

Karl Håkon Johansen Hegna

Matematisk identitet i ulike kontekster

En kvantitativ studie av sykehusansattes matematiske identitet

Masteroppgave i Matematikdidaktikk 5.-10. trinn

Veileder: Eivind Kaspersen

November 2020

Karl Håkon Johansen Hegna

Matematisk identitet i ulike kontekster

En kvantitativ studie av sykehusansattes matematiske identitet

Masteroppgave i Matematikdidaktikk 5.-10. trinn
Veileder: Eivind Kaspersen
November 2020

Norges teknisk-naturvitenskapelige universitet
Fakultet for samfunns- og utdanningsvitenskap
Institutt for lærerutdanning



Kunnskap for en bedre verden

Sammendrag

Matematikk er et tema som i større eller mindre grad spiller en rolle i alle personers liv. Det er blant annet et sentralt fag i skolen og mange yrker krever en matematikkompetanse, både i utdanningen til yrket, men også i den daglige praksisen. Hvordan personer forholder seg til matematikk er forskjellig fra person til person. Noen opplever å få til matematikk, mens andre mener de selv ikke er i stand til det. Personers selvbylde i møte med matematikken kan kalles deres matematiske identitet. En persons matematiske identitet kan være viktig for om de mestrer faget og er dermed av interesse for det fagdidaktiske forskningsfeltet.

Denne studien har som formål å bidra til å finne svar på om en persons matematiske identitet er avhengig av konteksten personen befinner seg i eller om den er uendret hos personen dersom personen endrer konteksten den er i. Mitt bidrag til dette forskningsspørsmålet er å finne ut av hvordan en kan måle matematisk identitet i en konkret kontekst og hvordan strukturen til den matematiske identiteten er i den konteksten sammenlignet med andre. For å gjøre det har jeg først målt personer i en konkret kontekst og deretter sammenlignet målene med andre mål gjort i andre kontekster for å finne fellestrekk. I denne oppgaven benytter jeg meg i hovedsak av identitetsteori av Deaux (1993) og målingsteori av Andrich (1978) og Thurstone (1959).

For å samle inn et datamateriale har jeg valgt å bruke en kvantitativ metode, nærmere bestemt en spørreundersøkelse som ble delt ut til ansatte ved et sykehus.

Undersøkelsen inneholdt utsagn som de ansatte skulle respondere på ved hjelp av en skala på hvor ofte utsagnet stemte for deres egen praksis. Et eksempel på et slikt utsagn er «Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tankerekker». Videre ble datamaterialet validert og analysert ved hjelp av analyseverktøyet WINSTEPS som baserer seg på Rasch-modellen. Deretter ble resultatene fra de sykehusansatte sammenlignet med resultater hentet fra to lignende studier gjort på henholdsvis ungdomsskoleelever og universitetsstudenter.

Resultatene av denne studien er at det lar seg gjøre å måle matematisk identitet hos ansatte ved et sykehus ved å gjennomføre en spørreundersøkelse med påstander tilknyttet matematisk identitet. Videre utpekte noen av påstandene i spørreundersøkelsen seg til å bli forstått likt av deltakerne i de tre kontekstene. Samtidig ble noen av påstandene for ulikt vektlagt til at vi kan si at de oppfattes likt på tvers av kontekstene. Dette tyder på at vi ikke kan utelukke at det finnes en felles struktur ved matematisk identitet som er uavhengig av konteksten den blir målt i. Det tyder samtidig på at noen utsagn får en annen betydning dersom en spør en sykehusansatt enn det den får dersom en spør en skoleelev.

Stikkord: Matematisk identitet, Rasch-modellen, måling, kvantitativ metode

Abstract

The subject of mathematics plays a role, in varying degrees, in every person's life. For instance, it is considered one of the essential subjects in school and many careers require competence in mathematics, both in the education leading up to the job as well as in the day to day work. How a people relate themselves to mathematics differ from person to person. Some experience mastery over the subject, while others view themselves as "not able". We can call a person's self-image relating to working with mathematics their mathematical identity. A person's mathematical identity can be of importance for their ability to master the subject and is therefore of interest to the research field of didactics of mathematics.

The aim of this study is to investigate whether a person's mathematical identity is dependent on the context the person resides within or if the mathematical identity remains unaltered even though the person moves between different contexts. To do this I have measured people in a concrete context and compared the measures with measures done in two other contexts to see if there were any similarities. In this study I rely mainly on theory of identity by Deaux (1993), theory of measuring by Andrich (1978) and Thurstone (1959).

To gather the data used in this study I have chosen the quantitative method, more precisely a survey which were dealt out to the staff of a hospital. The survey contained statements which the staff were to respond to via a scale of how often the statements were true for their own day to day practice. An example of these statements is «Mathematical ideas I read or hear about makes me think thoughts of my own". The data was further validated and analysed using the analysing tool WINSTEPS which is based upon the Rasch Model. After the analysis were the results from the staff of the hospital compared with results gathered from two similar studies done in the contexts of lower secondary school and university.

The results of this study are that it is possible to measure the mathematical identity of staff at a Norwegian hospital by doing a survey containing statements concerning mathematical identity. Some of the statements in the survey were shown to be understood in the same way by the participants of all three contexts. There were also some statements that were understood too differently to say they mean the same in every context. This means that we cannot exclude there being a common structure to the mathematical identity which is indifferent to the context it is being measured in. It also means that for some statements the meaning will differ depending on what context you are measuring in.

Keywords: Mathematical identity, the Rasch Model, measuring, quantitative method

Forord

Det å skrive en masteroppgave midt inne i en pandemi og ferdigstille den ved siden av full jobb viste seg å være en tøffere oppgave enn først antatt. Som en som gjennom skoleløpet og utdanningen aldri har opplevd å ikke bli ferdig med arbeidet i tide var det en tøff avgjørelse for meg å utsette innleveringen av denne oppgaven. Det har blitt mange tunge stunder og det har gått med mye tankekraft til denne prosessen, men nå som oppgaven nærmer seg å være ferdigstilt er det flere jeg ønsker å benytte denne muligheten til å takke.

Jeg ønsker å takke min veileder, Eivind Kaspersen, for alle tilbakemeldingene, tipsene og de oppløftende ordene han har kommet med i arbeidet med denne oppgaven.

Jeg ønsker også å rette en stor takk til alle sykepleierne, legene og alle andre sykehusansatte som har deltatt i denne studien. Uten deres bidrag ville det ikke vært mulig for meg å skrive denne oppgaven.

Videre vil jeg takke familien min for all støtten jeg har fått i form av heiarop, forsyningspakker og oppmuntrende samtaler. Jeg vil særlig takke min mor, Siri Johansen, både for å ha bidratt med norsklærertilbakemeldinger med rødpenn og for å ha vært der for meg da jeg trengte henne aller mest.

Jeg vil også benytte anledningen til å takke mine medstudenter for samholdet og støtten både på lesesalen og utenom. Studietiden er noe jeg kommer til å se tilbake på med et stort smil og det er mye takket være dem.

Karl Håkon Johansen Hegna
Hammerfest, november 2020

Innholdsfortegnelse

| | |
|--|-----|
| Sammendrag | I |
| Abstract | II |
| Forord | III |
| Innholdsfortegnelse | IV |
| Figuroversikt | VI |
| Tabelloversikt | VI |
| Forkortelser | VI |
| 1. Innledning | 1 |
| 1.1 Hva er skolens formål? | 1 |
| 1.2 Matematikkfagets relevans for arbeidslivet | 1 |
| 1.3 Problemstillingen | 2 |
| 2. Teori | 4 |
| 2.1 Definisjon av matematisk identitet | 4 |
| 2.2 Filosofisk aspekt | 5 |
| 2.3 Måling | 6 |
| 2.3.1 Endimensjonalitet | 6 |
| 2.3.2 Additivitet | 7 |
| 2.3.3 Invarians | 7 |
| 2.3.4 Relasjonelle mål | 7 |
| 2.4 Valg av rammeverk/definisjon | 8 |
| 2.4.1 Oppsummering | 8 |
| 3. Metode | 9 |
| 3.1 Kvantitativ metode | 9 |
| 3.2 Måling | 9 |
| 3.2.1 Classical Test Theory (CTT) | 9 |
| 3.2.2 Item-Response Theory (IRT) | 10 |
| 3.2.3 Rasch-modellen | 10 |
| 3.2.4 Ulike former for Rasch-måling | 12 |
| 3.3 Innsamling av data | 13 |
| 3.4 Validering av data | 15 |
| 3.4.1 Wolfe og Smiths åtte aspekter for validering | 15 |
| 3.4.2 Content Validity | 15 |
| 3.4.3 Substantive Validity | 16 |
| 3.4.4 Structural Validity | 18 |
| 3.4.5 Generalizability Validity | 18 |
| 3.4.6 External Validity | 18 |
| 3.4.7 Consequential Validity | 19 |

| | |
|--|----|
| 3.4.8 Responsiveness Validity | 19 |
| 3.4.9 Interpretability Validity | 19 |
| 3.7 Hvordan fungerer Winsteps? | 20 |
| 3.8 Etiske betraktninger | 20 |
| 4. Resultat | 22 |
| 4.1 Innholdsaspektet | 22 |
| 4.1.1 Infit og outfitverdier | 22 |
| 4.1.2 ICC-kurver | 23 |
| 4.2 Det substansielle aspektet | 25 |
| 4.3 Det strukturelle aspektet | 27 |
| 4.4 Generaliserbarhetsaspektet | 29 |
| 4.5 Det eksterne aspektet | 29 |
| 4.6 Resultat av målingen | 33 |
| 5. Diskusjon | 35 |
| 5.1 Hva peker resultatene mot? | 35 |
| 5.2 Implikasjoner | 36 |
| 5.2.1 Didaktiske implikasjoner | 37 |
| 5.2.2 Forskningsmessige implikasjoner | 37 |
| 5.3 Mangler ved og forbedringer av studien | 38 |
| 5.3.1 Utsagnenes formuleringer | 38 |
| 5.3.2 De åtte aspektene | 39 |
| 5.4 Videre forskning | 40 |
| 5.4.1 Etisk utfordring knyttet til gjenbruk av datamateriale | 40 |
| 5.4.2 Veien videre | 40 |
| 6. Litteraturliste | 42 |
| Vedlegg | 44 |
| Vedlegg 1: Svarskjema A | 44 |
| Vedlegg 2: Svarskjema B | 46 |
| Vedlegg 3: Svarskjema C | 48 |
| Vedlegg 4: Svarskjema D | 50 |

Figuroversikt

| | |
|--|----|
| Figur 1 - Rasch-modellen grafisk (Wright & Stone, 1979). Den grønne kurven representerer personens sannsynlighet for mestring langs y-aksen og personens evne langs x-aksen. Den røde, stiplede linjen markerer hvor det er like stor sannsynlighet for at personen «løser» oppgaven som at den ikke gjør det. Den blå, stiplede linjen representerer like stor sannsynlighet for mestring som for ikke mestring. | 11 |
| Figur 2 – ICC-kurve til utsagn 4. | 24 |
| Figur 3 – ICC-kurve til utsagn 9. | 25 |
| Figur 4 - Grafisk fremstilling av svarkategoriens sannsynlighet for å bli valgt gitt personers mål. Hvert siffer (1-4) tilhører en svarkategori og danner hver sin kurve. Kurvene representerer sannsynligheten for at den tilhørende svarkategori blir valgt langs y-aksen, gitt personene som har respondert med svarkategoriens mål langs x-aksen. | 26 |
| Figur 5 – kart over utsagnenes vekting mot deres mål. Hvert utsagn er her representert som en stor eller liten bokstav. x-aksen er utsagnenes mål og y-aksen er deres vekting. Utsagnene er delt inn i tre grupper (clusters) etter om de er over den grønne linjen, under den røde eller i midten..... | 28 |
| Figur 6 – Venndiagram over hvilke utsagn som er signifikant lik i de ulike kontekstene.30 | |
| Figur 7 – Scatterplot av sykehusansattes egne mål (x-aksen) mot deres mål med målestokk fra universitetsstudentene (y-aksen). Vi kan se at samtlige personer er innefor konfidensintervallet. | 32 |
| Figur 8–Scatterplot av sykehusansattes egne mål (x-aksen) mot deres mål med målestokk fra ungdomsskoleelevene (y-aksen). Også i dette tilfellet er samtlige innenfor konfidensintervallet. | 33 |
| Figur 9 - Wright-kart av sykehusansattes sosiale mi. Hver syg/syj er en sykehusansatt. De ansatte er til venstre og påstandene er til høyre..... | 34 |
| Figur 10 - Wright-kart av personer og utsagn i studien. De som har lavere mål enn målestokken er markert i rød boks. | 39 |

Tabelloversikt

| | |
|--|----|
| Tabell 1 - De 20 påstandene (Kaspersen, 2015)..... | 13 |
| Tabell 2 – Standardisert restvarians for utsagnene uttrykt i eigenverdier. Den mest relevante eigenverdien tilhører «unexplained variance in 1st contrast», markert i rødt. 27 | |
| Tabell 3 - De åtte utsagnene med tilnærmet lik vanskegrad uavhengig av kontekst. | 30 |

Forkortelser

| | |
|-------------|--|
| MI | Matematisk Identitet |
| CTT | Classical Test Theory |
| ICC | Item Characteristics Curves |
| IRT | Item Response Theory |
| NESH | Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora |
| PCA | Principal Component Analysis |
| DIF-analyse | Differential Item Functioning Analysis |

1. Innledning

1.1 Hva er skolens formål?

Opplæringsloven skiller mellom kunnskaper og holdninger. Vi finner i formålsparagrafen at: «Elevane og lærlingane skal utvikle **kunnskap**, dugleik og **holdningar** for å kunne meistre liva sine og for å kunne delta i arbeid og fellesskap i samfunnet. Dei skal få utfalde skaparglede, engasjement og utforskartrøng.» (Opplæringslova – oppl, 1998). I matematikkfaget kan vi altså skille mellom konkrete fagkunnskaper og abstrakte, affektive holdninger og verdier. De konkrete fagkunnskapene handler for eksempel om det som har med regning, generalisering og manipulering av tall å gjøre. Et typisk eksempel på en slik fagkunnskap er Pythagoras læresetning om hvordan sidelengdene i en rettvinklet trekant forholder seg til hverandre. Den fagkunnskapsorienterte siden ved matematikkfaget skal dermed være kjent av både lærere, elever og andre utenom skolen som en essensiell del av faget. Den siden ved faget som kanskje færre tenker på, dersom de blir utfordret til å forklare hva matematikkfaget handler om, er den affektive siden. Begreper som motivasjon, holdninger, verdier, angst og matematisk identitet (MI) er eksempler på temaer som faller inn under den affektive siden ved matematikkfaget. Denne masteroppgaven har til hensikt å undersøke MI nærmere.

MI er et tema det forskes på i økende grad, dette til tross for at mange har ulike definisjoner og perspektiver på MI. Noen ser på identitet som noe en har, andre mener det er noe som blir til i en kontekst, i samspill med andre. Felles for noen av definisjonene er at det handler om en slags personkarakteristikk i arbeidet med faget matematikk. Det er ikke snakk om i hvor stor grad en mestrer fagets innhold, men hvordan en opplever faget. En måte å se MI på, er å se på det som at det i grove trekk handler om ett sett med karakteristikk som personer kjenner seg igjen i, i større eller mindre grad. Noen eksempler på slike karakteristikk kan være: «Å synes det er vanskelig å legge vekk et uløst problem» eller «Å like å diskutere matematikk på fritiden» (Darragh, 2016).

Alle har hatt matematikkfaget på skolen og har dermed antakeligvis gjort seg opp noen meninger og erfaringer omkring det å tilegne seg matematikkunnskap. Jeg har selv i løpet av lærerstudiet og i arbeid som lærer møtt elever som har et sterkt og positivt forhold til matematikk. De liker å fundere på matematiske problemer og synes det er meningsfullt å diskutere ulike fremgangsmåter og metoder, i tillegg til å oppleve en form for positivitet rundt det å arbeide med matematikk. De ser matematikkens relevans i verden, og for seg selv. Jeg har også møtt elever som ikke kjenner seg igjen i slike beskrivelser og opplever matematikkfaget kun som noe de må igjennom fordi skolen krever det. De foretrekker å få en metode servert i fanget og bruke den uten å gå nærmere inn på hvorfor den kan brukes, hvordan den fungerer og i hvilke tilfeller den er relevant. Noen kjenner seg kanskje også litt igjen i begge disse beskrivelsene og vil plassere seg et sted mellom de to ytterpunktene. Dersom vi generaliserer er det altså snakk om et spekter av MI; noen har sterk MI, noen har svak MI og noen er et sted imellom.

1.2 Matematikkfagets relevans for arbeidslivet

En forutsetning for at skolen skal være i stand til å oppfylle sitt ansvar om å utvikle kunnskap, skikkethet og holdninger hos elevene slik at de er i stand til å delta i

arbeidslivet, er at skolen vet hvilke kunnskaper og holdninger som er relevante og nødvendige for å klare seg i en arbeidshverdag. Dette er en krevende oppgave for skolen ettersom det finnes et mangfold av ulike yrker som hver har sine spesifikke kompetansekrav. I tillegg er verden vi lever i, i konstant forandring og nye arbeidsoppgaver og yrker dukker stadig opp, samtidig som andre blir automatisert eller overflødige. Det medfører at vi ikke alltid har kunnskap om hvilke yrker vi kommer til å ha i fremtiden, eller hvilke kunnskaper og ferdigheter som kreves av de fremtidige verdensborgerne. Skolen har uansett som oppgave å forberede elevene på disse fremtidige yrkene, og forsøker å løse denne utfordringen ved å kontinuerlig endre seg i takt med samfunnet. Dette gjør Utdanningsdirektoratet blant annet ved å stadig revidere og utforme nye læreplaner. Høsten 2020 tredde det i kraft en ny læreplan i matematikk for 1-9, trinn og vg1. Skolen gjør altså det den kan for å oppfylle de krav og det ansvaret den er pålagt om å forberede elever på et liv som arbeidstakere, ved å gi dem de kunnskapene og holdningene som de trenger for å lykkes. Et viktig spørsmål å stille i denne sammenhengen er: lykkes skolen med det arbeidet?

Noss, Hoyles og Pozzi (2002) har sett på forholdet mellom hvilke kunnskaper elever lærer på skolen og hvilke kunnskaper som blir tatt i bruk i arbeidslivet. Deres arbeid melder seg dermed i den pågående debatten om hvorvidt det er en korrelasjon mellom skolens opplæring av elevene, og hva elevene benytter i sitt arbeid som voksen arbeidstaker. Noss et al. (2002) har funnet ut at i realiteten opplever en del personer at den kunnskapen de opparbeidet seg på skolen ikke er relevant for den de benytter i sitt arbeid. Enkelte rapporterer at den kunnskapen de benytter seg av på arbeidet fikk de først etter at de begynte på arbeidsplassen. Dette kan sees på som et problem i seg selv, men det betyr ikke nødvendigvis at den tiden og det arbeidet elevene legger i skolearbeidet er forgjeves.

Den affektive siden ved matematikkfaget kan være et nyttig fokusområde hvis vi ønsker at skolen skal være relevant med tanke på elevenes fremtidige arbeid. Dersom det er slik at mange av kunnskapene vi lærer oss på skolen ikke er relevante for arbeidslivet, betyr ikke det nødvendigvis at holdningene til fag og arbeidsmetoder også er irrelevante. For eksempel kan det være vanskelig å se for seg hvordan alle yrker skal ha bruk for Pythagoras læresetning, men mindre vanskelig å se hvordan yrker kan trenge arbeidstakere som ikke ønsker å legge fra seg et problem før det er løst.

1.3 Problemstillingen

Selv om den affektive siden ved matematikkfaget både vektlegges av opplæringsloven og forskes på i økende grad, er det lite forskning på om den er overførbar mellom ulike kontekster. De fleste studiene som ser på matematikkfagets overførbarhet fra utdanning til arbeid, fokuserer på de konkrete fagkunnskapene og ikke i like stor grad på de affektive sidene, dermed heller ikke MI. Det gjør at det er usikkerhet knyttet til om MI er felles for skolen og arbeidslivet. I tillegg til at det er lite forskning på overføringen av affektive sider ved matematikkfaget, er de studiene som handler om MI i hovedsak kvalitative. Det melder seg dermed også et behov for flere kvantitative studier på temaet MI (Darragh, 2016).

Hvorvidt personer «tar med seg» sin matematiske identitet fra utdanning til jobb er et empirisk spørsmål. En forutsetning for å måle dette er (1) at MI kan måles i ulike kontekster, og (2) at det gir mening å sammenlikne MI i ulike kontekster. Det er dette jeg ønsker å undersøke i denne masteroppgaven.

For å gjøre dette har jeg først sett nærmere på hvilke holdninger til faget matematikk en gruppe ansatte i en konkret kontekst har, og om det lar seg gjøre å måle MI i den. Konteksten jeg har valgt å undersøke er et sykehus i Norge. Sykehuset har om lag 600 ansatte og befinner seg i en mellomstor norsk by. Blant de ansatte er det både leger, sykepleiere, spesialsykepleiere, ingeniører, vaktmestere, økonomer, kokker, renholdere, samt en rekke andre profesjoner som alle har sine egne arbeidsoppgaver og utfordringer hvor det kreves kunnskaper og ferdigheter, deriblant også kunnskaper innen matematikk og regneferdigheter, i større eller mindre grad.

Deretter har jeg forsøkt å sammenligne strukturen til MI til de sykehusansatte med ungdomsskoleelever og ingeniørstudenter for å se om strukturen på MI legger seg likt. Ungdomsskoleelevenes og ingeniørstudentenes datagrunnlag er tidligere innsamlet og validert av andre deltagere i dette forskningsprosjektet, respektivt Ytterhaug (2019) og Kaspersen et al. (2017).

Konkret består min problemstilling i denne oppgaven av to underproblemstillinger:

- 1. Hvordan kan en måle MI hos ansatte ved et sykehus?*
- 2. Er strukturen til den matematiske identiteten til sykehusansatte invariant sammenlignet med ungdomsskoleelvers og ingeniørstudenters?*

Dersom det viser seg at det er tilnærmet lik struktur ved MI i de ulike kontekstene, vil dette peke mot at holdninger og arbeidsmetoder vi tilegner oss i skolen er relevante også i arbeidslivet. Det kan dermed tyde på at det vil være viktig for skolen både å vektlegge og styrke elevenes holdninger til matematikkfaget, i tillegg til helt konkret faginnhold.

Dersom det viser seg at MI i de ulike kontekstene ikke har lik struktur, kan dette tyde på at det også for det affektive aspektet er diskutabelt om det er mulig å se en overførbarhet fra skole og utdanning til arbeidsliv. En slik konklusjon vil bidra til å sette spørsmålstegn ved skolen, for hvis hverken kunnskapen eller de affektive sidene vi får fra skolen er relevante og nødvendige for arbeidslivet, burde vi diskutere hvordan vi skal få innholdet i skolen til å bli det. Det må imidlertid i denne sammenhengen også påpekes at å forberede elever til arbeid ikke er den eneste funksjonen/oppgaven til skolen. Skolen har i tillegg til de ovennevnte oppgaver også oppgaver som å styrke elevene som individer, gi dem innsikt i kulturellt mangfold, lære dem kritisk tenkning, hvordan å handle etisk og miljøbevisst og gjøre dem i stand til å rå over sitt eget liv (Opplæringslova – oppl, 1998).

Oppsummert er hensikten min med denne oppgaven å bidra til å finne ut om MI er kontekstuavhengig eller ikke. For å gjøre det vil jeg først i kapittel 2, teorikapittelet, legge frem hvilke teorier og definisjoner som omkranser temaet MI, og hvilke av dem jeg bygger mine argumenter på. I tillegg til definisjonene og det teoretiske rammeverket, trenger jeg et datamateriale å sammenligne tidligere funn med. I kapittel 3 om metode forklarer jeg hvordan jeg har gått frem for å samle mitt datamateriale, samt hvordan jeg har validert det for å sikre at datamaterialet jeg har samlet, er anvendelig til det som er min hensikt. I kapittel 4 om analyse og resultat vil jeg legge frem de konkrete grepene jeg har gjort for å sikre reliabiliteten til mitt datagrunnlag, samt hvilke funn og resultater analysen gav. Til slutt vil jeg i kapittel 5, diskusjonskapittelet, diskutere hva funnene impliserer, hvilke endringer som burde gjøres dersom studien skal repliseres, samt hva som er konklusjonen til problemstillingen min.

2. Teori

Denne masteroppgaven har som hovedmål å måle og undersøke strukturen på MI til personer i ulike kontekster. De sentrale begrepene som brukes gjennomgående er *matematisk identitet, måling og kontekster*. Begrepene kan ha en hverdagslig betydning i tillegg til de mer teoretiske. Det er også flere ulike definisjoner blant forskere innen dette feltet (Darragh, 2016). Det er derfor viktig å presisere og definere hvilken betydning begrepene er ment å ha i denne teksten. Jeg vil i dette kapittelet derfor forklare begrepene og plassere min bruk av dem i det overhengende forskningsfeltet. Det vil jeg gjøre ved å først presentere forskningsfeltet MI og et utvalg av ulike definisjoner som oftest blir benyttet av forskere i feltet. For å kunne velge hvilket rammeverk jeg støtter meg til, vil jeg videre forklare hvilke filosofiske aspekter jeg støtter meg til. Deretter tar jeg for meg begrepet måling og fokuserer videre på måling av latente trekk ved mennesker som MI faller under. Til slutt forklarer jeg definisjonen jeg benytter og hvordan den plasserer seg i forskningsfeltet.

2.1 Definisjon av matematisk identitet

Det er som tidligere nevnt en pågående diskusjon omkring definisjonene til begrepet MI (Darragh, 2016). Et annet problem Darragh (2016) oppfatter fra sin litteraturstudie, er at bruken av definisjonene ikke er konsekvent. I flere av artiklene Darragh undersøkte var den teoretiske definisjonen av identitet som ble fremlagt, ikke forenelig med definisjonen av identitet resultatene ble drøftet i lys av. Det er altså ikke bare et problem at det er uenighet rundt hva som skal ligge i begrepet identitet, det er også et problem at det i flere artikler forekommer bruk av begrepet identitet med to ulike betydninger. Dette kan sees på som en følgefeil av identitetsbegrepets manglende entydige definisjon, og kan igjen føre til svakere forskning da resultatenes teoretiske forankring og resultatenes fortolkning ikke samsvarer. For å komme videre anser jeg det derfor som nødvendig at jeg først finner en definisjon som jeg konsekvent forholder meg til. Definisjonen må romme det jeg har til hensikt å undersøke, og må være praktisk med tanke på det formål jeg har, nemlig å måle MI.

Det er foretatt flere studier som ser på publikasjoner om temaet MI, deriblant Darragh (2016) og Graven og Heyd-Metzuyanin (2019). Darragh (2016) tok for seg 188 publikasjoner tilknyttet begrepet og forsøkte å analysere og kode det for å se om det var noen fellestrekk i alt fra metode, resultat og definisjoner. I sin artikkel kom hun frem til at noe av det som det var mest uenighet om i forskningsfeltet, var nettopp definisjonen av selve begrepet *identitet*. Diskusjonen dreide seg ikke om at det ikke eksisterte en definisjon, men snarere at det var flere. Disse ulike definisjonene argumenterte hun for at var ulike nok til at en vil ende opp med å beskrive forskjellige ting, avhengig av hvilken definisjon en benytter. Konsekvensen av dette er at det gjør det vanskelig å skille mellom hvilke artikler som omhandler det samme og hvilke som egentlig ser på to forskjellige ting, men som begge går under den samme fanen identitet.

Darragh (2016) reduserer alle de ulike definisjonene ned til fem kategorier. De fem kategoriene har jeg oversatt navnet på og de er: deltagende (particiative), narrative (narrative), diskursive (discursive), psykoanalytiske (psychoanalytical) og fremførende (performative). Videre reduserer hun disse ned til to hovedkategorier: (1) Identitet er noe en har eller tilegner seg eller (2) noe en gjør, altså en handling. Denne todelingen er i tråd med synspunktene til respektivt Erikson (1968) og Mead (1934) som blir av Graven

og Heyd-Metzuyanin (2019) beskrevet som de første bidragsyterne innen feltet identitet.

En annen måte å formulere denne tilsynelatende dikotomien er om identiteten til en person er avhengig av konteksten personen er i, eller om den forholder seg uendret dersom personen beveger seg fra en kontekst til en annen. En kontekst kan her forstås som en sosial struktur hvor det foregår en aktivitet hvor flere personer kommer sammen for å delta og bidra (Kaspersen et al., 2017). Eksempler på en slik kontekst kan være en skole, en arbeidsplass eller en fritidsforening. Det presiseres også av Darragh (2016) at enkelte forskeres definisjoner ikke passer inn i denne todelingen, deriblant Martin (2000), men at mange gjør det. En av konklusjonene til Darragh (2016) er at det kan virke som at tendensen blant forskere innenfor feltet er å se på identitet som noe en gjør, fremfor noe personen tilegner seg.

Darragh (2016) trekker frem de forskningsartiklene som oftest har blitt referert til. Blant de finner vi Sfard og Prusak (2005), Gee (2000), Martin (2000) og Wenger (1998). En måte å se identitet på er som noe som oppstår i samspillet mellom mennesker. Wenger (1998) snakker om «communities of practice». Sfard og Prusaks (2005) definisjon av matematikk er ofte referert til i artikler om MI (Graven & Heyd-Metzuyanin, 2019). De forklarer begrepet identitet som historier personer forteller om seg selv. Slike historier mener de skal være mulig å omfavne, betydningsfulle og tingliggjørende (reifying), det vil si ta noe abstrakt og gjøre det mer konkret. Martin (2000) snakker om identitet som et sett med antagelser eller overbevisninger (beliefs), altså at identitet er noe iboende i individet. Samtidig snakker han om at identitet dreier seg om en forhandling og en deltagelse i det han kaller «masternarratives». Martin (2000) kan altså sies å tilhøre begge leirene; de som ser på identitet som noe individet har innad i seg selv, og som noe som skapes i et deltakende felleskap.

En annen måte å snakke om identitet på er slik Deaux (1993) definerer begrepet. Hun deler begrepet identitet inn i sosial identitet og personlig identitet. Deaux (1993) forklarer sosiale identiteter som «de rollene eller de medlemskapskategoriene som en person anser som representative» (Deaux, 1993, s. 6). Personlige identiteter er «de trekkene og måtene å oppføre seg på som personen mener er selvbeskrivende, karakteristikk som typisk er koblet opp mot en eller flere av identitetskategoriene.» (Deaux, 1993, s. 6). Identitet er altså, i følge Deaux (1993), ikke enten personlig eller sosial. Den er heller ikke litt av begge. De to delene er knyttet sammen og en del av hverandre. Personlige identiteter er avhengige og tilknyttet de sosiale identitetene som igjen er skapt av individene i den sosiale settingen. Kanskje kan forfatterens syn best illustreres med allegorien om høna og egget. Det gir ikke mening å snakke om opphavet til den ene uten å ta med den andre.

2.2 Filosofisk aspekt

Når vi snakker om begrepet identitet har vi altså flere måter å definere det på, og det gjør at vi må foreta et valg. Det melder seg dermed et behov for en måte å velge et av de mange rammeverkene. Hvilket rammeverk eller hvilken teori en støtter seg til, er et filosofisk og epistemologisk valg da det avhenger av hvilken tilnærming en har til kunnskap, og hvordan vi mennesker tilegner oss den.

Det finnes ulike måter å se teorier og teoretiske rammeverk på. Et alternativ er å tenke at det finnes kun én absolutt «sannhet» og dermed kun én korrekt teori som forklarer denne sannheten. En annen måte å tilnærme seg valg av teori på er pragmatisme.

Pragmatismen ble utviklet som en retning innen filosofien av blant andre Charles Sander Peirce og John Dewey på slutten av 1800-tallet og begynnelsen av 1900-tallet. Et av punktene til pragmatismen er teoripluralisme. Det går ut på at en har en grunnholdning om at det ikke er én teori som er riktig og som beskriver virkeligheten perfekt, men at det finnes flere og at ulike teorier tilbyr ulike innfallsvinkler eller perspektiver til det som undersøkes. Spørsmålet går dermed over fra å være «Hvilken teori er riktig?» til å være «Hvilken teori egner seg til min studie?».

I denne oppgaven har jeg valgt en pragmatisk tilnærming til valg av rammeverk. For å velge hvilket av de tidligere nevnte rammeverkene som egner seg til min studie, trenger jeg noen kriterier. Disse kriteriene er avhengige av forskningsspørsmålet mitt, som består i å kunne måle MI. Jeg trenger derfor noen kriterier for hva en måling må bestå av, før jeg kan ta stilling til hvilken definisjon av identitet jeg skal støtte meg til i denne oppgaven.

2.3 Måling

Måling er noe vi mennesker har en lang historie med å benytte oss av og utvikle. I tidlige sivilisasjoner ble det blant annet nødvendig å kunne måle avstander for å fordele jord, beregne avlinger og bygge hus. De første måleenhetene som ble benyttet var ulike kroppsdelene, som for eksempel fot, tommer og favn. Disse måleenhetene har den fordel at de er lett tilgjengelige, noe som har ført til at de fortsatt brukes (Burton, 2011).

MI er et abstrakt og psykologisk aspekt ved et menneske, noe som medfører utfordringer når det kommer til å måle den, da vi med våre fysiske målemetoder ikke direkte har tilgang til den. Problemstillingen om hvorvidt det er mulig å måle noe «ikke-fysisk» skapte en diskusjon blant forskere på 1900-tallet. Noen forskere, deriblant Campbell, hevdet at måling var forbeholdt konkrete objekter. Andre forskere, som for eksempel Stevens, var mer åpne for ideen om at psykometri, altså måling av psykologiske aspekter, var noe som lot seg gjøre, på lik linje med all annen måling (Linacre, 2012). I denne oppgaven støtter jeg meg til tankegangen til Stevens om at psykometri lar seg gjøre og at jeg derfor i teorien skal kunne måle MI hos personer.

Thurstone var en av pionerene når det kom til å «standardisere» måling. Han mente at for at noe skulle kunne karakteriseres som måling, så måtte det oppfylle et sett med krav, uavhengig av om det som skulle måles var konkrete eller abstrakte objekter (Andrich, 1989).

2.3.1 Endimensjonalitet

Det første prinsippet Thurstone snakket om er endimensjonalitet. Et eksempel som kan forklare endimensjonalitet kan være å måle høyden til en person. Spørsmålet da er om «hvor lang» personen er. Personens kropp er i tre dimensjoner, likevel er det høyden vi er ute etter, altså hvor mange meterstokker personen tilbakelegger fra føttene til toppen av hodet. Vi går fra en multidimensjonal variabel, kroppens størrelse, til en endimensjonell måling, høyden. Meterstokkene kan sees på som en linje med inndelte punkter som alle refererer til ett bestemt lengdemål. Personens lengde blir dermed det punktet som er nærmest toppen av hodet dens. Thurstone mente at en slik punktindelt linje skal kunne eksistere for alle typer målinger (Andrich, 1989). MI er et latent trekk hos mennesker som trolig er flerdimensjonale. Til tross for dette må vi likevel selektere én dimensjon å måle, dersom vi følger Thurstones prinsipp.

På lik linje med at høydemåling krever en form for meterstokk, krever måling av MI et instrument. En persons kropp er selvsagt ikke en nøyaktig rett strek og derfor blir måling av dens høyde med en meterstokk et spørsmål om «nøyaktig nok». Vi godtar en viss feilmargin. Dette kan overføres til måling av MI ved at vi kan godta at variabelen ikke er fullstendig endimensjonal, så lenge den er endimensjonal nok til at målingene vi får er meningsfulle. Årsaken til at endimensjonalitet er et krav til måling, er at dersom det er flere dimensjoner som påvirker det endelige målet, kan vi hverken være sikre på hvilke dimensjoner som er målt, eller om målet er et resultat av en kombinasjon av flere dimensjoner. Vi kan dermed ikke hente ut noen meningsfull informasjon fra målingen.

2.3.2 Additivitet

Additivitet handler om å kunne legge sammen målingene uten at det oppstår selvmotsigelser. Dersom en person måles til å være 5 cm høyere enn en annen og 12 cm høyere enn en tredje, så skal differansen mellom den andre og den tredje summert med differansen mellom den andre og den første være lik differansen mellom den første og den tredje. Det samme skal også gjelde dersom vi byttet ut høydene med MI. Vi vil at målingene skal kunne sammenlignes, ikke bare etter rekkefølge, men hvor mye sterkere grad av MI en person har i forhold til en annen. Et eksempel på noe som bryter additivitet er hvis det krever 12 poeng å gå fra karakter 3 til 4, men 20 poeng for å gå fra karakter 4 til 5. Dermed er det ikke det samme å legge til et karakterhopp til karakter 3 som til karakter 4 (Andrich,1989).

2.3.3 Invarians

Invarians er prinsippet som sier at en måling skal kunne avgi nøyaktige mål uavhengig av andre faktorer enn det som måles. For å konkretisere kan vi se for oss at vi har en meterstokk som er laget av et materiale som er svært påvirkelig av temperatur. Dersom det er kaldt trekker den seg sammen og hvis den blir varm utvider den seg. Høyden til en person som ble målt med en slik meterstokk ville dermed variert ut fra hvilken temperatur det var i rommet, selv om personens høyde ville vært tilnærmet lik gjennom hele prosessen. En slik meterstokk vil derfor ikke være produktiv å måle høyde med, fordi den motstrider prinsippet om invarians som altså handler om at instrumentet skal være standhaftig og ikke endre seg fra måling til måling.

Dette prinsippet skal også gjelde ved måling av latente trekk eller variabler hos mennesker. En metode som ofte brukes i målinger av slike variabler, er å gi noen påstander til de som skal måles, og videre analysere deres respons til disse påstandene. Måten invarians blir tilknyttet en slik måling på er om disse personene er påvirket av andre faktorer når de responderer. Konteksten de er i kan for eksempel ha innvirkning på hvordan de responderer. Et konkret eksempel jeg har sett fra egen praksis på dette er elever som, under en trivselsundersøkelse for klassen, unnlater å skrive hvem som er mobbere i klassen fordi de er redde for at mobberne skal se at de blir utpekt, ettersom undersøkelsen ble foretatt med alle elevene i samme klasserom. Et annet punkt som vektlegges er at de påstandene som brukes for å måle, ikke kan være påvirket av hvordan de som lager dem oppfatter måleobjektet eller påstandene (Andrich, 1989).

2.3.4 Relasjonelle mål

Det fjerde prinsippet handler om å relatere målingen til en struktur. La oss si at vi benyttet oss av et måleinstrument og målte en person. Målingen viste at personen var 1,8. Hva vil det si? 1,8 kg? 1,8 meter? Det gir ikke mening å snakke om et mål uten å relatere det til en struktur som det er tydelig definert hvor det vilkårlige nullpunktet er, og hvor mye en vilkårlig enhet i strukturen er (Bond & Fox, 2015). Jeg vil her vektlegge

ordet «vilkårlig» ettersom det ikke egentlig har noe å si hvor vi sier nullpunktet er eller hvor mye en enhet er, det viktigste er at det spesifiseres. Vi kunne i teorien satt en centimeter til å være det dobbelte av det den er nå og at vi begynner å måle høyde fra fem centimeter under bakken, vi må bare være enige om at det er slik vi måler høyde og forholde oss til disse definisjonene.

2.4 Valg av rammeverk/definisjon

Thurstones prinsipper for måling har blitt beskrevet som grunnleggende målingsteori. Jeg vil derfor argumentere for at Thurstones prinsipper for måling egner seg godt som krav til valg av teoretisk rammeverk og jeg velger derfor å støtte meg til disse prinsippene i denne masteroppgaven.

Blant de tidligere nevnte definisjonene av identitet kan det virke som at Deaux' (1993) er et mulig valg. Dette begrunner jeg med at prinsippet om relasjonelle mål, at målingen skal kunne relateres til en struktur, er forenelig med Deaux' (1993) definisjon. Jeg sikter her til hvordan personlig identitet relateres til den sosiale identiteten. En slik måte å definere MI er også forenelig med Kaspersen et al. (2017) og Ytterhaug (2019) som baserer seg på kultur-historisk aktivitetsteori (KHAT). Dette er ønskelig da det er deres datamateriale jeg benytter meg av i denne oppgaven for å sammenligne med mitt eget materiale. Valg av rammeverk faller derfor på Deaux (1993). Jeg vil samtidig poengtere at dette er et pragmatisk valg og at jeg derfor ikke anser eller argumenterer for at andre rammeverk og teorier er ugyldige eller upassende.

MI vil jeg, med Deaux' (1993) definisjon av personlig og sosial identitet som bakgrunn, dermed definere som: Et samlebegrep som omhandler relasjonen mellom sosial matematisk identitet og personlig matematisk identitet. Den sosiale matematiske identiteten kan forstås som et sett med karakteristikk som er strukturert etter hvordan personene i en valgt kontekst har respondert til dem. Den personlige matematiske identiteten er hvordan en person posisjonerer seg i forhold til den sosiale matematiske identiteten.

2.4.1 Oppsummering

I dette kapitlet har jeg presentert forskningsfeltet MI og hvordan det innad i feltet er flere definisjoner av identitet, noe som gjorde det nødvendig å avklare hvordan begrepet skal forstås i denne oppgaven. Ved hjelp av en pragmatisk tilnærming til forskningsteori, nærmere bestemt teoripluralisme, har jeg lagt Thurstones prinsipper om måling, til grunn for valg av et rammeverk som er egnet til min studie. Dette rammeverket baserer seg på Deaux (1993) som definerer identitet som en relasjon mellom sosial og personlig identitet. Urustet med denne definisjonen av hva MI er, vil jeg i det neste kapitlet presentere de metodiske valgene jeg har tatt for å samle og validere et datamateriale som skal danne grunnlaget for at jeg skal kunne svare på problemstillingene mine.

3. Metode

I forrige kapittel argumenterte jeg for mitt valg av definisjon på MI. I dette kapittelet vil jeg diskutere og argumentere for mine metodologiske valg for å måle MI. Disse valgene består i korte trekk av hvilken målingsmetode jeg pragmatisk har valgt, hvordan jeg har gått frem for å samle datamaterialet mitt og hvordan jeg har gått frem for å analysere det. Uavhengig av metode når en bedriver forskning, særlig i tilfeller hvor mennesker er involvert, følger det med etiske betraktninger. Jeg vil derfor mot slutten av dette kapittelet også legge frem hvilke etiske betraktninger jeg har gjort i mitt arbeid med denne studien.

3.1 Kvantitativ metode

I review-artikkelen til Darragh (2016) kommer hun frem til at forskningen innen MI i hovedsak består av kvalitative studier. Disse studiene baserer seg blant annet på intervjuer, observasjon og analyser av videoer og dokumenter. Av de totalt 188 artiklene Darragh har med i sin artikkel er det kun 17 (9%) som baserer seg på empiri fra mer enn 100 deltakere. Dersom vi ønsker oss et så nyansert bilde av MI som mulig, er det viktig å belyse saken fra flere ulike vinkler for å sikre oss mot at perspektiver blir utelatt eller oversett. Det kan dermed argumenteres for at flere kvantitative undersøkelser er ønskelig i forskningen på MI. Dette er et av argumentene mine for at jeg har valgt å gjøre en kvantitativ undersøkelse.

Et annet argumentet er at for å gjøre den type analyse jeg ønsker, kreves det måledata. Måledata er et eksempel på et kvantitativt datamateriale, som uttrykkes som tall eller andre mengdetermer. I denne studien benytter jeg meg av både måledata jeg selv har samlet inn, og måledata fra to andre studier. Dette legger dermed føringer for valg av forskningsmetode. Selv om denne studien i hovedsak bygger på den kvantitative metoden, har jeg også benyttet kvalitativ metode noe i forbindelse med arbeidet med validitet og reliabilitet.

3.2 Måling

De fleste har en oppfatning av hva det vil si å måle noe. De fleste har for eksempel brukt en linjal for å måle en sidelengde av et kvadrat på skolen, og mange har vært i butikken og målt hvor mange hektogram potet de skal betale for ved hjelp av en vekt. Vi kan se fra de to eksemplene at for å gjøre målinger trenger vi et instrument, et redskap vi kan lese av noen verdier som relaterer til den spesifikke egenskapen ved det vi måler. Måling er relativt enkelt å gjøre i eksempelet med linjalen, hvor vi måler en lengde, ettersom både linjalen og sidelengden er fysiske objekter vi kan helt konkret håndtere. Det oppstår derimot en utfordring når det vi ønsker å måle ikke er en håndfast ting, men noe abstrakt som vi ikke har direkte tilgang til. Det kan for eksempel være en følelse, en opplevelse eller en holdning. Denne utfordringen har flere forskere arbeidet med, deriblant Thurstone (1959) som la frem sine prinsipper om hva måling er.

3.2.1 Classical Test Theory (CTT)

En teori som ofte har vært brukt for psykometri, er CTT (Bond & Fox, 2015). Hovedprinsippet bak denne teorien er at en person har et mål når det kommer til et hvilken som helst latent trekk, for eksempel MI. Dette målet kan kalles det sanne målet til personen, derfor kalles denne teorien også True Score Theory (TST). Det er personens

sanne mål (true score, T) vi er ute etter å finne, men ettersom det er noe som vi ikke direkte har tilgang til, kan vi ikke vite hva det er helt uten videre. Vi er derfor nødt til å støtte oss på et annet mål, det observerbare målet (observed score, X), for å tilnærme oss det sanne målet. Videre sier teorien at ulikheten mellom det sanne og det observerte målet skyldes en feil (error, e) som oppstår tilfeldig i målinger. Forholdet mellom de to målene kan dermed beskrives som $X=T+e$ (Bond & Fox, 2015).

En antagelse denne teorien gjør er å anta at alle oppgavene/påstandene i et måleinstrument er like enkle/vanskelige. Et «riktig» svar i en slik test gir dermed like mye poeng til den totale scoren, uavhengig av hvilken oppgave svaret tilhørte. Det blir dermed et fokus på testen som en helhet, og målet til personen som tar testen avhenger da i hovedsak av hvor stor andel av testen personen greide å svare «korrekt» på. Årsaken til at jeg bruker anførselstegn på riktig og korrekt er at det ikke nødvendigvis er rette eller gale svar, bare mer eller mindre tilstedeværelse av den latente egenskapen som blir forsøkt målt. Dette gjelder også for andre målingsteorier som jeg redegjør for i dette kapittelet.

3.2.2 Item-Response Theory (IRT)

IRT er et paraplybegrep for flere av de mer moderne modellene for å drive med psykometri, derfor går det også under navnet modern test theory (Bond & Fox, 2015). I denne gruppen finner vi de tre modellene: ett-, to- og tre-parametermodellene. Felles for dem er at de fokuserer på oppgavene (the items) i en undersøkelse, i motsetning til CTT som har fokuset på den totale undersøkelsen. Det som skiller de tre modellene er at to- og tre-parametermodellene har et parameter som diskriminerer mellom oppgavene i testen, som ett-parametermodellen ikke har, og i tillegg har tre-parametermodellen et parameter som tar høyde for gjetning (Bond & Fox, 2015).

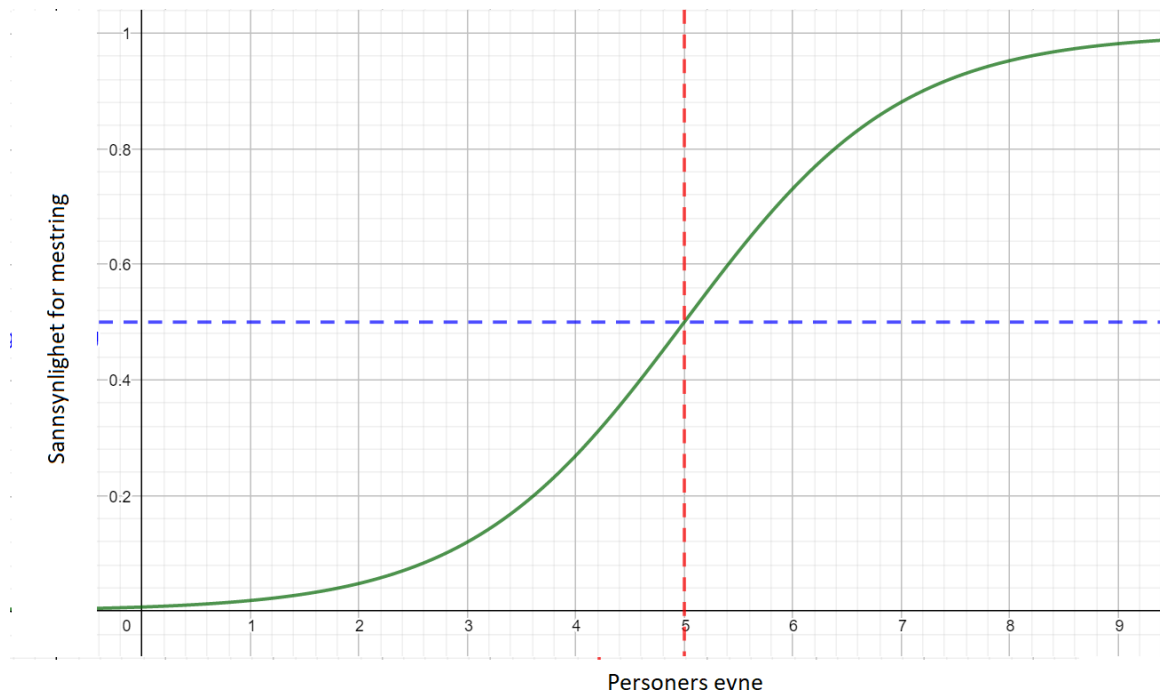
For å finne ut hvilken av teoriene som passer best for denne studien, er det viktig å se på ulikhetene i teoriene. Bond og Fox (2015) argumenterer for at IRT er et bedre valg enn CTT ettersom CTTs fokus på hele undersøkelsen gjør at den antar at alle oppgavene i studien er like vanskelige, en antagelse som sjeldent stemmer overens med virkeligheten. Et annet punkt som skiller CTT og IRT er konseptet om endimensjonalitet som kun modellene innen IRT har. Endimensjonalitet er også et av prinsippene til Thurstone som jeg i min studie anser som grunnleggende for måling. Utfra dette velger jeg å støtte meg på argumentet til Bond og Fox (2015) og velger derfor bort CTT.

Videre argumenterer Bond og Fox (2015) for at det er en vesensforskjell ved en av de tre modellene som gjør den særdeles bedre egnet til formålet om psykisk måling. Modellen det er snakk om er ett-parametermodellen som er bedre kjent som Rasch-modellen (Bond & Fox, 2015).

3.2.3 Rasch-modellen

En av de mest anerkjente modellene for psykometri er Rasch-modellen. Rasch-modellen baserer seg på at dersom en skal måle et latent trekk ved et menneske trenger en et instrument som består av oppgaver som personen skal respondere på. Hva personen responderer, for eksempel om den er i stand til å løse en oppgave eller ikke, er i hovedsak avhengig av to ting: personens evne, det vil si hvor mye av det latente trekket som vi ønsker å måle personen innehar. Det andre er oppgavens vanskegrad, med andre ord hvor mye av det latente trekket en person trenger for å få den til (Wright & Stone, 1979). For å eksemplifisere kan vi se for oss at en elev tar en matematikktest i algebra. Det latente trekket som måles er algebrakunnskapene til eleven. Om eleven klarer å svare riktig på en oppgave avhenger av hvor vanskelig oppgaven er, samt elevens evner

innen temaet algebra. Dette kan vi modellere med sannsynlighet. Dersom eleven ikke har lært om algebra vil han eller hun trolig få 0 poeng på slike oppgaver i de fleste tilfeller. Det kan hende at eleven gjetter og får riktig, men sannsynligheten er tilnærmet lik 0. På samme måte vil sannsynligheten være tilnærmet lik 1 for at eleven mestrer oppgaven dersom eleven mestrer alt ved algebra på det gjeldende nivået. Vi kan imidlertid ikke si at den er nøyaktig 1 da selv de aller mest høytpresterende av og til gjør en feil. Denne sammenhengen mellom evne, vanskegrad og sannsynlighet kan uttrykkes grafisk som en kurve. På midten av denne kurven kan vi lese av at dersom oppgavens vanskegrad og elevens evne er like stor, er det like stor sannsynlighet for at eleven skal greie oppgaven som at den ikke gjør det.



FIGUR 1 - RASCH-MODELLEN GRAFISK (WRIGHT & STONE, 1979). DEN GRØNNE KURVEN REPRERENTERER PERSONENS SANNSYNLIGHET FOR MESTRING LANGS Y-AKSEN OG PERSONENS EVNE LANGS X-AKSEN. DEN RØDE, STIPLEDE LINJEN MARKERER HVOR DET ER LIKE STOR SANNSYNLIGHET FOR AT PERSONEN «LØSER» OPPGAVEN SOM AT DEN IKKE GJØR DET. DEN BLÅ, STIPLEDE LINJEN REPRERENTERER LIKE STOR SANNSYNLIGHET FOR MESTRING SOM FOR IKKE MESTRING.

For å kunne gjøre utregninger trenger vi å omformulere tankegangen bak teorien til matematisk språk. Den vanligste måten å formulere det er slik: En person v har sin evne gitt ved β_v . En oppgave l har sin vanskegrad gitt ved δ_l . Ettersom hensikten med hele denne modellen er måling som skal følge prinsippene til Thurstone, blant annet endimensjonalitet og additivitet, vil vi at både β_v og δ_l skal ligge på samme, felles linje. Forholdet mellom personens evne og oppgavens vanskegrad uttrykkes dermed mest lettvis og naturlig, i følge Wright og Stone (1979, s. 12) av differansen mellom dem, altså $\beta_v - \delta_l$. Videre vil vi at differansen skal uttrykkes som en sannsynlighet. Dette for å modellere hvordan det kan oppstå resultater som bryter med den teoretiske baktanken. I følge teorien forventer vi for eksempel at en person med lav evne aldri vil klare en oppgave med høy vanskegrad, men i virkeligheten hender det noen ganger at personer har flaks og får til en vanskelig oppgave. Sannsynlighet uttrykkes som et tall mellom 0 og 1 og vi vil dermed at differansen alltid skal uttrykkes som et tall mellom 0 og 1. Det

gjøres ved å først gjøre differansen til eksponenten av e ($e^{\beta v - \delta l}$) for at resultatet alltid vil være mellom 0 og positiv uendelig. Videre deler en på $1 + e^{\beta v - \delta l}$ for å få verdier mellom 0 og 1.

Sannsynligheten for et positivt/riktig svar er dermed gitt ved uttrykket:

$$P \{X_{vl} = 1 \mid \beta_v, \delta_l\} = \frac{e^{\beta v - \delta l}}{1 + e^{\beta v - \delta l}}$$

Det er dette uttrykket som er Rasch-modellen i sin matematiske form (Wright & Stone, 1979).

Rasch-modellen følger Thurstones prinsipper. I forskning som omhandler måling av latente trekk hos mennesker, er det i moderne tid omtrent lik fordeling av de som benytter seg av IRT og de som benytter seg av Rasch-modellen. Resultatene av de to er også nokså like. Det er derfor ingen klar enighet blant forskere om den ene er å foretrekke fremfor den andre. De to modellene skiller lag når det gjelder sin filosofiske holdning til grunnleggende målingsteori, Thurstones prinsipper. IRT anser disse prinsippene som antagelser, mens Rasch-modellen anser dem som krav for å kunne gjøre Rasch-analyser (Bond & Fox, 2015). Ettersom jeg i denne oppgaven har brukt Thurstones prinsipper i valg av rammeverk har jeg allerede vektlagt dem, noe som gjør det naturlig for meg å anse dem som krav. Dette gjorde at mitt valg, med hensyn til hvilken målingsmodell jeg skulle benytte, ble Rasch-modellen.

3.2.4 Ulike former for Rasch-måling

Innenfor Rasch-modellen finnes det flere undermodeller som avhenger av hvilke typer responser en ønsker å samle inn. Originalt var modellen utformet til å være dikotom (dichotomous). Det vil si at respondenten blir stilt ovenfor en oppgave (item). Denne oppgaven kan for eksempel være å respondere på en påstand, et utsagn eller et spørsmål. Svaret respondenten avgir på denne oppgaven betraktes enten som riktig eller galt, eventuelt som tilstedeværelse eller fravær av det som måles. Dette svaret kan kodes som 0 og 1. Det må her vektlegges at 1 betyr mer av det som måles enn 0. Det kan ikke bare være et annet svar, som for eksempel gutt eller jente. Det er også muligheter for at det er mer enn to svaralternativer, men det vil da kodes som at et svar enten er 0 eller 1, altså helt riktig eller helt galt (Bond & Fox, 2015).

En annen modell innen Rasch er den som kalles Rating Scale Model (RSM). Den skiller seg fra den dikotome modellen ved at den har flere svaralternativer, det er en polytom modell hvor svarene har ulike grader av tilstedeværelse av det som måles. Disse svaralternativene er rangert etter en skala. Likerts skala er et eksempel på dette, hvor skalaen rangerer i hvor stor grad personen som svarer er enig eller uenig i utsagnet (Bond & Fox, 2015). I min studie valgte jeg RSM da det gjorde det enklest å sammenligne med de tidligere resultatene fra ungdomsskole og universitet, som også benyttet RSM.

En utfordring en møter på hvis en benytter seg av Likerts skala er at det ikke er et tydelig og definert sprang fra den ene svarkategorien til den neste. Det blir dermed opp til hver enkelt person som blir stilt ovenfor spørsmålet å definere hva som skiller svarkategoriene. Sagt med andre ord er det subjektivt. Den tradisjonelle metoden for å analysere Likerts skala ser bort fra at det er subjektivt, for å raskt kunne skape noen håndfaste resultater. Det mener Bond og Fox (2015) ikke er godt nok fordi en ikke vet om det er lik avstand mellom svarkategoriene og heller ikke mellom oppgavene. Dermed mener de at Likerts skala gjør en antagelse. Spørsmål som oppstår i denne

sammenhengen er: Er «noen ganger» nøyaktig like mye mer enn «sjeldent» som «ofte» er mer enn «noen ganger»? Vil «ofte» antyde like stor tilstedeværelse av det vi måler hos alle oppgaver? Antagelsen kan føre til at det blir usikkert om prinsippet om additivitet blir ivaretatt. Ved å bruke Rasch-modellen unngår en slike usikkerheter, ettersom Rasch-modellen ikke gjør denne antagelsen (Bond & Fox, 2015).

3.3 Innsamling av data

For å undersøke problemstillingen min var det nødvendig at instrumentet for å måle MI var likt med de tidligere studiene. Valget om å bruke de 20 påstandene som ble brukt i dem var dermed pragmatisk og nødvendig for å kunne sammenligne resultatene. De 20 påstandene hadde tidligere blitt validert av Kaspersen (2015). Likevel var det ikke sikkert at de ville gjelde når konteksten gikk fra å bli brukt på ungdomskoleelever/ingeniørstudenter til å måle sykehusansatte. Påstandene måtte derfor også valideres i den nye konteksten.

TABELL 1 - DE 20 PÅSTANDENE (KASPERSEN, 2015).

| | |
|-----|---|
| 1. | Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobb legger opp til. |
| 2. | Når jeg lærer en ny metode, bruker jeg tid på å se om jeg kan finne en bedre metode. |
| 3. | Når jeg lærer en ny metode, prøver jeg å finne situasjoner hvor denne ikke virker. |
| 4. | Jeg har problemer med å legge fra meg matematiske oppgaver. |
| 5. | Dersom jeg har glemt en formel/metode, prøver jeg å utlede den selv. |
| 6. | Jeg blir engasjert når noen starter en matematisk diskusjon. |
| 7. | Når jeg lærer noe nytt, stiller jeg meg selv egne spørsmål som jeg jobber med. |
| 8. | Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tankerekker. |
| 9. | Når jeg lærer en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gjøre. |
| 10. | Hvis jeg prøver på en metode som ikke fører frem, bruker jeg tid på å finne ut hvorfor denne ikke virker. |
| 11. | Når jeg lærer en ny metode/algoritme, prøver jeg å finne ut hvorfor den virker. |
| 12. | Når jeg kommer over et matematisk bevis/forklaring, studerer jeg det til det gir mening. |
| 13. | Når jeg møter et matematisk problem, tenker jeg over om det finnes flere måter å løse oppgaven på. |
| 14. | Når jeg jobber med et matematisk problem hopper jeg mellom ulike strategier. |
| 15. | Når jeg lærer noe nytt, fører det til at det er flere ting jeg ønsker å finne ut. |
| 16. | Når jeg jobber med en oppgave, stopper jeg opp underveis og reflekterer over hva jeg gjør. |
| 17. | Hvis jeg står fast, prøver jeg å visualisere problemet. |
| 18. | Jeg kan forklare hvorfor løsningen min er rett. |

| | |
|-----|---|
| 19. | Jeg prøver å koble det jeg lærer opp mot det jeg vet fra før. |
| 20. | Jeg fortsetter å prøve meg frem selv om jeg ikke får det til med en gang. |

Rekkefølgen på disse påstandene stammer fra den tidligere undersøkelsen hvor påstandene ble brukt på ingeniørstudenter. Der ble utsagn 1 ansett som det vanskeligste utsagnet å si seg enig i, og deretter ble utsagnene ansett som gradvis lettere helt til utsagn 20 som dermed ble ansett som det enkleste. Videre i denne oppgaven vil nummereringen av utsagnene i denne tabellen bli benyttet til å referere til utsagnene, selv om de viser seg å plassere seg annerledes i andre kontekster. I skjemaene som ble delt ut til de sykehusansatte i min studie, var utsagnene randomisert i fire ulike skjemaer, A-D, for å unngå at rekkefølgen utsagnene ble lest i skulle ha en innvirkning på besvarelsene, og dermed det endelige resultatet av undersøkelsen.

Svarkategoriene var aldri/nesten aldri, noen ganger, ofte og alltid/nesten alltid. Antall svarkategorier har en del å si for hvilken Rasch-analyse en benytter seg av. Valget om å ha fire ulike svarkategorier ble igjen et pragmatisk valg for å kunne sammenligne med tidligere studier i dette prosjektet. Et annet argument er elimineringsmetoden. En kan i utgangspunktet velge mellom et uendelig antall svarkategorier. Grunnen til at akkurat fire ble valgt er for det første at det ikke er produktivt å be personer å velge mellom for mange svarkategorier. I et foredrag av Benjamin D. Wright eksemplifiserer han dette med å spørre en person om å karakterisere en appelsins søthet med en skala fra 1 til 100. Dette ville resultere i noe som ikke ga særlig mening, for hva er det som skiller en appelsin med søthet 83 fra en appelsin med søthet 84? For mange svaralternativer kan dermed gjøre hver enkel respons mindre nøyaktig. Wright mener det ikke er produktivt å ha mer enn rundt fem alternativer. Videre kan vi argumentere for at et oddetall antall svarkategorier er ugunstig. Argumentet er at ved et oddetall antall svarkategorier vil det være et midtpunkt som er like langt fra det ene ytterpunktet som til det andre. Dette kan medføre unøyaktighet ettersom enkelte personer kan velge det midterste svaralternativet og mene at de ikke vil/greier å ta stilling til karakteristikken/spørsmålet når det midterste svaret er ment å være en bestemt grad av enighet/riktighet til karakteristikken/spørsmålet (John Michael Linacre, 2017).

Ved å eliminere antall svarkategorier over 5 samt alle oddetall sitter vi igjen med å ha to eller fire. I min studie har jeg valgt å ha fire svarkategorier da jeg anser det for å gi et mer nyansert og dermed mer eksakt bilde av hvor en person selv opplever å ligge på skalaen, samtidig som at det er mulig å klart distingvere mellom dem.

Innsamlingen av mitt datamateriale ble gjort ved først å utforme de 20 påstandene til et skjema hvor hver påstand skulle besvares med en av de fire svarkategoriene (vedlegg 1-4). I tillegg var det et femte svaralternativ som var «vet ikke», men i analysen ble slike responser ansett som manglende svar, med andre ord det samme som å ikke svare i det hele tatt.

Skjemaet ble printet ut og delt ut til ulike avdelinger, sengeposter og andre arbeidsområder. Å nå ut til respondenter viste seg å være mer utfordrende på et sykehus enn hva tilfelle kan være på en skole eller et universitet, da ikke alle de potensielle respondentene var samlet på ett sted til samme tid som følge av vaktordninger. Denne utfordringen ble løst ved å legge igjen en bunke med skjemaer slik at de som hadde vakt senere på dagen eller neste dag også fikk muligheten til å delta. Instruksen til deltagerer var å fylle ut skjemaene i ledig tid og samle dem i en folder

som ble hentet av meg etter en til to dager. Dette var mulig da informasjonen på skjemaet hverken var personidentifiserende eller av en sensitiv natur.

3.4 Validering av data

I en kvantitativ undersøkelse med datamateriale hentet fra flerfoldige respondenter er det ikke til å unngå at det oppstår små feil. Det kan være respondenter som ikke leser påstander godt nok til å kunne ta ordentlig stilling til dem. Det kan være avlesningsfeil av responser. Det kan også hende at det oppstår regnefeil eller avlesningsfeil i analyseringen av datagrunnlaget. Det er dermed relevant å stille spørsmålet: Kan vi være sikre nok på at dette datamaterialet stemmer og er representativt for det vi ønsker å undersøke? Det er i denne sammenhengen begrepet validitet blir brukt. Reliabiliteten til et datamateriale er høy/god dersom de tilfeldige feilene som oppstår nært sagt i et hvert datamateriale er små og få. Reliabiliteten er derfor et uttrykk for i hvor stor grad en undersøkelses resultater er til å stole på, eller sagt med andre ord hvor nøyaktig er undersøkelsen utført (Wolfe & Smith, 2007). I konteksten av å måle noe, som er det denne oppgaven har som mål, er måleinstrumentets nøyaktighet en sentral del av reliabiliteten. Om et instrument er til å stole på, er avhengig av hva formålet med undersøkelsen er, hvilket rammeverk en velger å støtte seg til og hva som anses som nøyaktig nok.

3.4.1 Wolfe og Smiths åtte aspekter for validering

I utviklingen av et instrument for å gjøre vitenskapelige undersøkelser, som for eksempel måling, er validering viktig (Wolfe & Smith, 2007). Vitenskapen er avhengig av undersøkelser og de resultater som kommer fra dem. For at vi skal kunne stole på vitenskapelige resultater stilles det strenge krav om pålitelighet. En må kunne legge frem gode valididetsargumenter for at metoden en har benyttet er relevant og gyldig. Slike argumenter krever naturligvis også bevis for at det de påstår stemmer. Wolfe og Smith (2007) legger i sin artikkel frem åtte aspekter for hvordan en analyse kan gjennomføres for å skaffe bevis som kan støtte argumenter om validitet, særlig rettet mot undersøkelser som benytter seg av Rasch. Jeg anser derfor denne artikkelen som et passende rammeverk for å analysere mitt datamateriale. Jeg skal nå gå nærmere inn på de ulike aspektene og forklare hva jeg har gjort for å sikre validiteten i min undersøkelse gjennom disse åtte. I analysekapittelet, kapittel 4, vil de konkrete resultatene av denne analysen bli lagt frem.

3.4.2 Content Validity

Det første aspektet Wolfe og Smith legger frem kan kalles innholdsvaliditet (content validity). Dette aspektet har, som tittelen antyder, fokus på innholdet i undersøkelsen. Spørsmål en kan stille er om innholdet er relevant for det som faktisk skal måles. Representerer dette datamaterialet det vi er ute etter å undersøke? Dette aspektet er sentralt for den første av mine problemstillinger: Hvordan kan en måle MI hos ansatte ved et sykehus?

Wolfe og Smith (2007) anbefaler, i tilfeller hvor en benytter Rasch, å se på den tekniske kvaliteten til utsagnene, nærmere bestemt to aspekter ved dem: item-measure correlation og mean-squared fit indices. Det første handler om at det skal være samsvar mellom hva som ønskes målt og hva påstandene i undersøkelsen handler om. Item-measure correlation skal i følge Wolfe og Smith (2007) være positiv, da det antyder at oppgaven (the item) har en i snitt lik poengsum med de andre, noe som er ønskelig ettersom vi vil at oppgavene våre skal handle om det samme. Dersom det oppstår en

negativ item-measure correlation, som vil si at oppgaven peker i motsatt retning av de andre oppgavene, kan det være et tegn på at det er en feil i kodingen til analyseprogrammet som gjør at besvarelser leses av som det motsatte av hva det skal leses. Et eksempel på en slik påstand kan være om en ville måle selvstendighet og hadde en påstand som var: «Jeg liker å få hjelp med ting». Selv om et ja i slike undersøkelser ofte blir kodet som «1» ville det i en undersøkelse om selvstendighet bli feil, ettersom et ja vil være et tegn på lav selvstendighet.

Dersom item-measure correlation er tilnærmet 0 kan en anslå at oppgaven enten ikke handler om det samme som de andre, eller at den er for langt unna de andre i vanskegrad, altså at den er veldig lett eller veldig vanskelig.

Spørsmålet blir om hvor en skal sette grensen for hva som er god korrelasjon og hva som blir for ulikt de resterende oppgavene. Wolfe og Smith (2007) foreslår en grense på 0,30 for dikotomt utformede undersøkelser og 0,40 for polytome for å «flagge» et utsagn. Dersom et utsagn blir flagget må det undersøkes kvalitativt om det er verdt å ha med eller om det tilfører for mye usikkerhet til datamaterialet.

Det andre underaspektet til innholdsaspektet Wolfe og Smith (2007) anbefaler å undersøke når en ser på den tekniske kvaliteten til utsagnene er mean-squared fit indices. De handler om hvor langt unna de målte verdiene er fra hva den teoretiske Rasch-modellen ville antatt. Desto større/mer positiv verdiene blir, desto større er skillet mellom teori og praksis. Dersom verdiene er negative betyr det at det er mindre variasjon enn det Rasch-modellen forventer. Negative verdier er som oftest ikke et stort problem, men kan være bekymringsverdig dersom absoluttverdien overstiger grenseverdien. Basert på simulasjoner gjort av Smith et al. (1998), anbefaler de en grenseverdi for å flagge utsagn/oppgaver som er over 2,00 for undersøkelser hvor det er færre enn 1000 respondenter, over 3,00 for undersøkelser hvor antall respondenter er mellom 1000 og 5000 og over 4,00 dersom det er mer enn 5000 respondenter. En annen verdi som også brukes når det gjelder å flagge oppgaver i denne sammenhengen er 1,3 (Bond & Fox, 2015). 1,3 er en strengere grense enn den Wolfe og Smith (2007) anbefaler og gir dermed økt validitet til studien. Jeg har derfor valgt å sette grenseverdien for flagging av utsagn til 1,3.

En annen måte å argumentere for innholdet i en studies validitet er ved hjelp av Item Characteristics Curves (ICC) (Bond & Fox, 2015). En ICC-kurve gir et bilde på hvor godt en oppgave treffer det latente trekket den er ment å måle. I en slik fremstilling blir det dannet grupper av personer etter deres mål. Disse gruppene blir omgjort til punkter som representerer gjennomsnittsmålet til gruppen og plasseres opp mot den forventede Rasch-modellen. Dermed kan en sammenligne de reelle målene med de forventede ved hjelp av sannsynlighetsberegning og konfidensintervall. Dersom den reelle kurven følger den forventede kurven, er det et tegn på at oppgaven egner seg til å måle det latente trekket. Dersom det er signifikant forskjell mellom de to kurvene tyder det derimot på at oppgaven er uegnet.

3.4.3 Substantive Validity

Den substantive validiteten til en undersøkelse kan bevises ved flere ulike metoder i følge Wolfe og Smith (2007). En av dem er å gjennomføre en pilotundersøkelse, som kan være å lage en prototype av undersøkelsen en har tenkt å gjennomføre, for deretter å se om det oppstår uforutsette problemer/komplikasjoner. Deretter kan en ta en prat med intervjuobjektet eller respondenten til spørreskjemaet og, dersom et problem oppstod, forsøke å avdekke hva som forårsaker et slikt problem og forsøke å forbedre

eller legge til noe i spørreskjemaet som kan løse det. Et eksempel kan være at et utsagn i et spørreskjema er ordlagt slik at respondenten er mer tilbøyelig til å svare noe annet enn det han eller hun egentlig ville svart dersom utsagnet var mer presist formulert. En pilotundersøkelse kunne i denne sammenhengen gjort forskeren oppmerksom på dette utsagnet og dets formulering, ved å først se at responsen på utsagnet ikke var i samsvar med det en ville forventet utfra hva respondenten svarte på andre lignende utsagn. Deretter, i dialog med respondenten, kunne forskeren oppdaget problemet og reformulert det slik at responsene i den egentlige undersøkelsen ble mer nøyaktige og i tråd med resten.

I min studie valgte jeg ut to sykehusansatte og fikk dem til å lese gjennom påstandene for så å gi meg tilbakemelding på om noen av utsagnene var utydelig formulert. Da de ikke hadde noen bemerkninger valgte jeg å gå videre uten å gjøre endringer på spørreskjemaet.

En annen måte en kan sikre substantiv validitet på er å undersøke i hvor stor grad responsen til respondentene på utsagnene er i sammenheng med hva en som utvikler av instrumentet har til hensikt. Det kan en gjøre i følge Wolfe og Smith (2007) ved å undersøke det de kaller «the functioning of distracters» eller graderingsskalaen (rating scale), alt etter hvilken form for instrument en utvikler. I min undersøkelse har jeg benyttet meg av graderingsskala og jeg vil derfor gå nærmere inn på det. Wolfe og Smith henviser til Linacre (2004) som legger frem flere måter å tyde hva ulike verdier i analysen betyr for validiteten. Linacre mener at det første en må se etter er om alle oppgavene «peker i samme retning», det vil si at alle punkt-målkorelasjonene (point-measure correlation) er positive. Videre legger han frem åtte punkter som er med på å sikre blant annet nøyaktigheten til målingen.

Det første punktet av Linacres åtte er at hver svarkategori må inneholde et minimum av ti responser. For det andre burde formen til hver rating scale-kurve være ensidig og jevn for å støtte argumentet om at svarkategoriene ikke blir oppfattet på en særegen måte. Det tredje punktet er at det gjennomsnittlige målet til personene som har svart på hver enkelt svarkategori, må være stigende i den rekkefølgen som skaperen av instrumentet hadde til hensikt. For eksempel hadde det vært et problem dersom respondenter med et høyt mål svarte «sjeldent» på noe hvor respondenter med lavt mål svarte «ofte» på, da «ofte» skal antyde mer eller større grad av det vi måler. Linacres fjerde punkt er at gjennomsnittlig kvadrert fit-statistikk (mean-square fit statistics) burde ligge under 2,00 da dette antyder at responsene er i samsvar med det vi ville forventet utfra Rasch-modellen. Tallet 2,00 er noe vilkårlig valgt, men kan forklares som at 1,00 er den perfekte verdien med tanke på Rasch-modellen og alt over det kan anses som støy. Dersom støyen blir høyere enn det vi ønsker, altså større enn 1,00, «hører» vi mer støy enn de resultatene vi ønsker å høre på, noe som ikke er ønskelig. For å forhindre for mye støy settes derfor den øvre grensen til 2,00

Videre er det femte punktet til Linacre å se på svarkategoriens kalibrering. Disse skal i følge han være i stigende rekkefølge. Det sjette punktet er at mål skal antyde svarkategori, samtidig som at svarkategori skal antyde mål. Sagt med andre ord skal det være en sammenheng mellom hva en person svarer og det målet den får. De som for eksempel får et høyt mål skal ofte svare alltid/nesten alltid, samtidig som det bare er de som ofte svarer alltid/nesten alltid som får et høyt mål. Punkt syv sier at differansen mellom to svarkategorier som er kalibrert til å ikke ha noen andre svarkategorier mellom seg, skal ikke være mindre enn 1,4 logits. Logits, eller log-odd units, er den matematiske enheten som brukes i Rasch-målinger (Wright & Stone, 1979). For

eksempel skal ikke «aldri/nesten aldri» være kalibrert til 0,2 logits hvis «noen ganger» er på 0,8 logits. Differansen mellom dem er da 0,6 som er langt under grensen på 1,4. Hvis dette var tilfelle ville det blitt vanskelig å meningsfullt skille personer som svarte det ene fra de som svarte det andre, da forskjellen mellom dem kan forklares med tilfeldigheter. Punkt åtte er nokså lik punkt syv, forskjellen er at differansen heller ikke skal overstige 5,0 logits.

3.4.4 Structural Validity

Den strukturelle validiteten er nært knyttet opp mot prinsippet om en-dimensjonalitet. Det handler i hovedsak om at datamaterialet er slik at vi kan legge oppgavens og personenes mål på en og samme linje og samtidig kunne si noe meningsfullt om dem. Dersom det ikke hadde vært en-dimensjonalt hadde en for eksempel ikke vært sikker på om en persons mål og plassering i forhold til andre var slik på grunn av personens MI, eller om det skyldtes en annen dimensjon som ikke kommer tydelig frem. Ifølge Wolfe og Smith (2007) kan en sikre seg dokumentasjon for datamaterialets strukturelle validitet ved å gjøre tre steg – steg som analyseverktøyet Winsteps gjør automatisk. Det første steget er å konvertere datamaterialet til å passe Rasch' endimensjonelle modell. Deretter gjør en en Principal Component Analysis (PCA) på det konverterte datamaterialet. I det tredje steget tar en eigenverdiene fra de Rasch-skalerte målene og resten (the residuals) av den prinsipielle komponentanalysen og plasserer de på en felles skala. Ved å se på eigenverdiene kan en fastslå om det er tilstrekkelig varians til å fortsette å utforske dimensjonaliteten.

3.4.5 Generalizability Validity

Dette aspektet handler om hvordan instrumentet kan generaliseres. Vil vi få det samme resultatet uansett om vi benytter instrumentet på en undergruppe av respondentene eller på samtlige? For å anskaffe slike beviser foreslår Wolfe og Smith (2007) å benytte seg av en DIF-analyse (differential item functioning analysis) av datamaterialet. En DIF-analyse gjøres ved at en sammenligner strukturen til en måling av en bestemt gruppe opp mot hvordan strukturen er hos en annen gruppe. Når det gjelder å benytte DIF-analyse for å sikre validitet skal en kunne forvente at disse to gruppene får tilnærmet like resultater, ettersom kriteriet for hvorfor en havner i den ene eller den andre gruppen ikke skal ha noe å si for det som måles. En vanlig måte å skille respondenter er ved å dele dem inn i kjønn, men det kan også være andre (arbitrære) kategorier som for eksempel om bursdagen til respondenten er et oddetall eller et partall, hvilken hårfarge respondenten har eller om forbokstaven i navnet deres er en vokal eller konsonant. Hensikten er å danne to grupper, basert på noe som ikke skal ha noe å si for det som måles. I min studie har jeg valgt å sammenligne menn og kvinner, i tillegg til å se på fire ulike yrkeskategorier en finner blant sykehusansatte. I min oppgave er DIF-analyse sentralt da det benyttes både i denne sammenhengen for å sikre validiteten/reliabiliteten til instrumentet innad i konteksten sykehus, men også i den større sammenhengen, hvordan strukturen til målingen av MI hos sykehusansatte er sammenlignet med ungdomsskoleelevers og ingeniørstudenters.

3.4.6 External Validity

I begrepet ekstern validitet ligger det at vi ser på noe som er utenfor den gjeldende undersøkelsen, for å få en pekepinn på om resultatene er til å stole på. En typisk metode for å gjøre dette er ved å sammenligne sine egne resultater med andres som har forsket på det samme, eller noe som er i nærheten av hva en selv undersøker. Dette aspektet ved validitet har et overlapp med den andre delen av min problemstilling som handler om å sammenligne målene fra sykehus med ungdomsskole (Ytterhaug, 2019) og

universitet (Kaspersen et al., 2017). Måten jeg har gjort det er ved en DIF-analyse for å se om forskjellene mellom de ulike målene er signifikant forskjellige eller om de har noen felles strukturelle likheter, det vil si at utsagnenes mål er tilnærmet like og plasserer seg i omtrent samme rekkefølge.

3.4.7 Consequential Validity

I dette aspektet til Wolfe og Smith (2007) diskuteres konsekvensene av bruken av måleinstrumentet. Et av fokusene her er på hvordan en har valgt grensene for hva som er innenfor det som er akseptabelt, og hva som ikke er det. Grensene skal helst være satt av troverdige forskere som har godt empirisk belegg og argumenter for å sette disse grensene til det de gjør. Et eksempel på dette aspektets relevans er når en lærer skal rette en prøve. Prøven er et forsøk på å måle elevenes kunnskap og ferdigheter, og når elevene har tatt den, har de fått et visst antall riktige svar. Læreren må da bestemme hva som skiller en sekser-elev fra en femmer-elev. Dette valget er viktig da det blant annet har konsekvenser for tilbakemeldingen elevene får om hvor stor grad de mestrer temaet. Dermed må det stilles krav til validiteten rundt hva som skiller karakterene.

Min hensikt i denne studien er å sammenligne målene og strukturen til målene ved tre ulike kontekster. Hva de faktiske målene er, er av liten til ingen relevans da det ikke er en del av forskningsspørsmålet mitt. Jeg har dermed ikke til hensikt å diskutere noen grenser for hva som for eksempel kan regnes som sterk MI og hva som ikke er det. Det medfører at jeg ikke har tatt med dette aspektet i analyseringen av mitt datagrunnlag.

3.4.8 Responsiveness Validity

Responsivitet i denne sammenhengen handler om hvor godt undersøkelsen kan respondere på endring, det vil si distingvere mellom før og etter en intervensjon i en gruppe er gjort. For eksempel har blant andre Pepin (2011) sett på ungdomsskoleelevers holdninger til matematikkfaget, og kommet frem til at de går noe i negativ retning fra 8. trinn til 9. trinn før de går i positiv retning igjen i 10. trinn. Dersom en foretok en måling av ungdomsskoleelevers holdninger til matematikkfaget og en kunne se en slik endring, ville det ha vært et argument for at målingen hadde validitet.

Ettersom konteksten jeg skal undersøke MI i er et sykehus, og det tidligere er gjort få lignende undersøkelser i denne konteksten, har jeg ingen tidligere funn å sammenligne en eventuell endring med. Det er heller ikke min hensikt i denne studien å måle den samme gruppen, før og etter en intervensjon. Dermed har jeg ingen forutsetninger for å argumentere for min studies responsive validitet og jeg velger derfor å ikke fokusere på dette aspektet.

3.4.9 Interpretability Validity

Det siste av de åtte aspektene, er et aspekt som går ut på om den eller de som skal benytte seg av, eller tolke en måling, forstår hva som menes med målene. Wolfe og Smith (2007) mener at for at et publikum, uten forkunnskaper om psykometri, skal ha den samme forståelsen av hva som ligger i målingen må en kvalitativ mening tillegges de kvantitative måledataene. Sagt med andre ord kreves det en referanseramme som kan mediere hvordan målene skal forstås. En slik ramme kan brukes for å vise hvor grensene for hva en anser som bestått eller hvor skillet mellom lav og middels grad av det som måles er.

I min studie er ikke hvilket mål hver enkelt person eller oppgave blir målt til sentralt, det er den helhetlige strukturen jeg er ute etter å undersøke. Det er derfor heller ikke

relevant for min studie å avgjøre hva som for eksempel karakteriserer en sykehusansatt med sterk MI. Jeg har derfor heller ikke vektlagt dette aspektet.

3.7 Hvordan fungerer Winsteps?

For å analysere datamaterialet mitt har jeg valgt å benytte meg av analyseverktøyet Winsteps, et analyseprogram som kan anvendes blant annet ved Rasch-analyser. Winsteps fungerer slik at det tar inn datamaterialet og tilegner hver person og hver oppgave et vilkårlig mål. Deretter beregner den sannsynligheten for at akkurat det faktiske datamaterialet ville dukket opp gitt disse antatte målene. Videre gjøres det en justering på de vilkårlige målene for å finne et sett med mål som gir en større sannsynlighet. Prosessen med å justere og regne på sannsynligheten gjentas til programmet ikke kan komme frem til noe som gir større sannsynlighet enn det allerede har kommet frem til. Slik går programmet frem for å finne de målene som gir datamaterialet størst sannsynlighet for å inntreffe. Disse målene er ikke nødvendigvis de faktiske målene til personene og oppgavene, men det er innenfor et konfidensintervall som er satt til 95% (dette kan justeres etter hva som er ønskelig/nødvendig). Nøyaktigheten er dermed ikke eksakt, men dette kan vi også forvente ved andre typer målinger. Dersom en måler høyden til en person med en meterstokk finner en som regel bare høyden i meter med to desimaler (cm) når den eksakte høyden kan i teorien måles til atomnivå.

Analyseverktøyet har ulike måter å fremstille ulike data, alt etter hva det er ønskelig å undersøke. Dersom en for eksempel ønsker å se hvordan oppgavene og personenes mål strukturerer seg i forhold til det latente trekket som måles, kan dette vises via et variabelkart (variable map). Måten det blir fremstilt er tilnærmet lik en målestokk hvor en linje er markert med punkter som indikerer mål. Langs denne meterstokken legger personene og oppgavene seg til sine respektive mål. Dermed kan en lese av både rekkefølgen på hvordan oppgavene og personene legger seg, og også avstanden fra et mål til et annet.

3.8 Etiske betraktninger

Forskning skal ikke gå på bekostning av menneskers liv og helse. Det er derfor viktig at personer som deltar i arbeid med forskning, enten som respondenter eller på andre måter, blir tatt hensyn til. Jeg har valgt å støtte meg til retningslinjene til Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH) for å ivareta hensynet til personene og institusjonene som deltar i min studie. Jeg har besluttet at i min studie skal derfor ingen personer eller institusjoner være identifiserbare. Dette ble sørget for hovedsakelig i utformingen av spørreskjemaet. Det ble blant annet ikke spurt om navn, fødselsdato eller andre personidentifiserbare opplysninger.

I følge NESH' retningslinjer skal personer som bidrar til forskningen ha tilstrekkelig innsikt i tematikken, formålet med undersøkelsen, hvem som får tilgang til informasjonen og hva resultatene er tiltenkt. I informasjonsskrivet, som fulgte med spørreskjemaet, ble en kort oppsummering av respondentenes rettigheter presentert, samt kontaktinformasjon til meg selv og veilederen min dersom ytterligere informasjon var ønskelig. I tillegg, for at det skal være mulig for respondentene å trekke seg og sin besvarelse fra studien, var det viktig at de hadde denne kontaktinformasjonen og at de fikk opplysninger om å ta kontakt dersom det var tilfelle. Dette ble det også opplyst om i

informasjonsskrivet. Respondentene ble også informert om at det var NTNU som var ansvarlig for studien og at datamaterialet ville bli overlevert til NTNU etter prosjektets slutt for lagring og videre bruk i nye forskningsprosjekt. Selv om personidentifiserende informasjon ikke ble innhentet er det viktig at respondentene i en slik studie aktivt godtar sin deltakelse. Derfor ble respondentene bedt om å krysse av for om de var inneforstått med studien og de tidligere nevnte rettighetene de hadde og for om de ga sitt samtykke til å delta i studien.

Selv om innsamlingen av datamaterialet og analyseringen av det er i samråd med retningslinjene til NESH er resultatene til en måling av denne typen diskuterbar, nærmere bestemt hva de blir brukt til. Hensikten med å måle personers MI er i denne oppgaven kun for å undersøke sammenhenger mellom ulike kontekster og dermed få en dypere innsikt i tematikken. Måling av personers MI for å vurdere personene, rangere dem, eller ta beslutninger som angår dem, er ikke min hensikt og noe jeg vil fraråde da forskningsfeltet enda er i en etableringsfase. Vi vet ikke hvor stort omfanget MI har hos en person eller en kontekst. Vi vet heller ikke om det finnes andre variabler som kan være like mye eller mer betydningsfull og avgjørende for personers affektive side til matematikkfaget.

4. Resultat

For å analysere dette datamaterialet har jeg som tidligere nevnt benyttet meg av Wolfe og Smiths (2007) inndeling av validitet i åtte aspekter. Jeg har valgt å utelate tre av dem da jeg har ansett dem som ikke relevante for min studie, ettersom de retter seg mot bruken av måleinstrumentet. I min studie er jeg kun interessert i å undersøke om instrumentet er egnet for å måle MI ved et sykehus. Jeg anser derfor hvordan målingene blir brukt som irrelevant. Jeg vil i dette kapitlet legge frem det analysen avdekket av relevante data slik de fremstår i de fem aspektene.

4.1 Innholdsaspektet

Innholdsaspektet tar for seg om innholdet i de oppgavene som blir benyttet i målingen er representativt for det vi ønsker å undersøke. I min studie er oppgavene de 20 påstandene som er benyttet i spørreskjemaet, og det jeg ønsker å undersøke er MI. Spørsmålet blir derfor om disse 20 påstandene handler om MI eller om de dreier seg om noe annet. I det tidligere studiet gjort av Kaspersen (2015) fant de frem til disse 20 påstandene ved å spørre matematikklærere og matematikkstudenter om hva som karakteriserer en person med sterk MI. De karakteristikkenes som gikk igjen i de fleste av svarene ble videre utarbeidet til de påstandene jeg senere også har benyttet i min studie. Det må her også presiseres at det ikke utelukkes at det finnes flere slike karakteristikker og at det ikke må antas at disse karakteristikkenes er de best egnede for å måle MI.

4.1.1 Infit og outfitverdier

Wolfe og Smith (2017) anbefaler, i tilfeller hvor en benytter Rasch, å se på den tekniske kvaliteten til utsagnene, nærmere bestemt to aspekter ved dem: item-measure correlation og mean-squared fit indices.

Som nevnt i metodekapitlet valgte jeg å benytte Bond og Fox' (2015) øvre grense på 1,3 for å flagge utsagn basert på deres infit og outfitverdier. Under analysen av datamaterialet ble utsagn 9 utpekt som det utsagnet som de ansatte jevnt over har besvart mest ulikt fra det Rasch-modellen forventet. Utsagnet var «Når jeg lærer meg en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gjøre.» Dette utsagnet viste seg etter å ha sett kvalitativt på det å være et negativt formulert utsagn. Det måtte dermed kodes motsatt, altså at de som hadde svart Aldri/nesten aldri (1) ble kodet til å ha svart 4 og de som hadde svart alltid/nesten alltid(4) ble ansett som å ha svart 1. Selv etter denne rettelsen var det fortsatt svært høye infit- og outfitverdier (2,08 i infit og 3,35 i outfit) som dermed førte til at jeg valgte å ta utsagnet ut av listen med utsagn. Årsaken til de høye verdiene kan være at ettersom det var det eneste utsagnet som var negativt formulert har det blitt ansett og besvart av en del av utvalget på samme måte som de andre utsagnene, mens den andre delen har lest mer nøye og sett at påstanden peker i motsatt retning. Slik kan det ha oppstått ulike tolkninger av påstanden og resultatet er at innholdet i påstanden ikke er entydig. En annen forklaring kan være at utsagnet handler om noe annet enn MI. Det ville også mest sannsynlig resultert i høye verdier. Uansett hva årsaken er valgte jeg å ikke ha med responsene til denne påstanden videre i analysen da de resterende påstandene ga infit og outfit-verdier som var mer i tråd med hva Rasch-modellen forventer.

Etter å ha fjernet utsagn 9 hadde samtlige av de resterende utsagnene en infit-verdi innenfor kravet og det var kun utsagn nummer 3 som hadde en outfitverdi over 1,3

(1,62). Ettersom outfitverdien er mer påvirkelig av enkeltutsagn valgte jeg å se på personene med høyst outfit. Person nummer 131 var den personen med høyst verdi, og etter å ha sett nærmere på dens svar, var det tydelig at det skyldtes nettopp utsagn 3. Utenom utsagn 3 hadde person 131 svart slik Rasch-modellen forventet og jeg valgte derfor å kun slette besvarelsen til person 131 på utsagn 3. Dette resulterte i at alle de resterende 19 utsagnene hadde en infit/outfit verdi på under 1,3, hvor den høyeste av dem var utsagn 5 med en infit på 1,27 og en outfit på 1,26 som i følge Bond og Fox (2015) er et akseptabelt nivå. Enkelte utsagn var også under 1, som vil si at de karakteriseres som overfit, men dette skyldes at snittet av alle verdiene skal være tilnærmet lik 1 og ettersom noen av utsagnene var noe over 1 må derfor enkelte utsagn ligge på noe under 1.

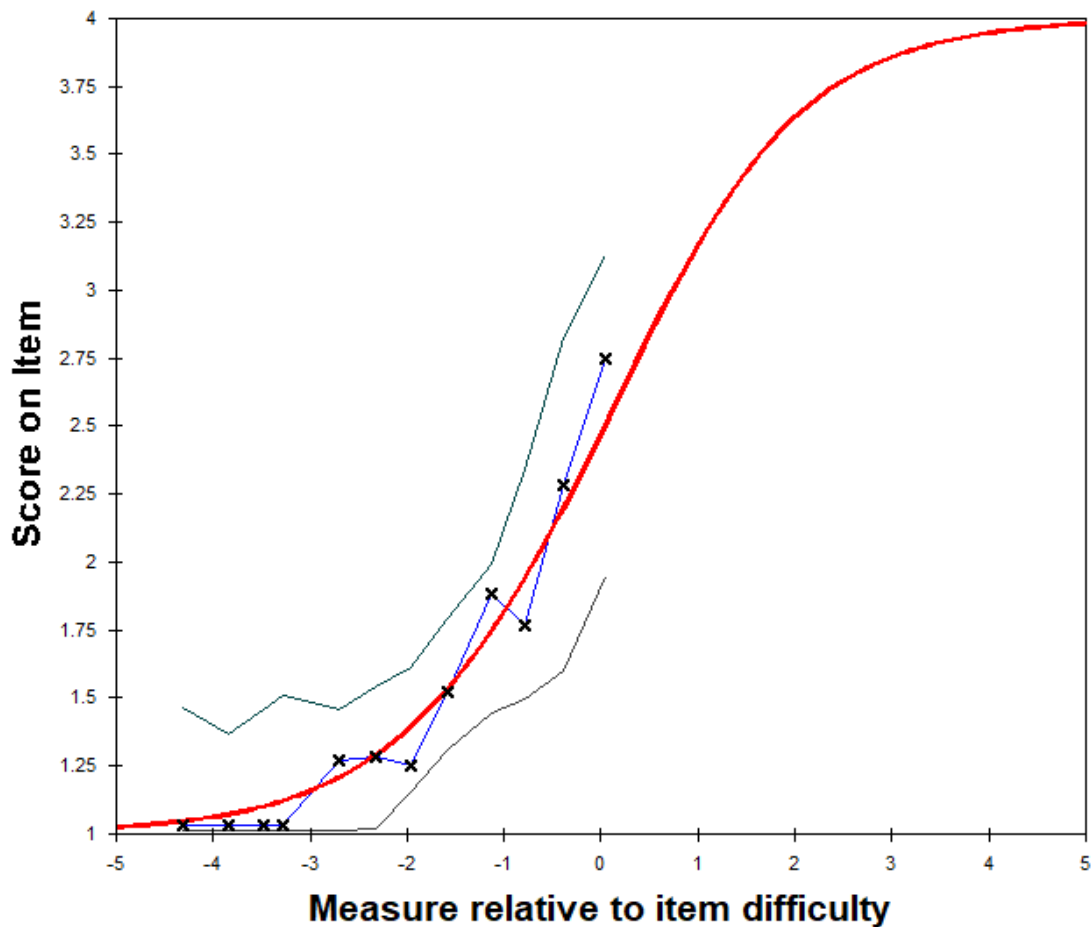
Videre i analysen så jeg på de ansattes infit/outfit verdier. I Wolfe og Smith (2007) henviser de til en annen studie. Basert på simulasjonene til Smith et al. (1998) anbefaler de en verdi på 2,00 som en grensescore for å flagge personer for misfit. Dette er et grep for å styrke det substantive aspektet av validitet. Det viste seg å være åtte personer; 90, 139, 96, 38, 51, 86, 27 og 68 som alle hadde en infitverdi, en outfitverdi eller begge som lå på over 2,00. Ved å lagre målene til utsagnene for så å ta vekk disse personene fra datamaterialet og til slutt sammenligne de gamle målene med de nye kunne jeg kvantitativt se om personene påvirket utsagnenes mål nevneverdig. Den største forskjellen mellom målene med samtlige ansatte og målene uten de 8 med høy fitverdi var utsagn 2 hvor forskjellen var 0,1 logits. Dette anså jeg som så liten forskjell at personene ikke påvirket målingen i nevneverdig grad og jeg valgte dermed å beholde dem i datamaterialet.

Etter å ha sett på enkeltpersoners verdier la jeg merke til at person 60 hadde svart 1 på samtlige påstander. Jeg valgte da å ta personen ut av undersøkelsen ettersom personens besvarelse ikke tilfører statistisk informasjon til målingen, den sier kun at personens «evne» ligger utenfor instrumentets rekkevidde. Sagt med andre ord har personen en lavere MI enn det «målestokken» måler og det har dermed ingenting for seg å ha personen med i denne sammenhengen.

4.1.2 ICC-kurver

For å støtte opp om resultatene og beslutningene fra å se på infit og outfitverdiene valgte jeg å undersøke ICC-kurvene av datamaterialet. Figur 2 og 3 er to eksempler på hvordan ICC-kurvene så ut. Den røde kurven i Figur 2 og 3 er den forventede Rasch-modellen og den blå er de faktiske responsene til de sykehusansatte. Hver «x» langs den blå linjen representerer en gruppe av respondenter. De to svarte linjene representerer konfidensintervall på 95%. Dersom responsene til de sykehusansatte faller utenfor konfidensintervallene betyr det at responsene er signifikant ulike fra Rasch-modellen. Y-aksen representerer hvilken svarkategori respondente har valgt og x-aksen er respondentenes mål relativt til utsagnets.

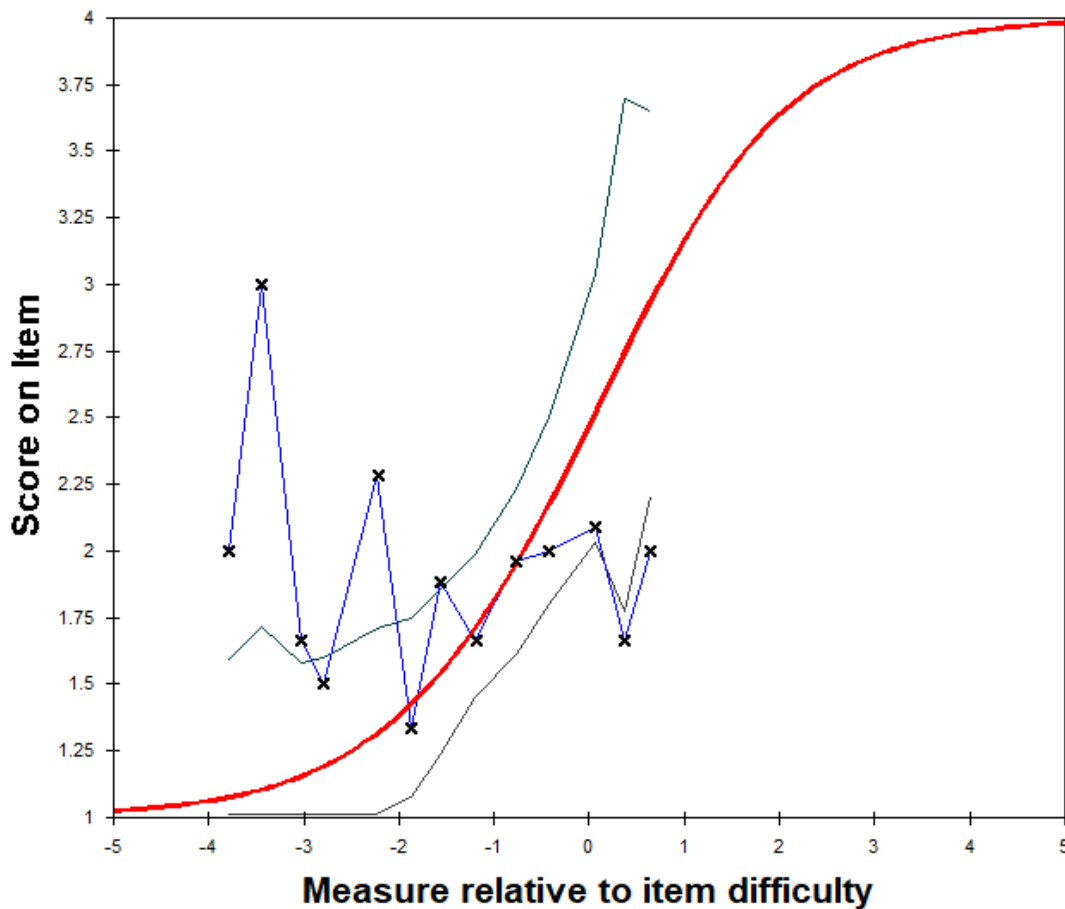
4. 4. Jeg har problemer med å legge fra meg matematiske oppgaver



FIGUR 2 – ICC-KURVE TIL UTSAGN 4.

Figur 2 viser hvordan de sykehusansatte har respondert til utsagn 4 og hvordan denne responsen er sammenlignet med Rasch-modellen. Vi kan se at den blå linjen hovedsakelig følger den røde. Enkelte punkter «bommer» noe, men ikke med så mye at de er utenfor konfidensintervallene, noe som tyder på at datamaterialet stemmer godt overens med Rasch-modellen. Dette var tilfellet med samtlige av de andre utsagnene, med unntak av utsagn 9 (Figur 3). En annen ting vi kan se er at den blå linjen stopper omtrent midt på den røde kurven. Dette tyder på at utsagn 4 er et relativt vanskelig utsagn som få av de sykehusansatte sier stemmer ofte eller alltid.

. Når jeg lærer en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gj



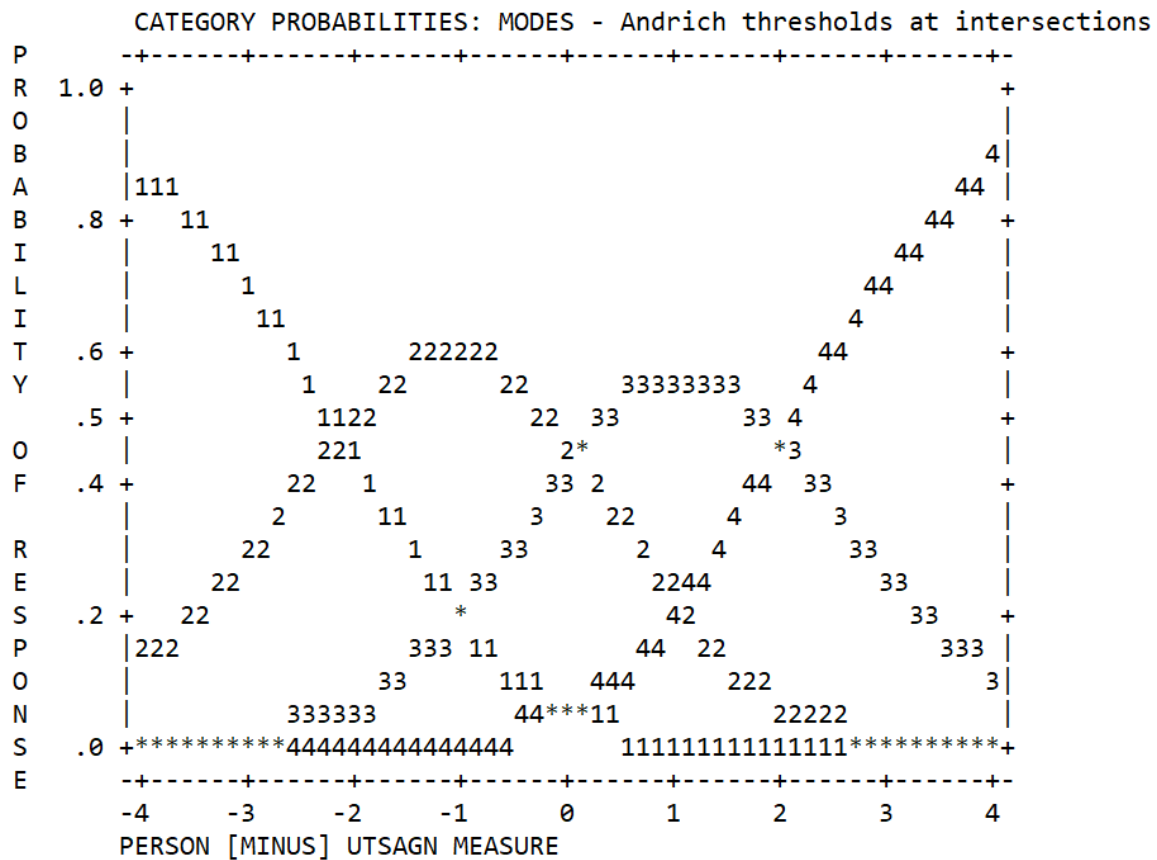
FIGUR 3 – ICC-KURVE TIL UTSAGN 9.

Det vi kan se fra Figur 3 er at de sykehusansatte har respondert på utsagn 9 ikke er særlig sammenfallende med Rasch-modellens forventninger. Vi kan også se at flere av punktene faller utenfor konfidensintervallet, noe som støtter valget om å ta utsagn 9 ut av studien.

4.2 Det substansielle aspektet

For å sikre validiteten til datamaterialet mitt i det substansielle aspektet fulgte jeg Linacres (2002) krav til svarkategoriene mine. Her kom det frem at hver svarkategori var besvart langt flere enn Linacres krav om ti ganger. Den kategorien som var mest sjeldent besvart var svarkategori 4, som var besvart 305 ganger, som utgjorde 12 % av de totalt registrerte, individuelle svarene. Det andre punktet til Linacre er at formen på kurvene til svarkategoriene må være jevn. I Figur 4 kan en se kurvene til hver svarkategori og legger seg i stigende rekkefølge fra 1-4 ettersom personenes mål øker. Dermed er også Linacres tredje punkt om at den gjennomsnittlige målet på personer som har svart hver kategori er i stigende rekkefølge. I det fjerde punktet til Linacres skal kategoriernes outfit-verdier være under 2,00. Igjen er alle kategoriene innenfor det som er ønskelig, hvor den med høyst verdi var svarkategori 4 med en verdi på 1,65. Det

femte punktet er at stegenes kalibrering skal være stigende. Dette var også tilfelle med mine data.



FIGUR 4 - GRAFISK FREMSTILLING AV SVARKATEGORIENES SANNSYNLIGHET FOR Å BLI VALGT GITT PERSONERS MÅL. HVERT SIFFER (1-4) TILHØRER EN SVARKATEGORI OG DANNER HVER SIN KURVE. KURVENE REPRESENTERER SANNSYNLIGHETEN FOR AT DEN TILHØRENDE SVARKATEGORIEN BLIR VALGT LANGS Y-AKSEN, GITT PERSONENE SOM HAR RESPONDERT MED SVARKATEGORIENS MÅL LANGS X-AKSEN.

Det sjette punktet er at mål skal antyde svarkategori, samtidig som at svarkategori skal antyde mål. I min studies tilfelle skal altså de som har blitt målt til å ha en sterk MI stort sett ha svart «ofte» eller «alltid/nesten alltid». De som havner midt på treet burde ofte svare «noen ganger» eller «ofte» og de som har blitt målt til å ha liten grad av MI skal som oftest ha svart «aldri/nesten aldri» eller «noen ganger». Hvis vi snur på det skal vi også kunne se at de som for eksempel stort sett har svart «aldri/nesten aldri» få et lavt mål. I Figur 4 kan en se at personer som har blitt målt til å ha rundt 3 logits, som er relativt høyt, har også størst sannsynlighet for å velge svarkategori 4, som tilsvarer «alltid/nesten alltid». Dersom vi snur på det og ser på hvor svarkategori 1 (aldri/nesten aldri) er mest sannsynlig at blir respondert, kan vi se i Figur 4 at det er lengst til venstre på den negative siden. Dermed antyder svarkategorien et lavt mål. Analysen bekrefter dermed at Linacres (2002) sjette punkt er innfridd.

Punkt syv og åtte sier at svarkategoriens vanskegrad burde stige med et minimum av 1,4 logits og et maksimum av 5,0 logits. Dersom avstanden i logits mellom to svarkategorier er under 1,4 kan det bli vanskelig å påvise en reell ulikhet mellom dem ved hjelp av sannsynlighetstester, som kan resultere i at vi ikke kan hente ut meningsfull

informasjon. Dersom avstandene mellom to kategorier er så store at de overstiger 5,0 logit er det en mulighet for at det skulle vært et svaralternativ til mellom de to som hadde tilført undersøkelsen mer nøyaktighet. Analysen avdekte at avstanden i logits mellom de fire målene var: 2,24 (fra «aldri/nesten aldri» til «noen ganger»), 2,14 (fra «noen ganger» til «ofte») og 2,12 (fra «ofte» til «alltid/nesten alltid»). Disse differensene er altså vel innenfor kravet til Linacre (2002).

Etter å ha gått gjennom datamaterialet mitt med Linacres (2002) åtte punkter viser de at datamaterialet er troverdig og har god substansiell validitet.

4.3 Det strukturelle aspektet

Det strukturelle aspektet dreier seg om dimensjonalitet. Dersom utsagnene i undersøkelsen måler flere dimensjoner/underdimensjoner ved det vi ønsker å måle, strider dette med Thurstones krav om endimensjonalitet. For å finne ut av om mitt datamateriale inneholdt flere dimensjoner benyttet jeg meg av en PCA. Ut fra denne analysen fikk jeg blant annet en eigenverdi. En eigenverdi kan benyttes til å undersøke den samlede summen av underdimensjonalitet til utsagnene. Dersom vi hadde 17 påstander som alle målte og tilhørte kun én dimensjon, for eksempel MI, og vi hadde to andre påstander som målte noe helt annet, for eksempel lesing, ville analysen rapportert undersøkelsen til å ha en eigenverdi på 2. I realiteten hender det at påstander ikke bare er knyttet til én dimensjon, men flere og de avgir derfor noe verdi til eigenverdien.

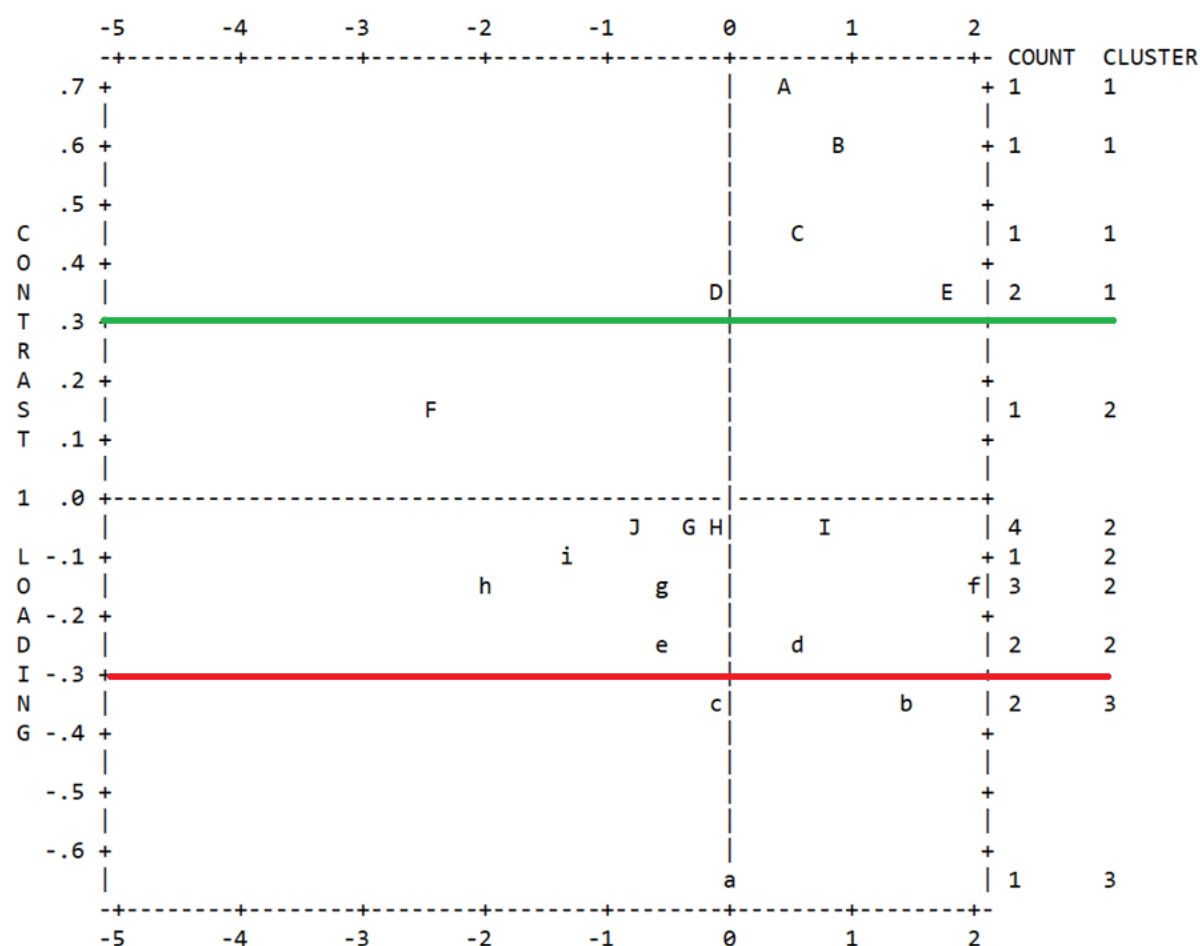
TABELL 2 – STANDARDISERT RESTVARIANS FOR UTSAGNENE UTTRYKT I EIGENVERDIER. DEN MEST RELEVANTE EIGENVERDIEN TILHØRER «UNEXPLAINED VARIANCE IN 1ST CONTRAST», MARKERT I RØDT.

| Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = UTSAGN information units | | | | |
|--|---------------|----------|----------|-------|
| | Eigenvalue | Observed | Expected | |
| Total raw variance in observations = | 37.0343 | 100.0% | 100.0% | |
| Raw variance explained by measures = | 18.0343 | 48.7% | 49.2% | |
| Raw variance explained by persons = | 5.7418 | 15.5% | 15.7% | |
| Raw Variance explained by items = | 12.2925 | 33.2% | 33.5% | |
| Raw unexplained variance (total) = | 19.0000 | 51.3% | 100.0% | 50.8% |
| Unexplned variance in 1st contrast = | 2.1608 | 5.8% | 11.4% | |
| Unexplned variance in 2nd contrast = | 1.7566 | 4.7% | 9.2% | |
| Unexplned variance in 3rd contrast = | 1.6770 | 4.5% | 8.8% | |
| Unexplned variance in 4th contrast = | 1.4208 | 3.8% | 7.5% | |
| Unexplned variance in 5th contrast = | 1.2756 | 3.4% | 6.7% | |

Da jeg gjorde en PCA avdekket winsteps at mitt datamateriale hadde en eigenverdi på 2,1608. Det tilsvarte at tilnærmet 2,2 utsagn handlet om noe annet enn MI. Selv om en forventer å få noe i eigenverdi ettersom få ting er helt én-dimensjonal var dette over grensen på 2,0 som Wolfe og Smith (2007) hevder er en god grense. Jeg måtte derfor undersøke nærmere om det var enkelte utsagn som skilte seg fra de resterende. Måten jeg da gikk frem var å se på hvilke utsagn programmet hadde vektlagt tyngst, og kvalitativt se på innholdet i påstandene. De to utsagnene med høyest vektleggelse i positiv retning, med andre ord over den grønne linjen i Figur 5, var utsagn 6 (Jeg blir engasjert når noen starter en matematisk diskusjon) med en vektning på 0,69 og utsagn 4 (Jeg har problemer med å legge fra meg matematiske oppgaver) med en vektning på 0,60. I Figur 5 er utsagn 6 representert som «A» og utsagn 4 som «B». Det jeg fant av fellestrekk mellom disse to utsagnene var at de handlet om affekt. Det å ha problemer

med å legge fra seg noe kan knyttes til at en opplever det som gøy, spennende eller engasjerende. Ordet engasjerende dukker også opp i utsagn 6. Jeg tenkte dermed at affekt var en annen dimensjon utsagnene målte i tillegg til MI.

Videre forsøkte jeg å se om noen av de andre aspektene kunne relateres til affekt. Da fant jeg at utsagn 18 (Når jeg lærer noe nytt, fører det til at det er flere ting jeg ønsker å finne ut) hvor ordet «ønsker» dukker opp og som jeg anser som også å tilhøre affekt. Dette utsagnet var derimot vektet negativt med en vekting på -0,10, noe som motstrider hypotesen om at affekt er en annen dimensjon utsagnene måler.



FIGUR 5 – KART OVER UTSAGNENES VEKTING MOT DERES MÅL. HVERT UTSAGN ER HER REPRESENTERT SOM EN STOR ELLER LITEN BOKSTAV. X-AKSEN ER UTSAGNENES MÅL OG Y-AKSEN ER DERES VEKTING. UTSAGNENE ER DELT INN I TRE GRUPPER (CLUSTERS) ETTER OM DE ER OVER DEN GRØNNE LINJEN, UNDER DEN RØDE ELLER I MIDTEN.

I den andre enden av vektingen, under den røde linjen i Figur 5, skilte utsagn 11 (Når jeg lærer en ny metode/algorithm, prøver jeg å finne ut hvorfor den virker) seg ut med en vekting på -0,63. Utsagnet kan finnes som «a» i Figur 5. Jeg forsøkte så å se kvalitativt på om det var noe med dette utsagnet som skilte seg fra de andre utsagnene. Det jeg da fant var at utsagn 11 handlet om forståelse, nærmere bestemt et ønske om å få forståelse. Videre forsøkte jeg å finne andre utsagn som handlet om det samme. Jeg fant da flere utsagn som passet til det å ønske en dypere forståelse, blant andre utsagn 7 (Når jeg lærer noe nytt, stiller jeg meg selv egne spørsmål som jeg jobber med.) og utsagn 1 (Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobber legger

opp til.) Deretter så jeg på vektingen av disse påstandene for å se om det var en tendens til at disse påstandene ble vektlagt negativt. En slik tydelig tendens fant jeg ikke. Utsagn 1 ble for eksempel vektlagt i positiv retning med 0,41. Jeg vil derfor ikke påstå at det er en tydelig underdimensjon i mitt datamateriale, selv om eigenverdien er noe over grensen til Wolfe og Smith (2007).

4.4 Generaliserbarhetsaspektet

For å sjekke om instrumentet er generaliserbart innad i konteksten trengte jeg å sammenligne undergrupper mot hverandre. Dette gjorde jeg for å se om det hadde noe å si for resultatene om jeg målte for eksempel bare de kvinnelige ansatte. Dersom det hadde vært stor forskjell mellom menn og kvinner i mål, hadde det vært et tegn på at instrumentet ikke er helt til å stole på da det er rimelig å forvente at menn og kvinner har noen lunde like forutsetninger for hva de svarer på en undersøkelse som denne.

Det er selvsagt rimelig å forvente en viss forskjell mellom gruppene som analyseres. Spørsmålet som melder seg i denne sammenhengen er hva en skal sette som grensen mellom hva som er forventede ulikheter i mål av to i utgangspunktet like grupper, og hva som skal karakteriseres som to ulike grupper. Linacre (u.å) benytter en grense på 0,64 logits. Jeg har derfor valgt å bruke den samme grensen.

Det er imidlertid ikke nok å bare se på forskjellen mellom to gruppers mål. Vi er ikke sikre på om denne forskjellen skyldes tilfeldigheter eller om det faktisk er en reell og markant forskjell. Derfor gjør også analyseverktøyet en sannsynlighetstest for å finne ut hvor stor sannsynligheten er for at resultatet skyldes tilfeldigheter. I denne sammenhengen er det vanlig å sette en kritisk grense for p -verdien til 5%.

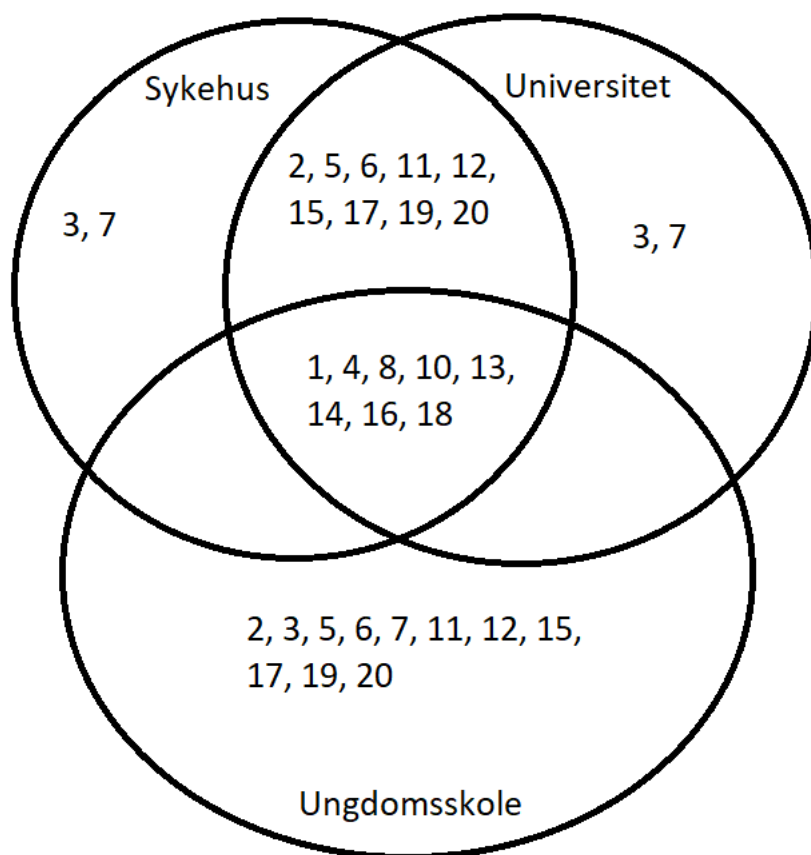
Med disse to grenseverdiene gikk jeg gjennom dataene fra analyseverktøyet og noterte ned tilfellene hvor begge verdiene ble overskredet. Da jeg først gikk gjennom kjønn ble to utsagn flagget: Utsagn 2 (Når jeg lærer en ny regnemetode, prøver jeg å finne en måte som fungerer bedre) og utsagn 11 (Når jeg lærer en ny formel, prøver jeg å finne ut hvorfor den virker). Utsagn 11 var 0,66 logits høyere for kvinner enn for menn og utsagn 2 var 0,79 logits høyere for kvinner enn for menn. De resterende 17 utsagnene var innenfor grensene. Etter å ha sett nærmere på de to flaggede utsagnene valgte jeg å gå videre uten å gjøre noe mer med det, da de to verdiene var så nær grenseverdiene, den ene bare 0.02 logits over, at jeg betraktet det som at det ikke hadde noe å si for min studie.

4.5 Det eksterne aspektet

Den andre problemstillingen i denne oppgaven handler om hvorvidt strukturen på MI er lik i ulike kontekster. Denne problemstillingen har som tidligere nevnt et overlapp med det eksterne aspektet til Wolfe og Smith (2007). Etter å ha validert resultatene fra de sykehusansatte, tydet det på at det var mulig å sammenligne dem med besvarelsene til de andre respondentene i de tidligere studiene gjort av Kaspersen et al. (2017) og Ytterhaug (2019). Derfor benyttet jeg meg på nytt av en DIF-analyse, denne gangen for å se om det var signifikante forskjeller mellom sykehusansatte, ungdomsskoleelever og universitetsstudenter.

For å først danne et helhetsbilde var det tydelig at ungdomsskoleelever var den av gruppene som skilte seg mest fra de to andre. Mellom sykehusansatte og

universitetsstudenter var det kun to utsagn som det ble rapportert om signifikante forskjeller: Utsagn 3 (Når jeg lærer en ny regnemetode, prøver jeg å finne situasjoner hvor denne ikke virker) og utsagn 7 (Når jeg lærer noe nytt i matematikk, lager jeg egne oppgaver). I denne analysene utpekte også enkelte påstander seg som signifikante forskjeller i hvordan de hadde blitt besvart i de ulike kontekstene. Disse var utsagn: 2, 3, 5, 6, 7, 11, 12, 15, 17, 19 og 20. Med unntak av utsagn 3 og 7 var alle disse utsagnene kun signifikant ulike hos ungdomsskoleelevenes mål og de andre to.



FIGUR 6 – VENNDIAGRAM OVER HVILKE UTSAGN SOM ER SIGNIFIKANT LIK I DE ULIKE KONTEKSTENE.

De resterende utsagnene (1,4,8,10,13,14,16 og 18) ble ikke utpekt som signifikant ulike. En annen måte å si det på er at disse åtte utsagnene hadde en tilnærmet lik vanskegrad i samtlige målte kontekster.

TABELL 3 - DE ÅTTE UTSAGNENE MED TILNÆRMET LIK VANSKEGRAD UAVHENGIG AV KONTEKST.

| | |
|----|---|
| 1. | Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobb legger opp til. |
| 4. | Jeg har problemer med å legge fra meg matematiske oppgaver. |

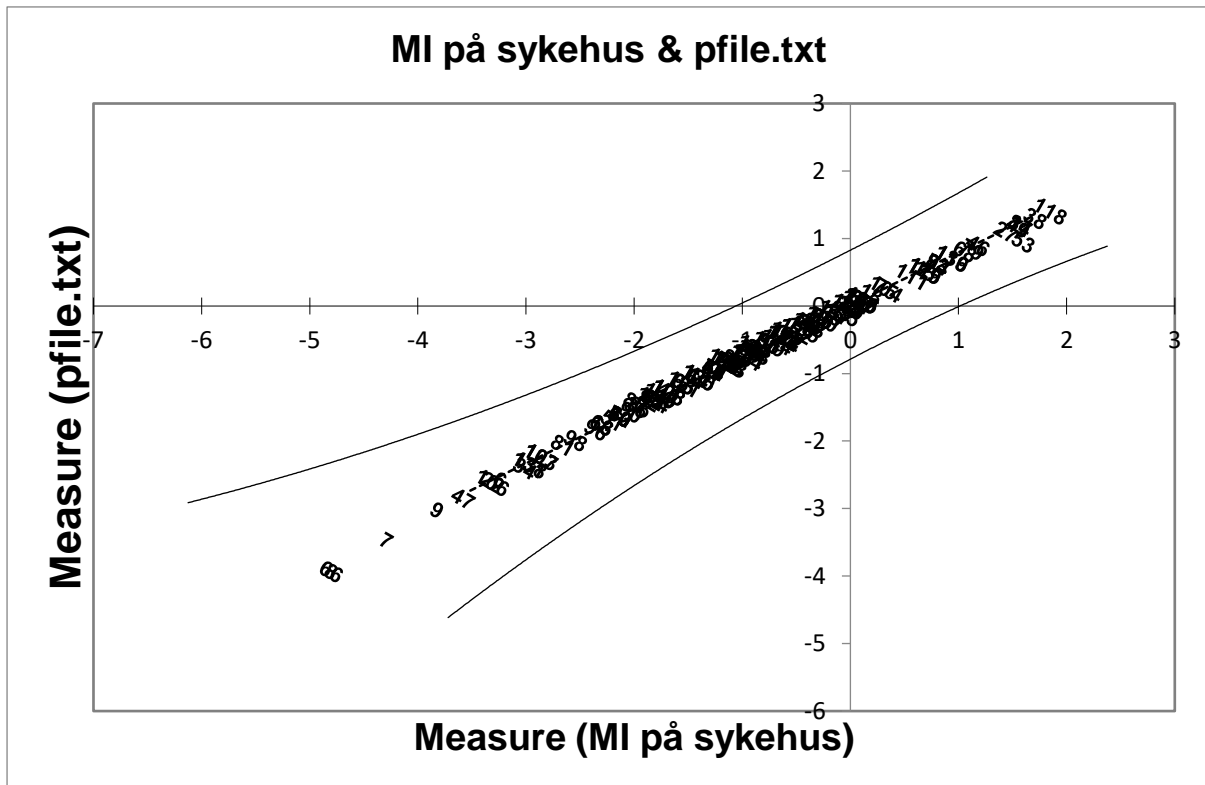
| | |
|-----|---|
| 8. | Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tankerekker. |
| 10. | Hvis jeg prøver på en metode som ikke fører frem, bruker jeg tid på å finne ut hvorfor denne ikke virker. |
| 13. | Når jeg møter et matematisk problem, tenker jeg over om det finnes flere måter å løse oppgaven på. |
| 14. | Når jeg jobber med et matematisk problem hopper jeg mellom ulike strategier. |
| 16. | Når jeg jobber med en oppgave, stopper jeg opp underveis og reflekterer over hva jeg gjør. |
| 18. | Jeg kan forklare hvorfor løsningen min er rett. |

Videre undersøkte jeg disse åtte utsagnene kvalitativt for å finne en eller flere fellesnevner med tanke på innholdet i dem. Den første sammenhengen jeg fant, var at flere av utsagnene var relatert til strategier i arbeid med matematiske oppgaver. Selve ordet «strategier» dukker opp i utsagn 14, men også relaterte ord som «metode», «måter», «løsningen» og «gjør» er å finne blant disse utsagnene. En annen sammenheng jeg oppdaget, var fokuset på dybdelære. Påstand 1 og 4 tolker jeg for eksempel til å handle om et ønske om å få en dypere innsikt i matematikken. Også andre påstander refererer til en videre forståelse av tematikken og en høyere grad av refleksjon. Utsagn 18 kan imidlertid argumenteres for at ikke handler om dypere forståelse, men mer om å ha rett. Samtidig kan det trekkes ut en dypere mening fra påstanden. «Løsningen min» kan antyde at det finnes andre løsningsforslag og det at personen kan forklare sin løsnings gyldighet tyder på en form for argumentasjon og refleksjon som krever forståelse.

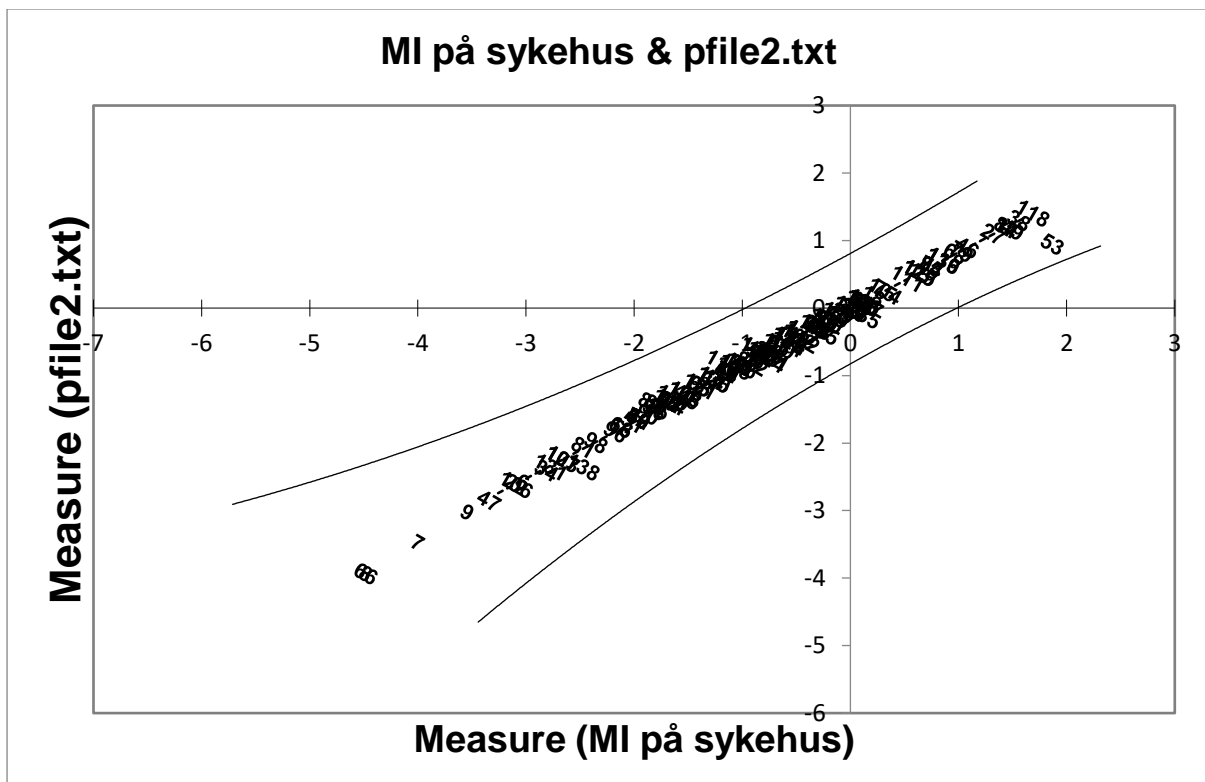
Et spørsmål en kan stille til at enkelte utsagn utpeker seg som signifikant ulike i de ulike kontekstene er om det har noe å si for målingen. Selv om noen utsagn har ulike mål i forskjellige kontekster trenger ikke det nødvendigvis å bety at resultatene av målingene av personene trenger å være ulike. Selv om logit-verdien til en person kan variere i ulike kontekster kan målet til personen fremdeles plassere seg tilnærmet likt i forhold til andre personer i konteksten, og på den måten har det ikke noen reell konsekvens for målingen om enkelte utsagn har ulike mål i ulike kontekster.

For å finne ut om det at noen utsagn får ulike mål i forskjellige kontekster har noe å si for personene i min studies mål relativt til hverandre har jeg valgt å gjøre en test som tester om forskjellene har praktiske konsekvenser når vi måler personenes MI. Testen sammenligner to mål av samme gruppe personer, en hvor de er målt med målestokken som er konstruert i sin egen kontekst, og en med en målestokk fra en annen kontekst. I mitt tilfelle gjorde jeg dette to ganger. Først sammenlignet jeg de sykehusansattes egne mål mot deres mål gjort med målestokken fra universitetsstudentenes, deretter mot målestokken fra ungdomsskolen. En måte å fremstille resultatene fra denne analysen er grafisk ved et «scatterplot». De to målene blir satt opp som hver sin dimensjon i en graf, og personene blir punkter i grafen. I tillegg kan vi se to linjer som er konfidensintervall. Dersom personene plasserer seg innenfor intervallet betyr det at det er liten grunn til å tro at ulikhetene skal ha noe å si for måling i de ulike kontekstene. Dersom personene

derimot er utenfor konfidensintervallet vil det si at målinger gjort i de ulike kontekstene vil være såpass ulike at de er signifikant forskjellige. Resultatene av mine analyser kan sees i Figur 7 og Figur 8.



FIGUR 7 – SCATTERPLOT AV SYKEHUSANSATTES EGNE MÅL (X-AKSEN) MOT DERES MÅL MED MÅLESTOKK FRA UNIVERSITETSSTUDENTENE (Y-AKSEN). VI KAN SE AT SAMTLIGE PERSONER ER INNEFOR KONFIDENSINTERVALLET.

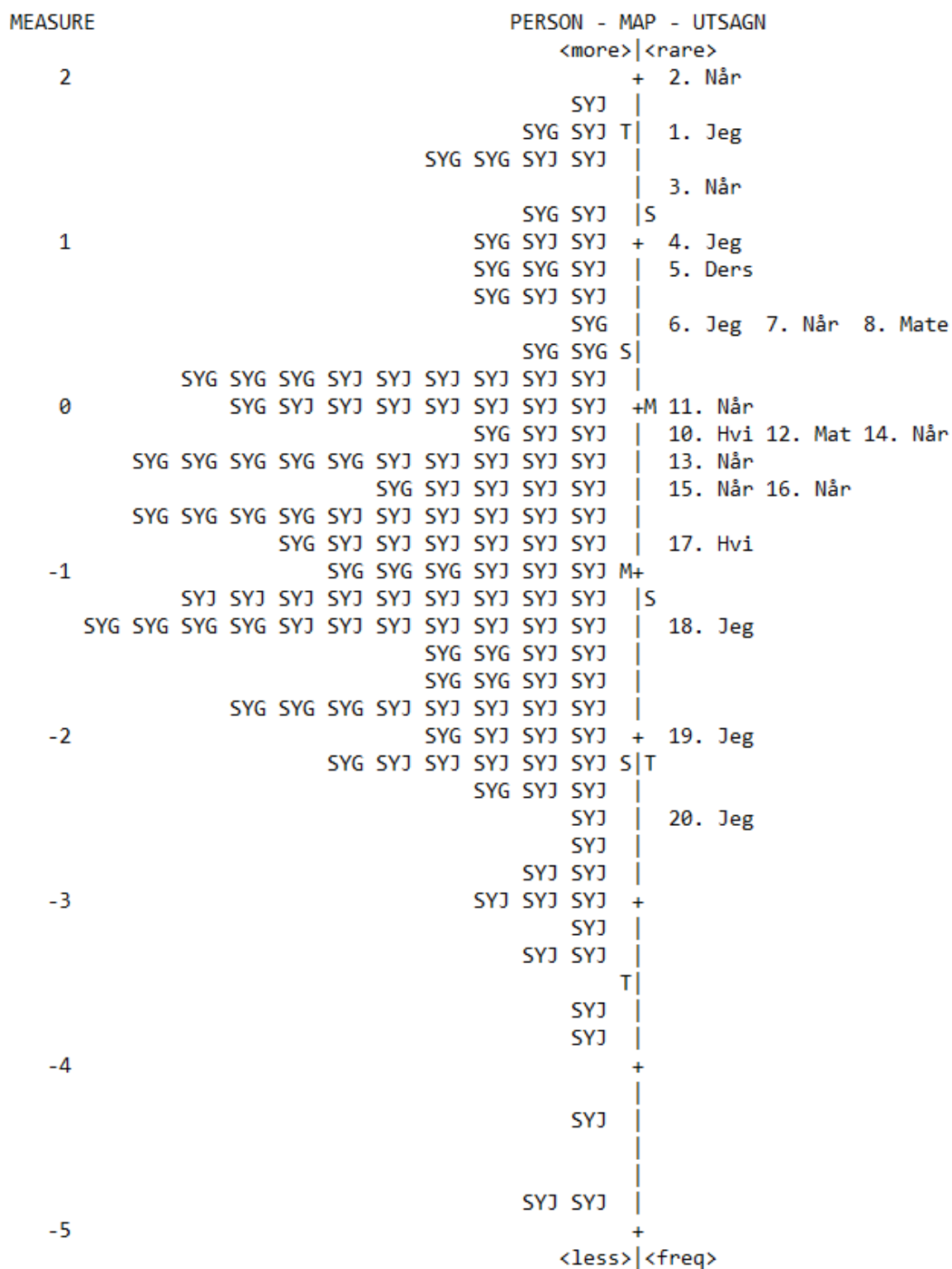


FIGUR 8—SCATTERPLOT AV SYKEHUSANSATTES EGNE MÅL (X-AKSEN) MOT DERES MÅL MED MÅLESTOKK FRA UNGDOMSSKOLEELEVENE (Y-AKSEN). OGSÅ I DETTE TILFELLET ER SAMTLIGE INNENFOR KONFIDENSINTERVALLET.

Slik det fremstår i de to «scatterplottene» er samtlige responser innenfor konfidensintervallet. Dermed er svaret på spørsmålet om det har noe å si for målingen at enkelte utsagn er signifikant ulike i to kontekster et tydelig nei. Personene får et tilnærmet likt resultat uavhengig av om de blir målt med en målestokk konstruert ved en ungdomsskole, et universitet eller et sykehus.

4.6 Resultat av målingen

Den sosiale MI til de sykehusansatte kommer til uttrykk i Wright-kartet i Figur 9. Kartet illustrerer påstandenes mål og deres plassering relatert til hverandre. Påstandenes nummer henger igjen fra hvordan de plasserte seg under målingen av ingeniørstudenter i Kaspersen et al. (2017). Den nummererte påstanden som har høyest mål og som dermed er ansett å være vanskeligst å være enig i, er plassert øverst og videre kommer resten av påstandene i synkende vanskegrad. Noen av påstandene har fått et så likt mål at de legger seg ved siden av hverandre i kartet. I tilfellene hvor det forekommer, er påstanden lengst til venstre på linjen vektlagt høyst, men det har lite å si for helheten da påstandene på linjen er tilnærmet likt vektlagt. Dersom vi ser på plasseringen til påstandene hos de sykehusansatte kan vi se at de legger seg nokså likt med plasseringen deres i den tidligere studien. Den personlige matematiske identiteten til de sykehusansatte kommer til uttrykk ved å se hvor de ansatte plasserer seg, hvilke påstander som er under dem og hvilke som er over dem. Vi kan blant annet se at den største andelen av sykehusansatte legger seg et sted mellom påstand 11 og 19 og har dermed et mål mellom cirka 0 og -2 logits.



FIGUR 9 - WRIGHT-KART AV SYKEHUSANSATTES SOSIALE MI. HVER SYG/SYJ ER EN SYKEHUSANSATT. DE ANSATTE ER TIL VENSTRE OG PÅSTANDENE ER TIL HØYRE.

5. Diskusjon

I dette kapitlet vil jeg først diskutere hva resultatene peker mot for den første av mine to problemstillinger; hvordan kan en måle MI hos ansatte ved et sykehus? Deretter vil jeg ta for meg den andre av de to; er strukturen til den matematiske identiteten til sykehusansatte invariant, sammenlignet med ungdomsskoleelvers og ingeniørstudenters? Videre diskuterer jeg noen didaktiske og forskningsmessige implikasjoner mine konklusjoner har. Gjennom arbeidet med denne studien har jeg gjort meg noen tanker om hva som kan forbedres dersom denne studien skal gjenskapes. Disse tankene vil jeg også diskutere i dette kapitlet. Jeg vil til slutt diskutere hvilke muligheter svarene jeg har kommet fram til åpner opp for videre forskning.

5.1 Hva peker resultatene mot?

For å svare på den første problemstillingen trengte jeg et eksempel på en metode for å måle MI hos ansatte ved et sykehus. Det eksempelet jeg har valgt å bruke er et instrument i form av et spørreskjema med 20 påstander. For å vise at dette instrumentet kan fungere som et eksempel på en måte å måle MI ved et sykehus, har jeg gjennomført en undersøkelse. Datamaterialet har jeg analysert ved å bruke Wolfe og Smiths (2007) rammeverk som legger frem aspekter som tar for seg ulike validitetsargument. De ulike aspektene har tilknyttede tester som ved å gjennomføres på et datamateriale, kan gi et innblikk i om datamaterialet er til å stole på.

Når det gjelder prinsippet om én-dimensjonalitet viser analysen at det er noe grunnlag for å mistenke at påstandene inneholder flere underdimensjoner. Eigenverdien på 2,2 er ikke langt over grensen på 2,0. Likevel bør det nevnes som et uromoment i validitetsspørsmålet og kravet til at instrumentet skal være én-dimensjonalt. Kravet om én-dimensjonalitet er dog et uoppnåelig mål da så godt som alle måleinstrumenter er påvirkelige av eksterne dimensjoner. Et termometer har for eksempel til hensikt å måle temperatur, men kan i noen tilfeller være påvirkelig av trykkforskjeller som gjør at det i den aller strengeste forståelsen av dimensjonalitet er flerdimensjonalt. Likevel anser vi et termometer for å være godt nok til å benytte i det daglige. Det blir et spørsmål om hva som er nøyaktig nok for den hensikten måleinstrumentet er tiltenkt.

Selv om det er antydning til flerdimensjonalitet, betyr ikke det at instrumentet ikke er mulig å benytte for å måle sykehuspersonell. En mulig løsning er å måle de ansatte to ganger, en med de påstandene som tilhører den ene underdimensjonen og en med de resterende påstandene. Dette vil gjøre at de ansattes personlige MI består av to mål i stedet for bare ett. Visuelt sett vil personenes mål da fremstilles som et punkt hvor kordinatene til punktet tilhører hver sin dimensjon.

Et annet uromoment i studien er at DIF-analysen på kjønn viste noen små tegn til forskjeller mellom kjønnene. Verdiene var som tidligere nevnt ikke langt over grenseverdien, men likefullt over. Årsaken til at denne ulikheten i analyseresultatet kan være tilfeldige ulikheter blant de ansatte ved akkurat dette sykehuset. En annen forklaring kan derimot være at min antagelse om at MI var tilnærmet uavhengig av kjønn var feil og at det i realiteten er en forskjell mellom kjønn på dette området. PISA-undersøkelsene rapporterer blant annet at norske jenter sliter signifikant oftere med matematikkangst enn gutter. I følge Kislenko (2009), som baserer seg på PISA-rapporten fra 2003, er norske gutter mindre redde for å gjøre feil og jentene engster seg mer for prøver i matematikk. På den andre siden presterer jenter jevnt over bedre enn

gutter i matematikkfaget i følge den nyeste PISA-rapporten fra 2018 (Jensen et al., 2019). Det er med andre ord grunn til å tro at det finnes kjønnsforskjeller i Norge når det kommer til matematikk, også når det gjelder affektive sider ved faget. Denne eventuelle forskjellen er noe som gjør at kjønn muligens ikke er en egnet måte å separere respondenter for å gjøre en DIF-analyse som skal begrunne validitet. I denne studien vurderte jeg at dette uromomentet var for lite til å påvirke resultatene i særlig stor grad, men det er et unødvendig moment da det muligens kan unngås ved å benytte en annen, mindre diskutabel måte å dele respondentene inn i to grupper.

Etter å ha gjennomgått testene på datamaterialet uten å oppdage større avvik enn tidligere nevnt fra hva Wolfe og Smith (2007) anbefaler som grenseverdier, har jeg dermed grunnlag for å påstå at instrumentet er et eksempel på hvordan en kan måle MI hos ansatte ved et sykehus. Jeg vil poengtere at jeg ikke hevder at dette instrumentet er det eneste eller det beste instrumentet til å måle MI hos sykehuspersonell. Antakeligvis finnes det flere som er like gode eller bedre, men i denne sammenhengen viser resultatene at dette instrumentet fungerer tilfredstillende til å måle MI ved et sykehus.

For å videre besvare den andre problemformuleringen min, gjennomførte jeg en DIF-analyse på datamaterialet mitt og de tidligere innsamlede datamaterialene fra ungdomsskole og universitet. Resultatene av denne undersøkelsen var at det til dels var liten forskjell mellom de sykehusansatte og universitetsstudentene, kun to av utsagnene var signifikant ulike. Forskjellene var derimot større mellom de sykehusansatte (og dermed også universitetsstudentene) og ungdomsskoleelevene. Disse resultatene peker dermed mot at de elleve karakteristikkene ikke er noe som er kontekstuavhengig. En ungdomsskoleelev har enten en annen vektlegging eller en annen oppfatning av hva det vil si å for eksempel studere et matematisk bevis/forklaring til det gir mening.

Noen av utsagnene forstås likt på tvers av kontekstene, andre ikke. Av de originale 20 utsagnene ble tolv av dem avdekket for å ikke være uavhengige av konteksten de ble brukt i. De åtte gjenstående utsagnene hadde heller ikke en nøyaktig lik måling på tvers av kontekstene, men de var like nok i forhold til hva vi forventer å se gitt tilfeldigheter. Det at de åtte utsagnene forholdt seg tilnærmet likt i de ulike kontekstene peker mot at innholdet i dem blir forstått likt av medlemmene i dem. Dette tyder på at en ungdomsskoleelev mener det samme som en sykehusansatt eller en ingeniørstudent når han eller hun for eksempel sier at de har problemer med å legge fra seg matematiske oppgaver. Dette funnet peker mot at det finnes både sider ved MI som forstås likt på tvers av kontekster og sider som forstås ulikt.

5.2 Implikasjoner

En implikasjon av at noen underdimensjoner av MI forstås likt på tvers av kontekster og andre ikke gjør det er at vi må være nøye med hvilke karakteristikker vi benytter for å gjøre en måling. Selv om de fleste av påstandene ble funnet til å være brukbare i samtlige kontekster, er målinger basert på samtlige av disse påstandene ikke nødvendigvis brukbare til å sammenligne målinger på tvers av kontekster. Grunnen til det er at betydningen av enkelte av påstandene er for ulike i de respektive kontekstene til at det skal være meningsfylt å sammenligne mål basert på dem. Dersom vi ønsker oss et instrument som er anvendbart uavhengig av kontekst, må vi benytte oss av påstander som vi vet forstås tilnærmet likt uansett hvem du spør.

De åtte påstandene som ikke var signifikant ulike på tvers av kontekstene, kan være et steg nærmere mot å finne en universell forståelse av hva MI innebærer. Et annet utfall

er at selv om disse viser seg å bli oppfattet likt i disse tre kontekstene denne oppgaven ser på, kan det være andre kontekster som ikke har samme fortolkning og vektlegging.

5.2.1 Didaktiske implikasjoner

Et mulig utbytte av denne studien i en didaktisk sammenheng er at vi nå vet om noen karakteristikk som er felles for ungdomsskoleelever, ingeniørstudenter og sykehusansatte. Det gjenstår enda å se om disse karakteristikene også gjelder for andre yrkesgrupper, men den foreløpige informasjonen vi har er at de åtte påstandene i alle fall gjelder på tvers av skole og arbeid. Under den kvalitative analysen av de åtte påstandene ble strategier og forståelse trukket frem som mulige fellesfaktorer. Jeg vil derfor argumentere for at disse to temaene er viktig å vektlegge i skolen og matematikkfaget, ettersom de oppleves som relevante både for elevene og for arbeidstakerene.

De to mulige fellesfaktorene forståelse og strategier er også noe som er vektlagt i fagfornyelsen. Utdanningsdirektoratet skriver blant annet: «I læreplanen er algoritmisk tenkning synliggjort fordi dette er en viktig problemløsningsstrategi. Når elevene bruker programmering til å utforske og løse problemer, kan det være et godt verktøy for å utvikle matematisk forståelse.» (Utdanningsdirektoratet, 2020). Algoritmisk tenkning og programmering er begreper som inngår i den nye læreplanen. Disse begrepene mener Utdanningsdirektoratet er viktige, både for å utvikle strategier for problemløsning og for å få en dyp og god forståelse av faget matematikk. Dermed vil jeg, ut ifra resultatene av denne studien, argumentere for at Utdanningsdirektoratet oppfyller sitt ansvar om å gjøre skolen relevant og forberede elevene på arbeidslivet. Dette fordi Utdanningsdirektoratet i revideringen av læreplanen vektlegger de samme begrepene som jeg i denne studien kommer fram til at er felles for ungdomsskoleelevene, ingeniørstudentene og de sykehusansatte.

5.2.2 Forskningsmessige implikasjoner

Resultatene av denne forskningen kan være med på å utvide forskningsfeltet MI. Analysen av datamaterialet avdekket blant annet at infit- og outfit-verdiene til utsagnene er innenfor det Rasch-modellen forventer, med unntak av utsagn 9. Dette impliserer at utsagnene egner seg til å måle MI ved det utvalgte sykehuset og til en viss grad at utsagnene egner seg for å måle MI ved konteksten sykehus generelt. Dersom vi generaliserer enda mer kan vi si at utsagnene tilsynelatende egner seg for å måle MI på arbeidsplasser generelt, men for å anta det trengs det nok flere studier som undersøker andre ulike kontekster.

Et annet resultat er dimensjonaliteten. Analysen avdekket en antydning til flerdimensjonalitet, noe som i følge Thurstone (1959) ikke er ønskelig. Denne antydningen var ikke å finne i de to andre kontekstene ungdomsskole og universitet. En mulig implikasjon av dette kan være at en må regne med at det oppstår flere underdimensjoner i målinger gjort på sykehus og kanskje også arbeidsplasser generelt. En måte å bøte på underdimensjonen er å si at MI består av to dimensjoner og videre måle hver dimensjon for seg med de utsagnene som tilhører dem. En persons mål vil da bestå av to tall som hver tilhører en dimensjon. På den måten opprettholder en kravet om én-dimensjonalitet, fordi en da kun betrakter en dimensjon av gangen. Ulempen ved det er at det kan bli vanskeligere å sammenligne personer ettersom de ikke lenger legger seg på en rett meterstokk, men som punkter i en graf.

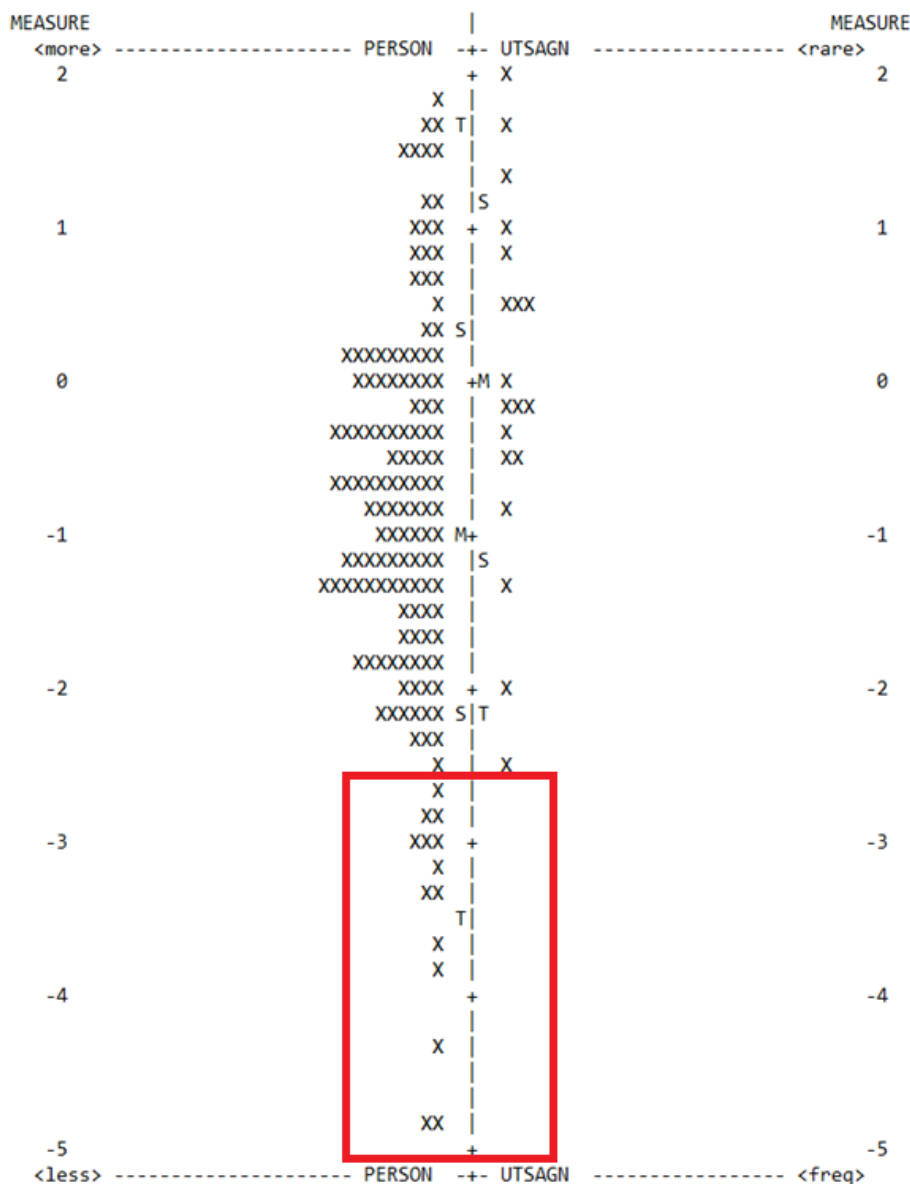
En mulig årsak til denne flerdimensjonaliteten kan være at i undervisningssammenheng har den sosiale MI en mer universell og endimensjonal fremtoning ettersom

undervisningsinstitusjoner gjerne har et felles mål og en felles oppfatning av hvordan matematikk skal gjøres og dermed hvordan den oppleves. I yrkessammenheng kan denne fremtoningen derimot være ulik fra person til person, ettersom hovedmålet med matematikk ikke er å lære, men å dra nytte av den for å utføre arbeidsoppgaver som kan variere mye ut ifra hvilken stilling en har.

5.3 Mangler ved og forbedringer av studien

5.3.1 Utsagnenes formuleringer

I analysen av datamaterialet fra de sykehusansatte kom det frem at en nokså stor andel av gruppen falt under målet til det enkleste utsagnet. Det at enkelte havner under der målestokken opererer er nesten umulig å unngå uten at det skal få konsekvenser for validiteten til målingen, og dermed noe en må regne med ved en slik undersøkelse. Det er likevel et uromoment at flere havner under denne grensen da det medfører at en andel av gruppen ikke får en nøyaktig måling. Fra tidligere studier, deriblant Kaspersen et al. (2017), kan vi se at det er en tendens til at dette skjer på den lave enden av skalaen, men stort sett ikke på den høye med akkurat disse utsagnene. Dette kan tyde på at utsagnene «sikter» litt for høyt, eventuelt for snevert, for å få en tilnærmet optimal måling.



FIGUR 10 - WRIGHT-KART AV PERSONER OG UTSAGN I STUDIEN. DE SOM HAR LAVERE MÅL ENN MÅLESTOKKEN ER MARKERT I RØD BOKS.

En mulig måte å forhindre at så mange havner under målestokken kan være å erstatte enkelte av utsagnene med noen som det kan antas at er enkle for de aller fleste å være enige i, eventuelt å legge til noen slike påstander. På den måten forlenger vi muligens målestokken nok til at så godt som alle kan få en nøyaktig måling av sin MI, som igjen vil medføre en mer nøyaktig og representativ statistikk.

5.3.2 De åtte aspektene

De åtte aspektene som Wolfe og Smith (2007) presenterer i sine artikler er alle ment som bidragsyttere til å sikre kvantitative måledatas validitet og reliabilitet. I min studie har jeg valgt å fokusere på fem av dem. Min begrunnelse for dette er at jeg ikke anså de tre jeg har utelatt som særlig relevante for min studie. Dette kan sees på som en svakhet ved min studie da det gjør mitt datagrunnlag, og dermed også mine funn, mindre pålitelig. Jeg vil samtidig få påpeke at vi i forskning basert på måling som regel ikke kan være 100% sikre på noe som helst (Kuhn, 1961), slik at selv om jeg hadde

vektlagt alle åtte aspekter hadde vi ikke kunnet garantere at datagrunnlaget var valid og helt til å stole på. Det ville kun potensielt økt validiteten og muligens gjort oss mer sikre. Min anbefaling til eventuell videre forskning innenfor denne tematikken er i planleggingsfasen av forskningsmetoden å tenke på hvordan en kan sørge for at alle åtte aspekter kan tas i bruk.

I tillegg til de tre utelatte aspektene vil jeg nevne valget av kjønn som måte å lage to undergrupper av respondentene til å anskaffe bevis for generaliserbarhetsaspektet som en svakhet ved studien. Det medførte at verdiene oversteg den fastsatte grenseverdien og er dermed ikke et like sterkt bevis som jeg hadde ønsket. Dersom jeg skulle gjenskapt studien ville jeg erstattet spørsmålet om kjønn i spørreundersøkelsen med noe som nok har mye mindre å si for en persons identitet, for eksempel om datoen en er født er et partall eller et oddetall.

5.4 Videre forskning

5.4.1 Etisk utfordring knyttet til gjenbruk av datamateriale

I informasjonsskrivet som fulgte med spørreskjemaet ble det blant annet opplyst om at det var NTNU som var ansvarlig for studien og at datamaterialet ville bli overlevert til NTNU etter prosjektets slutt for lagring og videre bruk i nye forskningsprosjekt. Et diskuterbart moment når det gjelder gjenbruk av datamaterialet er at respondentene skal ha innsikt i hva deres responser blir benyttet til, og de skal også ha muligheten til å trekke sin besvarelse dersom de ombestemmer seg og trekker sitt samtykke til å være med. Ettersom spørreskjemaet ikke samlet inn personidentifiserende data ville det vært svært utfordrende å finne tilbake til en spesifikk respons dersom en respondent ønsket å trekke den. Respondenten måtte i så tilfelle huske oppimot samtlige av alle sine responser på skjemaet, men i et slikt tilfelle kunne det ha vært gjort. Det hadde imidlertid vært umulig å fjerne en respons dersom datamaterialet ble gitt videre til NTNU eller NSD slik at det lå tilgjengelig ute for enhver som måtte ønske å laste det ned. Jeg kan ikke tilby noen løsning på dette problemet, om det i det hele tatt er et problem, men jeg ønsker å påpeke det til videre forskning.

5.4.2 Veien videre

Denne oppgaven er ment som et bidrag til et større spørsmål innen forskning på MI: Er MI kontekstuavhengig eller ikke? Vi vil nok aldri være helt sikre på om det er en struktur på MI som er felles for alle, ettersom vi ikke kan måle alle personer og alle kontekster. Det kan imidlertid oppdages bevis for at de kontekstuavhengige påstandene jeg har funnet i denne studien, ikke er kontekstuavhengige likevel. Dersom en gjorde lignende studier i andre kontekster hvor det viste seg at de åtte påstandene som ble analysert til å være kontekstuavhengige i denne oppgaven, ikke ble forstått og ansett som like vanskelig som i min studie, ville det antydning at MI ikke er kontekstuavhengig. Dersom disse åtte påstandene fikk det samme målet i andre kontekster er det derimot større belegg for å anta at det eksisterer en felles forståelse av MI som ikke er avhengig av konteksten den blir målt i. Flere lignende studier kan altså være til nytte for å besvare spørsmålet om kontekstuavhengighet. Dersom lignende studier skulle bli gjort i andre kontekster, vil jeg anbefale å forsøke å måle i en kontekst som har minst mulig til felles med de allerede målte kontekstene. En slik kontekst kan skille seg ut for eksempel ved at deltakerne i den er i en annen alder, har en annen type utdanning eller har andre typer arbeidsoppgaver. Grunnen til at jeg vil anbefale å studere så ulike kontekster som mulig er for å skape et så komplett bilde av MI som overhodet mulig. Et mer helhetlig

bilde vil gjøre oss mer i stand til å svare på spørsmål relatert til MI og spørsmålet om det er kontekstuavhengig.

Fremtidig forskning burde gjøre tiltak for å sikre de validitetsaspektene jeg ikke har gjort noe/nok for å sikre. De utelatte aspektene er konsekvensaspektet, responsivitetsaspektet og fortolkningsaspektet. Disse er mer knyttet til bruken av måleinstrumentet og hvordan det fungerer når en måler samme gruppe over tid. Dermed kan en annen mulig retning å ta forskningen på MI være å forsøke å se på overgangen fra en kontekst til en annen.

I min studie sammenligner jeg ungdomsskoleelever med sykehusansatte. Disse to gruppene har ikke i utgangspunktet noe med hverandre å gjøre. En annen studie kunne for eksempel målt studenter ved et universitet eller en høyskoles MI og deretter målt de samme studenter etter at de var ferdig studert og i arbeid. En slik studie kunne hatt et nærmere blikk på den personlige matematiske identiteten til deltakerne og sammenlignet den med de to sosiale matematiske identitetene i de to kontekstene. Studien ville også vært et godt supplement til spørsmålet om kontekstuavhengighet da en forhåpentligvis ville kunne se om personenes MI forble uendret, eller om de tilpasset seg den nye sosiale strukturen i den nye konteksten. Gitt sistnevnte utfall ville det vært interessant å undersøke hvilke eventuelle endringer som oppstod og hvor store de eventuelle endringene var.

6. Litteraturliste

Andrich, D. (1989). Distinctions between assumptions and requirements in measurement in the social sciences. *Mathematical and theoretical systems*, 4, 7-16.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (3. Utg.). New York: Routledge.

Burton, D. M. (2011). *The history of mathematics: an introduction* (7. Utg.). McGraw-Hill.

Darragh, L. (2016). Identity research in mathematics education. *An International Journal*, 93(1), 19-33.

Deaux, K. (1993). Reconstructing social identity. *Personality and social psychology bulletin*, 19(1), 4-12.

Graven, M. & Heyd-Metzuyanım, E. (2019). Mathematics identity research: the state of the art and future directions. *Mathematics Education*, 51(3), 361-377.

Jensen, F., Pettersen, A. Frønes, T. S., Kjærnsli, M., Rohatgi, A., Eriksen, A. & Narvhus, E.K. (2019). PISA 2018. *Norske elevers kompetanse i lesing, matematikk og naturfag*. Oslo: Universitetsforlaget.

John Michael Linacre. (2017,02. februar). *Introduction to the Rasch Model. Benjamin D. Wright, 1994. All - Parts 1+2+3*. YouTube. <https://youtu.be/K8fdBLGfhxc>

Kaspersen, E. (2015). Using the Rasch Model to Measure the Extent to which Students Work Conceptually with Mathematics. *Journal of applied measurement*, 16(4), 336-352.

Kaspersen, E., Pepin, B. & Sikko, S. A. (2017). Measuring student teachers' practices and beliefs about teaching mathematics using the Rasch model. *International Journal of Research & Method in Education*, 40(4), 421-442.

Kislenko, K. (2009). An investigation of Norwegian students' affective domain in mathematics. *Nord Stud Math Educ*, 14(4), 33-64.

Kuhn, T. S. (1961). The function of measurement in modern physical science. *Isis*, 52(2), 161-193.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of applied measurement*, 3(1), 85-106.

Linacre, J. M. (2012, juni). *Winsteps Tutorial 1*. Winsteps.
<https://www.winsteps.com/a/winsteps-tutorial-1.pdf>

Linacre, J.M. (2020). Winsteps® (Version 4.5.1) [Computer Software]. Beaverton, Oregon: Winsteps.com. Hentet Januar 1, 2020.

Linacre, J.M. (u.å.). *Table 30.1 Differential item functioning DIF pairwise*. Winsteps.
https://www.winsteps.com/winman/table30_1.htm

Noss, R., Hoyles, C. & Pozzi, S. (2002). Working knowledge: Mathematics in use. I A. Bessot & J. Ridgeway (Red.), *Education for Mathematics in the Workplace. Mathematics Education Library* (1. utg., s. 17-35). Springer Netherlands: Imprint: Springer.
https://doi.org/10.1007/0-306-47226-0_3

Opplæringslova – oppl. (1998). *Lov om grunnskolen og den vidaregåande opplæringa (opplæringslova)*. (LOV-1998-07-17-61). Lovdata. <https://lovdata.no/lov/1998-07-17-61>

Pepin, B. (2011). Pupils' attitudes towards mathematics: a comparative study of Norwegian and English secondary students. *ZDM*, 43(4), 535-546.

Thurstone, L. L. (1959). *The measurement of values*. University of Chicago Press.

Utdanningsdirektoratet. (2020, 03. september). *Hva er nytt i matematikk?*.
<https://www.udir.no/laring-og-trivsel/lareplanverket/fagspesifikk-stotte/nytt-i-fagene/hva-er-nytt-i-matematikk/>

Wolfe, E. W., & Smith, E. E. (2007a). Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part I-Instrument Development Tools. *Journal of Applied Measurement*, 8(1), 97-123

Wolfe, E. & Smith, J. E. (2007b). Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part II-Validation Activities. *Journal of applied measurement*, 8(2), 204-234.

Wright, B. D. & Stone, M. H. (1979). *Best test design*. Mesa Press.

Ytterhaug, B. O. (2019). *Matematisk identitet i ungdomsskolen* [Masteroppgave]. Norges teknisk-naturvitenskapelige universitet.

Vedlegg

Vedlegg 1: Svarskjema A

| | Aldri/nesten aldri | Noen ganger | Ofte | Alltid/nesten alltid | Vet ikke |
|--|--------------------|-------------|------|----------------------|----------|
| 1. Jeg har problemer med å legge fra meg matematiske oppgaver. | | | | | |
| 2. Dersom jeg har glemt en formel/metode, prøver jeg å utlede den selv. | | | | | |
| 3. Når jeg lærer en ny metode/algoritme, prøver jeg å finne ut hvorfor den virker. | | | | | |
| 4. Hvis jeg prøver på en metode som ikke fører frem, bruker jeg tid på å finne ut hvorfor denne ikke virker. | | | | | |
| 5. Hvis jeg står fast, prøver jeg å visualisere problemet. | | | | | |
| 6. Jeg prøver å koble det jeg lærer opp mot det jeg vet fra før. | | | | | |
| 7. Jeg fortsetter å prøve meg frem selv om jeg ikke får det til med en gang. | | | | | |
| 8. Når jeg lærer noe nytt, stiller jeg meg selv egne spørsmål som jeg jobber med. | | | | | |
| 9. Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobber legger opp til. | | | | | |
| 10. Når jeg lærer en ny metode, prøver jeg å finne situasjoner hvor denne ikke virker. | | | | | |
| 11. Når jeg jobber med et matematisk problem hopper jeg mellom ulike strategier. | | | | | |
| 12. Når jeg lærer en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gjøre. | | | | | |
| 13. Jeg blir engasjert når noen starter en matematisk diskusjon. | | | | | |
| 14. Når jeg jobber med en oppgave, stopper jeg opp underveis og reflekterer over hva jeg gjør. | | | | | |
| 15. Når jeg kommer over et matematisk bevis/forklaring, studerer jeg det til det gir mening. | | | | | |
| 16. Når jeg lærer en ny metode, bruker jeg tid på å se om jeg kan finne en bedre metode. | | | | | |
| 17. Når jeg lærer noe nytt, fører det til at det er flere ting jeg ønsker å finne ut. | | | | | |
| 18. Når jeg møter et matematisk problem, tenker jeg over om det finnes flere måter å løse oppgaven på. | | | | | |
| 19. Jeg kan forklare hvorfor løsningen min er rett. | | | | | |
| 20. Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tankerekker. | | | | | |

Kjønn: _____

Yrkestittel (sett kryss):

| | | | |
|--|--|---|--|
| Sykepleier | | Lege | |
| Sykepleier med spesialisering/etterutdanning | | Service/støttefunksjon (Eks. Administrasjon, Økonomi, Kjøkken etc.) | |

Vil du delta i forskningsprosjektet «Matematisk identitet i ulike kontekster»?

Formål

Jeg er masterstudent ved NTNU, Institutt for Lærerutdanning, og jeg skal gjennomføre et forskningsprosjekt om matematisk identitet i ulike kontekster. Disse påstandene handler om hva du tenker eller gjør når du arbeider med noe knyttet til matematikk.

Hvem er ansvarlig for forskningsprosjektet?

NTNU, Institutt for Lærerutdanning

Hvorfor får du spørsmål om å delta?

Fordi du er ansatt i helsevesenet som er en annen kontekst enn skolevesenet.

Hva innebærer det for deg å delta?

Det innebærer å ta stilling til og krysse av for hvor ofte 20 påstander gjelder deg selv samt å oppgi kjønn og yrke.

Det er frivillig å delta

Det er selvsagt frivillig å delta og dermed har du også rett til å trekke deg og din besvarelse fra studien når som helst.

Hvordan vi oppbevarer og bruker dine opplysninger

Dette skjemaet vil bli samlet inn, digitalisert og lagret hos NTNU.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- NTNU, Institutt for Lærerutdanning ved Eivind Kaspersen [eivind.kaspersen@ntnu.no].
- Masterstudent Karl Hegna [karl.hegna.95@gmail.com]

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Matematisk identitet i ulike kontekster» og har fått anledning til å stille spørsmål.

Jeg samtykker til:

- å delta i spørreundersøkelsen.
- at svarene i spørreskjema lagres etter prosjektslutt hos NTNU, til senere forskningsprosjekter.

Med vennlig hilsen Karl Hegna

Vedlegg 2: Svarskjema B

| | Aldri/nesten aldri | Noen ganger | Ofte | Alltid/nesten alltid | Vet ikke |
|--|--------------------|-------------|------|----------------------|----------|
| 1. Når jeg møter et matematisk problem, tenker jeg over om det finnes flere måter å løse oppgaven på. | | | | | |
| 2. Når jeg lærer en ny metode, bruker jeg tid på å se om jeg kan finne en bedre metode. | | | | | |
| 3. Hvis jeg prøver på en metode som ikke fører frem, bruker jeg tid på å finne ut hvorfor denne ikke virker. | | | | | |
| 4. Når jeg lærer noe nytt, stiller jeg meg selv egne spørsmål som jeg jobber med. | | | | | |
| 5. Når jeg lærer en ny metode/algoritme, prøver jeg å finne ut hvorfor den virker. | | | | | |
| 6. Når jeg jobber med et matematisk problem hopper jeg mellom ulike strategier. | | | | | |
| 7. Når jeg lærer en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gjøre. | | | | | |
| 8. Når jeg kommer over et matematisk bevis/forklaring, studerer jeg det til det gir mening. | | | | | |
| 9. Jeg blir engasjert når noen starter en matematisk diskusjon. | | | | | |
| 10. Dersom jeg har glemt en formel/metode, prøver jeg å utlede den selv. | | | | | |
| 11. Jeg fortsetter å prøve meg frem selv om jeg ikke får det til med en gang. | | | | | |
| 12. Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tanker. | | | | | |
| 13. Når jeg lærer en ny metode, prøver jeg å finne situasjoner hvor denne ikke virker. | | | | | |
| 14. Jeg har problemer med å legge fra meg matematiske oppgaver. | | | | | |
| 15. Jeg prøver å koble det jeg lærer opp mot det jeg vet fra før. | | | | | |
| 16. Jeg kan forklare hvorfor løsningen min er rett. | | | | | |
| 17. Hvis jeg står fast, prøver jeg å visualisere problemet. | | | | | |
| 18. Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobber legger opp til. | | | | | |
| 19. Når jeg jobber med en oppgave, stopper jeg opp underveis og reflekterer over hva jeg gjør. | | | | | |
| 20. Når jeg lærer noe nytt, fører det til at det er flere ting jeg ønsker å finne ut. | | | | | |

Kjønn: _____

Yrkestittel (sett kryss):

| | | | |
|--|--|---|--|
| Sykepleier | | Lege | |
| Sykepleier med spesialisering/etterutdanning | | Service/støttefunksjon (Eks. Administrasjon, Økonomi, Kjøkken etc.) | |

Vil du delta i forskningsprosjektet «Matematisk identitet i ulike kontekster»?

Formål

Jeg er masterstudent ved NTNU, Institutt for Lærerutdanning, og jeg skal gjennomføre et forskningsprosjekt om matematisk identitet i ulike kontekster. Disse påstandene handler om hva du tenker eller gjør når du arbeider med noe knyttet til matematikk.

Hvem er ansvarlig for forskningsprosjektet?

NTNU, Institutt for Lærerutdanning

Hvorfor får du spørsmål om å delta?

Fordi du er ansatt i helsevesenet som er en annen kontekst enn skolevesenet.

Hva innebærer det for deg å delta?

Det innebærer å ta stilling til og krysse av for hvor ofte 20 påstander gjelder deg selv samt å oppgi kjønn og yrke.

Det er frivillig å delta

Det er selvsagt frivillig å delta og dermed har du også rett til å trekke deg og din besvarelse fra studien når som helst.

Hvordan vi oppbevarer og bruker dine opplysninger

Dette skjemaet vil bli samlet inn, digitalisert og lagret hos NTNU.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- NTNU, Institutt for Lærerutdanning ved Eivind Kaspersen [eivind.kaspersen@ntnu.no].
- Masterstudent Karl Hegna [karl.hegna.95@gmail.com]

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Matematisk identitet i ulike kontekster» og har fått anledning til å stille spørsmål.

Jeg samtykker til:

- å delta i spørreundersøkelsen.
- at svarene i spørreskjema lagres etter prosjektslutt hos NTNU, til senere forskningsprosjekter.

Med vennlig hilsen Karl Hegna

Vedlegg 3: Svarskjema C

| | Aldri/nesten aldri | Noen ganger | Ofte | Alltid/nesten alltid | Vet ikke |
|---|--------------------|-------------|------|----------------------|----------|
| 1. Når jeg lærer en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gjøre. | | | | | |
| 2. Når jeg jobber med en oppgave, stopper jeg opp underveis og reflekterer over hva jeg gjør. | | | | | |
| 3. Når jeg jobber med et matematisk problem, hopper jeg mellom ulike strategier. | | | | | |
| 4. Når jeg lærer en ny metode/algoritme, prøver jeg å finne ut hvorfor den virker. | | | | | |
| 5. Når jeg lærer noe nytt, fører det til at det er flere ting jeg ønsker å finne ut. | | | | | |
| 6. Når jeg lærer en ny metode, prøver jeg å finne situasjoner hvor denne ikke virker. | | | | | |
| 7. Jeg har problemer med å legge fra meg matematiske oppgaver. | | | | | |
| 8. Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobber legger opp til. | | | | | |
| 9. Når jeg møter et matematisk problem, tenker jeg over om det finnes flere måter å løse oppgaven på. | | | | | |
| 10. Jeg prøver å koble det jeg lærer opp mot det jeg vet fra før. | | | | | |
| 11. Jeg kan forklare hvorfor løsningen min er rett. | | | | | |
| 12. Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tankerekker. | | | | | |
| 13. Dersom jeg har glemt en formel/metode, prøver jeg å utlede den selv. | | | | | |
| 14. Når jeg lærer en ny metode, bruker jeg tid på å se om jeg kan finne en bedre metode. | | | | | |
| 15. Jeg fortsetter å prøve meg fram selv om jeg ikke får det til med en gang. | | | | | |
| 16. Jeg blir engasjert når noen starter en matematisk diskusjon. | | | | | |
| 17. Når jeg kommer over et matematisk bevis/forklaring, studerer jeg det til det gir mening. | | | | | |
| 18. Hvis jeg prøver på en metode som ikke fører fram, bruker jeg tid på å finne ut hvorfor denne ikke virker. | | | | | |
| 19. Hvis jeg sitter fast, prøver jeg å visualisere problemet. | | | | | |
| 20. Når jeg lærer noe nytt, stiller jeg meg selv egne spørsmål som jeg jobber med. | | | | | |

Kjønn: _____

Yrkestittel (sett kryss):

| | | | |
|--|--------------------------|---|--------------------------|
| Sykepleier | <input type="checkbox"/> | Lege | <input type="checkbox"/> |
| Sykepleier med spesialisering/etterutdanning | <input type="checkbox"/> | Service/støttefunksjon (Eks. Administrasjon, Økonomi, Kjøkken etc.) | <input type="checkbox"/> |

Vil du delta i forskningsprosjektet «Matematisk identitet i ulike kontekster»?

Formål

Jeg er masterstudent ved NTNU, Institutt for Lærerutdanning, og jeg skal gjennomføre et forskningsprosjekt om matematisk identitet i ulike kontekster. Disse påstandene handler om hva du tenker eller gjør når du arbeider med noe knyttet til matematikk.

Hvem er ansvarlig for forskningsprosjektet?

NTNU, Institutt for Lærerutdanning

Hvorfor får du spørsmål om å delta?

Fordi du er ansatt i helsevesenet som er en annen kontekst enn skolevesenet.

Hva innebærer det for deg å delta?

Det innebærer å ta stilling til og krysse av for hvor ofte 20 påstander gjelder deg selv samt å oppgi kjønn og yrke.

Det er frivillig å delta

Det er selvsagt frivillig å delta og dermed har du også rett til å trekke deg og din besvarelse fra studien når som helst.

Hvordan vi oppbevarer og bruker dine opplysninger

Dette skjemaet vil bli samlet inn, digitalisert og lagret hos NTNU.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- NTNU, Institutt for Lærerutdanning ved Eivind Kaspersen [eivind.kaspersen@ntnu.no].
- Masterstudent Karl Hegna [karl.hegna.95@gmail.com]

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Matematisk identitet i ulike kontekster» og har fått anledning til å stille spørsmål.

Jeg samtykker til:

- å delta i spørreundersøkelsen.
- at svarene i spørreskjema lagres etter prosjektslutt hos NTNU, til senere forskningsprosjekter.

Med vennlig hilsen Karl Hegna

Vedlegg 4: Svarskjema D

| | Aldri/nesten aldri | Noen ganger | Ofte | Alltid/nesten alltid | Vet ikke |
|--|--------------------|-------------|------|----------------------|----------|
| 1. Når jeg lærer noe nytt, stiller jeg meg selv egne spørsmål som jeg jobber med. | | | | | |
| 2. Jeg blir engasjert når noen starter en matematisk diskusjon. | | | | | |
| 3. Når jeg jobber med et matematisk problem hopper jeg mellom ulike strategier. | | | | | |
| 4. Når jeg lærer en ny metode, bruker jeg tid på å se om jeg kan finne en bedre metode. | | | | | |
| 5. Hvis jeg prøver på en metode som ikke fører frem, bruker jeg tid på å finne ut hvorfor denne ikke virker. | | | | | |
| 6. Matematiske ideer jeg leser eller hører om setter meg på sporet av egne tankerekker. | | | | | |
| 7. Når jeg lærer en ny matematisk metode, liker jeg å bli fortalt nøyaktig hva jeg skal gjøre. | | | | | |
| 8. Jeg kan forklare hvorfor løsningen min er rett. | | | | | |
| 9. Når jeg møter et matematisk problem, tenker jeg over om det finnes flere måter å løse oppgaven på. | | | | | |
| 10. Når jeg lærer en ny metode, prøver jeg å finne situasjoner hvor denne ikke virker. | | | | | |
| 11. Jeg prøver å koble det jeg lærer opp mot det jeg vet fra før. | | | | | |
| 12. Jeg fortsetter å prøve meg frem selv om jeg ikke får det til med en gang. | | | | | |
| 13. Når jeg lærer en ny metode/algoritme, prøver jeg å finne ut hvorfor den virker. | | | | | |
| 14. Når jeg lærer noe nytt, fører det til at det er flere ting jeg ønsker å finne ut. | | | | | |
| 15. Jeg har problemer med å legge fra meg matematiske oppgaver. | | | | | |
| 16. Dersom jeg har glemt en formel/metode, prøver jeg å utlede den selv. | | | | | |
| 17. Når jeg jobber med en oppgave, stopper jeg opp underveis og reflekterer over hva jeg gjør. | | | | | |
| 18. Jeg tar initiativ til å lære mer om et matematisk emne enn skole/jobben legger opp til. | | | | | |
| 19. Hvis jeg står fast, prøver jeg å visualisere problemet. | | | | | |
| 20. Når jeg kommer over et matematisk bevis/forklaring, studerer jeg det til det gir mening. | | | | | |

Kjønn: _____

Yrkestittel (sett kryss):

| | | | |
|--|--------------------------|---|--------------------------|
| Sykepleier | <input type="checkbox"/> | Lege | <input type="checkbox"/> |
| Sykepleier med spesialisering/etterutdanning | <input type="checkbox"/> | Service/støttefunksjon (Eks. Administrasjon, Økonomi, Kjøkken etc.) | <input type="checkbox"/> |

Vil du delta i forskningsprosjektet «Matematisk identitet i ulike kontekster»?

Formål

Jeg er masterstudent ved NTNU, Institutt for Lærerutdanning, og jeg skal gjennomføre et forskningsprosjekt om matematisk identitet i ulike kontekster. Disse påstandene handler om hva du tenker eller gjør når du arbeider med noe knyttet til matematikk.

Hvem er ansvarlig for forskningsprosjektet?

NTNU, Institutt for Lærerutdanning

Hvorfor får du spørsmål om å delta?

Fordi du er ansatt i helsevesenet som er en annen kontekst enn skolevesenet.

Hva innebærer det for deg å delta?

Det innebærer å ta stilling til og krysse av for hvor ofte 20 påstander gjelder deg selv samt å oppgi kjønn og yrkestittel.

Det er frivillig å delta

Det er selvsagt frivillig å delta og dermed har du også rett til å trekke deg og din besvarelse fra studien når som helst.

Hvordan vi oppbevarer og bruker dine opplysninger

Dette skjemaet vil bli samlet inn, digitalisert og lagret hos NTNU.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- NTNU, Institutt for Lærerutdanning ved Eivind Kaspersen [eivind.kaspersen@ntnu.no].
- Masterstudent Karl Hegna [karl.hegna.95@gmail.com]

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Matematisk identitet i ulike kontekster» og har fått anledning til å stille spørsmål.

Jeg samtykker til:

- å delta i spørreundersøkelsen.
- at svarene i spørreskjema lagres etter prosjektslutt hos NTNU, til senere forskningsprosjekter.

Med vennlig hilsen Karl Hegna

