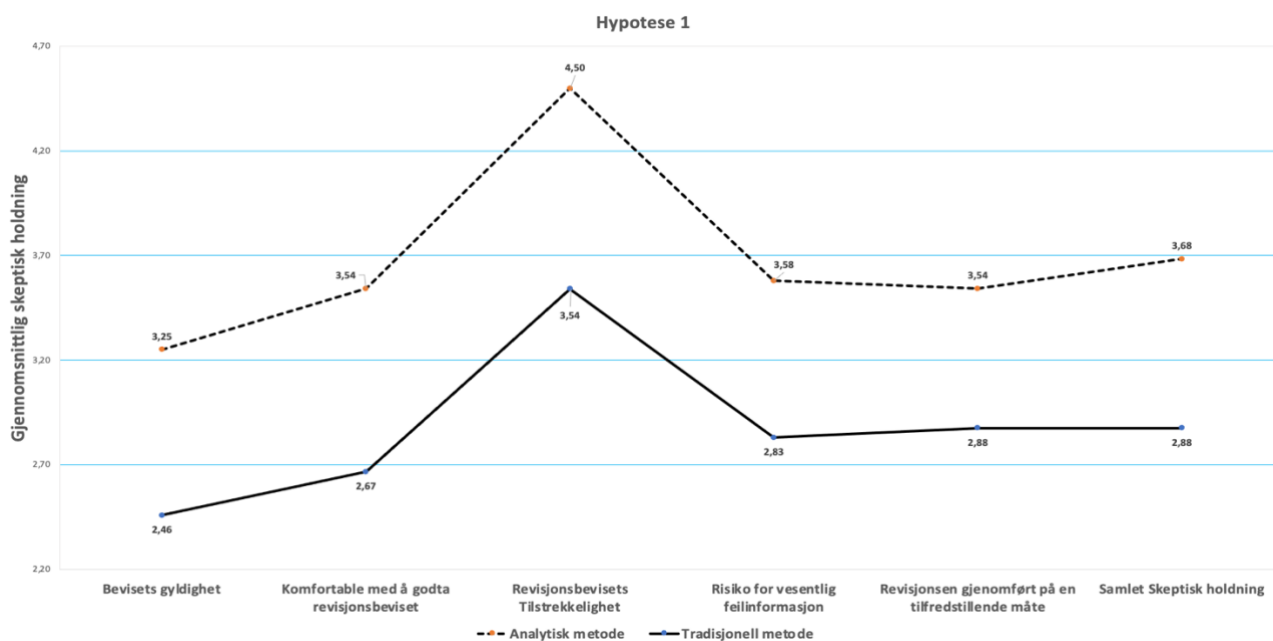
5.3.2 Resultater hypotese 1:

Formålet med hypotese 1 er å undersøke om bruk av analytisk metode i revisjon av varelager vil føre til høyere nivåer av profesjonell skepsis sammenlignet med bruk av tradisjonelle metoder. For å gjøre dette vil vi sammenligne svarene fra gruppe 1 som bruker tradisjonelle metoder, og gruppe 2 som bruker analytisk metode, med utgangspunkt i de fem påstandene om profesjonell skepsis.

¹⁴ Se vedlegg 2 for fullstendig tabell

¹⁵ Se tabell for Shapiro-Wilk test i vedlegg 4

¹⁶ Se tabell for Crohnbachs-Alfa test i vedlegg 3



Figur 6: Hypotese 1 - Skeptisk holdning

Figur 6 viser at gruppe 2 som benyttet analytisk metode har en høyere skeptisk holdning enn gruppe 1 som benyttet tradisjonell metode i revisjonen. Dette gjelder både påstandene individuelt sett, og samlet i profesjonell skepsis-indeksen. Gjennomsnittet viser at respondentene har en lavere profesjonell skepsis ved bruk av de tradisjonelle verktøyene, kontra bruk av ADA, noe som var forventet med teorien. Vi gjennomførte så en *Mann-Whitney U-test*¹⁷ for å kartlegge om observasjonen var signifikant.

Grupper	Gjennomsnitt	Standardavvik	Mann-Whitney U Test
Gruppe 1: Tradisjonell metode	2,875	1,521	0,139
Gruppe 2: Analytisk metode	3,683	0,900	

Tabell 5: Hypotese 1 (Mann-Whitney-U-Test)

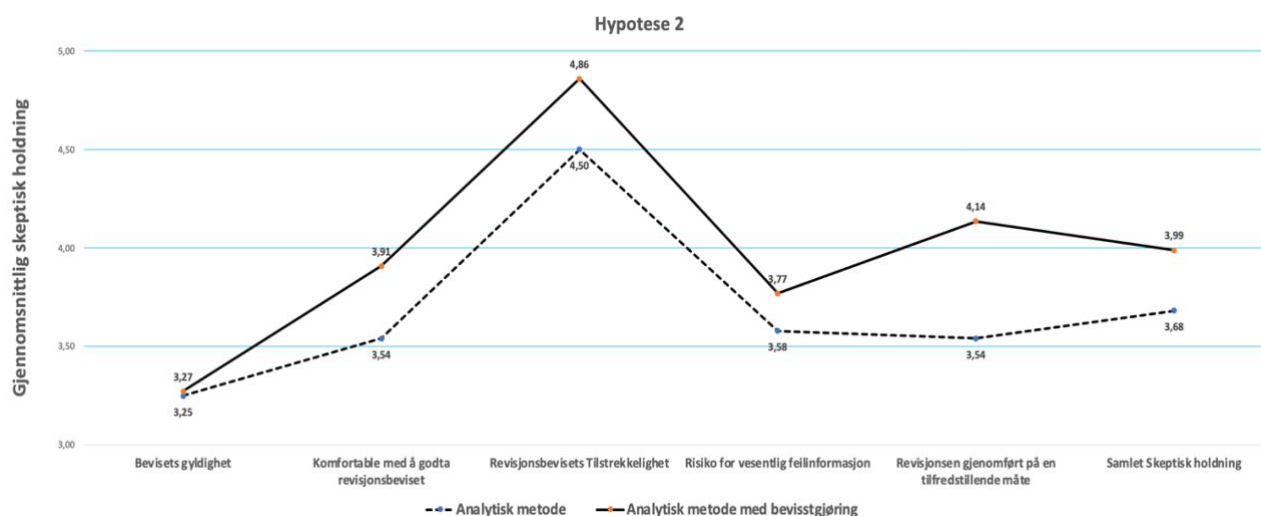
¹⁷ Se vedlegg 5 for fullstendig tabell fra SPSS

Testen fikk en p-verdi på $0,139$, noe som er høyere enn signifikansnivået på $0,05$. Dette betyr at resultatet ikke er signifikant, og vi kan dermed ikke forkaste null-hypotesen.

H_0 beholdes.

5.3.3 Resultater hypotese 2:

Hypotese 2 har som mål å undersøke om økt bevissthet om risikoene ved bruk av analytiske metoder vil påvirke vurderingene til respondentene i eksperimentet. For å gjøre dette sammenligner vi svarene fra gruppe 2 som bare bruker analytisk metode, med svarene fra gruppe 3 som også ble utsatt for bevisstgjøring om risikoen. Vi vurderer svarene mot de fem påstandene om profesjonell skepsis.



Figur 7: Hypotese 2 – Skeptisk holdning

Figur 7 viser at gruppen analytisk metode med bevisstgjøring av risiko har høyere skeptisk holdning enn gruppen analytisk metode. Gjennomsnittet for den samlede profesjonelle skepsisen viser at gruppen analytisk metode med bevisstgjøring er mer skeptisk enn gruppen analytisk metode.

	Bevisets gyldighet	Komfortabel godta revisjonsbeviset	Revisjonsbevisets tilstrekkelighet	Risiko ikke inneholder feilinformasjon	Revisjon gjennomført tilfredsstillende måte	Samlet skeptisk holdning
Gruppe 2: analytisk						
Gjennomsnitt	3,250	3,542	4,500	3,580	3,542	3,683
Standardavvik	0,897	1,030	1,532	1,349	0,977	0,900
Gruppe 3: Analytisk metode m/bevisstjøring						
Gjennomsnitt	3,273	3,909	4,860	3,770	4,136	3,991
Standardavvik	1,077	1,192	1,552	1,541	1,207	1,011
T-Test						
Levene's Test						0,643
Sig. (One sided)						0,141
Sig. (Two sided)						0,281

Tabell 6: Hypotese 2 (T-Test)

Tabell 6 gir en detaljert oversikt over de forskjellige påstandene. Den samlede skeptiske holdningen i gruppe 2 er 3,68, mens i gruppe 3 er den 3,99. Det er derfor ingen store avstander i de forskjellige gruppene skeptiske holdninger. Det er forskjellene i påstand 5 som utgjør den største forskjellen mellom gruppene.

Videre så gjennomførte vi Levene's test¹⁸, for å teste variansens likhet for gruppene. Testen ga en p-verdi som var høyere enn 0,05, som betyr at testen ikke er signifikant, og at man ikke kan forkaste null-hypotesen om varians-ulikhet. Dette betyr at populasjonsavviket er like og variansen er homogen.

Vi utførte en T-Test¹⁹ for uavhengige grupper og benyttet en ensidig test. P-verdien er på 0,141, som betyr at resultatet ikke er signifikant, og vi kan dermed ikke forkaste null-hypotesen.

H₀ beholdes.

¹⁸ Se vedlegg 6, Hypotese 2 for fullstendig tabeller fra SPSS

¹⁹ Se vedlegg 6 for fullstendige tabeller fra SPSS

5.4 Andre funn

5.4.1 Profesjonell skepsis som en tilstand

I eksperimentet benyttet vi også Robinson-skala for å måle profesjonell skepsis som en tilstand hos respondentene. Dette for å måle respondentenes generelle skepsis under caset, og ikke bare skepsisen tilknyttet bevisene. Vi ønsket derfor å sammenligne rammeverket til Robinson mot Nelson for å se om respondentenes generelle skeptiske holdning har effekt på vurderingen av revisjonsbevisene. Vi antar videre at skalaen kan kvantifiseres på samme måte som i (Hurt, 2010), med verdier fra 1 til 7. Dermed vil skalaen variere fra 10 til 70, etter vi har multiplisert den. Hvis en deltaker scorer høyt på skalaen, indikerer det en større grad av skepsis sammenlignet med en deltaker som scorer lavt. Robinson-skalaen består av 12 spørsmål, som vi samler i en indeks. Indeksen er testet ved Crohnbachs-alfa²⁰ som gav en verdi på 0,922.

Robinson score	Antall	Gjennomsnitt	Standardavvik	Min.	Maks.
Gruppe 1					
Tradisjonell metode	24	51,670	11,839	26,670	69,170
Gruppe 2					
Analytisk metode	24	48,093	10,443	22,500	70,000
Gruppe 3					
Analytisk m/bevisstgjøring	22	49,886	12,728	27,500	66,670
Totalt	70	49,883	11,670	25,557	68,613

Tabell 7: Deskriptiv statistikk for Robinson score

Robinson score gav oss et totalt gjennomsnitt på 49,883. Vi brukte deretter det totale gjennomsnittet til å dele gruppene inn i to, enten over gjennomsnittet eller under gjennomsnittet. Gruppen med score under gjennomsnitt, blir sammenlignet med deres gjennomsnitt svar i de 5 påstandene om profesjonell skepsis i rammeverket til Nelson. Samme

²⁰ Se vedlegg 7 «Andre funn» for fullstendig tabell fra SPSS

gjøres også ved de i gruppen med høy score. Dette gjør vi for å kunne sammenligne om det er en sammenheng mellom høy/lav skeptisk tilstand og gjennomsnittlig skeptisk holdning.

Skeptisk tilstand		Antall	Skeptisk holdning gjennomsnitt	Standardavvik	Min.	Maks.
Lav score	Tradisjonell metode	12	3,18	0,359	1	4
	Analytisk metode	17	3,45	0,221	2	6
	Analytisk metode m/bevisstgjøring	12	3,57	0,296	2	5
	Totalt	41	3,40	0,292	2	5
Høy score	Tradisjonell metode	7	2,57	0,507	1	5
	Analytisk metode	10	4,26	0,221	3	5
	Analytisk metode m/bevisstgjøring	10	4,50	0,239	3	6
	Totalt	27	3,77	0,322	2	5

Tabell 8: Sammenheng mellom skeptisk tilstand og skeptisk holdning

Tabellen viser at gjennomsnittet for de som scorer høyt i Robinson skala, også har en høyere gjennomsnittlig skeptisk holdning, enn de som scorer lavt i Robinson skala. I tråd med utfallet av resultatene i hypotesen så viser tabellen at tradisjonell metode med høy og lav score i Robinson skala, at skeptisk holdning er lavere enn analytisk metode. Det samme gjelder analytisk metode med bevisstgjøring, som har høyere skeptisk holdning enn analytisk metode, uavhengig om det er høy eller lav score Robinson skala.

Videre så viser tabellen at de i gruppen tradisjonell metode som scorer høyt i Robinson skala, har en lavere gjennomsnittlig skeptisk holdning, enn respondentene i gruppen tradisjonell metode som scorer lavt i Robinson skala.

5.4.2 Positiv til bruk av drone i revisjon og vektleggelse av bevisstgjøring

Etter at respondentene hadde vurdert revisjonsbevisene, ble det stilt tilleggsspørsmål til respondentene i både gruppe 2 og 3. Disse spørsmålene handlet om graden av positivitet til bruk av drone og kunstig intelligens i revisjonssammenheng, samt hvor mye de vektlagt

kommentaren fra Finanstilsynet. Det var imidlertid kun respondentene i gruppe 3 som ble spurt om det siste spørsmålet.

Gruppe	Spørsmål	Antall	Gjennomsnitt	Standardavvik	Min.	Max.
Analytisk metode	Positiv til bruk av drone i revisjonen	24	5,417	0,240	2	7
Analytisk metode m/bevisstgjøring	Positiv til bruk av drone i revisjonen	22	5,364	0,305	1	7
	Grad vektleggelse FT	22	5,000	0,302	3	7

Tabell 9: Sammenligning positivitet og grad av vektlegging av kommentar fra FT

Tabellen presenterer forskjellen mellom graden av positivitet i gruppe 2 og gruppe 3 når det gjelder bruk av drone i revisjonen. Resultatene viser at gjennomsnittet for gruppen som brukte analytisk metode var 5,417, mens gruppen som brukte analytisk metode med bevisstgjøring hadde en litt lavere gjennomsnittsverdi på 5,364. Forskjellen mellom gruppene var imidlertid så liten at det er vanskelig å konkludere med om kommentaren hadde noen påvirkning på graden av positivitet i gruppe 3. Selv om tabellen viser at denne gruppen la relativt stor vekt på Finanstilsynets vurdering, med en gjennomsnittsverdi på 5.

6. Konklusjon og diskusjon

I dette kapittelet vil vi diskutere resultatene og analysen opp mot tidligere teori og forskning. Ut fra resultatene vil vi utarbeide en konklusjon på våre funn, og til slutt diskutere begrensningene med oppgaven, for å så komme med forslag til videre forskning.

6.1 Diskusjon av hypoteser

6.1.1 Hypotese 1

H_0 beholdes: *Ved bruk av ADA i revisjonen av varelageret vil revisors profesjonelle skepsis ikke bli høyere enn ved bruk av tradisjonelle revisjonsmetoder.*

Vår studie om at revisors profesjonelle skepsis blir høyere ved bruk av analytiske revisjonsmetoder ved vurdering av innhentete revisjonsbevis av varelager, kontra de tradisjonelle metodene, gav ikke et signifikant statistisk grunnlag å konkludere på denne hypotesen. P-verdien på 0,139 gav oss et grunnlag for å beholde nullhypotesen.

Revisjonsbransjen er på lik linje med andre bransjer avhengige av å følge den teknologiske utviklingen for å holde tritt. Siden introduksjonen av CAATs, til forlengelsen til ADA, til nå introdusering av droner under pandemien, har praktiserende revisorer, revisjonsforskere, standardsettere og offentlige tilsyn vært uenige om revisjonsbevisene som innhentes ved bruk av digitale verktøy er gode nok til å vurderes hensiktsmessige og tilstrekkelige nok. Dette gir derfor en uenighet om revisors profesjonelle skepsis bør økes eller ikke, ved bruk av disse verktøyene. Vår studie viser at i gjennomsnitt ved målt av indeks at revisors profesjonelle skepsis er høyere ved bruk av de analytiske metodene mot de godt implementerte tradisjonelle måtene å gjøre det på. Resultatet viser at de som benyttet analytisk metode hadde 28% høyere grad av profesjonell skepsis. Dette er i tråd med teori fra blant annet (Nearon, 2005), som tidlig mente at ved bruk av analytiske metoder må revisor innhente revisjonsbevis på den «vanlige» måten i tillegg, og at revisor må akte høyere grad av skepsis. Dette støttes videre om man ser på profesjonell skepsis påstandene individuelt sett, hvor tilstrekkelighet, altså nødvendighet for å innhente flere bevis, er den største differansen mellom analytisk metode og tradisjonell

metode. Bevisets gyldighet har også en markant forskjell, hvor det er større tvil om man kan stole på beviset som er innhentet av den analytiske metoden, noe som støttes av revisjonsteori.

Det må imidlertid kommenteres at standardavviket i gruppen som benyttet tradisjonell metode er høyt (1,531). Spredningen i hva respondentene svarte varierte derfor mye, noe som kan skyldes forskjellige faktorer. En mulighet er at oppgaven var for kompleks eller uoversiktlig, som kan være grunnet for mye informasjon i caseoppgaven. En annen årsak kan være at respondentene som benyttet tradisjonell metode stort sett ikke har erfaring med å gjøre dette i praksis, da de fleste er studenter. Vi ser fra tabellen i vedlegg 5 at respondentene uten arbeidserfaring har en lavere skepsis (2,9), men med et høyere standardavvik (1,57). Det samme gjelder de med under 2 års arbeidserfaring. De med 2-5 års erfaring har høyere skepsis, og en lavere spredning, men de er et fåtall respondenter. Spredningen blant de med ingen eller mindre erfaring, er nok grunnen til at gruppen tradisjonell metode ikke ble normalfordelt, og som gjør at resultatet vårt fra analysen ikke hadde en signifikant påvirkningskraft.

Funnene fra testingen viser et mønster som forteller at det er en forskjell mellom gruppene, selv om resultatet ikke er signifikant med en p-verdi på 13,9%. Nullhypotesen kan da ikke forkastes, og det kan ikke statistisk sett påvises her at profesjonell skepsis øker ved bruk av ADA til bevisinnhenting. Det kan tenkes at resultatet hadde vært signifikant ved en større populasjon, og/eller at en større andel av populasjonen hadde arbeidserfaring, i tråd med teori fra (Nearon, 2005). Dette er imidlertid en kjent og akseptert risiko ved å gjennomføre et forskningseksperiment som en masteroppgave.

6.1.2 Hypotese 2

H₀ beholdes: Ved bruk av ADA og bevisstgjøring i revisjonen av varelageret vil revisors profesjonelle skepsis ikke bli høyere enn når man ikke har bevisstgjøring.

Hypotese to undersøker om det oppstår en effekt av bevisstgjøring rundt risikoene som kan påfalle revisor ved bruk av analytiske verktøy for å innhente revisjonsbevis, og dermed øke den profesjonelle holdningen. Resultatet fra testen gav ikke et statistisk signifikant grunnlag, og vi kan dermed ikke konkludere at bevisstgjøringen øker revisors skeptiske holdning. P-verdien på 0,141 gir oss grunnlag til å beholde nullhypotesen.

Selv om teknologien er til stede, er potensialet til bruk av analytiske metoder i revisjonen langt fra utnyttet. Dette er grunnet en såkalt «First-mover»-situasjon mellom revisorer, standardsettere og tilsynsmyndigheter. Spørsmålet er hvem som skal ta det første steget til å bruke dette når retningslinjene er uklare og risikoen stort sett sitter på revisor (Christ et al., 2021; Eilifsen et al., 2014). Resultatet fra undersøkelsen viser at respondenter som ble bevisstgjort av risikoene ved bruk av analytisk revisjonsmetodikk hadde en høyere grad av skepsis totalt sett, en respondenter som ikke fikk dette. Vi ser at respondentene som hadde bevisstgjøringseffekt hadde en gjennomsnittlig skeptisk holdning på 3,991, mot de som ikke hadde bevisstgjøring på 3,661, en gjennomsnittlig forskjell på 9,01%. Standardavvikene var henholdsvis (1,011) mot (0,9) på de respektive gruppene.

Enkeltpåstanden «tilfredsstillende revisjon» hadde den største forskjellen, hvor respondentene som fikk bevisstgjøring av risikoene en høyere skeptisk holdning på 4,136 i snitt mot 3,542. Påstanden «komfortabel med å godta revisjonsbeviset», viser også en større forskjell enn resten med et snitt på 4,136 med bevisstgjøring, mot 3,542 for de som ikke fikk påminnelse. En mulig forklaring på dette er at med økt grad av bevissthet kan føre til at man tolker risikoene annerledes, og reagerer annerledes ut fra automatisk respons (LeadersharpGroup, 2019). Ut fra Prospekt-teorien av (Kahneman & Tversky, 1979), vurderer man risiko forskjellig ut fra tap og gevinst, og at man legger større vekt på tapet en gevinsten. Forskjellene i gruppene kan dermed være at man vurderer risikoen for at man kan bli erstatningsansvarlig høyere ved bruken av disse verktøyene når man har blitt bevisst på risikoen, og at det ikke er verdt å risikere i og med at man kan bli erstatningsansvarlig selv.

I og med at resultatet ikke er signifikant kan det være andre faktorer som avgjør. En forklaring kan være at bevisstgjørings-notisen ble bortgjemt i resten av informasjonen, og ble derfor ikke lagt merke til i så stor grad som ønskelig. Videre igjen som i forrige hypotese kan mangel på erfaring blant respondentene som fikk denne gruppen en årsak, dette var kun 1 (over 5 år). Det er likevel ikke sikkert at mer erfaring tilsier at bevisstgjøringen hadde hatt effekt, men at det er en relativt kjent problemstilling blant revisorer at revisjonsbevisene må være tilstrekkelig og hensiktsmessig, og at analytiske metoder ikke nødvendigvis gir dette, kunne det vært annerledes med en større populasjon med mer erfaring.

Våre funn sier igjen at det er et mønster mellom bevisstgjøring av risikoene for revisor tilknyttet uregulert bruk av analytiske revisjonsmetoder, men testene ikke er statistisk signifikant med en p-verdi på 14,1%. Nullhypotesen beholdes følgelig.

6.1.3 Andre funn

I undersøkelsen inkluderte vi 12 tilleggsspørsmål som alle respondentene fikk spørsmål om. Disse 12 spørsmålene måler Robinson-skala, og skal kunne definere revisors profesjonelle skeptisk som en tilstand (Robinson et al., 2018).

Resultatet fra testen viste at den gruppen som scorer høyt i Robinson-skala viser en gjennomsnittlig profesjonell skepsis ovenfor caset på 3,6414, mot de som scorer lavt i Robinson-skala som viser et gjennomsnitt på 3,4049 i profesjonell skepsis. Dette betyr at de med høy profesjonell skepsis som en tilstand også utøvde en høyere gjennomsnittlig skepsis ved vurderingen av de ulike påstandene i caset. Denne trenden kan tyde på at det er en sammenheng mellom profesjonell skeptisk som en tilstand og de skeptiske vurderingene som ble gjort i undersøkelsen.

Det er likevel usikkert hvor mye vi kan stole på at det er en gjennomgående trend med at høy profesjonell skepsis som en tilstand, fører til at revisor viser til høyere profesjonell skepsis ovenfor caset. Resultatet viser også at gruppen tradisjonell metode viser lavere profesjonell skepsis når den har høy profesjonell skepsis som en tilstand enn når den har profesjonell skepsis som en tilstand som lav. Dette kan kanskje skyldes at de i gruppen tradisjonell metode ikke har forstått spørsmålene eller har fått med seg all relevant informasjon, noe som kan forklares med at de fikk mer informasjon i caset enn de andre gruppene.

Vi kan konkludere med at det er en tydelig sammenheng mellom graden av skeptisk holdning som en tilstand, og den gjennomsnittlige profesjonelle skepsisen ovenfor påstandene i caset for både gruppen som bruker analytisk metode, og gruppen som bruker analytisk metode med bevisstgjøring. Vi kan derimot ikke konkludere noe tegn til en trend i gruppen tradisjonell metode, ettersom det er et betydelig avvik og ingen trend å vise til.

Vi inkluderte også spørsmålet «*hvor positiv er du til bruk av drone i revisjonen*» til gruppene analytisk metode og analytisk metode med bevisstgjøring i undersøkelsen. Dette spørsmålet

ble stilt for å kunne se om det var noen forskjeller mellom gruppene ved deres syn om bruk av drone i revisjonen.

Resultatet fra testen²¹ viser at gruppen analytisk metode har et gjennomsnitt på 5,417, mot gruppen analytisk metode med bevisstgjøring, som har et gjennomsnitt på 5,364. Dette resultatet viser at gruppen analytisk metode er mer positiv til bruk av drone og teknologi, enn gruppen analytisk metode med bevisstgjøring i revisjonen. Dette resultatet bidrar derfor til å støtte oppunder hypotese 2, nemlig at det finnes en trend at gruppen analytisk metode med bevisstgjøring har en høyere profesjonell skepsis enn gruppen analytisk metode.

For å kunne konkludere at utfallet i spørsmålet om «hvor positiv til bruk av drone i revisjonen», så stilte vi også et kontrollspørsmål til gruppen analytisk metode med bevisstgjøring. Der stilte vi spørsmålet om «i hvilken grad vektlegger de finanstilsynets utsagn». Resultatet²² viser et gjennomsnitt på 5, som viser tydelig at det ikke er tilfeldig at analytisk metode med bevisstgjøring er mer skeptiske til drone og teknologi i revisjonen enn gruppen analytisk metode.

6.2 Konklusjon

Vårt eksperiment hadde som formål å undersøke sammenhengen mellom valg av revisjonsmetode og nivå av revisors profesjonelle skepsis, i tillegg om bevisstgjøring av risikoene til revisor hadde en effekt. Vi målte revisors profesjonelle skepsis gjennom Nelsons (2009) rammeverk når respondentene vurderte revisjonsbevisene fra henholdsvis tradisjonelle revisjonsmetoder med stikkprøver, og analytisk metode med bruk av drone og kunstig intelligens. Våre funn indikerer at ved bruk av analytisk metode er revisors skeptiske holdninger høyere enn ved bruk av tradisjonelle metoder. Likevel ga ikke resultatene statistisk signifikante verdier som gir grunnlag for å konkludere dette.

²¹ Se vedlegg 3 for fullstendig test resultat

²² Se vedlegg 3 for fullstendig tabell.

Videre undersøkte vi om bevisstgjøring av risikoene rundt det å være «first mover» av de analytiske metodene hadde effekt for revisors skeptiske holdning. Av resultatene ser vi at det er en liten forskjell mellom revisors skeptiske holdning når man får informasjon om risikoene før man benytter de analytiske metodene, mot når man ikke får dette. Forskjellen er her heller signifikant, noe som betyr at vi ikke kan konkludere med at bevisstgjøring av risikoene har noe effekt.

6.3 Begrensninger

Vår eksperimentelle studie har som alle studier innenfor denne forskningsmetoden sine begrensninger. Dette ettersom man undersøker noen få aspekter som er til stede ved en reell setting. Studien er en masteroppgave hvor det er begrenset med ressurser og tid, noe som betyr at både utvalget og størrelsen av utvalget kunne vært bedre. Det gunstige utvalget ved denne type undersøkelse, hvor man tester revisjonsmetoder og revisjonsbevis, ville vært å benytte allerede praktiserende revisorer.

Videre er både profesjonell skepsis og effekten av bevisstgjøring vanskelig å måle. Førstnevnte fordi det er vanskelig å definere hva dette faktisk er, og i tillegg til at metoden for måling av bevisene benyttet i undersøkelsen er relativt ny sammenlignet med tidligere teori. Dette er da en begrensning, selv om vi benyttet to målemetoder i undersøkelsen for å styrke troverdigheten av målingene. Bevisstgjøringseffekten er utfordrende å måle, selv om vi ser fra resultatet at respondentene la vekt på notisen i relativ stor grad. Det kan hende at notisen forsvant i mengden av omfanget av caset, noe som gjorde at resultatet ikke er like valid som ønsket. Vi ønsket å lage en undersøkelse som ikke var for lang, men at den virket realistisk nok til å kunne ligne en ekte setting, hvor man har mye informasjon og beslutningsgrunnlag.

6.4 Videre forskning

Vår studie er innenfor et nokså nytt område innenfor revisjonsfaget. Bruk av droner og automatisert teknologi er fortsatt i en tidlig fase, og ikke så utbredt i revisjonspraksis. Det finnes derfor mange interessante områder som kan undersøkes videre. For eksempel ville det vært interessant å gjennomføre studien med praktiserende revisorer for å få potensielt mer realistiske resultater, siden de har mer erfaring og kan ha foretatt andre vurderinger enn studenter. Det ville også vært interessant å gjennomføre intervjuer med selskaper som allerede har brukt droner og kunstig intelligens i revisjonsprosessen for å høre deres synspunkter på temaet.

En annen spennende problemstilling er "Remote-auditing". Etter Covid-19 og de påfølgende nedstengningene har flere virksomheter sett at det ikke nødvendigvis er behov for å være fysisk til stede på arbeidsplassen for å utføre sine oppgaver. Remote-auditing innebærer at revisor ikke trenger å være til stede hos revisjonsklienten, og at alle revisjonsbevisene kan lastes opp og deles i en skybasert løsning (Teeter et al., 2010). Dette er en løsning som sannsynligvis vil bli mer vanlig i fremtiden, og det ville vært interessant å undersøke om dette kan forbedre revisjonen, eller ikke.

Litteraturliste

Abdolmohammadi, M., & Wright, A. (1987). An Examination of the Effects of Experience and Task Complexity on Audit Judgments. *The Accounting Review*, 62(1), 1–13.

Banker, S. (2016, juni 24). *Drones and Robots in the Warehouse*. Forbes. <https://www.forbes.com/sites/stevebanker/2016/06/24/drones-and-robots-in-the-warehouse/>

Bloomberg. (2020, januar 2). *Big Four Invest Billions in Tech, Reshaping Their Identities*. <https://news.bloombergtax.com/financial-accounting/big-four-invest-billions-in-tech-reshaping-their-identities>

Brataas, E. (2022, mai 12). *Data analytics as an audit tool – advantages and some experiences*. IIA. <https://iaa.no/data-analytics-as-an-audit-tool-advantages-and-some-experiences/>

Charness, G., Gneezy, U., & Kuhn, M. A. (2011). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8. <https://doi.org/10.1016/j.jebo.2011.08.009>

Christ, M. H., Emett, S. A., Summers, S. L., & Wood, D. A. (2021). Prepare for takeoff: Improving asset measurement and audit quality with drone-enabled inventory audit procedures. *Review of Accounting Studies*, 26(4), 1323–1343. <https://doi.org/10.1007/s11142-020-09574-5>

DataConsulting. (2009, september 1). *Knowledge base – What are CAATs?* <http://www.dataconsulting.co.uk/knowledge-base-what-are-caats/>

Deloitte. (2019, november 3). *Teknologien snur tradisjonelt revisorarbeid på hodet*. Deloitte Norway. <https://www2.deloitte.com/no/no/innsikt/okonomi-og-finans/revisjon-regnskap/analytics-revisjon.html>

Eilifsen, A., Kinserdal, F., Messier, W. F., & McKee, T. E. (2020). An Exploratory Study into the Use of Audit Data Analytics on Audit Engagements. *Accounting Horizons*, 34(4), 75–103. <https://doi.org/10.2308/HORIZONS-19-121>

Eilifsen, A., Messier Jr, William F., Glover, Steven M., & Prawitt, Douglas F. (Red.). (2014). *Auditing and assurance services* (3rd internasjonale ed.). McGraw-Hill Education.

Finanstilsynet. (2017, februar 16). *Revisor*. Finanstilsynet. <https://www.finanstilsynet.no/tilsyn/revisor/>

Grønmo, S. (2016). *Samfunnsvitenskapelige metoder* (2. utg.). Fagbokforlaget.

Hayes Jr., Arthur A. (2014). Improper payments—How Technology Can Negatively Impact the Exercise of Professional Skepticism. *The Journal of Government Financial Management*, 63(2), 54–55.

Hurt, R. K. (2010). Development of a Scale to Measure Professional Skepticism. *AUDITING: A Journal of Practice & Theory*, 29(1), 149–171. <https://doi.org/10.2308/aud.2010.29.1.149>

Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation. *Journal of Emerging Technologies in Accounting*, 13(2), 1–20. <https://doi.org/10.2308/jeta-10511>

IAASB. (2009a). *ISA 200 Overordnede mål for den uavhengige revisor og gjennomføringen av en revisjon i samsvar med de internasjonale revisjonsstandardene* (s. 26). IFAC.

IAASB. (2009b). *ISA 501 Revisjonsbevis—Særlig hensyn knyttet til utvalgte poster* (s. 10). IFAC.

IAASB. (2009c). *ISA 530 Stikkprøver i revisjonen* (s. 15). IFAC.

IAASB. (2009b). *ISA 500 Revisjonsbevis* (s. 69). IFAC.

Jacobsen, Dag Ingvar. (2018). *Hvordan gjennomføre undersøkelser?: Innføring i samfunnsvitenskapelig metode* (3.utg). Cappelen Damm Akademisk.

Jans, M., Alles, M., & Vasarhelyi, M. (2013). The case for process mining in auditing: Sources of value added and areas of application. *International Journal of Accounting Information Systems*, 14(1), 1–20. <https://doi.org/10.1016/j.accinf.2012.06.015>

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>

Kinserdal, F. (2017). NHH skal forske på digitalisering i revisjonsbransjen. *Magma*, 79–86.

KPMG. (2011). *Elevating Professional Judgement in Auditing and Accounting: The KPMG Professional Framework*.
https://www.academia.edu/35020990/Elevating_Professional_Judgment_in_Auditing_and_Accounting_The_KPMG_Professional_Judgment_Framework

Kulset, E. (2020). Revisors bruk av «Computer Assisted Audit Techniques (CAATs)» for å identifisere transaksjoner med høy risiko for feil og misligheter. *Kommunerevisoren*, 2, 19–21.

Laerd Statistics. (u.å.-a). *Independent t-test in SPSS Statistics—Procedure, output and interpretation of the output using a relevant example* | Laerd Statistics. Hentet 19. april 2023, fra <https://statistics.laerd.com/spss-tutorials/independent-t-test-using-spss-statistics.php>

Laerd Statistics. (u.å.-b). *Mann-Whitney U Test in SPSS Statistics | Setup, Procedure & Interpretation* | Laerd Statistics. Hentet 24. april 2023, fra <https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php>

Larsen, A. K. (2017). *En enklere metode* (2. utg.). Fagbokforlaget.

LeadersharpGroup. (2019). *What is conscious leadership and why does it matter?* <https://www.leadersharp.com/publications/what-is-conscious-leadership-and-why-does-it-matter/>

Libby, R. (1981). *Accounting and human information processing: Theory and applications*. Prentice-Hall, Englewood Cliffs, N.J.

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, 27, 775–810. [https://doi.org/10.1016/S0361-3682\(01\)00011-3](https://doi.org/10.1016/S0361-3682(01)00011-3)

Mortensen, T., Fisher, R., & Wines, G. (2012). Students as surrogates for practicing accountants: Further evidence. *Accounting Forum*, 36(4), 251–265. <https://doi.org/10.1016/j.accfor.2012.06.003>

Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics*, 167(2), 209–234. <https://doi.org/10.1007/s10551-019-04407-1>

Nearon, B. H. (2005). Foundations in Auditing and Digital Evidence. *The CPA Journal*, 75(1), 32–33.

Nehmer, R. A., & Appelbaum, D. (2017). Using Drones in Internal and External Audits: An Exploratory Framework—Journal of Emerging Technologies in Accounting. *American Accounting Association*, 14(1), 99–113. <https://doi.org/10.2308/jeta-51704>

Nelson, M. J. (2009). A Model and Literature Review of Professional Skepticism in Auditing. *Auditing: A Journal of Practice & Theory*, 28(2), 1–34. <https://doi.org/10.2308/aud.2009.28.2.1>

Nhh-arsrapport-2021.pdf. (u.å.). Hentet 24. april 2023, fra <https://www.nhh.no/globalassets/om-nhh/nhh-arsrapport-2021.pdf>

Nolder, C. J., & Kadous, K. (2018). Grounding the professional skepticism construct in mindset and attitude theory: A way forward. *Accounting, Organizations and Society*, 67, 1–14. <https://doi.org/10.1016/j.aos.2018.03.010>

PCAOB. (2002). *AS 1015: Due Professional Care in the Performance of Work*. PCAOB - Public Company Accounting Oversight Board. <https://pcaobus.org/oversight/standards/auditing-standards/details/AS1015>

PCAOB. (2020). *Data Technology Project Spotlight*. <https://pcaobus.org/Documents/Data-Technology-Project-Spotlight.pdf>

Pedersen, Jo Sigurd. (2016). *Dataanalyse i revisjon: Den nye historie fortelleren*. Regnskap & revisjon. <https://www.revregn.no/asset/pdf/2016/7-30-1d.pdf>

Peecher, M. E., & Solomon, I. (2001). Theory and Experimentation in Studies of Audit Judgments and Decisions: Avoiding Common Research Traps. *International Journal of Auditing*, 5(3), 193–203. <https://doi.org/10.1111/1099-1123.00335>

PricewaterhouseCoopers. (u.å.-a). *Big data*. PwC. Hentet 6. mars 2023, fra <https://www.pwc.no/no/teknologi-omstilling/digitalisering-pa-1-2-3/big-data---data-lake.html>

PricewaterhouseCoopers. (u.å.-b). *Continuous audit and monitoring*. PwC. Hentet 8. mars 2023, fra <https://www.pwc.com/vn/en/services/consulting/continuous-audit-monitoring.html>

PricewaterhouseCoopers. (u.å.-c). *Digitale revisjonsverktøy*. PwC. Hentet 8. mars 2023, fra <https://www.pwc.no/no/tjenester/revisjon/innovasjon-og-verifikasjonstjenester.html>

PricewaterhouseCoopers. (u.å.-d). *Hva er kunstig intelligens?* PwC. Hentet 14. mars 2023, fra <https://www.pwc.no/no/teknologi-omstilling/digitalisering-pa-1-2-3/kunstig-intelligens.html>

PricewaterhouseCoopers. (2016). *Clarity from above*. 40.

PricewaterhouseCoopers. (2018). *GL.ai: PwC's anomaly detection for the general ledger*. <https://www.pwc.com/m1/en/events/socpa-2020/documents/gl-ai-brochure.pdf?fbclid=IwAR1nrZBc7okgu6aTyeNJfG32imXFQlmOA20-hG0eWD9nTpfWsBNe6XEaGS8>

PricewaterhouseCoopers. (2019, januar). *PwC completes its first stock count audit using drone technology*. PwC. <https://www.pwc.co.uk/press-room/press-releases/pwc-first-stock-count-audit-drones.html>

Redaksjonen. (u.å.). *Automatisering*. Hentet 6. mars 2023, fra <https://www.cegal.com/no/ordbok/automatisering>

Revisors erstatningsansvar for vesentlige feil i årsregnskapet | Revisjon og Regnskap. (u.å.). Hentet 5. mai 2023, fra https://www.revregn.no/journal/2020/7/m-1207/Revisors_erstatningsansvar_for_vesentlige_feil_i_arsregnskapet

Robinson, S. N., Curtis, M. B., & Robertson, J. C. (2018). Disentangling the Trait and State Components of Professional Skepticism: Specifying a Process for State Scale Development. *AUDITING: A Journal of Practice & Theory*, 37(1), 215–235. <https://doi.org/10.2308/ajpt-51738>

Sjanifer. (2020, april 29). *Remote Auditing During Pandemic: Zoom and Drones*. Thomson Reuters Tax & Accounting News. <https://tax.thomsonreuters.com/news/remote-auditing-during-pandemic-zoom-and-drones/>

StatisticsHowTo. (u.å.). *Cronbach's Alpha: Definition, Interpretation, SPSS*. Statistics How To. Hentet 19. april 2023, fra <https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/cronbachs-alpha-spss/>

Svartdal, F. (2023). Bevissthet – psykologi. I *Store norske leksikon*. https://snl.no/bevissthet_-_psykologi

Techtarget.com. (u.å.). *What is RFID and how does it work?* IoT Agenda. Hentet 12. mai 2023, fra <https://www.techtarget.com/iotagenda/definition/RFID-radio-frequency-identification>

Teeter, R. A., Alles, M. G., & Vasarhelyi, M. A. (2010). The Remote Audit. *Journal of Emerging Technologies in Accounting*, 7(1), 73–88. <https://doi.org/10.2308/jeta.2010.7.1.73>

Tidemann, A., & Elster, A. C. (2023). Maskinl ring. I *Store norske leksikon*. <http://snl.no/maskinl%C3%A6ring>

Trochim, William M., Donnelly, James P., & Arora, Kanika. (2014). *Research Methods: The Essential Knowledge Base* (2. utg). Cengage Learning.

Trotman, K. T. (2001). Design Issues in Audit JDM Experiments. *International Journal of Auditing*, 5(3), 181–192. <https://doi.org/10.1111/1099-1123.00334>

Vedlegg

Vedlegg 1: spørreundersøkelsen

Intro tekst

Hei! Takk for at du tar deg tid til å delta i undersøkelsen i forbindelse med vår masterutredning. Svarene du gir er anonyme og dataen behandles med full konfidensialitet. Estimert tidsbruk er 7 minutter. Vi ber vennlig om at informasjonen leses nøye og at du svarer ærlig og selvstendig.
Lykke til!

Case tekst

Din rolle

Du er partner og revisor i revisjonsselskapet Nexus Revisjon AS, og er i gang med å revidere årsregnskapet til bedriften TechZone AS for 2022. I den forbindelse skal varetellingen gjennomføres, og du er pliktig til å være til stede på denne. Du har vært revisor for selskapet i 5 år og kjenner derfor selskapet godt.

Informasjon om TechZone AS

TechZone er en middels stor elektrokjede som har butikker flere steder i Norge. Den største butikken er lokalisert i Fyllingsdalen, hvor de også har et stort varelager som de sender varer videre ut til resten av landet. TechZone selger alt fra mobiltelefoner og datamaskiner til TV-er og spillkonsoller. Selskapet har bare eksistert i 7 år, og har hatt en

voldsom vekst de siste årene hvor de har tatt markedsandeler.

	2022	2021	2020
Salgsinntekter	308,000,000	265,000,000	207,000,000
Driftresultat	10,800,000	9,400,000	6,600,000
Varelager	42,291,000	36,693,000	29,497,000
Omløpshastighet varelager i dager	54	52	52

Varelageret

Ledelsen i TechZone er veldig positiv til selskapets fremtid og tror på fortsatt vekst og suksess. De ser utfordringene med det store varelageret som en mulighet til å forbedre driften og øke lønnsomheten. Ledelsen har tatt initiativ til å optimalisere varelageret og har investert i teknologi og samarbeid med leverandører for å øke effektiviteten og redusere kostnadene. De har dermed satt seg et hårete mål om å redusere omløpshastigheten fra 54 dager til 25 dager.

Selskapet har et integrert ERP system hvor bestillinger til varelageret og fakturaer mottas gjennom samme system. Ved ankomst av varer blir alle varer kontrollert, deretter godkjenner lagersjefen fakturaen. Når dette er godkjent av lagersjefen blir varene bokført i egen varelinje i varelageret til kostpris. Det er kun lagersjefen som kan gjennomføre dette, og har ikke mulighet til å gjøre det utenfor åpningstid.

Done + AI bevisstgjøring

Nexus Revisjon er et fremoverlent revisjonsselskap som i de siste årene har investert stort i teknologi som kan gjøre revisjonen mer sikker og effektiv. Et av systemene de har utviklet er "DroneAudit.AI" (DAA). Dette er et digitalt verktøy som i samarbeid med drone bruker kunstig intelligens til å inspisere og evaluere hele varelageret på en effektiv og nøyaktig måte.

Programvaren fungerer ved at en drone flyr rundt varelageret, tar høyoppløselige bilder og skanner alle produkter som ligger inne. Bildene og dataen blir deretter behandlet og analysert av programvaren ved hjelp av maskinlæring og kunstig intelligens-algoritmer. Dette gjør det mulig å gjenkjenne forskjellige produkter og beregne mengder og verdier på en nøyaktig måte. DAA kan også oppdage avvik og anomalier i varelageret, for eksempel ødelagte eller manglende produkter.

Programmet kan deretter generere en detaljert rapport som gir en oversikt over varelageret, og gir verdifulle innsikter som kan hjelpe revisor i vurderingen av flere revisjons handlinger.

Du har besluttet å benytte dette verktøyet i revisjonen av Techzone ettersom du mener selskapets ERP-systemer og kontroller er tilstrekkelige nok til at man kan stole på dataen dronen henter inn. Både du og dronen er fysisk til stede, og DAA bruker informasjon fra denne, i tillegg til inputen du har lagt inn, for å vurdere varelageret. For at DAA skal kunne ta beslutning på riktig grunnlag har du tatt en stikkprøve av inputene og kan med betryggende sikkerhet si at den ikke inneholder vesentlige feil.

ISA`ene gir ingen informasjon om det er lov eller ikke til å ta i bruk drone eller kunstig intelligens i revisjonen av varelager. Finanstilsynet har heller ikke tatt stilling til spørsmålet om å ta i bruk teknologien. Revisor kan oppleve av at det er helt greit å ta i bruk ny teknologi fra standardsettere og tilsyn, men man kan også risikere å få erstatningsansvar ved å bruke teknologi som ikke er godkjent.



Fellesspørsmål

Hvor bekymret er du for at det er vesentlig feilinformasjon i varelageret med hensyn til

	Veldig liten					Veldig stor	
	1	2	3	4	5	6	7
Eksistens	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verdi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Har TechZone hatt en positiv vekst de siste årene?

Ja

Nei

I hvilken grad

	I svært liten grad					I svært stor grad	
	1	2	3	4	5	6	7
Mener du at revisjonsbeviset er gyldig?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Er du komfortabel med å godta dette som revisjonsbevis?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mener du det finnes behov for å innhente flere revisjonsbevis?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Basert på innhentet revisjonsbevis mener du at det er risiko for at varelageret inneholder vesentlig feilinformasjon?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Syns du revisjonen er gjennomført på en tilfredstillende måte?

Drone + AI bevisstgjøring spørsmål

I hvor stor grad

	Svært liten grad					Svært stor grad	
	1	2	3	4	5	6	7
La du vekt på finanstilsynets stilling ved vurderingen du gjorde?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Er du positiv til å bruke drone og AI i revisjonen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Drone + AI

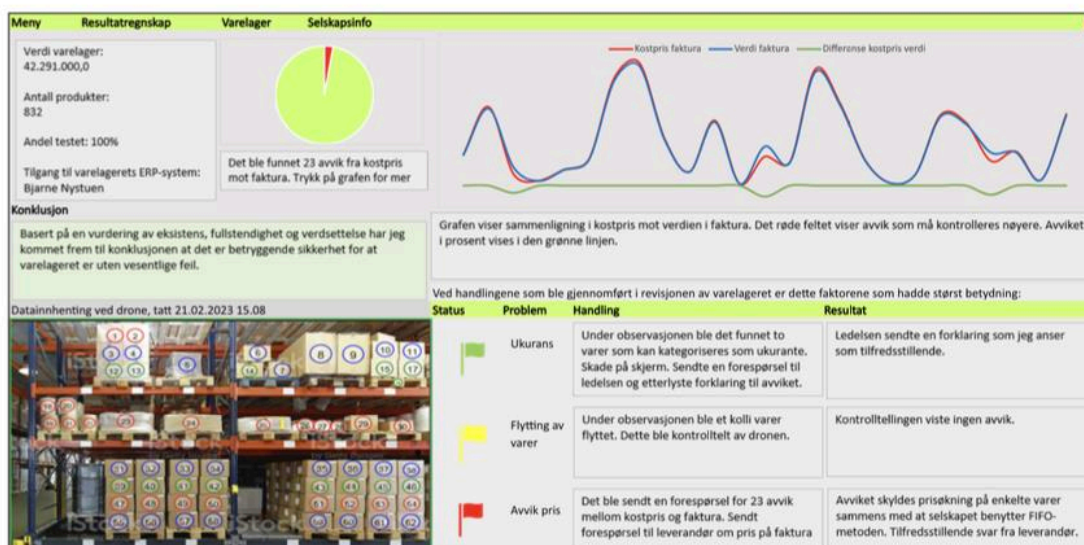
Nexus Revisjon er et fremoverlent revisjonsselskap som i de siste årene har investert stort i teknologi som kan gjøre revisjonen mer sikker og effektiv. Et av systemene de har utviklet er "DroneAudit.AI" (DAA). Dette er et digitalt verktøy som i samarbeid med drone bruker kunstig intelligens til å inspisere og evaluere hele varelageret på en effektiv og nøyaktig måte.

Programvaren fungerer ved at en drone flyr rundt varelageret, tar høyoppløselige bilder og skanner alle produkter som ligger inne. Bildene og dataen blir deretter behandlet og analysert av programvaren ved hjelp av maskinlæring og kunstig intelligens-algoritmer. Dette gjør det mulig å gjenkjenne forskjellige produkter og beregne mengder og verdier på en nøyaktig måte. DAA kan også oppdage avvik og anomalier i varelageret, for eksempel ødelagte eller manglende produkter.

Programmet kan deretter generere en detaljert rapport som gir en oversikt over

varelageret, og gir verdifulle innsikter som kan hjelpe revisor i vurderingen av flere revisjonshandlinger.

Du har besluttet å benytte dette verktøyet i revisjonen av Techzone ettersom du mener selskapets ERP-systemer og kontroller er tilstrekkelige nok til at man kan stole på dataen dronen henter inn. Både du og dronen er fysisk til stede, og DAA bruker informasjon fra denne, i tillegg til inputen du har lagt inn, for å vurdere varelageret. For at DAA skal kunne ta beslutning på riktig grunnlag har du tatt en stikkprøve av inputene og kan med betryggende sikkerhet si at den ikke inneholder vesentlige feil.



Block 9

I hvor stor grad

I svært liten grad

I svært stor grad

1 2 3 4 5 6 7

Er du positiv til å bruke drone og AI i revisjonen?

Tradisjonell metode

Du og revisjonsteamet har valgt ut varelager som en vesentlig post i regnskapet, og har valgt påstandene verdsettelse, eksistens og nøyaktighet. Revisjonshandlinger som er valgt er fysisk deltakelse på varetelling iht. ISA 501 og kontroll av kostpris mot inngående faktura. Utvalgsmetode er MUS som er varelagerliste som er avstemt mot regnskapslinjen uten avvik. MUS gir revisjonsteamet ett utvalg på 25 stykker som igjen utgjør 10% av regnskapslinjen "Varelager". Forutsette at du har alle revisjonsbevis tilgjengelig som underdokumentasjon (varelagerliste + testede faktura).

Formål:	Kontrollere nøyaktigheten av varelagerets priser og eksistens
Planlagt revisjonshandling:	Kontrollere et tilfeldig utvalg av produkter som representerer 10% av varelageret
Utførte revisjonshandlinger:	Inspeksjon av produktene og tatt utvalg av kostpriser i varelageret

Varenr	Varenavn	antall	antall observasjoner	kostpris liste	kostpris faktura	Kostpris faktura	verdi i faktura	diff i kostpris	diff i verdi	Forklaring
3737128	HP ProBook 445 GB 14" Full HD	17	17	6,990	118,830	6,990	118,830	0	0	OK
5382809	Lenovo ThinkCentre M90q G3 Tiny	24	24	12,490	299,760	12,490	299,760	0	0	OK
4078813	Swive 34" Gamingskjerm Pxx 34D901 (sort)	22	21	3,199	67,179	3,199	70,378	0	-3,199	OK, ukurans
5823097	Philips 27" skjerm 273V7QDS8	15	15	1,199	17,985	1,199	17,985	0	0	OK
9313685	Samsung Galaxy A53 5G 256GB	19	19	3,196	60,724	3,196	60,724	0	0	OK
3678609	Samsung Galaxy Tab S6 lite 4G 64GB	23	23	4,392	101,016	4,392	101,016	0	0	OK
3696790	Macbook Pro 16 (2023) 512GB	15	15	27,192	407,880	27,192	407,880	0	0	OK
1610837	Samsung Galaxy Book3 Ultra 16" AMOLED WQXG/	14	14	33,352	466,928	33,352	466,928	0	0	OK
3157174	Samsung Galaxy S23 Ultra 1TB	11	11	16,392	180,312	16,392	180,312	0	0	OK
3041847	Apple AirPods Pro 2 trådløse ørepropper	22	22	2,392	52,624	2,392	52,624	0	0	OK
5211686	Mac Mini (2023) 512GB	17	17	14,559	247,503	14,566	247,622	-7	-119	OK, prisøkning i år
8284512	Apple HomePod mini	7	7	999	6,993	999	6,993	0	0	OK
1231828	DJI Mini 3 Fly More Combo	18	18	8,472	152,496	8,472	152,496	0	0	OK
6469906	Sony WH-1000XM5 trådløse hodetelefoner	25	25	3,592	89,800	3,592	89,800	0	0	OK
3656619	iPad Pro 12,9" (2022) 512GB WiFi	28	28	15,762	441,336	15,762	441,336	0	0	OK
7469993	iPhone 14 Pro Max 512GB	20	20	15,592	311,840	15,592	311,840	0	0	OK
9040112	Samsung 65" Q70B 4K QLED TV	12	12	7,992	95,904	7,992	95,904	0	0	OK
6418655	Trust Irs 4K konferansekamera + mikrofon	3	3	3,229	9,687	3,229	9,687	0	0	OK
6453240	DeLonghi Dedicca Arte ECB85.GY Manual Espressor	25	25	1,759	43,975	1,759	43,975	0	0	OK
1481643	LG 65" OLED 4K TV OLED65C56LA	21	21	12,792	268,632	12,792	268,632	0	0	OK
5070125	Philips 65" 4K OLED Ambilight 65OLED807/12	15	15	15,992	239,880	15,992	239,880	0	0	OK
1578290	Samsung The Freestyle DLP LED projektor	23	22	5,592	128,024	5,592	128,616	0	-5,592	OK, ukurans
1252476	Philips 55" 4K OLED Ambilight 55OLED707/12	13	13	9,952	129,376	9,952	129,376	0	0	OK
1887296	Corsair AX1600i, 1600W PSU	5	5	4,392	21,960	4,392	21,960	0	0	OK
3238166	Lenovo ThinkPad T14s G3 14" Full HD	16	16	16,990	271,840	16,990	271,840	0	0	OK
Sum				248,460	4,227,484		4,236,394			
Andel testet					10.0%					

Resultat

I revisjonen fant man to avvik mot siste faktura. Det første avviket gjaldt ukurans på to produkter som hadde fått skader på skjermen. Du har fått avklart med ledelsen hva dette gjelder, og de har gitt en tilfredsstillende forklaring. Det andre avviket gjaldt prisøkning på nye innkjøp i løpet av året. Disse ble sjekket mot tidligere fakturaer, og gjennomsnittsprisen er vurdert OK. Videre ble ingen vesentlige avvik oppdaget, og varelagerets kostpris blir

vurdert som tilstrekkelig nøyaktig.

Profesjonell skepsis

I hvilken grad er du enig i påstanden.

	Svært uenig		Middels enig			Svært enig	
	1	2	3	4	5	6	7
På generell basis hadde jeg en tendens til å stille spørsmål ved informasjon gitt i caset?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ved gjennomføring av Caset, stilte jeg spørsmål om elementer jeg så eller leste?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ved gjennomføring av spørreundersøkelsen hadde jeg en tendens til å ikke godta informasjon underveis, som ikke var bevist at det var sant?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Underveis i caset, så tok jeg meg godt tid til å ta beslutningene?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	1	2	3	4	5	6	7
Under eksprimentet likte jeg ikke å ta beslutninger uten å ha sett all tilgjengelig informasjon?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Jeg likte ikke å måtte ta raske beslutninger under caset	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Når jeg jobbet med caset, forsøkte jeg å forsikre meg om at jeg							

hadde vurdert all tilgjengelig informasjon før jeg fattet en beslutning?

Mens jeg jobbet med caset, så ventet jeg med å ta beslutning frem til jeg fant informasjonen jeg trengte?

1 2 3 4 5 6 7

Jeg føler at det ville øke mine muligheter til å ta riktig beslutninger i caset, dersom jeg hadde hatt mulighet til å se underliggende revisjonsbevis

Jeg forsøkte å lete etter mer revisjonsbevis for å styrke mine muligheter til å finne korrekt svar i caset

Jeg søkte aktiv etter all informasjon som var tilgjengelig ved gjennomføring av caset

Jeg benyttet alle ressurser som var tilgjengelig tilgjengelig for å få mest mulig informasjon ved gjennomføring av caset

Kontrollspørsmål

Kjønn?

Mann

Dame

Annet

Alder?

Under 20

20 - 25

26 - 30

Over 30

Arbeidserfaring innen revisjon?

Ingen arbeidserfaring

Under 2 år

2 - 5 år

Mer enn 5 år

Hvilken av disse studiene tilhører du eller graden du har tatt tidligere?

MRR - Master i regnskap og revisjon

MØA - Master I økonomi og administrasjon

Bachelor i regnskap og revisjon

Bachelor I økonomi og administrasjon

Annet

Har du tatt MRR453(digital revisjon), MRR411/ACC402 (revisjon 1) eller evt noe lignende fag?

Ja

Nei

Lvert av Qualtrics

Vedlegg 2: Trimmet gjennomsnitt

Descriptives				
	Hvilken gruppe		Statistic	Std. Error
Profesjonell skepsis Index	Tradisjonell metode	Mean	2.8750	.31040
		5% Trimmed Mean	2.8444	
	Analytisk metode	Mean	3.6833	.18373
		5% Trimmed Mean	3.6444	
	Analytisk metode med bevisstgjøring	Mean	3.9909	.21561
		5% Trimmed Mean	4.0232	

Vedlegg 3: Cronbach's alpha:

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.901	.907	5

Vedlegg 4: Normalfordeling – Shapiro-wilk test:

Hvilken gruppe	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Profesjonell skepsis Index	.218	24	.004	.854	24	.003
Tradisjonell metode	.116	24	.200*	.956	24	.371
Analytisk metode	.185	22	.048	.936	22	.164

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Vedlegg 5: Hypotese 1

	N	Mean	Std. Deviation	Minimum	Maximum
Profesjonell skepsis Index	70	3.5029	1.25559	1.00	5.80
Hvilken gruppe	70	1.9714	.81599	1.00	3.00

Mann-Whitney Test

Hvilken gruppe	N	Mean Rank	Sum of Ranks
Profesjonell skepsis Index			
Tradisjonell metode	24	21.52	516.50
Analytisk metode	24	27.48	659.50
Total	48		

Test Statistics^a

	Profesjonell skepsis Index
Mann-Whitney U	216.500
Wilcoxon W	516.500
Z	-1.480
Asymp. Sig. (2-tailed)	.139

a. Grouping Variable: Hvilken gruppe

Vedlegg 6: Hypotese 2

T-Test

Group Statistics

	Hvilken gruppe	N	Mean	Std. Deviation	Std. Error Mean
Profesjonell skepsis Index	Analytisk metode	24	3.6833	.90008	.18373
	Analytisk metode med bevisstgjøring	22	3.9909	1.01132	.21561

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
Profesjonell skepsis Index	Equal variances assumed	.218	.643	-1.091	44	.141	.281	-.30758	.28182	-.87554	.26039
	Equal variances not assumed			-1.086	42.236	.142	.284	-.30758	.28328	-.87916	.26400

Independent Samples Effect Sizes

		Standardizer ^a	Point Estimate	95% Confidence Interval	
				Lower	Upper
Profesjonell skepsis Index	Cohen's d	.95479	-.322	-.903	.262
	Hedges' correction	.97146	-.317	-.887	.258
	Glass's delta	1.01132	-.304	-.886	.285

- 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.

Vedlegg 7: Andre funn

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.922	.923	12

Population Descriptive Statistics

	N	Mean	Std. Deviation	Variance
Robinson profesjonelle skepsis	70	49.883	11.519	132.676
Valid N (listwise)	70			

Std. Deviation and Variance use N rather than N-1 in denominators.

Descriptive Statistics

Robinson profesjonelle skepsis	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean		Std. Deviation Statistic	Variance Statistic
					Statistic	Std. Error		
Tradisjonell metode	24	42.50	26.67	69.17	51.6698	2.41654	11.83857	140.152
Analytisk metode	24	47.50	22.50	70.00	48.0928	2.13176	10.44343	109.065
Analytisk metode med bevisstgjøring	22	39.17	27.50	66.67	49.8864	2.71352	12.72755	161.990

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
				Std. Error		
Profesjonell skepsis						
Lav Score	41	1.00	5.80	3.4049	.16162	1.03488
Høy Score	29	1.00	5.60	3.6414	.28291	1.52351

Descriptive Statistics

Profesjonell Skepsis	Hvilken gruppe	N	Minimum	Maximum	Mean	Std. Deviation	Variance
					Std. Error		
Lav Score	Tradisjonell metode	12	1.00	4.40	3.1833	.35884	1.24304
	Analytisk metode	17	2.40	5.80	3.4471	.22131	.91249
	Analytisk metode med bevisstgjøring	12	1.80	5.20	3.5667	.29627	1.02632
Høy Score	Tradisjonell metode	12	1.00	5.40	2.5667	.50682	1.75568
	Analytisk metode	7	3.40	5.00	4.2571	.22131	.58554
	Analytisk metode med bevisstgjøring	10	2.80	5.60	4.5000	.23898	.75572

Positiv til bruk av drone/Finanstilsynet

Descriptive Statistics^a

Hvilken gruppe		N	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance
						Std. Error		
Analytisk metode	Positiv til bruk av drone i revisjonen	24	2.00	7.00	130.00	5.4167	.24014	1.17646
	I hvor stor grad – La du vekt på finanstilsynets stilling ved vurderingen du gjorde?	0						
Analytisk metode med bevisstgjøring	Positiv til bruk av drone i revisjonen	22	1.00	7.00	118.00	5.3636	.30540	1.43246
	I hvor stor grad – La du vekt på finanstilsynets stilling ved vurderingen du gjorde?	22	3	7	110	5.00	.302	1.414

a. No statistics are computed for one or more split files because there are no valid cases.

Gjennomsnitt profesjonell skepsis mot arbeidserfaring

Descriptive Statistics								
Hvilken gruppe	Arbeidserfaring innen revisjon?	N	Minimum	Maximum	Mean		Std. Deviation	Variance
Tradisjonell metode	Ingen arbeidserfaring	12	1.00	5.40	2.9167	.45558	1.57817	2.491
	Under 2 år	8	1.00	4.80	2.3750	.60526	1.71193	2.931
	2 - 5 år	4	3.40	4.00	3.7500	.12583	.25166	.063
Analytisk metode	Ingen arbeidserfaring	17	2.40	5.00	3.6353	.18727	.77212	.596
	Under 2 år	5	2.40	5.00	3.4800	.46303	1.03537	1.072
	2 - 5 år	2	3.40	5.80	4.6000	1.20000	1.69706	2.880
Analytisk metode med bevisstjøring	Ingen arbeidserfaring	16	1.80	5.60	4.0375	.26091	1.04363	1.089
	Under 2 år	3	3.40	5.20	4.2000	.52915	.91652	.840
	2 - 5 år	2	2.20	4.20	3.2000	1.00000	1.41421	2.000
	Mer enn 5 år	1	4.20	4.20	4.2000	.	.	.