



Uit

NORGES
ARKTISKE
UNIVERSITET

Fakultet for humaniora, samfunnsvitenskap og lærerutdanning

Lærernes vurderingspraksis i matematikkfaget

En kvalitativ studie fra Nord-Troms av matematikklæreres oppfatninger av vurdering i faget

—

Silje Kristiansen

Masteroppgave i lærerutdanning 5.-10. trinn, 2017

LRU-3903 Matematikdidaktikk



Forord

Å levere denne mastergradsoppgaven er for meg et mål som til tider har virket ganske fjernt. Et femårig utdanningsløp er fullført, og jeg kan endelig påta meg tittelen som lektor.

Arbeidet bak har vært krevende og frustrerende, men samtidig noe jeg anser som berikende både personlig og profesjonsmessig. Avhandlingen ville dog ikke vært et faktum foruten mine kjære informanter - en stor takk rettes til dere! Dere har bidratt med kunnskap som jeg ut over dette prosjektet vil ta med meg videre inn i yrkeslivet.

Jeg vil også rette en takk til min veileder Ove Drageset. Du har vært til stor hjelp og bidratt med eksemplarisk veiledning i en tid som har vært hektisk for oss begge.

Til sist ønsker jeg å takke for den moralske støtten fra både familie, venner og kjæreste. Dere har vist forståelse når jeg har vært fraværende, og vært positive i tider hvor jeg selv har sett mørkt på det. Med dette kom jeg i mål.

Tromsø, 2017

Silje Kristiansen

Sammendrag

Denne mastergradsoppgaven bygger på et konstruktivistisk verdenssyn, og er følgelig av kvalitativt preg. Jeg har gjennom åtte semistrukturerte intervjuer med matematikklærere på ungdomstrinnet ønsket å finne ut av:

Hva er matematikklærernes oppfatninger av vurdering i faget?

Et underliggende begrepsapparat utgjorde teorigrunnlaget mitt. Teorigrunnlaget og innsamlet data ble gjenstand for en tematisk analyse, som ved en delvis deduktiv tilnærming resulterte i 5 temaer jeg mener dekker lærernes oppfatninger av vurdering:

- Matematikk er et lett fag å vurdere i
- Karakterer kan ha en fremmede effekt på faglig sterke elever, og motsatt effekt på svake elever
- Relasjoner er avgjørende for en god vurderingsprosess
- Feedback avhenger av elevens faglige ståsted
- Summativ vurdering kan brukes formativt

Resultatene fra undersøkelsen viser en nær sammenheng mellom høyt og lavt presterende elevers læringsutbytte når det kommer til vurdering i både formativ og summativ forstand. Hvorvidt en vurderingspraksis med karakter fører til økt læringsutbytte, mener lærerne å oppfatte som avhengig av elevens faglige ståsted. Det samme gjelder for tilbakemeldingene og deres nivå av ros. Det viser seg at de fem funnene kan ses i lys av hverandre, i en prosess mot det forskning omtaler som en fullverdig formativ vurderingsprosess.

Innhold

1. Innledning	1
1.1. Bakgrunn for valg av tema	1
1.2. Formål og forskningsspørsmål	1
1.3. Oppgavens oppbygning	2
2. Teori.....	3
2.1. Oppfatninger	3
2.1.1. Orientering og kunnskapssyn	4
2.2. Vurdering.....	6
2.3. Feedback	8
2.4. Formativ vurdering	10
2.4.1. Tre nøkkelprosesser	10
2.4.2. TRU Math og didaktiske implikasjoner	11
3. Metode	15
3.1. Konstruktivistisk verdenssyn.....	15
3.2. Metodisk tilnærming	15
3.2.1. Intervju som forskningsmetode	16
3.3. Intervjuguide.....	18
3.4. Prøveintervjuene.....	19
3.5. Valg av informanter.....	19
3.6. Analyseprosessen	20
3.7. Studiens begrensninger.....	23
3.7.1. Reliabilitet.....	23
3.7.2. Troverdighet, generalisering og viktighet.....	24
3.8. Forskningsetiske overveielser.....	26
4. Analyse og funn	29
4.1. Matematikk er et lett fag å vurdere i.....	29
4.2. Karakterer kan ha en fremmede effekt på faglig sterke elever, og motsatt effekt på svake elever	31
4.3. Relasjoner er avgjørende for en god vurderingsprosess	33
4.4. Feedback avhenger av elevens faglige ståsted.....	35
4.5. Summativ vurdering kan brukes formativt	36

5. Diskusjon.....	39
5.1. Matematikk er et lett fag å vurdere i.....	39
5.2. Karakterer kan ha en fremmede effekt på faglig sterke elever, og motsatt effekt på svake elever.....	39
5.3. Relasjoner er avgjørende for en god vurderingsprosess	40
5.4. Feedback avhenger av elevens faglige ståsted.....	41
5.5. Summativ vurdering kan brukes formativt	42
6. Avslutning	45
6.1. Oppsummering.....	45
6.2. Konklusjon.....	46
6.3. Veien videre.....	46
Litteraturliste.....	49
VEDLEGG.....	II
Vedlegg 1: Intervjuguide.....	II
Vedlegg 2: Godkjenning fra NSD	IV
Vedlegg 3: Informasjonsskriv og samtykkeskjema.....	VI

Figuroversikt

Figur 1 Aspekter av “teaching-in-context” (Schoenfeld, 1998, s. 13).....	5
Figur 2 Effektiv tilbakemelding (Butler, illustrert i Wiliam, 2007, s. 1073).....	9
Figur 3 Aspekter av formativ vurdering (Wiliam, 2007, s. 1064)	11
Figur 4 Måleinstrument for kvalitet i undervisninga (Schoenfeld, 2014, s. 23)	12

1. Innledning

Denne mastergradsoppgaven undersøker matematikklæreres oppfatninger av vurdering i faget. Innsamlet data analyseres ut fra semistrukturerte intervjuer med åtte ungdomsskolelærere. I dette kapitlet introduserer jeg til selve avhandlingen gjennom å belyse bakgrunn for valg av tema, undersøkelsens formål, samt presentasjon av forskningsspørsmål og oppgavens oppbygning.

1.1. Bakgrunn for valg av tema

Etter levert oppgave sikter jeg meg inn mot en ungdomsskole, hvilket impliserer hyppig karakterbruk. Studieløpets praksisperioder har for det lengste vart opp til fire uker, noe jeg mener ikke er tilstrekkelig for å verken kartlegge elevenes forståelse og jobbe videre ut fra, eller opparbeide kunnskap om hva som kreves for å få den enkelte karakter innenfor faget. Vurderingspraksisen vår er mye omdiskutert og jeg fikk nå benyttet anledningen til å snakke med lærerne ute i skolen om hva deres oppfatninger av vår vurderingskultur er.

Vurdering, både i form av måle- og veiledningsfunksjon er noe jeg som lærer må jobbe med i henhold til å gi en best mulig tilpasset opplæring. Hvordan fungerer de hver for seg? Hva er hensikten bak? Og hvordan brukes disse for å oppnå økt læringsutbytte?

Frank K. Lesters *Second Handbook of Research on Mathematics Teaching and Learning* (2007) tar for seg hvordan vi skal “keeping learning on track” (kapittel 23, Wiliam, 2007) og påpeker viktigheten av en god vurderingspraksis. Rammeverket TRU Math som er et måleinstrument for kvalitet i undervisningen tar for seg hvordan læreren best gir en «formative assessment» og etterspør et «cognitive demand» (Schoenfeld, 2016). På bakgrunn av dette og at forskning synes å tale ulikt om type feedback og dens nytte for skolens formål (Butler, 1988; Hattie & Timperley, 2007; Smith & Gorard, 2005; Wiliam, 2007), ønsket jeg å gå inn å undersøke lærernes oppfatninger av den praktiseringen de hele tiden gjør.

1.2. Formål og forskningsspørsmål

Formålet med denne undersøkelsen er å få en bredere forståelse for vurderingspraksisen vår – hvordan den er i skolen, og hvordan de som praktiserer den (lærerne) anser den. Dette gjøres gjennom intervjuer hvor jeg etterspør lærernes *oppfatninger*. Disse er komplekse og sammensatte, og resultatene i prosjektet bør ses i lys av dette.

Vurdering er et område det har blitt forsket mye på, spesielt de siste tiårene. Ulike studier og metastudier (f.eks. Butler, 1988; Hattie & Timperley, 2007) fra området har fokusert på

elevenes læringsutbytte, og på bakgrunn av interessante funn ønsker jeg å finne ut hvordan lærerne selv oppfatter sin vurderingspraksis, og med dette få frem den variasjonen vi ser i skolen i dag.

På bakgrunn av undersøkelsens formål, har jeg formulert følgende forskningsspørsmål:

Hva er matematikklærernes oppfatninger av vurdering i faget?

1.3. Oppgavens oppbygning

Teorigrunnlaget mitt utgjorde intervjuguiden min, og dataene mine kan på denne måten fremstå som en deduktiv analyse med forhåndsbestemte kategorier. Samtidig møtte jeg analyseprosessen med en (delvis) induktiv tilnærming (se kapittel 3.6.) da jeg her oppdaget interessante temaer innenfor vurdering som bidro til at jeg valgte å formulere forskningsspørsmålet slik. Betydningen samt begrepsavklaring til forskningsspørsmålet blir gjort rede for i teoridel, og hvordan det skal undersøkes blir utdypet i metodedel. Her presenteres også studiens teoretiske perspektiv og hvordan jeg analyserte innsamlet datamateriale. I kapittel 4 presenteres informantenes utsagn med begrunnelse om hvorfor disse utgjør et felles tema. I kapittel 5 ser jeg disse funnene i lys av tidligere forskning og teori (som mitt kapittel 2 utgjør). Kapittel 5 er derfor mer tolkende enn kapittel 4. Funnene oppsummeres i det avsluttende kapittel 6, med konklusjon av problemstillingen og forslag til videre forskning.

2. Teori

Teoridelen for denne masteravhandlingen vil fokusere på hovedbegrepene *oppfatning* og *vurdering* som eksplisitt nevnes i forskningsspørsmålet.

2.1. Oppfatninger

Oppfatninger, eller *beliefs*, er et nokså komplekst område som har fått mye og varierende uttalelser i forskningslitteraturen. Philipp (2007) definerer:

Beliefs might be thought of as lenses through which one looks when interpreting the world, and affect might be thought of as a disposition or tendency one takes toward some aspect of his or her world; as such, the beliefs and affect one holds surely affect the way one interacts with his or her world (Philipp, 2007, s. 257-258)

Kort handler beliefs om hvordan en tolker og forstår noe - sann er det, uten at det kan bevises eller avkreftes. I dette forskningsprosjektet er jeg interessert i å finne ut av matematikklæreres oppfatninger av vurdering. Dette er altså ingen kunnskapstest, men en gråsoner jeg er interessert i å avdekke.

Tidligere forskning tyder på at lærernes oppfatninger og verdier rundt undervisning og læring påvirker deres undervisningspraksis (Philipp, 2007; Thompson, 1992). Lærerens atferd i klasserommet er ikke kun en konsekvens av deres kunnskap om temaet, men også påvirket av den aktuelle lærerens holdninger og oppfatninger av faget, og tilhørende didaktikk (Akinsola, 2009). Elever tar med seg forskjellige former for matematisk kunnskap og oppfatninger av faget som den har bygget seg opp fra erfaringer (Daskalogianni & Simpson, 2000), og på samme måte så er lærernes oppfatninger av matematikk så dypt inngrodd at forandringer i form av læreplaner eller undervisningsmateriale ikke påvirker dem (Furinghetti & Pehkonen, 2002). Randolph A. Philipp har forsket på matematikklæreres *beliefs* og *affects*, og peker på det faktum at slik en person oppfatter dens verden, ikke bare definerer den personen for verden, men også verden for den personen (Philipp, 2007). Philipp (2007) definerer *affekt* som en disposisjon eller følelse knyttet til en ide eller et objekt, og at dette påvirker følelser, holdninger eller oppfatninger (s. 259). McLeod (1992) har delt de affektive sidene inn i tre deler: *emotions*, *attitudes* og *beliefs*. Forskning synes å omtale begrepene definisjoner ulikt, men jeg har i denne oppgaven valgt å betegne disse som følelser, holdninger og oppfatninger. McLeod (1992) gjorde følgende distinksjoner: *følelser* er den mest intense av disse tre, men samtidig den minst stabile i forhold til endring; *oppfatninger* er den minst intense, men mest stabile i forhold til endring da den er utviklet over en lengre periode; mens *holdninger* var en plass midt imellom, altså er den sterkere enn en følelse, men vil endre seg raskere enn en

oppfatning. Altså er oppfatning ansett som kognitiv og den mest stabile av disse tre, mens følelser er de som endres raskest. Imidlertid, mens *følelser* er en bevissthetstilstand; *holdninger* er oppførsel, følelser, tanker som viser ens disposisjon eller mening; er *oppfatning* en psykologisk tolkning, begrunnelse eller påstand om at den omverden som vi lærer er sann (McLeod, 1992). Senere forskning har også supplert med et fjerde element, *values* (verdier) (DeBellis & Goldin, 1997). Dette er dog det minst anvendte elementet av disse fire. Bishop, Seah & Chin (referert i Philipp, 2007) identifiserer distinksjonene mellom oppfatninger og verdier som at oppfatninger ofte er assosiert med sanne/falske motsetninger, mens verdier ofte er assosiert med en ønsket/uønsket motsetning. De argumenterer altså for at oppfatning er mer kontekstavhengig da en sann/falsk bedømming må referere til et objekt, mens ønskede/uønskede motsetninger er assosiert med mer generelle og mindre kontekstavhengige attributter (s. 265).

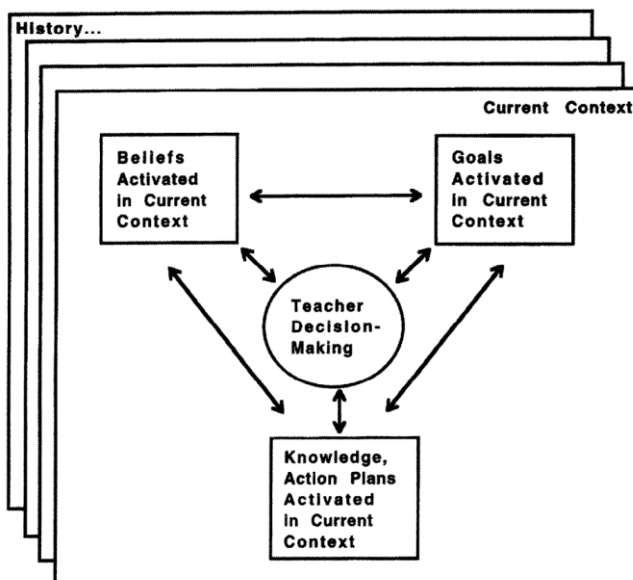
Thompson (1992) understreker også at det er viktig å se på sammenhengen mellom kunnskap og oppfatning; en studie av lærernes oppfatning uten forskning rundt deres kunnskap vil kunne resultere i et ufullstendig bilde (s. 131). Oppfatning alene kan verken bekrefte eller avkrefte, den er noe du har gjort deg opp personlig. Kunnskap, derimot, anses som en sikker oppfatning (Philipp, 2007, s. 259). Dette til tross for at oppfatninger ofte anses som kunnskap av den som oppfatter. Gettier (1963) sine eksempler i sin korte artikkel "Is Justified True Belief Knowledge?" er dog et klassisk filosofisk moteksempel til definisjonen av kunnskap som «sann begrunnet oppfatning». Kunnskap er også en sosial konstruksjon, og det å skille mellom lærernes kunnskap og oppfatning blir derfor en falsk motsetning (Op't Eynde, De Corte, & Verschaffel, 2002). Ut fra dette, er det tydelig at kunnskap, holdninger, følelser og verdier har sin påvirkning på vår oppfatning.

2.1.1. Orientering og kunnskapssyn

En annen avgjørende faktor for hvordan matematikken undervises i skolen, er lærernes oppfatninger av den matematikken de underviser. Etter at Thompson (1992) omtalte lærerens oppfatning og kunnskap som *conceptions*, refererte hun senere sammen med kolleger til lærernes syn på matematikken som *orientation*. De skilte mellom *conceptual orientation* og *calculational orientation* (Thompson, Philipp, Thompson, & Boyd, 1994), som jeg har valgt å oversette til konseptuell og prosedyrebasert orientering. Begrepene kan anses som to ulike kunnskapssyn. Selv om begge har som mål å utvikle problemløsningsferdigheter, er tilnærmingene ulike; en lærer med konseptuell orientering fokuserer mer på matematiske

sammenhenger og er opptatt av begreper og kontekst, mens en lærer med prosedyrebasert orientering fokuserer mer på tall, beregninger og resultat (s. 6-7). Thompson et al. (1994) definerer orienteringene som kunnskapssyn, og kan derfor på mange måter være sammenfallende med Hiebert & Lefevres (1986) prosedyrekunnskap og begrepsmessig kunnskap. Hvilket igjen er nært knyttet til Skemp (1976) beskrivelse av instrumentell og relasjonell forståelse. Til tross for at norske elever synes å altfor sjelden få mulighet til å utvikle begrepsmessig kunnskap eller relasjonell forståelse (Matematikksenteret, 2015), sier Skemp (1976) at det er flere fordeler med denne type undervisning: *“just because knowledge is involved, one can often get the right answer more quickly and reliably by instrumental thinking than relational. This difference, is so marked that even relational mathematicians often use instrumental thinking”* (s. 8). Likevel vil lærerens orienteringer eller kunnskapssyn avgjøre hvordan matematikken undervises i skolen.

I likhet med Thompson et al. (1994) og Brown, Harris & Harnett (2012) som hevder: *“the beliefs that teachers have about the nature and purpose of feedback will likely affect the quality of AfL [assessment for learning] implementation”* (s. 969), så mener også Schoenfeld (1998) at bl.a. oppfatninger er med på å forme lærerens valg i forhold til undervisningspraksisen. Han trekker inn beliefs, goals og knowledge som tre komponenter for å oppnå en undervisning med «teaching-in-context»:



Figur 1 Aspekter av “teaching-in-context” (Schoenfeld, 1998, s. 13)

De tre komponentene oppfatninger, mål og kunnskap påvirker hverandre, og har ikke nødvendigvis prioritet foran en av de andre komponentene. Som vist i figuren ovenfor er hver

boks merket med «Activated in current context», dette for å understreke det faktum at lærerne kan ha flere ulike oppfatninger som former undervisningen deres gjennom året; *«however, at any particular time some of these beliefs are strongly activated, either because of prior planning or because their activation level was influenced by an event that has recently taken place; others are inactive or activated at low levels»* (Schoenfeld, 1998, s. 14). Altså, på et bestemt tidspunkt, er noen av disse oppfatningene aktivert i større grad, enten på grunn av forberedende planlegging eller fordi aktiveringsnivået var påvirket av en nylig hendelse; de andre oppfatningene er inaktive eller i liten grad aktive. Det samme vil gjelde for mål og kunnskap. Dette impliserer at lærernes oppfatninger avhenger av intervjuetidspunktet, og det er mulig de vil ha en annen oppfatning på et annet tidspunkt. Hvilket impliserer at studiens reliabilitet vil utfordres med dette begrepet til grunn.

Feilkilder rundt mine intervjuer uten observasjon kan være at det som blir sagt ikke nødvendigvis er tilfellet i praksis (Lev-Zamir & Leikin, 2013). Etter Thompson (1984) argumenterte for sammenhengen mellom lærernes oppfatninger og påfølgende undervisning, har det kommet flere studier som peker i den retning av inkonsistens mellom oppfatninger og praksis. Eksempelvis presenterer Fang (1996) flere studier der forskere fant lite konsistens mellom lærernes oppfatninger og deres praksis, og konkluderte med at kontekstuelle faktorer forstyrret lærernes evne til å konsekvent anvende sin tro i praksis. Idet jeg søker mine informanternes oppfatninger, vil jeg være bevisst på dette, samt det faktum at oppfatningene kan endres i form av aktiveringsnivå (Schoenfeld, 1998).

2.2. Vurdering

Kunnskap om vurdering er et sentralt kompetanseelement hos lærere. Internasjonal forskning har bidratt til områdets større fokus de siste årene. National Council of Teachers of Mathematics [NCTM] presenterer i «Principles and Standards for School Mathematics» to hovedidéer: (1) Vurdering burde fremme elevens læring, og (2) vurdering er et verdifullt verktøy i prosessen av læring (NCTM, 2000, s. 22). På denne måten argumenterer de for en vurderingspraksis med formativ funksjon. Wiliam (2007) bekrefter dette ved å hevde at: *«assessment is a powerful servant but a bad master»* (s. 1056). Han tar opp hensikten med vurdering, og skiller mellom formativ, summativ og evaluerende vurdering. Evaluerende vurdering omhandler kvalitetssikring av skolesystemet (Wiliam, 2007, s. 1056), og står derfor litt utenfor de to andre vurderingsformene. Formativ vurdering støtter læring, og omtales derfor som *vurdering for læring* eller *undervisningsvurdering*. Summativ vurdering, derimot,

fokuserer mer på resultat enn prosess, og kartlegger den nåværende forståelsen/oppnåelsen innenfor et bestemt emne, ofte i form av poeng eller karakterer. Det er mye diskusjon rundt hvordan sluttvurdering i matematikkopplæringen bør utformes, både vurderingsformer og matematisk innhold. Matematikksenteret (2017) skriver at et hovedprinsipp for all vurdering er at elevene skal bli vurdert ut fra kriterier som er kjent for elever og lærere. Disse kriteriene skal utformes i henhold til kompetansemålene i læreplanen. Utdanningsdirektoratet (2016) mener kjennetegn på måloppnåelse er utformet på tvers av hovedområdene i faget for å uttrykke en helhetlig kompetanse i faget. Det er derimot ikke selve vurderingsformen og teknikken som brukes som avgjør hvorvidt vurderingen er formativ eller summativ, men heller formålet bak. Black & Wiliam (1998) bekrefter at vurderingen ikke er formativ før den blir brukt til å tilpasse undervisningen etter elevenes behov. Birenbaum (1996) mener at formativ og summativ vurdering gjenspeiler to ulike kulturer, en vurderingskultur og en testkultur. Ludvigsen-utvalget argumenterer for at vi gjennom fagfornyelse og dybdelæring skal skape «Fremtidens skole», og foreslår bl.a. hvordan den summative vurderingen kan virke formativt på lærerne og deres praksis; standpunkt-karakterer kan sammenlignes med eksamens-karakterer i faget for å vurdere om det er systematiske avvik mellom dem over tid- og på denne måten bruke eksamen som en kvalitetssikring av skolens standpunkt-vurderingspraksis (NOU 2015: 8, 2015, s. 83)

Likevel er kritikken av karakterbruk stor, og frykten for en undervisningspraksis basert på «teaching to the test» (Suurtamm, et al., 2016) er mye omtalt. Svært mange av de spørsmålene som skal besvares i summative tester, er spørsmål som allerede er besvart i lærebøkene, og tester dermed i stor grad elevenes hukommelse. Schoenfeld (2007a) tar for seg en studie der Silicon Valley Mathematics Assessment Collaborative ga to ulike prøver til ulike klasser og klassetrinn for å kunne anslå eventuelle distinksjoner mellom disse. Den ene prøven SAT-9 var en skills-oriented test i tråd med California mathematics standards. Den andre, Balanced Assessment test, var i henhold til NCTMs principles and standards og testet større områder som fokuserte på kompetanser og problemløsning. Elevenes resultat ble kategorisert ut fra hvorvidt svaret var kyndig eller ikke. Det viste seg at mer enn 90% av elevene som var erklært kyndige på Balance Assessment-testen også var erklært kyndig på SAT-9. Dette gikk derimot ikke andre veien. Omtrent en tredjedel av elevene som var erklært kyndige på SAT-9 var erklært ikke-kyndig på Balance Assessment-testen (s. 63). Både «teaching to the test»-fenomenet og Schoenfelds (2007a) funn kan ses i sammenheng med instrumentell forståelse og den ikke-eksisterende praksisen med å skape et intellektuelt behov (Harel, 2013).

Informasjonen om elevens prestasjoner i testene blir heller ikke nødvendigvis anvendt i videre arbeid, og karakterene får følgelig kun en målefunksjon:

[..] if the formal assessments deliver just scores or percentile ratings, that information is of negligible use to teachers. If the results are returned weeks or months after the test is taken, as is often the case with high-stakes assessments, the results are of even less value (Schoenfeld, 2007a, s. 9)

En forutsetning for å lære er gode tilbakemeldinger. Verken karakter eller poengsum vil gi elevene tilstrekkelig informasjon for en aktuell progresjon. Som lærer vil det være en utfordring i det å være motiverende, men samtidig kritisk til noe en elev har produsert.

William (2007) tar opp det med overfladisk skryt og argumenterer for at tilbakemeldinger til elevene bør fokusere på det som trengs å forbedre, fremfor hva som er gjort bra: «*Feedback to learners should focus on what they need to do to improve, rather than on how well they have done, and should avoid comparison with others*» (s. 1085). Matematikksenteret (2017) derimot, mener at vi bør legge vekt på å synliggjøre for eleven hva den kan og mestrer i større grad enn hva den ikke mestrer. Dette er ikke nødvendigvis ei motsetning til William, da en kan være formativ samtidig som en forsterker positivt. Shute (2008) mener at tilbakemeldingen bør være elevsensitiv og basert på tillit. Kluger & DeNisi (1996) mener at effekten på ei fremovermelding vil minke dersom den består av ros, straff eller belønning. St.meld. 22 (2010-2011) sier at høy vurderingskompetanse med tilhørende evne om å gi eleven realistiske utfordringer, god faglig veiledning og mestringsopplevelser kan forstås som virkningsfulle motivasjonelle faktorer. På samme måte forslår Kunnskapsdepartementet (2017) i sitt høringsutkast til ny generell del av læreplanen at kartlegging og observasjon av hver enkelt elev har liten verdi dersom det ikke følges opp med konstruktive tiltak. Overdreven vekt på dokumentasjon og rangering kan svekke den enkeltes selvbilde eller hindre utviklingen av et godt læringsmiljø (Kunnskapsdepartementet, 2017, s. 16).

2.3. Feedback

I forskrift til opplæringsloven (2009, § 3-11) står det om undervisvurdering at elevene skal få tilbakemeldinger som forteller dem om kvaliteten på arbeidet eller prestasjonen. Videre under samme paragraf står det også at vurderingen skal inneholde informasjon om hvilken kompetanse eleven innehar samt rettleiding til hvordan eleven kan utvikle kompetansen i faget.

Bruken av ulike tilbakemeldingsmetoder er mye omdiskutert, spesielt karakterbruk satt opp mot kommentarer (Hattie & Timperley, 2007; Smith & Gorard, 2005; William, 2007). Butler

(1988) har forsket på denne vurderingspraksisen og ser på tre ulike tilbakemeldingsformer: kommentar alene, karakter alene og kommentar og karakter sammen. Han konkluderer her med at kommentarer blir overflødig dersom karakter også gis:

Table 23.3 Impact of Feedback Interventions on Achievement and Attitude (Butler, 1988)

	Achievement gain	Attitude toward subject
Scores only	none	High achievers: positive Low achievers: negative
Comments only	30%	All positive
Scores and comments	none	High achievers: positive Low achievers: negative

Figur 2 Effektiv tilbakemelding (Butler, illustrert i Wiliam, 2007, s. 1073)

Når det ble gitt både karakter og kommentarer, så elevene først på karakter for deretter å spørre etter sidemannens karakter. Han konkluderer med at det derfor er bortkastet tid for læreren å plassere eleven på en karakterskala når vurderingen kombineres med kommentar. Studien ble gjort på en slik måte at elevene ble sammenlignet fra hver klasse ut fra deres holdninger og prestasjoner innenfor faget. Tabellen viser at de høyt-presterende elevene var positive til alle tre vurderingsformene, mens de lavt-presterende var negative til både karakter alene og karakter og kommentar kombinert, altså kan det se ut til at de var negative til karakterbruk generelt. Konklusjonen bekreftes og likestilles med Page (1958) sin studie som også kom frem til at tilbakemelding i form av korte kommentarer har større effekt enn bare karakterer på prøveresultater. Hun kom også frem til, i likhet med Butler (1988), at tilbakemelding gjennom kun kommentarer førte til økt læringsutbytte, mens karakterer alene eller sammen med kommentarer ikke gjorde det. Smith & Gorard (2005) derimot, har forsket seg frem til å snakke den summative vurderingens sak. Studien «They don't give us our marks» foretar en evaluering av en endring i vurderingspraksisen på en skole i Wales. Skolen delte 104 «year 7 pupils» (tilsvarer 11-12 år) med varierende kompetanse inn i fire grupper. En av disse gruppene ble gitt formative tilbakemeldinger på arbeidet deres i ett år, men uten karakter. Disse utgjorde altså tiltaksgruppa som ble brukt til å undersøke effekten av kommentarer alene. De resterende i kontrollgruppa fikk karakterer med kommentarer i mindre grad, slik den vanligste praktiseringen der faktisk er. Resultatet av undersøkelsen viste at mesteparten av elevene ikke følte tilbakemeldingene/kommentarene ga tilstrekkelig informasjon eller instruksjoner på veien til forbedring. Samtidig hevdet flere av elevene at de ønsket karakterene tilbake, da de uten disse visste lite om forventet innsats videre. Det var

altså ingen klare bevis på at denne intervensjonen hadde en positiv effekt, men ble heller oppfattet skadelig for elevene. Forskerne oppsummerte med at formative tilbakemeldinger gir lite fremgang på en læringsprosess. Kommentarene fokuserte dog mer på å fremme elevens selvbilde, enn på hva som faktisk trengtes å jobbes med, og hvordan (Smith & Gorard, 2005, s. 36). Hvilket strider imot bl.a. Wiliams (2007) definisjon på type tilbakemelding. En annen ting å merke seg i forholdet til denne studien er at slik det er i Wales hvor lærerne lenge har praktisert karakterbruk og dermed vent elevene til en slik praksis, vil det følgelig ha stor innvirkning på hva elevene forventer og forholder seg til i en vurderingsprosess. Elevene her var avhengige av å få en karakter for å forstå hvorvidt arbeidet var bra eller ikke. ¹

2.4. Formativ vurdering

En gruppe anerkjente britiske forskere som utgjør Assessment Reform Group (2002), omtaler formativ vurdering eller vurdering for læring som: *“the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there”* (s. 2). Ifølge Stobart (2008, s. 146) var det denne forskergruppa som ønsket å betegne underveisvurdering som *vurdering for læring*, fremfor *formativ vurdering*; med argumentasjon om at formativ vurdering hadde en for vid betydning og ofte blir tillagt «mini-summative» praksiser. Black, Harrison, Lee, Marshall & Wiliam (2004) mener vurdering for læring er hensikten (purpose), mens formativ vurdering er funksjonen (function), altså når den faktisk blir brukt for å møte elevenes læringsbehov. Jeg anvender *formativ vurdering* i mitt forskningsspørsmål da jeg ønsker å finne ut av funksjonen og dens prosesser for å fremme læring. Hattie (2008) har i sin metastudie «Visible Learning» målt flere ulike faktorer for å illustrere i hvilken grad disse fører til økt læringsutbytte. Her finner vi bl.a. «providing formative evaluation» som en av de sterkeste faktorene.

2.4.1. Tre nøkkelprosesser

Forskere synes å ha en felles mening om at formativ vurdering består av tre nøkkelprosesser: *“1) Establishing where the learners are in their learning, 2) Establishing where they are going and 3) Working out how to get there”* (Wiliam, 2007, s. 1064). Black & Wiliam (2009) trekker paralleller til et tidligere teoretisk grunnlag som kom fra fem type aktiviteter for formativ vurdering: *«1) Sharing success criteria with learners, 2) Classroom questioning, 3)*

¹ Dette avsnittet ble i stor grad benyttet i en annen sammenheng, eksamen LRU-3351 Matematikdidaktikk 1.SEM. (H2016): Refleksjonstekst: «Hvordan bør den formative vurderingen i faget være for å oppnå sitt formål?» (Kristiansen, 2016, s. 3-5).

Comment-only marking, 4) Peer- and self-assessment, 5) Formative use of summative tests» (s. 7-8). Det har vært mye diskusjon rundt hvem som skal gi tilbakemeldinger. Hattie & Timperley (2007) definerte tilbakemelding som: «*information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding»* (s. 81). Wiliam (2007) har laget en modell ut fra hans tre nøkkelprosesser der han i henhold til å oppnå de tre prinsippene setter krav til nettopp både lærer, medelever og eleven selv:

	Where the learner is going	Where the learner is right now	How to get there
Teacher	Clarifying learning intentions and sharing and criteria for success	Engineering effective classroom discussions and tasks that elicit evidence of learning	Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	Activating students as the owners of their own learning	

Figur 3 Aspekter av formativ vurdering (Wiliam, 2007, s. 1064)

Disse aspektene kan ifølge Black & Wiliam (2009) ses i sammenheng med de tidligere fem aktivitetene som utgjorde grunnlaget for de nåværende tre nøkkelprosessene i formativ vurdering. Eksempelvis kan det å stadfeste «where the learner is right now» gjennom «engineering effective classroom discussions and tasks that elicit evidence of learning» vise til «classroom questioning». Under prosessen «how to get there» vil «providing feedback that moves learners forward» linkes til «comment-only marking». Dette igjen, kan linkes til studiene utført for å verifisere effekten av de ulike tilbakemeldingsformene. For eksempel Butlers (1988) studie som kom frem til at kommentarer alene ga størst læringsutbytte i formativ vurdering, altså «how to get there». ²

2.4.2. TRU Math og didaktiske implikasjoner

Wiliams (2007) overnevnte prosess ved formativ vurdering er å finne igjen ved rammeverk som måler kvalitet i undervisningen. «Teaching for Robust Understanding» (heretter: TRU-Math) er et slikt rammeverk. Her presenteres fem dimensjoner av god undervisning, hvorav en omhandler formativ vurdering - å møte elevene der de er, og forme undervisningen deretter:

² Kapittel 2.4.1.: «Tre nøkkelprosesser» ble i stor grad benyttet i en annen sammenheng, eksamen LRU-3351 Matematikdidaktikk 1.SEM. (H2016): Refleksjonstekst: «Hvordan bør den formative vurderingen i faget være for å oppnå sitt formål?» (Kristiansen, 2016, s. 6-7).

[Formative Assessment]: The extent to which classroom activities elicit student thinking and subsequent interactions respond to those ideas, building on productive beginnings and addressing emerging misunderstandings. Powerful instruction “meets students where they are” and gives them opportunities to deepen their misunderstandings (Schoenfeld, 2016, s. 11)

Videre kategoriserer han bruken av vurdering på tre nivå:

Uses of Assessment	
1	Student reasoning is not actively surfaced or pursued. Teacher actions are limited to corrective feedback or encouragement.
2	The teacher refers to student thinking, perhaps even to common mistakes, but specific students' ideas are not built on (when potentially valuable) or used to address challenges (when problematic).
3	The teacher solicits student thinking and subsequent instruction responds to those ideas, by building on productive beginnings or addressing emerging misunderstandings.

Figur 4 Måleinstrument for kvalitet i undervisninga (Schoenfeld, 2014, s. 23)

Slik figuren ovenfor viser, indikerer det første og laveste nivået at lærerens praksis er svært begrenset. Her kan læreren eksempelvis «hjelp» eleven med et regnestykke i den grad til at de ikke bidrar med å skape en indre kognitiv konflikt, og dermed eliminerer de kognitive utfordringene som skaper læring. Denne problemstillingen blir tatt opp i «Standards» (NCTM, 1991) hvor viktigheten av at læreren tar i bruk vurderingsstrategier som fokuserer på konseptuell forståelse fremfor prosesskunnskap understrekes. Harel (2013) tar i denne sammenheng for seg et annet rammeverk for god undervisning, *DNR* (*duality, necessity* og *repeated reasoning*), som setter fokus på å fremkalle elevens intellektuelle behov for å lære matematikk. DNR argumenterer for et psykologisk behov for å motivere, og et intellektuelt behov for å forstå hensikten bak (Harel, 2013). Begge disse faktorene er avgjørende for den læringsprosessen som ønskes å få frem hos eleven. En kombinasjon av disse vil resultere i at eleven selv ønsker å lære, og idealet i TRU-Math vil også dermed bli lettere å oppnå for læreren.

Det andre nivået er litt forbedret, men det er fremdeles ingen klar indikator på at læreren rettleder eleven. Det høyeste nivået, derimot, indikerer at læreren tilpasser undervisningen etter den formative vurderingen som er foretatt. Viktig å merke seg er at alle nivåene stiller krav til læreren. Likevel, er det først når vurderingen ligger på høyeste nivå (3) at læreren har vært innom alle de tre punktene Wiliam (2007) omtaler som en fullverdig formativ

vurderingsprosess. Wiliam (2007, s. 1075) sammenligner dette med å be en mislykket komiker om å bli morsommere, altså uten rettleiding til videre handling. At vurderingen skal være formativ er en lang prosess som krever at det jobbes opp mot det høyeste punktet i dette rammeverket.

I sammenheng med «formative assessment»-dimensjonen, er en annen aktuell dimensjon innenfor rammeverket «cognitive demand». Den ligner veldig på prinsippet om å skape et intellektuelt behov, da begge handler om at elevene skal utfordres (Schoenfeld, 2016; Harel, 2013). Denne dimensjonen foreslår følgende aktivitet: *“Teachers can actively support students in individual work, group work, and whole class discussions by asking clarifying questions and providing scaffolds, instead of moving directly to suggesting overly specific ways to go about assigned tasks”* (Schoenfeld, 2016, s. 5). Denne måten kan brukes for å finne ut “hvor eleven er”. Her er det derfor viktig å tenke på, slik det også presiseres, at man ikke eliminerer de kognitive utfordringene. I denne dimensjonen er det avgjørende at læreren får forståelse for hvordan eleven forstår noe, og dermed får en indikator på hvordan eleven kan lære videre, man lærer nye ting gjennom det man kan fra før. Schoenfeld (2016) sier at: *“Researchers use the term «cognitive demand» to describe the level of difficulty, relative to what they know, of the work that students are asked to engage in”* (s. 5). Dette tyder på at dimensjonen “formative assessment” kan ligge til grunn for lærerne for å justere nivået av “cognitive demand”. Både her og i Wiliams (2007) tre nøkkelprosesser er relasjonen til Vygotskys proksimale utviklingszone tydelig.

Utdanningsforbundet (2017) tar opp den mye omtalte diskusjonen om karakterer på barnetrinnet. De viser til nasjonal forskning og mener at:

elever som motiveres av å oppnå best mulig karakterer, ikke vil oppnå like god kunnskap som dem som har den indre motivasjonen til å lære. Hvis elevene virkelig skal bli flinke, er det viktig at de motiveres ut fra et eget *ønske om å forstå*. Det er også viktig for videre utdanning at de har en base for forståelse som er motivert ut fra ønske om å lære framfor å få en god karakter (Utdanningsforbundet, 2017)

Dette peker i den retning at det å vurdere etter et lavt nivå i rammeverket eliminerer også muligheten for elevene til å skape et intellektuelt behov, «ønske om å forstå». Og det å belønne gjennom karakterer kan virke hemmende i en slik læringsprosess. Motivasjon kan deles i to kategorier: *indre motivasjon* hvor elevene blir motiverte av at de ønsker å lære og forbedre sine ferdigheter; *ytre motivasjon* der elevene vil demonstrere sin kompetanse, og vise at de er flinkere enn andre (Utdanningsforbundet, 2017). Altså kan karakterer være

hemmende for de faglig svake elevene i forhold til demotivasjon, men også for de faglig sterke (og de svake) i forhold til å skape et intellektuelt behov.

3. Metode

I dette kapittelet presenterer jeg forskningsdesignet og dets overbyggende plan for hvordan forskningsspørsmålet mitt skal undersøkes. Den tar meg fra underliggende filosofiske premisser som konstruktivisme og kognitiv psykologi, via forskningsstilen generisk kvalitativ metode, til datainnsamlingsmetoden intervju. Jeg vil også ta for meg studiens kvaliteter og begrensninger hvor reliabilitet og validitet er slutninger gjennom hele forskningsprosessen. Til sist presenterer jeg de forskningsetiske overveielene som er foretatt i dette prosjektet.

3.1. Konstruktivistisk verdenssyn

Creswell (2014) tar opp verdenssyn, og mener at denne informasjonen vil forklare hvorfor vi i forskning velger en kvalitativ, kvantitativ eller blandet tilnærming. Han presenterer fire ulike verdenssyn eller paradigmer: postpositivisme, konstruktivisme, deltakende og pragmatisme (Creswell, 2014). Jeg ønsket å finne ut av lærernes oppfatninger av vurdering og deres generelle vurderingspraksis. En konstruktivistisk epistemologi gir meg rom til å tolke den informasjonen jeg innhenter ut fra deltakernes tolkning av virkeligheten (Creswell, 2014). Grbich (2013) mener at konstruktivisme ikke har noen objektiv kunnskap uavhengig av vår kognisjon. Kunnskap konstrueres av individet i bestemte kontekster. Målet er å forstå hvordan mennesker tolker og forstår deres egne erfaringer (s. 7). På samme tid som studiet mitt har et konstruktivistisk verdenssyn, er jeg også innenfor det teoretiske perspektivet kognitiv psykologi. Kognitiv psykologi ser på hvordan individer tilegner seg kunnskap gjennom konklusjoner basert på tolkningen av deres egen erfaring og forståelse (Cobb, 2007). Hvordan individet tenker og resonnerer, går på *oppfatningene* jeg er interessert i å avdekke. Fra mine intervjuer skal jeg kunne trekke ut informasjon som bygger opp empirisk vitenskap. Dette være, i motsetning til eksperimentell psykologi, mangfoldig og fleksibel kunnskap som kan bidra til utvikling av undervisningen i klasserommet (Cobb, 2007, s. 20)

3.2. Metodisk tilnærming

Ifølge Creswell (2014, s. 98) vil metodologien påvirkes av teoretisk paradigme. Selv om de fleste studier som gjennomføres i (her: kognitiv) psykologien har kvantitative forskningsdesign (Martin, Carlson, & Buskist, 2010, s. 65), har ikke jeg behov for tall og måling i min forskning. Kvantitative data ville vært ideelt om jeg allerede visste hva slags type oppfatninger som finnes, og ønsket å kartlegge omfanget av hver type. Denne tilnærmingen krever mange informanter slik at de er representative for en populasjon (Cohen,

Manion & Morrision, 2007). Kvalitative studier, derimot, går ut på å studere hva som finnes, ikke hvor ofte eller vanlig fenomenet er (Cohen et al., 2007). Det viktigste med forskningsdesignet er at det skal hjelpe meg å svare på forskningsspørsmålet mitt. Jeg har behov for et rikere datamateriale som gir meg dypere innsikt i lærernes oppfatninger og tanker, og følgelig har forskningsdesignet mitt et kvalitativt preg.

Analysen min hadde både deduktive og induktive element. Det deduktive kan ses av intervjuguiden (vedlegg 1), som delvis styrer analysen. Det induktive er også tydelig, da jeg var åpen for nye temaer. Med en delvis deduktiv analyseprosess, som Creswell (2014) mener er typisk for kvantitative metoder, ser jeg behovet for en fleksibel metodisk tilnærming. Caelli, Ray & Mill (2003) presenterer *generisk kvalitativ metode* som en tilnærming som gir meg muligheten til å tilpasse prosjektet underveis. Av samme grunn klassifiserer Caelli et al. (2003) denne forskningsstilen som den minst krevende. Jeg hadde ikke et ferdig utformet forskningsspørsmål ved datainnsamling. Dette sett sammen med min manglende erfaring som forsker, utgjorde mitt behov for de fleksible rammene denne forskningsstilen tillater.

Generisk kvalitativ metode åpner for muligheten til å bruke kvalitative tilnærminger som intervju og observasjon for datainnsamling. For å styrke en undersøkelse om «den generelle vurderingspraksisen i skolen» ville en kombinasjon av intervju og observasjon vært ideelt. En slik kombinasjon gjør at forskeren unngår feilkilder som at det som blir sagt ikke er tilfellet i praksis (Lev-Zamir & Leikin, 2013). Jeg ønsker derimot å finne ut av lærernes *oppfatninger* av vurdering, og har følgelig ikke et behov for observasjon. Hvilket bekreftes av Christoffersen & Johannessen (2012) som sier at en forsker vil møte utfordringer idet han vil observere hva mennesker tenker og føler. Samtaler er viktig for at mennesker skal forstå hverandre, svare på hverandres spørsmål, kommentere hverandres utsagn eller handlinger, samt beskrive hvilke intensjoner de har, hva de tenker, føler og mener. Hvis jeg også skulle undersøke hvorvidt det er en inkonsistens mellom lærernes oppfatninger og deres praksis, kunne jeg brukt flere metoder for å supplere intervju. Men i dette tilfellet har jeg valgt å avgrense det til intervju og oppfatninger for at datamengden skal være håndterbar for et masterprosjekt.

3.2.1. Intervju som forskningsmetode

Innenfor kvalitative studier er det intervju som er mest aktuelt, da jeg her kommer i samtale med lærerne og får innsikt i hva de tenker om vurdering i faget. Kvale & Brinkmann (2015) sier intervju er gunstig som metode når forskeren har behov for å gi informantene større frihet

til å uttrykke seg, samtidig som interessefeltet er menneskers erfaringer og oppfatninger.

Kvale & Brinkmann (2015) definerer det kvalitative forskningsintervjuet som en samtale med en struktur og et formål, som søker å forstå verden sett fra intervjupersonens side. Målet er å få frem betydningen av folks erfaringer og å avdekke deres opplevelse av verden (s. 22). På bakgrunn av disse skildringene vil et kvalitativt forskningsintervju være mest passende for at jeg skal få innsikt i lærernes oppfatninger.

Creswell (referert i Turner, 2010) tar opp ulike intervjudesign som kan brukes for å oppnå et rikt datamateriale innenfor et kvalitativt perspektiv; (a) informal conversational interview, (b) general interview guide approach, and (c) standardized open-ended interview (s. 752). Den første uformelle tilnærmingen krever ingen spesifikke spørsmål fra forskeren, men stoler heller på interaksjonen på deltakerne imellom. Denne type intervju er for mange fordelaktig i form av lite strukturerte rammer som gir rom for fleksibilitet. For en uerfaren forsker som meg medfører dette både fordeler og ulemper; intervjusituasjonen ville vært lettere, men en konsekvens er samtidig et ustrukturert datamateriale som blir langt mer utfordrende å kode. Samtidig, med mine *individuelle intervjuer* (se kapittel 3.5.) ville ikke dette vært like gjennomførbart som ved for eksempel fokusgrupper. Likevel sammenligner Cope (referert i Clifford, French & Valentine, 2010) semistrukturerte intervju med fokusgrupper, på den måten at begge har en nokså uformell tone. Dette ville likevel ikke være aktuelt for meg da en karakteristikk i fokusgrupper er interaksjon mellom medlemmene, fremfor intervjuer og informant. En mulig fallgrube ved å samle lærerne ville dessuten være at de påvirker hverandre. En annen ting er praktiske årsaker som at det sjeldent ville latt seg gjøre å samle et tilstrekkelig antall lærere samtidig. Cohen et al. (2007) tar også opp disse ulike intervjutypene. Forskerne kaller disse intervjuene hvor forskerne ikke enda er bevisst på hva de ikke vet og stoler på deltakerne til å fortelle dem, for ustrukturerte (s. 354). Et delvis strukturert, eller semistrukturert intervju, tilsvarer (b) general interview guide approach, i form av at intervjuet har en noenlunde bestemt rekkefølge, men som samtidig tillater fleksibilitet i deltakers favør. Jeg hadde ikke en ferdigstilt problemstilling da jeg startet min innsamling av data. Intervjuguiden var utformet på den måten at jeg ville ha rom for oppfølgings spørsmål dersom informantene syntes å ha en interessant eller felles oppfatning av noe jeg ikke hadde forutsett. Med dette til grunn, falt mitt metodevalg på et semistrukturert intervju, som da ga meg muligheten til å ha kontroll på at alle de forhåndsbestemte temaene ble gjennomgått, samt gå dypere inn i fenomenet, da fleksibiliteten er større. Ifølge McNamara (sitert i Turner, 2010) er nettopp dette styrken ved et slikt type intervju; for forskeren “...to

ensure that the same general areas of information are collected from each interviewee; this provides more focus than the conversational approach, but still allows a degree of freedom and adaptability in getting information from the interviewee” (s. 755). Jeg hadde behov for fleksibilitet i intervjusituasjonen, noe et strukturert intervju ville satt begrensninger for. Et slikt intervju, som Turner (2010) klassifiseres som standardized open-ended interview medfører nemlig lite rom for å avvike fra det utarbeidede manuskriptet. Martin et al. (2010, s. 65) mener at et semistrukturert intervju tillater meg å innta en mer utforskende rolle i intervjusituasjonen. Semistrukturerte intervju kan på denne måten gi en god balanse mellom standardisering og fleksibilitet, hvilket er passende for min studie med både deduktive og induktive element.

3.3. Intervjuguide

Med intervju som metode, er utforming av intervjuguide nødvendig (Cohen et al., 2007; Kvale & Brinkmann, 2015; Turner, 2010). En intervjuguide vil vanligvis ha en bestemt rekkefølge på temaene, men min semistrukturerte tilnærming tillater at denne kan endres på dersom informanten bringer nye temaer på bane. Kvale & Brinkmann (2015, s. 162-163) mener at man kan vurdere et intervju spørsmål både med hensyn til en tematisk og en dynamisk dimensjon: tematisk med hensyn til produksjon av kunnskap, og dynamisk med hensyn til den interpersonlige relasjonen i intervjuet. Da jeg er interessert i rike data om lærernes oppfatninger, er den dynamiske dimensjonen avgjørende for at jeg får stimulert intervjupersonene til å snakke om dette. Christoffersen & Johannessen (2012) oppfordrer til å først stille fakta- og introduksjonsspørsmål for å etablere en relasjon og tillitsforhold til informanten. Jeg valgte derfor å først stille innledende spørsmål om informanten selv (se vedlegg 1). Neste punkt før hoveddelen i intervjuet (hvor informanten har medbestemmelse på rekkefølge av temaene) vil jeg bevege meg lenger inn mot overgangsspørsmål i form av: «hva er vurdering?» og «hvorfor vurderer vi?». Clifford et al. (2010) argumenterer for en kombinasjon av ulike typer spørsmål: *“researchers often start with a question that participants are likely to feel comfortable answering. More difficult or, sensitive or thought-provoking questions are best left to the second half of the interview or focus group when participants are feeling more comfortable”* (s. 107). Når det kommer til det tematiske bidraget og produksjon av kunnskap unngikk jeg ledende spørsmål for å påvirke informantene mine minst mulig. Åpne spørsmål som «hva tenker du om å vurdere i matematikkfaget?» har ingen svaralternativer, informantene svarer med egne ord og på den måten får deltakerne delt sine synspunkt (Cohen et al., 2007). Å anvende slike åpne spørsmål er typisk innenfor et

konstruktivistisk verdenssyn som har intervju som forskningsstil (Creswell, 2014). Til tross for en semistrukturert tilnærming på intervjuet som kan gjøre analyseprosessen mer utfordrende (se kapittel 3.2.1.), har jeg likevel fastsatte temaer jeg ønsker å gå gjennom. Denne mer fleksible tilnærmingen med temaer fremfor spørsmål kan føre til at temaet om karakterbruk virker mer passende som et følgespørsmål til informantens utsagn. Mens i andre tilfeller vil temaet «karakterer» bli tatt opp som en mindre naturlig overgang.

3.4. Prøveintervjuene

En ulempe med intervju som metode er at det krever trening og erfaring i ulike spørreteknikker for å få frem den data man er ute etter fra informantene. Kvale & Brinkmann (2015, s. 36) kritiserer seg selv ved å hevde at intervjuferdighetene læres gjennom intervju praksis; et dobbelt paradoks er når læringen av muntlige ferdigheter skal læres i skriftlig form. Som uerfaren forsker så jeg selv behovet for å gjennomføre prøveintervjuer før jeg skulle innhente data. Her ble jeg bevisst på både stemmebruk og hastighet, men kanskje enda viktigere oppdaget jeg viktigheten av begrepsavklaringer. Mitt høyst relevante begrep *formativ vurdering* kan uten oppfriskning fort blandes med *summativ vurdering*, noe som potensielt kunne gi meg veldig misvisende data i de reelle intervjuene. Jeg opplevde det å ha et semistrukturert intervju som betryggende; jeg kunne la (prøve)informantene mine snakke om det de var komfortable med, når jeg samtidig visste at vi ville bevege oss inn på rett spor senere. Med andre ord var det deltakerne som bestemte rekkefølgen på temaene, men jeg som intervjuet hadde likevel kontroll over at alle temaene ble gjennomgått. I prøveintervjuene oppdaget jeg altså i praksis at forskningsintervjuet ikke er en konversasjon mellom likeverdige deltakere, ettersom jeg som forsker kontrollerer samtalen. Dette gjorde meg bevisst på det asymmetriske maktforholdet et forskningsintervju innebærer (Kvale & Brinkmann, 2015, s. 52). I denne studien vil det kunne oppstå asymmetri både på grunn av aldersforskjell mellom informant og forsker, samt intervjuers rolle som forsker i prosjektet (Thagaard, 2009, s. 103-104).

3.5. Valg av informanter

Mine utvalgsstrategier baseres først og fremst på hva som var praktisk og mest hensiktsmessig å gjennomføre. Min problemstilling om vurderingspraksisens oppfatninger stilte ikke noen kriterier utover at: (1) lærerne må undervise på et karakterpraktiserende trinn som samtidig er relevant innenfor min utdanning, altså ungdomsskole, (2) lærerne må ha undervist i matematikk i minimum 3 år, dette for erfaringens del, og (3) antall informanter må

være høyt nok for å svare på problemstillingen. På bakgrunn av disse kravene kan jeg si jeg har anvendt en kriteriebasert utvelgelse (Christoffersen & Johannessen, 2012). Å få tilgang til informanter - og enda til et tilstrekkelig antall, var utfordrende. Men trolig siden jeg ikke hadde behov for verken et ekstremt eller avvikende utvalg, fikk jeg tilgang på 8 informanter. Det at intervjuet var såpass kortvarig (ca. 30 min) og ikke krevde noen form for forberedelse eller etterarbeid for deltakerne tror jeg også bidro til at jeg fikk tak i det utvalget jeg gjorde. Forskere synes å ikke ville fastslå et bestemt antall informanter, men i studentprosjekter som mitt med begrenset økonomi og tid til rådighet sier Christoffersen & Johannessen (2012) at det er rimelig å begrense seg til færre enn 10 informanter. Graden av homogenitet i målgruppa tilsier også at dette vil være et tilstrekkelig antall for å nå en *teoretisk metning*, i motsetning til om jeg hadde sammenlignet barne- og ungdomsskolelæreres oppfatninger av vurdering og dermed hadde hatt en mer heterogen målgruppe i form av både praktiserende og ikke-praktiserende karakterbruk.

3.6. Analyseprosessen

Kvalitative studier har ofte en induktiv tilnærming, hvilket impliserer at forskeren samler data, for deretter å konstruere temaer eller kategorier ut ifra det innsamlede datamaterialet (Merriam, 2009, s. 15). På det området jeg skal undersøke har jeg opparbeidet en god del teorikunnskaper og har derfor gått motsatt vei med en deduktiv tilnærming. Samtidig var jeg åpen for nye temaer dersom lærerne skulle ha noen felles eller interessante temaer jeg ikke hadde forutsett.

Datamaterialet mitt ble gjenstand for en tematisk analyse. Braun & Clarke (2006) omtaler denne analysemodellen som fleksibel, som også presenterer de grunnleggende egenskapene en forsker kan ta med seg innenfor kvalitativ analyse. Til sammenligning tar Strauss & Corbin (1990) for seg grounded theory. I likhet med Schwarz' (2015) qualitative content analysis går også denne kodingen ut på å finne mønstre i data. Men i grounded theory er datainnsamling, dataanalyse og teoriutvikling i stor grad parallelle prosesser (Strauss & Corbin, 1990), og derfor utfordrende for en uerfaren forsker som meg. Selv om jeg inntar forskningen med et begrepsapparat og ikke er ute etter å utvikle teori som er «grounded» i datamaterialet, vil jeg likevel kode og kategorisere data på samme måte som i grounded theory.

Jeg transkriberte innsamlet data fra lyd til tekst, for deretter å kode teksten. Da jeg tematiserte kodingen min, tok jeg hensyn til Braun & Clarkes (2006) anbefalinger om å ikke knytte dette for nært opp mot intervjuguiden. Dette ga meg muligheten til å avvike fra de allerede fastsatte

temaene, og være åpen for nye. Som uerfaren forsker har jeg valgt å følge Braun & Clarkes (2006) seks faser i min analyse:

Fase 1 handler om å *bli kjent med datamaterialet*. Da jeg har stått for gjennomføring av intervjuene og transkriberingen selv, har det å bli kjent med datamaterialet vært en selvstendig prosess. Etter alle åtte intervjuene var gjennomført, hørte jeg gjennom lydopptakene for å få et helhetlig inntrykk av hva informantene sa, både hver for seg og som en enhet. Dette gjorde selve transkriberingen lettere, hvor jeg nok en gang fikk godt innsyn i datamaterialet. På denne måten dannet jeg meg allerede da et inntrykk av hva som var felles oppfatninger hos lærerne.

Fase 2 tar for seg den *innledende kodingen*. Denne fasen var utfordrende for meg som ikke hadde utformet et spesifikt forskningsspørsmål. Selv om Holloway & Todres (sitert i Braun & Clarke, 2006) sier: “*what is important is choosing a method that is appropriate to your research question, rather than falling victim to ‘methodolatry’, where you are committed to method rather than topic/content or research questions*” (s. 97), så mener Braun & Clarke (2006) at: “*you can either code for a quite specific research question (which maps onto the more theoretical approach) or the specific research question can evolve through the coding process (which maps onto the inductive approach)*” (s. 84). Da jeg på forhånd satt med et begrepsapparat og tilhørende teorikunnskap om tidligere forskning gjorde dette likevel analyseprosessen min lettere. Dette til tross for at jeg kodet intervjuene uten å sette de opp mot intervjuguiden. Jeg utførte da en *meningsfortetting* (in vivo) som medfører en forkortelse av intervjupersonenes uttalelser til kortere formuleringer. Lange setninger komprimeres til kortere, hvor den umiddelbare mening i det som blir sagt, gjengis med få ord (Kvale & Brinkmann, 2015, s. 232). Et eksempel er Heidis uttalelse som jeg har valgt å komprimere til «*summativ vurdering kan brukes formativt*»:

Heidi: Fremovermeldingene i elevsamtaler og vurderingssamtaler blir jo gjerne knyttet opp mot en karakter – sånn «på det siste så hadde du den karakteren – men jeg vet du har muligheten til å komme opp et nivå».

Jeg lokaliserte relevante utsagn knyttet til studiens tematikk. Mine analysestrategier kan trekkes paralleller ved til studien «To see the Woods for the Trees» (Vollstedt, 2015) da jeg har valgt å sammenligne intervju for å nå en *teoretisk metning*. Jeg startet med to intervju som var relativt like og rike på relevante data, for deretter å sette disse opp mot et annet som var mest mulig forskjellig fra de to andre. Analyseprosessen var inspirert av en *constant comparison* hvor kategorier og temaer konstant ble justert ved å kontinuerlig sammenlignes

med nye data (Percy, Kostere, & Kostere, 2015). Til slutt ble ingen nye kategorier eller temaer generert, og forholdet mellom disse var godt etablert. Jeg har dermed gjennom inspirasjon fra grounded theory brukt *åpen koding* med to analytiske prosesser hvor jeg: 1) stiller spørsmål til data, og 2) konstant sammenligner hendelser (Glaser, 1965).

I fase 3 *søker jeg etter temaer*, da kodene fra fase 2 igjen må grupperes i temaer eller kategorier for at datamengden skal bli håndterlig. Dette skjedde gjennom en *aksial koding* der kategorier blir relatert til sine subkategorier slik at forklaringene til fenomenet blir mer presise og fullstendige (Nilssen, 2012). Denne kodingen krever da at jeg har noen kategorier (fase 2), men at jeg allerede under den åpne kodingen begynner å danne meg en oppfatning av hvordan disse forholder seg til hverandre. Strauss A. L. (1987, s. 32) foreslår å stille spørsmål om hvordan ting henger sammen og finne ut dette gjennom deduktive og induktive analyser. Jeg har gjennom bruk av meningsfortetting og åpen koding slått sammen følgende koder fra forrige fase: «*karakterer kan virke fremmende på faglig sterke elever*», «*faglig sterke elever har en indre motivator*» og «*karakterer virker hemmende for faglig svake elever*» til «*karakterer kan ha en fremmende effekt på faglig sterke elever, og motsatt effekt på svake elever*».

Under fase 4 hadde jeg en *gjennomgang av temaer*. Etter sammenligning av koder og temaer på tvers av datasettet kom jeg frem til følgende 5 temaer: matematikk er et lett fag å vurdere i; karakterer kan ha en fremmende effekt på faglig sterke elever, og motsatt effekt på svake elever; relasjoner er avgjørende for en god vurderingsprosess; feedback avhenger av elevens faglige ståsted; summativ vurdering kan brukes formativt.

Fase 5 og 6 besto av *produksjon og ferdigstillelse av resultater*. Her skrev jeg om hvert tema, og hvordan disse henger sammen. Dette utgjorde rapporteringen av resultatet. Utenfor et eksisterende rammeverk har dette likevel en begrenset tolkningsverdi. Adler & Adler (sitert i Saldaña, 2009) hevder at “*..your level of personal involvement as a participant observer – as a peripheral, active, or complete member during fieldwork – filters how you perceive, document, and thus code your data*” (s. 7). Dette kan også ses i sammenheng med denne studien som baseres på oppfatninger. Typisk for kvalitative studier er utfordringen med å skille tydelig mellom observasjoner og egne tolkninger. Det betyr at idet jeg presenterer funnene mine må jeg begrunne og dokumentere at disse temaene faktisk oppsummerer/dekker lærernes oppfatninger av vurdering.

3.7. Studiens begrensninger

Jeg hadde ikke et ferdig utformet forskningsspørsmål før etter endt analyseprosess. Grunnen til at jeg endte opp med det noe brede forskningsspørsmålet «hva er matematikklæreres oppfatninger av vurdering i faget?» er for å begrense inkonsistens mellom datainnsamling og resultater. Dataene jeg satt igjen med etter innsamling utgjør like fullt et resultat av lærernes oppfatninger av vurdering i faget. En annen begrensning er intervjuguiden. Her er det sannsynlig at spørsmålene jeg har valgt har påvirket hvilke kategorier jeg har fått ut av analysen, og at andre spørsmål kunne ført til andre kategorier. Dette har jeg likevel tatt tak i ved å vise fram data i analysen. Samtidig har jeg valgt spørsmål for å dekke temaet så godt som mulig, og det er alltid slik at spørsmålene vil påvirke informasjonen man får inn. Dessuten har jeg i analyseprosessen fulgt Braun & Clarkes (2006) anbefalinger om å ikke knytte kodingen for nært opp mot intervjuguiden, men heller sammenlignet intervju ved en «constant comparison» (se kapittel 3.6.). Videre er generaliserbarheten min utfordrende (se videre kapittel 3.7.2.), da mine åtte deltakere utgjør et nokså begrenset utvalg. Jeg har likevel forsøkt å nærme meg en metning ved å spisse dette utvalget gjennom en kriteriebasert utvelgelse (se kapittel 3.5.).

3.7.1. Reliabilitet

Et viktig aspekt ved forskning er at data er avhengig av forståelse. Forskere kan ha ulike forhåndsoppfatninger av samme fenomen og følgelig ulik tilnærming eller fokus (Christoffersen & Johannessen, 2012). Forskjellige lesninger av et intervju kan resultere i ulike fortolkninger – for eksempel hvordan et ledende spørsmål fører til pålitelig eller upålitelig kunnskap. Kvale & Brinkmann (2015, s. 238) omtaler dette som *fortolkningsmangfold*. Dette impliserer at den dataen jeg sitter med nå er et resultat av min personlige interesse eller fokus. På samme måte bruker informantene mine sine egne personlige kunnskaper, forståelse og oppfatninger for å tolke både omverdenen og spørsmålene jeg stiller. Det er også viktig å merke seg at en kan ha mange oppfatninger, men i ulike kontekster vil det variere hvilke av disse som er «activated» (Schoenfeld, 1998, s. 14). Jeg vil også påstå at den fremkommende dataen ikke kun er et resultat av egne preferanser, da jeg har underkastet meg forskningsetiske prosedyrer. Kvale & Brinkmann (2015, s. 238) slår tilbake mot kritikken av at intervju ikke fremstår som en vitenskapelig metode dersom ulike fortolkere finner ulike meninger i det samme intervjuet; en uttalelse har ikke bare én riktig og objektiv mening, og min oppgave som fortolker er heller ikke å finne fram til denne ene sanne meningen. Altså er ikke kunnskap produsert via intervju objektiv, men det er heller ikke et

krav. Reliabilitet handler om hvorvidt, eller i hvilken grad, en ny etterprøvende studie vil produsere de samme resultatene. Jeg har vært klar over mine egne begrensninger som uerfaren forsker, og vært tydelig på hvordan jeg har samlet inn data, og hvordan disse ble analysert. På grunn av mine tydelige metodiske redegjørelser, styrkes reliabiliteten. Derimot utfordrer mine semistrukturerte intervjuer reliabiliteten. I motsetning til strukturerte intervju som sikrer samme spørsmålsstilling i hvert tilfelle og derav gir informasjon som er lett å produsere, har ikke jeg i like stor grad fulgt et manuskript. Jeg har dog fått tilgang på et kriteriebasert utvalg som impliserer at informantene mine har kunnskap på feltet. Også i forhold til transkripsjonen kan det stilles spørsmål ved påliteligheten. Kvale & Brinkmann (2015, s. 211) foreslår å foreta en kvantifisert reliabilitetssjekk: å la to personer hver for seg skrive ned den samme uttalelsen i et intervjuopptak, for deretter å la et dataprogram lage en liste over og telle antall ord som er ulike i de to transkripsjonene. Likevel, når det er rom for tolkninger både gjennom bruken av semistrukturert intervju der spørsmålstillingene varierer og det følgelig kan gi andre svar enn det som forventes, eller at jeg på samme måte som informantene mine kan ha en ulik oppfatning av samme utsagn, så vil ikke et slikt program være løsningen i forhold til både mitt og mine informanternes fortolkningsmangfold.

3.7.2. Troverdighet, generalisering og viktighet

Schoenfeld (2007b) trekker frem tre punkter for kvalitet i forskning: troverdighet, generalisering og viktighet. Med troverdighet mener han vi bør stille spørsmål ved blant annet repliserbarheten og hvorvidt en triangulering er foretatt. Dette fordi en *konteksteffekt* kan oppstå grunnet at menneskers respons på stimuli påvirkes av kontekst (s. 83). Innenfor kvalitativ metode benyttes ofte triangulering til å finne frem til de forskjellige aspektene i undersøkelsen (Clifford et al., 2010, s. 106). I forskning skiller man mellom intern og ekstern validitet, hvor nettopp triangulering bidrar til å styrke den *interne validiteten* gjennom å bruke ulike kilder til informasjon. Med studiens konstruktivistiske verdenssyn (se kapittel 3.1.) hvor kunnskap konstrueres av menneskers tolkning av virkeligheten, mener Golafshani (2003) at ulike datainnsamlingsmetoder må brukes for å oppnå gyldige data (s. 604). Min interne validitet blir altså påvirket gjennom studiens mangel på praktisering av flere forskningsmetoder. Dette er imidlertid ikke nødvendigvis en svakhet ved den praktiserte metoden, men snarere en konsekvens av min tidsfrist vedrørende innlevering av oppgaven. I forhold til repliserbarheten kan vi snakke om teoretisk og statistisk generalisering, som vil si at resultatene vil kunne være repliserbare gitt bestemte utvalg i gitte kontekster som kjennetegnes av bestemte egenskaper (Schoenfeld, 2007b, s. 87). Den *eksterne validiteten*

omhandler nettopp generalisering (Cohen et al., 2007). Mine anvendte teorikunnskaper i denne studien baserer seg for det meste på anerkjent internasjonal forskning. Med disse kunnskapene som grunnlag var jeg skeptisk til overføringsverdien til mine innsamlede data. Dette med tanke på at landene forsket på i studiene jeg setter forskningen min opp mot innebærer et helt annet skolesystem, og dermed vurderingskultur. Kvale & Brinkmann (2015, s. 290), i likhet med Bogdan & Biklen (referert i Cohen et al., 2007, s. 137) mener at hvis vi er interessert i generalisering, må vi ikke spørre om intervjuresultatene kan generaliseres globalt, men om den kunnskapen som produseres i en spesifikk intervjusituasjon kan overføres til andre relevante situasjoner. Schoenfeld (referert i Cohen, et al., 2007) påpeker derav viktigheten av *“in qualitative research to provide a clear, detailed and in-depth description so that others can decide the extent to which findings from one piece of research are generalizable to another situation”* (s. 137). På denne måten har jeg sikret validitet i form av tydelige og detaljerte beskrivelser av studiens design (se kapittel 3.2.). Creswell & Miller (2000) foreslår en annen måte å styrke validiteten på, “member checking”: *“it consists of taking data and interpretations back to the participants in the study so that they can confirm the credibility of the information and narrative account”* (s. 127). I mitt tilfelle vil dette tilsi å ettersende intervjutranskripsjonene til informantene mine med formål om en bekreftelse. Dette ville resultert i sikrere data, men med rammene satt for denne avhandlingen tillot jeg ikke meg selv å gjøre det. Når lærerne i tillegg hadde konstatert at de anså det å delta i studien som problemfritt da jeg ikke satte krav til noe for- eller etterarbeid, ønsket jeg ikke be om dette heller. Jeg har dog substituert med å høre gjennom intervjuene flere ganger for å sikre meg at transkripsjonene samsvarte med uttalelsene.

Schoenfelds (2007b) siste punkt for kvalitet i forskning omhandler viktighet og hvorfor vi skal bry oss om undersøkelsen. Vurdering, med særlig vekt på *vurdering for læring* er et stort satsningsområde i skolen. Ved bruk av alle mine åtte informanternes uttalelser var det etter endt analyse lite nye temaer som ble generert, og det kan derfor virke som jeg oppnådde en teoretisk metning. Dette kan igjen tyde på at resultatene er oppfatninger som er typiske i en norsk skole. Forskningsprosjektet er dog utført i Nord-Troms, og jeg kan derfor være teoretisk generaliserende, uten å generalisere statistisk. Dette fordi antall informanter, til tross for en tilsynelatende metning, gir et begrenset empirisk utvalg, noe som medfører mindre statistiske generaliserbare resultater.

3.8. Forskningsetiske overveielser

Forskning forutsetter at de ulike etiske retningslinjene ligger til grunn. Studier som mitt kvalitative forskningsintervju innebærer nær kontakt mellom forsker og de som studeres. Gjennom dette vil forskeren sitte med innsamlet data som kan knyttes til prosjektets deltakere (Thagaard, 2009). Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH) har vedtatt etiske retningslinjer i form av tre typer hensyn som forskeren må ta: (1) informantenes rett til selvbestemmelse og autonomi, (2) forskerens plikt til å respektere informantenes privatliv og (3) forskerens ansvar for å unngå skade (Nedrum referert i Christoffersen & Johannessen, 2012). Dette innebærer at jeg som forsker hadde et etisk ansvar ovenfor mine informanter, altså de intervjuede lærerne.

Før jeg kunne starte forskningen måtte jeg få prosjektet mitt godkjent av Norsk senter for forskningsdata (NSD). Personvernombudet vurderte prosjektet og fant behandlingen av personopplysninger meldepliktig i henhold til personopplysningsloven § 31. Med forutsetning av at prosjektet gjennomføres i tråd med opplysningene jeg oppga i meldeskjemaet, kunne behandlingen av personopplysninger settes i gang (se vedlegg 2).

Ved utvelging av informanter hadde jeg på forhånd skrevet *forespørsel om deltakelse i forskningsprosjektet* i form av et informasjonsskriv (se vedlegg 3). Her ba jeg om samtykke til deltakelse i studien, samt informerte om mitt formål med studien og hva eventuell deltakelse innebærer. De aktuelle lærerne skrev under på samtykkeskjemaet, og med det hadde jeg et informert samtykke (Kvale & Brinkmann, 2015) Samtidig gjorde jeg både en muntlig og mer formell skriftlig avtale med skolens ledelse, hvor også de ble informert om min bakgrunn og hensikt.

I intervjusituasjonen var det både viktig og avgjørende at informantene mine skulle føle seg trygge og komfortable. Dette både med tanke på de forskningsetiske overveielser og kvaliteten på innsamlet data. Jeg informerte om intervjuets varighet (noe som var avgjørende da intervjuet foregikk i arbeidstiden) og informantenes rett til å når som helst kunne trekke seg uten noen bestemt grunn, både fra intervjuet og selve prosjektet. Det var også viktig for meg å understreke at jeg ikke var ute etter å teste deres matematikkunnskaper på noe som helst nivå. Dette ble også presisert før selve intervjuet, da jeg da mottok flere forespørsler angående hvorvidt de måtte lese seg opp på noe i forkant. Jeg hadde på informasjonsskrivet informert om bruk av lydopptaker, men understrekte da nok en gang at det kun var jeg (og eventuelt veileder) som ville ha tilgang på materialet. Jeg påminnte om at kodenøkkelen som

gjør det mulig å koble lydopptaket opp mot intervjuobjektet ville til enhver tid være overvåket av meg, adskilt fra andre data, og være innelåst i et skap når jeg ikke arbeider med opptakene. Oppsummert har alle personopplysninger blitt behandlet konfidensielt, og taushetsplikten forblir overholdt.

4. Analyse og funn

Fra den tematiske analysen av datamaterialet oppstod det 5 temaer som sier noe om lærernes oppfatninger av vurdering i matematikkfaget. Clement (2000) snakker om analyser i flere nivå, der han går fra mer direkte observasjoner til mer spekulative eller teoretiske tolkninger. Jeg har valgt strukturere kapittel fire og fem på lignende måte.

Jeg har gitt informantene mine fiktive navn: Lise, Karl, Rune, Tor, Lars, Thomas, Jonas og Heidi. Jeg henviser til meg selv som intervjuer.

Etter en tematisk analyseprosess (se kapittel 3.6.) som kan minne om grounded theorys åpne og aksiale koding, presenteres de sammenslåtte kodene som følgende fem temaer:

4.1. Matematikk er et lett fag å vurdere i

I en samtale med Lise om vurdering i faget sa hun følgende:

Lise: [...] Matematikkoppgaver og svaret er så umiddelbart imot en norsktekst hvor flere faktorer spiller inn, så får du i matematikk et ene rett svar; det tror jeg skaper entusiasme og glede hos eleven.

Her ser vi at Lise synes det er en vesentlig forskjell mellom å vurdere i matematikkfaget og norskfaget. Hun nevner at mens hun i norskfaget må ta flere faktorer i betraktning, vil vi i matematikk få kun et riktig svar. Lise sier at dette fører til entusiasme og glede i matematikk. Fokuset Lise har på umiddelbare svar og ikke prosess er kjennetegn på instrumentell tenking (Skemp, 1976) og prosedyrekunnskap (Hiebert & Lefevre, 1986). En annen lærer, Tor, sier noe lignende:

Tor: Når jeg setter karakterer tar jeg utgangspunkt i maks poengsum.. det er jo ei grense for bere med 95% riktig, også går det nedover til 75, 55, 35, 15 osv – slik at en har en øvre og nedre grense – et intervall man må holde seg innenfor altså. Dette i motsetning til norskfaget hvor du må se på både skrivning, setningsbygning og innhold ikke minst.

Tor snakker om vurdering ut fra poengsum. Dette kan minne om Lises uttalelser om et ene rett svar i dette faget. Han sammenligner også med norskfaget for å underbygge påstanden om at matematikk er et lett fag å vurdere i: der må en se på både skrivning, setningsbygning og innhold. Dette ligner Lises første uttalelse om at man i norskfaget har flere faktorer som spiller inn.

Et fokus på å telle poeng for å vurdere tyder på at Tor vektlegger summativ vurdering. Samtidig vurderer han detaljer som skrivning, setningsbygning og innhold i norskfaget som gir muligheter for å vurdere formativt. Tor og Lise har flere tydelige likheter i sitatene sine.

Begge sitatene peker på at norskfaget er komplekst, og matematikkfaget er enkelt å vurdere i. En annen lærer, Rune, har et lignende syn:

Rune: Det å jobbe med matematikk er jo veldig konkret. Hvis du har et gangestykke så vet du jo at $9 \times 9 = 81$ – men får du en tekst eleven skal svare på i norsk for eksempel, så er det ikke noe du kan sette to streker under- for svaret kan være alt mulig, ikke sant. Hvis en elev gjør hjemmeregning og prøver og sånt så vet du hele tiden hvilken karakter du skal sette ut ifra den normen du har- ut ifra det du har jobbet med – for da har du jo et poeng til slutt, et tall – og det tallet gir da en femmer for eksempel.

Rune trekker også inn at matematikk er et konkret fag med et tall, poengsum, som gir karakteren fem. Dette kan tyde på at svaret, poengsum og karakter vektlegges mer enn prosess. Dette ligner Lises summative fokus med et ene rett svar og Tors fokus på poengsum. Runes sier at matematikk er konkret og at vi her i motsetning til norskfaget kan sette to streker under svaret. Dette minner veldig om et instrumentelt syn. Mens Lise, Tor og Rune begrunner at matematikk er lett å vurdere nokså likt, kommer Karl med et annet utsagn:

Karl: Jeg synes det er enkelt og mye raskere. Ei formativ vurdering i matematikkfaget går mye lettere enn språkfag. Og nå inkluderer jeg prosessen deres; når jeg vurderer og ser hvor langt de har gjort rett – for eksempel hvor langt de er inne på en tanke at de kan komme frem til rett svar – så har de for eksempel kommet halvveis inn i oppgaven og har tenkt på en eller annen måte at de er på tur mot et svar og så får de feil – så vurderer jeg jo frem til dit de har tenkt rett; her kunne dere ha tatt et sånt eller sånt valg og fortsatt kommet frem til rett svar – og da er min vurdering at de har delvis rett på oppgaven.

Karl mener også at vurdering er lettere å utføre i matematikk enn i språkfag. Han forklarer hvordan han utfører (formativ) vurdering underveis, og ikke bare basert på svaret. Karl fokuserer på prosess og ikke produkt, og utsagnet kan derfor kobles til relasjonell tenking og begrepsmessig kunnskap (Hiebert & Lefevre, 1986). Karls syn står altså i motsetning til de andre lærerne, da dette synet peker på at formativ vurdering er lettere i matematikk fordi han kan gå inn i prosessene og veilede. Tor og Rune mener det er lett fordi en kan sette poengsum og viser til det faktum at det er regler for hvor mange poeng som gir de ulike karakterene. Lise fokuserer på et ene rett svar og hvor umiddelbart dette er. Slike uttalelser peker mot en karakter- og poengfokusert praksis, som kan ligne et summativt fokus. At produkt vektlegges fremfor prosess peker også mot et instrumentelt syn i matematikkopplæringen. Karl skiller seg dermed fra de tre andre.

Det som er felles er at alle fire lærerne oppfatter matematikk som et lett fag å vurdere i. Jeg har i analysen oppdaget to ulike syn på hvorfor det er lett å vurdere i matematikk. Det ene handler om at det er et rett og et galt svar, telle poeng og sette karakter. Dette er et summativt syn som også kan knyttes til instrumentell tenking. Det andre handler om at det er lett å gå inn i prosessen og forme (formativ vurdering), hvilket kan kobles til relasjonell tenking.

4.2. Karakterer kan ha en fremmede effekt på faglig sterke elever, og motsatt effekt på svake elever

I en samtale med Lise om karakterenes effekt på elevene, sa hun følgende:

Lise: Karakterer virker fremmede for de aller fleste. Også har du noen som, kanskje særlig i matematikk og engelsk, de kommer ikke lenger enn til karakteren 3. Det er kanskje det lengste de kan komme. For dem vil dette være en negativ bekreftelse. Karakterene er jo der på ungdomsskolen, og jeg tror det at for de aller flinkeste- de som ligger i toppsjiktet- så er karakterene viktig- de er en motiverende faktor for at de faktisk skal være der oppe og gjøre den jobben.

Lise sier at karakterer virker fremmede for de aller fleste elevene. Og da spesielt viktige for de aller flinkeste elevene. Samtidig nevner hun elevene som ikke kommer lenger enn til lav eller middels måloppnåelse, og hvordan karakterbruk vil være en negativ bekreftelse for dem. Etter denne uttalelsen presiserer Lise igjen at karakterene er en motiverende faktor for de flinkeste elevene for å fortsette å gjøre den jobben de gjør. En annen lærer, Heidi, sier noe lignende:

Heidi: For de fleste tror jeg [karakterer]virker fremmede – de fleste blir litt sånn at de setter inn et ekstra gir og ja- det tror jeg. For de svake så tror jeg det er litt demotiverende- for noen av de synes selv de jobber veldig godt – og de kan jo være både flittig og arbeidsom uten å oppnå de faglige resultatene.

Heidi sier, i likhet med Lise, at karakterer vil virke fremmede for de fleste elevene. I motsetning til Lise, hevder Heidi at grunnen til dette er at elevene setter inn et ekstra gir. Lise mente at de var en motiverende faktor for de aller flinkeste elevene, og for at de skal fortsette å gjøre den jobben de gjør. Samtidig sier Heidi at karakterer vil virke demotiverende på de svake elevene. De kan være flittige og arbeidsomme uten å oppnå gode karakterer. I sammenheng med karakterbruk er dette ene utsagnet til Heidi mer nøytralt enn det ene utsagnet Lise kommer med. Dette til tross for hennes uttalelse om at karakterer virker fremmede for de fleste elevene. En annen lærer, Lars, omtaler karakterbruken og dens effekt på følgende måte:

Lars: [...] Karakterer er jo motiverende for elever som gjør det bra, men ikke for de som gjør det dårlig. Derfra kunne jeg egentlig godt ha tenkt meg at man ikke hadde brukt karakterer.

Lars mener at karakterer er fremmende for faglig sterke elever, men samtidig såpass hemmende for faglig svake, at han kunne tenkt seg en praksis uten denne type måling. Mens Lise har et utsagn om at karakterer er nødvendige for de flinke elevene, synes Heidi sine utsagn å forholde seg mer nøytral. Til tross for både hemmende og fremmende effekter av karakterbruk, viser altså Lars en tredje holdning til karakterbruk. I samtale med Karl om en vurderingspraksis uten karakter, sier han følgende:

Karl: En del elever vil nok synes det er trasig at han ikke fikk denne målingen

Intervjuer: Hvilken elev er det?

Karl: Det er nok en elev som er litt over snittet og opp som er veldig glad i å få den målingen og bekreftelsen på at den er så god som den trodde den var, og «jeg kan bli bedre – jeg har det i meg at jeg kan bli bedre». Også har du de andre da som sliter i faget, som synes det er vanskelig å få ting til og så får de en bekreftelse på at de faktisk er så dårlige.

Karl snakker om en vurderingspraksis uten karakterer og hvordan en slik praksis ville vært en uheldig situasjon for de sterke elevene. Han skyter inn at en karaktersettende praksis også gir de svake elevene en bekreftelse på at de er dårlige. Karls utsagn viser i likhet med Lise og Heidis, at de sterke elevene har behov for den bekreftelsen og målingen en (god) karakter medfører. På samme måte vil dette virke hemmende for de svake som vil få en negativ bekreftelse. Dette kan tyde på at han også, i likhet med Heidi, forholder seg nøytral til hvordan vurderingspraksisen vår er lagt opp. En annen lærer, Rune, sier følgende:

Rune: For de flinke elevene er karakterene en motivator. Men så er det de som er svak, de som hele tiden får 2ere eller 3ere og aldri ser lyset i tunnelen for å komme opp i 4 og 5. Da kan faktisk karakterene være destruktive- du gir opp – du kommer ingen vei – uansett hva du gjør så får du 2. En karakter er ikke motivator i seg selv, det må være andre ting som motiverer en elev for å gjøre det bra og være flink på skolen.

Rune sier her at de flinke elevene kan oppleve karakterer som en motivator. Han nevner også hvordan disse samtidig kan være destruktive for de svake elevene. Med bruk av ordet destruktiv utdyper han videre at eleven vil kunne gi opp fordi den ser at uavhengig av innsats vil den alltid få en dårlig karakter. Han legger til at en karakter ikke kan være en motivator i seg selv, og at det må være andre ting som motiverer en elev for å være flink på skolen. Rune tar altså denne konklusjonen til tross for at han mente karakterene er en motivator for de

flinke elevene. Samtidig presiserer han hvilken destruktiv effekt karakterene har på de svake elevene. Dette kan tyde på at Rune ikke ser behovet for denne målekulturen.

Fra tre ulike koder om at «karakterer kan virke fremmende på faglig sterke elever» (Lise, Heidi, Lars, Karl, Rune); «faglig sterke elever har en indre motivator» (Rune); og at «karakterer virker hemmende for faglig svake elever» (Lise, Heidi, Lars, Karl, Rune) oppsto disse sammenslått (aksial koding) til temaet «karakterer kan ha en fremmende effekt på faglig sterke elever, og motsatt effekt på svake elever». Dette fordi lærerne som mente karakterene kunne ha en fremmende effekt på de sterke elevene, også mente at karakterene kunne ha en hemmende effekt på de faglig sterke elevene. Likevel synes lærerne å konkludere ulikt rundt denne observasjonen. Mens Lise virker til å ha et syn om at karakterene er såpass viktige for de sterke elevene, kunne Lars tenkt seg en vurderingspraksis uten karakterer fordi det er såpass destruktivt for de svake elevene. Heidi og Karl kommenterer ikke noe ytterligere, og synes derfor å forholde seg nøytral til denne vurderingspraksisen. Rune har en litt annen uttalelse, som kan minne om Lars' konklusjon med at vi ikke trenger karakterer. Grunnlaget er dog forskjellig. I motsetning til Lars som oppfatter karakterer såpass demotiverende for svake elever, sier Rune at en karakter ikke er en motivator i seg selv. Oppsummert taler én lærer for karakterbruk, to er nøytrale, én taler mot, og den siste taler også mot, dog på et annet grunnlag.

4.3. Relasjoner er avgjørende for en god vurderingsprosess

I samtale med Lise om hennes tanker rundt prosessen for å finne ut hvor eleven er, hvor den skal, og hvordan den kommer seg dit, sa hun følgende:

Lise: Når man kjenner eleven etter hvert så vet du jo omtrent hvor potensialet ligger, så da vet du jo omtrent hvor langt du klarer å få dem.

Her trekker Lise paralleller mellom *kjennskap til eleven* (relasjoner) og *kunnskap om elevens potensiale*. Dette tyder på at relasjoner i form av å «kjenne eleven» er en faktor i Lises vurderingsprosess. Kunnskap om elevens potensiale inngår i Williams (2007) tre punkter i det han betegner som en fullverdig formativ vurderingsprosess. Lises uttalelse «når man kjenner eleven..» kan knyttes til «hvor eleven er» (Wiliam, 2007), «..så vet du jo omtrent hvor potensialet ligger» kan knyttes til «hvor eleven skal» (Wiliam, 2007). En annen lærer, Heidi, sier noe lignende:

Heidi: Den 8.klassen jeg har nå skal jeg følge ut tiende, tanken bak er å gjennom hele prosessen vite hvor de står. Veldig grei løsning.

Heidi tar opp hvordan kontinuitet i faget hjelper henne å holde oversikt over hvor elevene er i læringsprosessen. Når Heidi skal følge hennes klasse ut tiende klasse, vil hun hele veien ha en formening om elevens faglige ståsted. Dette tyder på at tid er en faktor som spiller inn på hvor godt hun kjenner eleven. Og med kontinuitet i faget vil læreren også ha oversikt over «hvor eleven er» (Wiliam, 2007). Altså kan Heidis uttalelse om kontinuitet i faget minne om Lises uttalelse om at kjennskap til eleven gir kunnskap om dens potensiale. Begge i den form av at disse faktorene gir kunnskap om «hvor eleven er» og kan knyttes til det Wiliam (2007) betegner som en fullverdig formativ vurderingsprosess.

I en samtale med Lars om relasjonsbygging og læring, sier han følgende:

Lars: Det har jo med kjemien mellom lærer og elev å gjøre. Vi har jo drevet med relasjonskartlegging på skolen med oss selv ovenfor elevene, det har med «Ungdomstrinn i Utvikling» å gjøre – forbedre miljøet litt, og da er relasjonen mellom lærer og elev en viktig faktor for å bedre læring. [...] Etter hvert så kjenner man jo elevene, jeg vet jo hvem som må ha hjelp og hvem som ikke må ha hjelp i klassen.

Lars sier relasjonen mellom lærer og elev er en viktig faktor for å bedre læring. Han hevder at gjennom kjennskap (relasjoner), så får han vite hvor elevene er i læringsprosessen, og derav har et godt grunnlag for en vurderingsprosess. Han sier at etter hvert som han kjenner elevene, så vil han ha en oversikt over hvem som treng hjelp og hvem som ikke gjør det. Dette peker i den retning av at han gjennom kjennskap vil vite hvor «eleven er» og «hvor den skal» (Wiliam, 2007). Hvilket resulterer i at hans relasjonsfokus (kjennskap) kan minne om Lises fokus på kjennskap og «hvor eleven skal». Utsagnet om at han vet hvem som må ha hjelp og ikke, forutsetter også at han vet «hvor eleven er». Hvilket punkt også Heidi fokuserer på gjennom kontinuitet i faget. En annen lærer, Jonas, har et litt annet fokus enn disse tre lærerne:

Jonas: Relasjon er veldig viktig, det har sammenheng med hvor godt du kjenner eleven og hvis eleven og du har en god relasjon så vil kanskje eleven i større grad ta til seg en tilbakemelding enn om dere ikke hadde en god relasjon: «nei han sa nå det, men jeg bryr meg ikke om det- han sier så mye rart». Så det med relasjoner er kjempeviktig – at elevene har tro på at det du sier er sant og riktig- at du ikke sier det bare for å si det- da blir du gjennomskuet av elevene.

Jonas peker på sammenhengen mellom relasjoner og kjennskap til eleven. Han hevder at en god relasjon er avgjørende i en vurderingsprosess der tillit fra elevene er essensielt. Dette kan ses i sammenheng med hva jeg i teoridelen nevnte om overfladisk skryt (Wiliam, 2007), og

Shutes (2008) mening om at tilbakemeldingen bør være basert på tillit. En annen lærer, Karl, sier noe lignende:

Karl: Relasjonsbygging er veldig i vinden. Vi har blant annet en kartlegging i forhold til vår relasjon til elevene. Det er for at vi skal bygge relasjonen vår og at du skal ønske å få vurdering fra meg, ønske å få veiledning fra meg.

Karl snakker om viktigheten av relasjonsbygging for at elevene skal ønske å få ei vurdering fra nettopp han. Hans utsagn tyder på at relasjoner ikke bare er viktig med tanke på oversikt over hvor elevene er i læringsprosessen (Lise, Heidi og Lars), men også for at eleven skal være mottakelig for å få vurdering fra aktuell lærer. Hvilket kan også minne om Jonas' utsagn angående tillitsforhold til læreren.

I forhold til å gjennomgå det Wiliam (2007) betegner som en fullverdig formativ vurderingsprosess (se kapittel 2.4.1.), synes lærerne å ha en felles oppfatning av at gode relasjoner er avgjørende i en vurderingsprosess. Fra de fem lærernes utsagn ser jeg dog tre ulike aspekter av relasjonsbygging som lærerne mener er viktig i en vurderingsprosess. Heidi fokuserer på *kontinuitet* i faget, og mener slik prioritering av tid hjelper godt i en vurderingsprosess. Lise og Lars snakker om relasjon i den form av *kjennskap*, og hvordan dette hjelper i å vite hvor langt du kan få eleven (i tillegg til å vite hvor eleven er). Til sist snakker Karl og Jonas om *tillit* og hvordan denne relasjonen er avgjørende i vurderingsprosessen.

4.4. Feedback avhenger av elevens faglige ståsted

I samtale med Heidi om kommentarfokus, sier hun følgende:

Heidi: Kommer litt an på hvor eleven er. Er det en sterk elev så bør en kanskje ikke ha hovedfokus på det de kan, men de få tingene de kanskje trenger å jobbe litt ekstra med. Men en svak elev kan kanskje like gjerne fokusere på det som er positivt – de trenger jo stort sett mere positive tilbakemeldinger, og det får de ikke i karakterene.

Her sier Heidi at en faglig sterk elev kan få kommentarer rettet mot videre arbeid, mens en faglig svak elev kan få kommentarer som vektlegger det positive. Denne kombinasjonen er derfor delvis i tråd med hva Wiliam (2007) sier om at vurdering må gi råd om videre arbeid for at den skal kunne kalles formativ. Dette referert til Heidis utsagn om den sterke eleven. På annen side er den også delvis i tråd med Matematikksenteret (2017) som mener at det eleven kan og mestrer bør vektlegges i større grad enn hva den ikke mestrer. Dette referert til Heidis utsagn om den svake eleven. En annen lærer, Thomas, sier noe lignende:

Intervjuer: I forhold til det å gi kommentarer – hvor mener du fokuset bør være, på det som var bra eller det som bør jobbes mer med?

Thomas: Begge deler. Jeg prøver hvis det er noen med prestasjoner som er bra eller at jeg ser en «ekstrem go» på en elev som kanskje ikke er en sånn topp elev, så kan jeg godt rose vedkommende.

Thomas sier at om «en elev som kan ikke er sånn topp elev» presterer bra, så kan han godt rose vedkommende, og dermed ha et kommentarfokus rettet mot det positive. Han sier også at han ville fokusert på både det positive og det som bør jobbes mer med. Dette er forskjellig fra det Heidi sier om at svake elever bør på kommentarer som vektlegger det positive. Og Thomas synes derfor å vise eleven rettledning videre, slik Wiliam (2007) mener en fullverdig formativ vurderingsprosess bør være. Dette skiller seg fra utsagnet til Heidi. Tor har enda et perspektiv:

Tor: Jeg har mange ganger hatt elever som har vært kjempeflinke, hatt kontroll på alt. Da er det jo litt sånn, hvor kritiske skal en være – skal en plukke på en liten detalj eller skal en se generelt på det. Jeg tenker på de som sliter både meg gangning og deling – det høres feil ut i mine ører å ha fokus på det negative.

Tor forholder seg usikker til hvor kritisk det er nødvendig å være med elever som «har vært kjempeflinke». Samtidig sier han at det virker feil at elever som sliter med både gangning og deling (faglig svake elever), skal få kommentarer som fokuserer på det negative.

Lærerne synes å ha en felles oppfatning av at kommentarfokuset bør være i henhold til «hvor eleven er», altså elevens faglige ståsted. Heidi, Tor og Thomas argumenterer for at svake elever har behov for oppmuntrende kommentarer. Mens Heidi samtidig mener at de sterke elevenes kommentarfokus bør være mer konstruktivt, mener Tor at det ikke alltid er nødvendig å plukke på detaljer. Sammenlagt synes en felles oppfatning å være at en svak elev og en sterk elev ikke ville fått samme type kommentar, til tross for like prestasjoner.

4.5. Summativ vurdering kan brukes formativt

I samtale med Lise og hennes tanker rundt formativ vurdering, sier hun følgende:

Lise: [...] Vi bruker for eksempel før vi skal ha prøver så har læreverket sånn oppsummering «prøv deg selv», og vi kjører hjemmeregning med det først så har de prøvd seg gjennom alle variantene så får de det tilbake så retter de og så er det prøve. Jeg tenker det er ei formativ vurdering.

Lise foretar altså en (summativ) vurdering av elevenes hjemmeregning, for deretter å gi den rettede prøven tilbake til elevene slik at de kan rette seg selv i forkant av prøven. Denne form for selvregulert læring omtaler Lise som formativ vurdering. En summativ vurdering kan ha

en formativ funksjon så lenge den gir rettleiding til videre arbeid. En forutsetning er dog (som ikke blir presisert her) at disse rettede prøvene kommer med konstruktive tilbakemeldinger. I samtale om kapittelprøver og deres formål og eventuell oppfølging, sier Thomas følgende:

Thomas: Jeg sender kapittelprøvene hjem til underskrift. Jeg skriver ofte en kommentar: «gå gjennom disse oppgavene på nytt hjemme, se læreboka s. 46» Foreldrene har et ansvar her altså, det må vi ikke glemme.

Thomas sier han retter elevenes kapittelprøver (summativ), og bruker hjemmet for at elevene skal jobbe mer med de oppgavene de ikke hadde prestert tilstrekkelig på. Denne måten å jobbe med prøvene i etterkant minner om Lises utsagn og hennes ene nevnte måte å drive formativ vurdering på. Her vil den samme forutsetning med konstruktive tilbakemeldinger gjelde. I samtale med Lars om formativ vurdering, sier han noe lignende:

Lars: Mine elever har hjemmelekse hver dag – de skal levere inn hjemmeregning til meg hver dag på skolen. De skal gjøre minst to stykker hjemme som de har i lekse. Jeg retter det, og så får de det tilbake hver dag for så å rette sine feil i tillegg til å gjøre to nye oppgaver. Jeg kaller det prosessorientert matematikk, og det er min måte å drive formativ vurdering på. Jeg får god oversikt over hvor hver enkelt står. I tillegg får du inn det her med mengdetrening i faget – de må regne litt hver dag. Stort sett er det oppgaver som de bør kunne, det de får i hjemmeregning. Her ser jeg også hvem som gjør noe og hvem som ikke gjør noe.

Lars snakker også om hjemmelekse i form av hjemmeregning. Mens Lise og Thomas snakker om at hjemmeregning gjøres i forbindelse med prøver og prøveretting, sier Lars at dette er noe som gjøres hver dag. I likhet med Lise med Thomas, sier Lars at han retter oppgavene og gir dem tilbake til elevene. Han sier også at han gir to nye oppgaver i tillegg til å rette. På denne måten sier at han får inn mengdetrening, og at slik prosessorientert matematikk er hans måte å drive formativ vurdering på. Lars' form for mengdetrening og prosessorientert matematikk er i likhet med Thomas og Lises metoder, egentlig summative hensikter gjort om til formative. Dette igjen med den forutsetning at «rettingen» hans er i tråd med det som kjennetegner formativ vurdering. En annen lærer, Heidi, tar opp et annet perspektiv:

Heidi: Fremovermeldingene i elevsamtaler og vurderingssamtaler blir jo gjerne knyttet opp mot en karakter – sånn «på det siste så hadde du den karakteren – men jeg vet du har muligheten til å komme opp et nivå».

Heidi snakker om hvordan fremovermeldingene i elevsamtaler ofte baseres på karakterer, samt hvordan disse kan brukes som rettleiding til videre læring. Karakterene som egentlig har

en summativ funksjon blir her brukt som et sammenligningsgrunnlag for videre arbeid. En vurdering sett fra «hvor eleven er» (slik Heidi tar utgangspunkt i) kan likevel kun kalles formativ om hun også forteller «hva som skal til for at eleven kommer dit». Snakk om «fremovermelding» tyder på dette.

En summativ vurdering er karakterisert ved at den kartlegger den nåværende forståelsen innenfor et bestemt emne (se kapittel 2.2.). På denne måten kan blant annet hjemmeregning betegnes som summativ. Lise snakker om hjemmeregning i forkant av prøve. Thomas snakker om hjemmeregning i etterkant av prøve. Og Lars snakker om hjemmeregning og prosessorientert matematikk. Felles for disse tre lærernes utsagn er at elevene deres utfører hjemmeregning/prøve, for deretter å få ei summativ vurdering hos læreren. Vurderingen stopper likevel ikke der. Lærerne sier elevene får tilbake de rettede oppgavene, for så å jobbe videre ut fra det som er rettet. Fellen i dette temaet er min forutsetning om at «de rettede oppgavene» kommer med konstruktive tilbakemeldinger, slik Wiliam (2007) definerer ei formativ vurdering. Heidi tar opp et litt annet perspektiv. Hun sier at hun kan bruke elevens karakterer til å gi ei fremovermelding på hvor eleven skal. Felles for disse utsagnene er at karakterer som egentlig har en summativ funksjon, her blir brukt formativt. Dog i en viss grad, og med forbehold om at «de rettede prøvene» gir rettleiding til videre arbeid.

5. Diskusjon

I kapittel 4 har jeg tolket funnene mine og sett hva som generelt preger temaene og hvordan disse henger sammen. I dette kapitlet vil jeg drøfte funnene jeg gjorde i kapittel 4 opp mot teorigrunnlaget mitt fra kapittel 2.

5.1. Matematikk er et lett fag å vurdere i

Lærerne sier matematikk er lett å vurdere i, men begrunnet det på to ulike måter. Lise, Tor og Runes begrunnelse handler om hvordan faget har et rett og et galt svar, samt poenggrenser og karaktersetting. Jeg relaterer dette til en instrumentell forståelse (Skemp, 1976) med summative hensikter. Når lærerne er mer opptatt av svar og resultat, kan det også minne om et kunnskapssyn som er prosedyrebasert (Thompson et al., 1994; Hiebert & Lefevre, 1986). Læreren Karl skiller seg fra Lise, Tor og Rune ved at han vektlegger prosess fremfor produkt. Han snakker om matematikk som et fag det er lett å gå inn i prosessen og forme. Hans perspektiv er derfor formativt og kan kobles til relasjonell forståelse (Skemp, 1976). Hans fokus på begreper og kontekst minner om et konseptuelt kunnskapssyn (Thompson et al., 1994). Lærernes ulike orientasjoner, eller kunnskapssyn, resulterer i ulike tilnærminger, slik det også viser seg hos disse lærerne. Selv om lærerne har ulike syn, så er det ikke nødvendigvis verken rett eller galt da begge perspektivene er viktige (Hiebert & Lefevre, 1986). Skemp (1976) snakker også om instrumentelle tilnærminger som en tidsmessig fordel, noe som kan forklare hvorfor majoriteten av disse lærerne utøver en slik praksis.

5.2. Karakterer kan ha en fremmede effekt på faglig sterke elever, og motsatt effekt på svake elever

Oppsummert handler dette temaet om at karakterer er fremmede for de flinke elevene, men ikke for de svake elevene. Fra analysen var det én lærer som talte for karakterbruk, to var nøytrale, to talte mot – dog på ulike grunnlag. Dette funnet synes å samsvare med Butlers (1998) studie og hans funn om at sterke elever var positive til «scores only», mens svake elever ikke var det. Butlers (1988) studie viser også at lavt-presterende (svake) elever var negative til karakterer generelt, da elevene var positive til kommentarer, men ikke i kombinasjon med karakterer. Page (1958) fant også noe tilsvarende. Hun mente at karakterbruk generelt ikke førte til økt læringsutbytte, og her uavhengig av elevens nivå. Smith & Gorards (2005) studie «They don't give us our marks» viste derimot ingen klare indikasjoner på at å fjerne karakterene hadde en positiv effekt. Verdt å nevne er dog elevenes

ordinære vurderingspraksis hvor alt arbeid ble målt. Et slikt system vil trolig resultere i, slik som vist i studien, en vurderingssituasjon hvor elevene er avhengige av å få en karakter for å ha begrep om kvaliteten på arbeidet deres.

I snakk om karakterbruk, sier Utdanningsforbundet (2017) at elever som motiveres av å oppnå best mulig karakter, ikke vil oppnå like god kunnskap som den som har den indre motivasjonen til å lære. Dette var også et argument mot karakterbruk som skilte seg ut blant lærerne. Samtidig som Rune omtalte karakterbruk som hemmende for svake elever og fremmede for sterke, sa han også at karakteren ikke kan være en motivator i seg selv. Utdanningsforbundet (2017) sier videre at om elevene skal bli flinke, er det viktig at de motiveres ut fra et eget ønske om å forstå. Jeg knytter dette «ønske om å forstå» opp mot et intellektuelt behov (Harel, 2013), og mener at det å belønne gjennom karakterer kan virke hemmende i en slik læringsprosess. I følge Runes uttalelser kan altså karakterer være hemmende for de faglig svake elevene i forhold til demotivasjon, men også for de faglig sterke (og svake) i forhold til å skape et intellektuelt behov.

Et annet interessant funn er hvordan svake elever ikke ønsker karakterer i faget, og at svake elever synes å omtales svak uavhengig av temaer i faget. Dette funnet minner mye om Schoenfelds (2007a) studie hvor det ble gitt en SAT-9 test og en Balanced Assessment-test for å se hva prøvene som brukes i skolen faktisk tester. Det viste seg at mer enn 90% av elevene som ble erklært kyndige på sistnevnte test, også var erklært kyndige på førstnevnte. Inkonsistensen ligger i at kun en tredjedel av elevene som var erklært kyndige på førstnevnte test ikke var erklært kyndige på siste testen (s. 63). Dette kan overføres til denne stigmatiseringen av «svake elever», og tyde på at prøvene våre fremstår homogene og ikke legger opp til testing av ulike kompetanser i faget.

5.3. Relasjoner er avgjørende for en god vurderingsprosess

En felles oppfatning synes å være at relasjoner er avgjørende i en vurderingsprosess. Tre ulike aspekter fremkom i analysen, hvorav en omhandler *kontinuitet* i faget, en annen *kjennskap* og den tredje *tillit*.

Forskere synes å ha en felles mening om at formativ vurdering består av tre nøkkelprosesser: “1) *Establishing where the learners are in their learning*, 2) *Establishing where they are going and* 3) *Working out how to get there*” (William, 2007, s. 1064). På denne måten vil Williams (2007) punkt 1 og 2 komme med kontinuitet i faget, da læreren her vil ha en større oversikt over eleven og dens læringsprosess. Kjennskap er på mange måter en konsekvens og

et mål ved kontinuitet, og er da også en fordel for en vurderingsprosess, med spesielt punkt 1 og 2 etter Wiliam (2007). Fra Wiliams (2007) betegnelse på en god vurderingsprosess viser det seg altså at kontinuitet og kjennskap begge fyller to av disse tre punktene. Dette tyder på at relasjoner faktisk er avgjørende for en god vurderingsprosess. Et tredje aspekt omhandlet tillit til læreren. To lærere nevnte at tillit fra elevene var av essensiell betydning. Jonas sier: «Så det med relasjoner er kjempeviktig – at elevene har tro på at det du sier er sant og riktig – at du ikke sier det bare for å si det – da blir du gjennomskuet av elevene». Shute (2008) sier noe lignende med at tilbakemeldingen bør være elevsensitiv og basert på tillit. Karl sier at relasjon er viktig for at eleven skal være mottakelig for å få vurdering fra nettopp han. I Hattie (2008) sin metastudie «Visible Learning» står relasjoner mellom lærer og elev frem som viktig for elevenes prestasjoner.

Relasjoner er altså viktig, noe vi ser hos både Wiliam (2007) og Hattie (2008). Og lærerne peker tillit, kontinuitet og kjennskap som faktorer innenfor temaet.

5.4. Feedback avhenger av elevens faglige ståsted

Lærerne synes å ha en felles oppfatning av at kommentarfokuset bør være i henhold til elevens faglige ståsted. Samtidig som det fremkom av analysen at de sterke elevenes kommentarfokus bør være mer konstruktivt, hevdet også en lærer at det ikke alltid var nødvendig å plukke på detaljer. Et felles utsagn hos lærerne var likevel at en svak elev får og behøver oppmuntrende feedback. Forskning synes også å tale ulikt om kommentarfokuset. Mens Matematikksenteret (2017) mener vi bør legge vekt på å synliggjøre for eleven hva den kan og mestrer, i større grad enn hva den ikke mestrer, mener Wiliam (2007) at feedback bør fokusere på hva eleven trenger å forbedre fremfor hva som er bra (s. 1015). “Principles and Standards for School Mathematics” presenterer «vurdering burde fremme elevs læring» som én av to hovedideer (NCTM, 2000, s. 22). På samme måte foreslår Kunnskapsdepartementet (2017) sitt høringsutkast til ny generell del av læreplanen at kartlegging og observasjon av hver enkelt elev har liten verdi dersom det ikke følges opp med konstruktive tiltak (s. 16). I forskrift til opplæringsloven (2009, § 3-11) står det om undervisvurdering at elevene skal få tilbakemeldinger som forteller dem om kvaliteten på arbeidet eller prestasjonen. Videre under samme paragraf står det også at vurderingen skal inneholde informasjon om hvilken kompetanse eleven innehar samt rettleiding til hvordan eleven kan utvikle kompetansen i faget. Hvilket har klare likheter til Wiliams (2007) tre punkter for en fullverdig formativ vurderingsprosess. Smith & Goards (2005) studie underbygger på mange måter Wiliams

påstand om riktig feedback. Studien kom frem til at en vurderingspraksis uten karakterer ga lite eller ingen effekt, og ble heller oppfattet skadelig for eleven. I studien fremkom det at kommentarene fokuserte mer på å fremme elevenes selvbylde, enn på hva som måtte jobbes med, og hvordan (s. 36). Dette funnet er altså i tråd med hva forskning sier om at det ikke holder med oppmuntring.

Med dette så viser det seg at mens litteraturen ikke vektlegger forskjell mellom sterke og svake elever, men heller kvaliteten på tilbakemeldingene, så legger lærerne vekt på at det er en forskjell på hvordan feedback gis til sterke og svake elever. Dette utgjør en interessant forskjell som bør undersøkes videre.

5.5. Summativ vurdering kan brukes formativt

Felles for lærernes utsagn er at karakterer/poengsummer som egentlig har en summativ funksjon, her (med forbehold) blir brukt formativt. Felles for disse utsagnene er også den forutsetning at det lærerne har «rettet» er i tråd med hva forskning sier om formativ vurdering og konstruktive tilbakemeldinger. Forskere mener en formativ vurdering består av tre nøkkelprosesser: “1) *Establishing where the learners are in their learning*, 2) *Establishing where they are going and* 3) *Working out how to get there*” (Wiliam, 2007, s. 1064). Og i forskrift til opplæringsloven (2009, § 3-11) står det om underveisvurdering at elevene skal få tilbakemeldinger som forteller dem om kvaliteten på arbeidet eller prestasjonen. Videre under samme paragraf står det også at vurderingen skal inneholde informasjon om hvilken kompetanse eleven innehar samt rettleiding til hvordan eleven kan utvikle kompetansen i faget.

Wiliams (2007) tre punkter ved formativ vurdering er å finne igjen ved Schoenfelds (2016) rammeverk TRU-Math. Den ene av fem dimensjoner omhandler formativ vurdering; å møte elevene der de er og forme undervisningen deretter:

- 1) Student reasoning is not actively surfaced or pursued. Teacher actions are limited to corrective feedback or encouragement
- 2) The teacher refers to student thinking, perhaps even to common mistakes, but specific students' ideas are not built on (when potentially valuable) or used to address challenges (when problematic).
- 3) The teacher solicits student thinking and subsequent instruction responds to those ideas, by building on productive beginnings or addressing emerging misunderstandings (Schoenfeld, 2014, s. 23)

Det første nivået i denne dimensjonen indikerer at lærerens praksis er svært begrenset. Hvilket kan kobles til Smith & Gorards (2005) studie der lærernes tilbakemeldinger per definisjon ikke var formative. Og følgelig en intervensjon som ikke fungerte. Mine intervjuede lærere har alle allerede «rettet» elevenes feil, hvilket indikerer at «*the teacher refers to student thinking...*» og befinner seg på nivå 2. Hvorvidt “*the teachers solicits student thinking and subsequent instruction responds to those ideas..*” avhenger likevel av lærernes tilbakemeldinger (opplæringsloven, 2009, § 3-11), og avgjør hvorvidt uttalelsene deres samsvarer med det høyeste nivået i rammeverket, og er i tråd med Williams (2007) tre punkter i en fullverdig formativ vurderingsprosess. Verdt å nevne her er funnet om at lærerne mener «5.4. Feedback avhenger av elevens faglige ståsted». Hvilket indikerer at dette funnet kun vil gjelde for de faglig sterke elevene.

6. Avslutning

6.1. Oppsummering

Jeg har i dette mastergradsprosjektet undersøkt hvilke oppfatninger åtte matematikklærere har av vurdering i faget. Til tross for at lærernes oppfatninger må ses i kontekst med deres kunnskapssyn eller orientering (Thompson et al., 1994), og ta i betraktning at dette er et resultat av de oppfatningene som er «activated in current context» (Schoenfeld, 1998), så har jeg funnet fem funn jeg mener dekker lærernes oppfatninger av vurdering.

Den første kategorien var at lærerne synes matematikk er et fag som er lett å vurdere i. Bare en av lærerne fokuserte på prosessen på en måte som kan relatere til relasjonell tenking (Skemp, 1976), mens de andre lærerne fokuserte på resultatet på måter som kan kobles til et instrumentelt (Skemp, 1976), calculational (Thompson, et al., 1994) eller prosedyrebasert (Hiebert & Lefevre, 1986) syn.

Den andre kategorien handler om at karakterer virker fremmende på sterke elever, og hemmende på svake elever. Funnet synes å samsvare med tidligere forskning. Butlers (1988) studie kom frem til at faglig svake elever var negative til to av tre tilbakemeldingsformer: bare karakterer og kommentarer i kombinasjon med karakterer. I samme studie var de sterke elevene positive til alle tilbakemeldingsformene. Én av de intervjuede lærerne trakk også frem at en karakter ikke kan være en motivator i seg selv. Utdanningsforbundet (2017) sier noe lignende, og understreker viktigheten av et «ønske om å forstå», eller et *intellektuelt behov* (Harel, 2013).

Den tredje kategorien omhandler relasjoner og hvordan disse er avgjørende i en vurderingsprosess. Lærerne peker på tillit, kontinuitet og kjennskap som faktorer innenfor temaet. Kontinuitet og kjennskap ser begge ut til å fylle to av Williams (2007) tre punkter for en god vurderingsprosess, noe som viser viktigheten av relasjoner i en slik prosess. Hattie (2008) viser også i sin metastudie «Visible Learning» at relasjoner mellom lærer og elev står frem som viktig for elevenes prestasjoner.

Den fjerde kategorien tar for seg hvordan lærerne sier at en svak elev og en sterk elev ikke ville fått samme type kommentar, tross like prestasjoner. Mens Smith & Gorards (2005) studie viste at feedback med oppmuntring ikke fungerte, sier annen litteratur og forskning (William, 2007; NCTM, 2000; Kunnskapsdepartementet, 2017; Forskrift til opplæringsloven,

2009, § 3-11) at elevene skal få tilbakemeldinger som forteller dem om prestasjon samt rettledning til utvikling i faget.

Den femte kategorien handler om hvordan summative hensikter som prøver og karakterer, *kan* brukes til et formativt formål. Wiliams (2007) formative vurderingsprosess har likheter med *formative assessment*-dimensjonen i Schoenfelds (2016) rammeverk TRU-Math. Hvorvidt lærerne underviser (vurderer) etter det høyeste nivået i rammeverkets dimensjon avhenger likevel av om feedbacken deres samsvarer med det forskning sier om formativ vurdering. Etter andre funn å bedømme, tyder mye på at «feedback avhenger av elevenes faglige ståsted» og at dette funnet følgelig kun vil gjelde for de sterke elevene.

6.2. Konklusjon

Med disse fem kategoriene har jeg svart på forskningsspørsmålet mitt. De viser til ulike oppfatninger som lærerne hadde om deres vurderingspraksis. Det viste seg blant annet at majoriteten av lærerne innehar en instrumentell forståelse til matematikkundervisningen. Dette synes å samsvare med at norske elever altfor sjelden får mulighet til å utvikle begrepsmessig kunnskap eller relasjonell forståelse (Matematikksenteret, 2015). En annen felles oppfatning hos lærerne var at karakterer har ulik effekt på ulike elever. Et annet tema var at også ulike elever får ulik feedback. Svake elever ønsker ikke karakterer, og tilsvarende elevgruppe får sjelden konstruktive tilbakemeldinger. Jeg stiller spørsmål ved om det er en sammenheng her. En annen felles oppfatning lærerne hadde var at summativ vurdering kan brukes formativt. Men uten rettledning til videre arbeid (slik forskning sier om formativ vurdering) til de svake elevene, vil de komme til kort, også her. En gjennomgående sak i denne forskningen er at en svak elev omtales svak uavhengig av tema innenfor faget. På bakgrunn av mine funn kan denne stigmatiseringen rettferdiggjøres gjennom lærernes vurderingspraksis i dette faget.

6.3. Veien videre

Det viser seg at «hvor eleven er» er avgjørende både for ønsket type feedback (kommentar, karakter), samt lærerens kommentarfokus til aktuell elev. Forskning sier likevel at vurdering skal inneholde rettledning til utvikling av kompetanse, uavhengig av elevens faglige ståsted. Altså, mens litteraturen ikke vektlegger forskjell mellom sterke og svake elever (men heller

kvaliteten på tilbakemeldingene), så mener lærerne at feedback skal justeres basert på elevene sine kunnskaper. Dette utgjør en interessant forskjell som bør undersøkes videre.

Litteraturliste

- Akinsola, M. K. (2009). Comparison of Prospective and Practicing Teachers' Mathematics Efficacy Beliefs Regarding Mathematics Teaching and Classroom Management. *Beliefs and Attitudes in Mathematics Education: New Research Results*, ss. 135-148.
- Assessment Reform Group. (2002). Assessment for Learning: 10 principles. I *Research-based principles to guide classroom practice*. Nuffield Foundation.
- Birenbaum, M. (1996). Assessment 2000: Towards a Pluralistic Approach to Assessment. *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge*, 42, ss. 3-29.
- Black, P., & Wiliam, D. (1998, Mars). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1).
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, 21, ss. 5-31.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86(1), ss. 8-21.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), ss. 77-101.
- Brown, G. T., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an Assessment for Learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, ss. 968-978.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British journal of educational psychology*, 58(1), ss. 1-14.
- Caelli, K., Ray, L., & Mill, J. (2003). 'Clear as Mud': Toward Greater Clarity in Generic Qualitative Research. *International Journal of Qualitative Methods*.
- Christoffersen, L., & Johannessen, A. (2012). *Forskningsmetode for lærerutdanningene*. Abstrakt forlag AS.
- Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. I R. Lesh, & A. Kelly, *Handbook of research methodologies for science and mathematics education* (ss. 341-385). Hillsdale, NJ: Lawrence Erlbaum.
- Clifford, N., French, S., & Valentine, G. (2010). *Key Methods in Geography* (2. utg.). SAGE.
- Cobb, P. (2007). Putting Philosophy to Work: Coping with Multiple Theoretical Perspectives. I F. K. Lester, *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 1, ss. 3-38). Charlotte, NC: Information Age Publishing.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6. utg.). London: Routledge.
- Creswell, J. W. (2014). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches* (4. utg.). Thousand Oaks, CA: SAGE Publications, Inc.

- Creswell, J. W., & Miller, D. L. (2000). Determining Validity in Qualitative Inquiry. *Theory Into Practice*, 39(3), ss. 124-130.
- Daskalogianni, K., & Simpson, A. (2000). Towards a definition of attitude: The relationship between the affective and the cognitive in pre-university students. *PME CONFERENCE*, 2, ss. 2-217.
- DeBellis, V. A., & Goldin, G. A. (1997). The affective domain in mathematical problem-solving. *PME CONFERENCE*, 3, ss. 2-209.
- Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, 38(1), ss. 47-65.
- Forskrift til opplæringslova. (2009). *Kapittel 3. Individuell vurdering i grunnskolen og i vidaregåande opplæring, § 3-11. Underevgsvurdering*. Hentet fra https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_4#KAPITTEL_4
- Furinghetti, F., & Pehkonen, E. (2002). Rethinking characterizations of beliefs. *Beliefs: A hidden variable in mathematics education?*, ss. 39-57.
- Gettier, E. L. (1963). Is Justified True Belief Knowledge? *Analysis*, 23(6), ss. 121-123.
- Glaser, B. G. (1965). The Constant Comparative Method of Qualitative Analysis. *Social Problems*, 12(4), ss. 436-445.
- Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Researchn. *The Qualitative Report*, 8(4), ss. 597-606.
- Grbich, C. (2013). *Qualitative Data Analysis: An Introduction* (2. utg.). London: SAGE Publications Ltd.
- Harel, G. (2013). Intellectual Need. I K. R. Leatham, *Vital Directions for Mathematics Education Research* (ss. 119-151). Springer New York.
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie, J., & Timperley, H. (2007, Mars). The Power of Feedback. *Review of Educational Research*, 77(1), ss. 81-112.
- Hiebert, J., & Lefevre, P. (1986). Conceptual and Procedural Knowledge in Mathematics: An Introductory Analysis. *Conceptual and Procedural Knowledge: The Case of Mathematics*, ss. 1-27.
- Kluger, A. N., & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2), ss. 254-284.
- Kunnskapsdepartementet. (2017, Mars 10). *Overordnet del - verdier og prinsipper*. Hentet fra regjeringen.no: <https://www.regjeringen.no/contentassets/ac9720408d464a83a7926f33dbcb7616/horingsutkast-fra-kunnskapsdepartementet-10.03.17--overordnet-del---verdier-og-prinsipper.pdf>

- Kvale, S., & Brinkmann, S. (2015). *Det kvalitative forskningsintervju* (3. utg.). (T. M. Anderssen, & J. Rygge, Overs.) Gyldendal Norsk Forlag AS.
- Lev-Zamir, H., & Leikin, R. (2013). Saying versus doing: teachers' conceptions of creativity in elementary mathematics teaching. *ZDM Mathematics Education*, 45(2), ss. 295-308.
- Martin, G. N., Carlson, N. R., & Buskist, W. (2010). *Psychology* (4. utg.). Great Britain: Pearson Education.
- Matematikksenteret. (2015). *Sentrale kjennetegn på god læring og undervisning i matematikk*.
- Matematikksenteret. (2017, April 1). *Vurdering i matematikk*. Hentet fra <http://www.matematikkcenteret.no/content/2383/Vurdering-i-matematikk>
- McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. *Handbook of research on mathematics teaching and learning*, ss. 575-596.
- Merriam, S. B. (2009). *Qualitative Research: A Guide to Design and Implementation*. USA: Jossey-Bass.
- National Council of Teachers of Mathematics (NCTM). (1991). *Professional standards for teaching mathematics*. Reston: VA: Author.
- National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston: VA: NCTM.
- Nilssen, V. (2012). *Analyse i kvalitative studier. Den skrivende forskeren*. Oslo: Universitetsforlaget.
- NOU 2015: 8. (2015). *Fremtidens skole: Fornyelse av fag og kompetanser*. Hentet fra <http://nettsteder.regjeringen.no/fremtidensskole/files/2015/06/NOU201520150008000DDDPDFS.pdf>
- Op't Eynde, P., De Corte, E., & Verschaffel, L. (2002). Framing Students' Mathematics-Related Beliefs. *Beliefs: A Hidden Variable in Mathematics Education?*, 31, ss. 13-37.
- Page, E. B. (1958). Teacher comments and student performance: A seventy-four classroom experiment in school motivation. *Journal of Educational Psychology*, 49(4), ss. 173-181.
- Percy, W. H., Kostere, K., & Kostere, S. (2015, Februar 16). Generic Qualitative Research in Psychology. *The Qualitative Report*, 20(2), ss. 76-85.
- Philipp, R. (2007). Mathematics Teachers' Beliefs and Affect. I F. K. Lester, *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 7, ss. 257-315). Charlotte, NC: Information Age Publishing.
- Saldaña, J. (2009). An Introduction to Codes and Coding. I J. Saldaña, *The Coding Manual for Qualitative Researchers* (ss. 1-32). SAGE Publications.
- Schoenfeld, A. H. (1998). Toward a theory of teaching-in-context. *Issues in Education*, 4(1), ss. 1-94.

- Schoenfeld, A. H. (2007a). *Assessing Mathematical Proficiency*. Cambridge University Press.
- Schoenfeld, A. H. (2007b). Method. I F. K. Lester, *Second Handbook of Research on Mathematics Teaching and Learning* (Vol. 3, ss. 69-107). Charlotte, NC: Information Age Publishing.
- Schoenfeld, A. H. (2016). *An introduction to the Teaching for Robust Understanding (TRU) Framework*. Berkeley, CA: Graduate School of Education.
- Schwarz, B. (2015). A Study on Professional Competence of Future Teacher Students as an Example of a Study Using Qualitative Content Analysis. I A. Bikner-Ahsbals, C. Knipping, & N. C. Presmeg, *Approaches to Qualitative Research in Mathematics Education* (ss. 381-399). Springer.
- Shute, V. J. (2008, Mars). Focus on Formative Feedback. *Review of Educational Research*, 78(1), ss. 153-189.
- Skemp, R. R. (1976). Relational Understanding and Instrumental Understanding. *Mathematics Teaching*, 77, ss. 20-26.
- Smith, E., & Gorard, S. (2005, Mars). 'They don't give us our marks': the role of formative feedback in student progress. *Assessment in Education*, 12(1), ss. 21-38.
- St.Meld. 22. (2010-2011). *Motivasjon - mestring - muligheter - ungdomstrinnet*. Hentet fra <https://www.regjeringen.no/no/dokumenter/meld-st-22-2010--2011/id641251/sec1>
- Stobart, G. (2008). *Testing Times: The uses and abuses of assessment*. Routledge.
- Strauss, A. L. (1987). *Qualitative Analysis for Social Scientists*. Cambridge University Press.
- Strauss, A., & Corbin, J. (1990). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (2. utg.). SAGE Publications.
- Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., . . . Vos, P. (2016). Assessment in Mathematics Education. I *Large-Scale Assessment and Classroom Assessment* (ss. 1-38). Springer International Publishing.
- Thagaard, T. (2009). *Systematikk og innlevelse. En innføring i kvalitativ metode* (3.. utg.). Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, 15(2), ss. 105-127.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. *Handbook of research on mathematics teaching and learning*, ss. 127-146.
- Thompson, A. G., Philipp, R. A., Thompson, P. W., & Boyd, B. A. (1994). Computational and conceptual orientations in teaching mathematics. I D. B. Aichele, & A. F. Coxford, *Professional development for teachers of mathematics* (ss. 79-92). Reston, VA: National Council of Teachers of Mathematics.

- Turner, D. W. (2010). Qualitative Interview Design: A Practical Guide for Novice Investigators. *The Qualitative Report*, 15(3), ss. 752-760.
- Utdanningsdirektoratet. (2015). *Utdanningsdirektoratet*. Hentet fra <https://www.udir.no/laring-og-trivsel/vurdering/om-vurdering/underveisvurdering/>
- Utdanningsdirektoratet. (2016). *Utdanningsdirektoratet*. Hentet fra Matematikk: kjenneteikn på måloppnåing: <https://www.udir.no/laring-og-trivsel/vurdering/sluttvurdering/matematikk-rettleiande-nasjonale-kjenneteikn-pa-maloppnaing-for-standpunktvurdering-etter-10.-trinn/>
- Utdanningsforbundet. (2017). *Utdanningsforbundet.no*. Hentet fra <https://www.utoanningsforbundet.no/Hovedmeny/Vi-mener/Karakterer-pa-barnetrinnet/>
- Vollstedt, M. (2015). To See the Wood for the Trees: The Development of Theory from Empirical Interview Data Using Grounded Theory. I A. Bikner-Ahsbaks, C. Knipping, & N. C. Presmeg, *Approaches to Qualitative Research in Mathematics Education* (ss. 23-48). Springer.
- Wiliam, D. (2007). Keeping learning on track: classroom assessment and the regulation of learning. I F. K. Lester, *Second Handbook of Research on Mathematics Teaching and Learning* (ss. 1053-1098). Information Age Publishing.

VEDLEGG

Vedlegg 1: Intervjuguide

Intervjuguide

Innledning

- Presentasjon av meg selv og studien (varighet osv.)
- Opptak, transkribering
- Anonymitet
- Informere om informantens rett til å når som helst kunne trekke seg fra intervjuet og/eller studien uten grunn

Innledende spørsmål om informanten selv

1. Hvilken utdanning har du?
2. Hvor lenge har du jobbet som lærer? Hvilke fag og hvilke årstrinn?
3. Hvilke holdninger har du til faget?

Vurdering

4. Hva er vurdering? Hvorfor vurderer vi?
5. Hva tenker du om det å vurdere i matematikkfaget?
6. Hva tenker du om formativ vurdering- altså vurdering for læring?
7. Ifølge tidligere forskning er det tre prosesser som må jobbes med for å få til en formativ vurdering:
 - i. hvor er eleven?
 - ii. hvor skal den?
 - iii. hva skal til for at eleven kommer seg dit?

Hva tenker du om dette?

8. Karakterbruk?

9. Kommentar som eneste feedback?

10. Tilbakemeldingsfokus?

11. Karakterer vs kommentarer

12. «Å *snu* dagens antatt summative vurderingspraksis til å bli gjennomført formativt er sannsynligvis den største utfordringen til dagens ungdomstrinns lærere» (UDIR, 2016).
Hva tenker du om dette?

13. Eventuelle spørsmål fra informant

Vedlegg 2: Godkjenning fra NSD



Per Øystein Haavold
Institutt for lærerutdanning og pedagogikk UiT Norges arktiske universitet
9006 TROMSØ

Vår dato: 16.01.2017

Vår ref: 51793 / 3 / KH

Deres dato: ✓

Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 24.12.2016. Meldingen gjelder prosjektet:

51793	<i>Hva er lærernes oppfatning av formativ vurdering</i>
Behandlingsansvarlig	<i>UiT Norges arktiske universitet, ved institusjonens øverste leder</i>
Daglig ansvarlig	<i>Per Øystein Haavold</i>
Student	<i>Silje Kristiansen</i>

Personvernombudet har vurdert prosjektet og finner at behandlingen av personopplysninger er meldepliktig i henhold til personopplysningsloven § 31. Behandlingen tilfredsstiller kravene i personopplysningsloven.

Personvernombudets vurdering forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, <http://www.nsd.uib.no/personvern/meldeplikt/skjema.html>. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, <http://pvo.nsd.no/prosjekt>.

Personvernombudet vil ved prosjektets avslutning, 31.05.2017, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen
Katrine Utaaker Segadal

Kjersti Haugstvedt

Kontaktperson: Kjersti Haugstvedt tlf: 55 58 29 53

Vedlegg: Prosjektvurdering

Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

Vedlegg 3: Informasjonsskriv og samtykkeskjema

Forespørsel om deltakelse i forskningsprosjektet

"Hva er lærernes oppfatninger av formativ vurdering"

Bakgrunn og formål

Jeg heter Silje Kristiansen og skal nå til våren avslutte mitt 5-årige masterstudie *integrert master i lærerutdanning 5-10kl* ved UiT Norges Arktiske Universitet, med en mastergradsoppgave innenfor matematikdidaktikk. Denne avhandlingen omhandler formativ vurdering innenfor faget, noe jeg per nå har lite kunnskap om i praktisk sammenheng. Jeg ønsker med dette å finne ut hva lærerne tenker rundt temaet og hvordan det praktiseres, for deretter å knytte det opp mot relevant teori.

Jeg begynte min skolegang i Nord-Troms og har i senere tid jobbet her. Å forske i distriktet konkretiseres i oppgaven min, og vil derav ha et spesifikt fokus. Eneste kravet jeg har er at informanten(e) underviser i ungdomsskolen, da karaktersetting er vesentlig innenfor min oppgave.

Hva innebærer deltakelse i studien?

Deltakelse i studien vil innebære et semi-strukturert intervju på ca. 30 min hvor jeg stiller spørsmål som vil omhandle vurdering og prosessen rundt. Data registreres gjennom lydopptak.

Som deltaker i denne studien, vil det ikke bli foretatt noen form for innsamling av personopplysninger fra informanten annet enn at oppgavebesvarelsen vil bli merket med en form for kodenøkkel slik at det vil være mulig å finne frem til den spesifikke oppgaven om informanten velger å trekke seg. Samtidig vil jeg presisere at oppgaven vil ta for seg Nord-Troms og distriktet som datainnsamlingssted, noe som blir presisert i forskningen min. Informantene vil likevel forbli anonymisert.

Hva skjer med informasjonen om deg?

Alle personopplysninger vil bli behandlet konfidensielt. All informasjon som blir innhentet i dette intervjuet vil kun bli behandlet av undertegnende (Silje) og veilederen min (Per Øystein Haavold). Kodenøkkel som gjør det mulig å koble lydopptaket opp mot intervjuobjektet vil til enhver tid være overvåket av meg, adskilt fra andre data, og vil være innelåst i et skap når jeg ikke arbeider med opptakene. Prosjektet skal etter planen avsluttes og leveres 15. mai 2017, og da vil kodenøkkel og alle intervju bli slettet og makulert.

Frivillig deltakelse

Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert.

Dersom du ønsker å delta eller har spørsmål til studien, ta kontakt med Silje Kristiansen tlf:....

I dette studentprosjektet er Per Øystein Haavold min veileder med tlf:...

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

Samtykke til deltakelse i studien

Jeg har mottatt informasjon om studien, og er villig til å delta

(Signert av prosjektdeltaker, dato)