Sammendrag

Denne masteroppgaven undersøker hvilke aspekter ved matematisk kompetanse som er gjenstand for vurdering på skriftlig eksamen i matematikk for 10. trinn (MAT0010). Analysen av eksamensoppgavene baserer seg på nivårammeverket for matematisk kompetanse utviklet av *the PISA Mathematics Expert Group* (MEG). Sammen med meg, til å analysere eksamensoppgavene, hadde jeg en gruppe bestående av tre erfarne ungdomsskolelærere. Hensikten med studien er å få innsikt i hvilke sider og på hvilket nivå matematisk kompetanse testes på eksamen, for på grunnlag av dette å kunne si noe om intensjonene i Kunnskapsløftet (LK06) gjenspeiles i eksamensoppgavene i matematikk. For å kunne si noe om hensiktsmessigheten av å anvende MEG-skjemaet på norske eksamensoppgaver, innbefatter denne studien også en evaluering av dette oppgaveanalyseskjemaet, der samsvarsindikatoren *intraclass correlation coefficient* (ICC) er benyttet for å undersøke kodesamsvaret.

Masteroppgaven presenterer resultatene av analysen av samtlige eksamensoppgaver i matematikk for 10. trinn i perioden 2017 – 2019. Til sammen er 165 oppgaver kategorisert fra nivå 0 til 3 for de syv kompetansene *receptive communication*, *constructive communication*, *devising strategies*, *mathematising*, *representation, using symbols, operations and formal language* og *reasoning and argument* av fire kodere. I tillegg til eksamensoppgavene er tilhørende veilednings- og vurderingsmateriell for eksamen og gjeldende læreplan i matematikk analysert, tolket og vurdert.

Ut fra kompetanse- og nivåkategoriseringen utviklet av MEG-studien, viser mine studier at elevene på skriftlig eksamen i matematikk for 10. trinn i perioden 2017 – 2019 hovedsakelig blir testet på de to laveste nivåene for alle kompetansene som inngår i MEG-skjemaet. Kompetansen *representation* skiller seg ut ved hovedsakelig å bli testet på laveste nivå, en nivåbeskrivelse som kjennetegnes av fravær eller svært lave krav til anvendelse, mens kompetansen *using symbols, operations and formal language* er den kompetansen som har lavest andel på nivå 0 og samlet sett blir testet på høyest nivå. Kategorisering av kompetanser til øverste nivå forekommer bare i 12 av totalt 4620 kodinger.

Analysen av resultatene i denne studien indikerer at elevene på eksamen i matematikk ikke får anledning til å vise et så bredt spekter av matematiske kompetanse som formålsbeskrivelsen for faget og rammeverket for eksamen i matematikk legger opp til. Fakta og prosedyrekunnskap er fortsatt det som testes på høyest nivå på eksamen, mens problemløsing og modellering, som vektlegges i formålsbeskrivelsen til faget, alt overveiende testes på nivå 0 og 1.

Analysen av kodesamsvaret viser at gruppa som helhet lå på *almost perfect agreement* for alle de syv kompetansene som ble kodet, og at MEG-skjemaet kan være et hensiktsmessig verktøy for å kartlegge i hvilken grad eksamensoppgavene i matematikk på 10. trinn tester ulike aspekter ved elevenes matematiske kompetanse.

Abstract

This master thesis examines the aspects of mathematical competence that are subject to assessment on the mathematics written exam for Grade 10 (MAT0010). The analysis of the exam exercises is based on the mathematical competence level framework developed by *the PISA Mathematics Expert Group* (MEG). With me, to analyse the exam tasks, I had a group consisting of three experienced lower secondary school teachers. The purpose of the study is to gain insight into what aspects and at what level mathematical competence is tested on the examination, for on the basis of this to be able to say something about the intentions of the Knowledge Promotion Reform (LK06) are reflected in the examination tasks in mathematics. In order to say anything about the appropriateness of applying the MEG scheme to Norwegian exam assignments, this study also includes an evaluation of this task analysis scheme, in which *the intraclass correlation coefficient* (ICC) compliance indicator has been used to investigate the extent of code match.

The master thesis presents the results of the analysis of all mathematics exam assignments for Grade 10 in the period 2017 – 2019. In total, 165 tasks are categorized at level 0 – 3 for the seven competences *receptive communication*, *constructive communication*, *devising strategies*, *mathematising*, *representation, using symbols, operations and formal language* and *reasoning and argument* of four encoders. In addition to the exam tasks, the accompanying examination and assessment materials for the exam and the current curriculum of mathematics have been analysed, interpreted and evaluated.

Based on the competence and level categorization developed by the MEG study, my studies show that students on the written exam in mathematics for Grade 10 in the period 2017 – 2019 are mainly tested at the two lowest levels for all the competencies included in the MEG scheme. The competence *representation* stands out primarily to be tested at the lowest level, a level description characterized by the absence or very low requirements of application, while the competence *using symbols, operations and formal language* is the competence that has the lowest proportion of level 0 and overall is tested at the highest level. The categorisation of competences to the top level occurs only in 12 of a total of 4620 encodings.

The analysis of the results in this study indicates that students in mathematics will not have the opportunity to demonstrate such a wide range of mathematical competence as the purpose description for the subject in LK06 and the framework for the mathematics exam intend. Fact and procedural knowledge are still what are tested at the highest level of examination, while problem solving and modelling, which are emphasized in the purpose description of the LK06, are all predominantly tested at level 0 and 1.

The analysis of code matching shows that the group as a whole was on the *almost perfect agreement* for all the seven competences that were coded, and that the MEG scheme may be an appropriate tool for mapping out the degree to which the exam exercises in mathematics for Grade 10 test various aspects of the students' mathematical competence.

Forord

Denne masteroppgaven startet med en telefonsamtale i mars 2016 fra daværende leder for Kompetansesenteret i Drammen kommune, Hilde Schjerven. Hun ville ha meg med på en pilotstudie for lærerspesialister i matematikk i regi av NTNU, og jeg måtte svare raskt! Dette studiet var så inspirerende at jeg ikke klarte å stoppe etter endte to år på universitetsbenken. Dokumentet du nå leser er et håndfast bevis på resultatet av denne hektiske telefonsamtalen.

Svein Arne Sikko, som jeg ble kjent med på lærerspesialiststudiet, har vært min veileder på masteroppgaven. Han har gitt tydelige og konstruktive kommentarer underveis i prosessen, i tillegg til å være ekstremt kjapp med tilbakemeldingene. Jeg retter deg en stor takk!

Takksigelser går også til min arbeidsgiver Drammen kommune og Kompetansesenteret ved Ingvild Waage, som har gitt med tid og rom mellom andre arbeidsoppgaver for å fullføre denne masteroppgaven. Takk også til min eminente kodegruppe, som i en ellers travel skolehverdag satte av tid og rom for å lære seg MEG-verktøyet og kode 165 oppgaver fordelt på tre eksamenssett. Dere har vært solide diskusjonspartnere og en viktig brikke i mitt arbeid.

Takk, Anne-Valérie Sickinghe, for hjelp til oversettelse av sammendraget til engelsk.

Sist, men ikke minst, vil jeg få takke min kone Nina for korrekturlesing og et utall antall ferdige middager og oppmuntrende ord undervegs: Livet er ikke det samme uten…

Takk til alle som heiet på meg!

Drammen, 23. november 2019

Svend Kristian Eidsten

Innhold

1 Innledning

1.1 Bakgrunn for valg av tema

Eksamen måler elevenes kompetanse ved avslutningen av skoleåret, basert på innhold- og kompetansebeskrivelser i fagplanen. Sentralt gitt eksamen i matematikk på 10. trinn er en prøve med viktig utfall. Den vektes på linje med standpunktkarakteren og kan få store konsekvenser både med tanke på skolens omdømme og for den enkelte elevs videre skolegang. For kommunale og sentrale myndigheter kan eksamen sees på som en viktig sluttkontroll, der oppgavene gitt på eksamen måler forventet kompetanse i forhold til målene i den intenderte læreplanen.

Eksamen i matematikk skal måle et bredt spekter av matematiske ferdigheter og prosesser, og inneholder derfor oppgaver som krever ulik kompetanse (Utdanningsdirektoratet, 2019a). På grunnskolenivå skal imidlertid skriftlig eksamen avvikles på fem timer. Dette setter begrensninger for hvor mange læreplanmål man praktisk sett rekker å teste, med det til følge at eksamen tester et smalere spekter av elevenes matematiske kompetanse enn standpunktvurderingen. Det hviler derfor et tungt ansvar på fagnemndene «for at oppgavene er i tråd med læreplanverket for Kunnskapsløftet og relevante bestemmelser i forskrift til Opplæringsloven.» (Utdanningsdirektoratet, 2017g, s. 3) og at eksamen i matematikk for 10. trinn samlet sett prøver «kandidatene bredt i kompetansemål fra alle hovedområder i læreplanen» (Utdanningsdirektoratet, 2019a, s. 6)

På bakgrunn av stor variasjon og svake resultater på sentralt gitt eksamen i enkelte matematikkfag, bl.a. eksamen i matematikk for 10. trinn (MAT0010), ba Utdanningsdirektoratet Matematikksenteret om en ekstern gjennomgang og vurdering av eksamensoppgaver og eksamensresultater. Oppdraget gikk bl.a. ut på å vurdere om eksamensoppgavene i matematikk var i samsvar med læreplanen i faget og om vanskegraden var tilpasset kompetansemålene og elevgruppen. I tillegg ble prosjektgruppa bedt om å gi råd om hvordan utarbeidelsen av eksamen kunne gjøres for på best mulig måte å måle elevenes kompetanse i matematikk.

Bestillingen fra Utdanningsdirektoratet munnet ut i rapporten «Vurdering av eksamen i matematikk» (Matematikksenteret, 2015), og peker på at det i perioden 2009 – 2014 var store svingninger i anvendelsen av kompetansemål fra år til år i sentralt gitt skriftlig eksamen for 10. trinn. Spesielt gjaldt dette hovedområdet tall og algebra, men de fant også store svingninger innenfor de andre hovedområdene. I tillegg påpeker prosjektgruppa bak rapporten nokså stor variasjon i vanskegraden til de ulike eksamenssettene for 10. trinn. For perioden rapporten dekker, er med andre ord oppgave-kompleksiteten og krav til kognitive ferdigheter for å løse eksamensoppgavene ikke stabil: Noen kompetansemål testes på høyt nivå, mens andre blir utelatt eller kun testet på et lavt nivå.

For å forbedre eksamensoppgavenes kvalitet, anbefaler rapporten fra Matematikksenteret å pilotere og systematisk analysere eksamensoppgaver og eksamenssett i matematikk, slik det f.eks. gjøres for nasjonale prøver i regning. En slik pilotering vil kunne gi en verdifull pekepinn på oppgavenes vanskelighetsgrad og «finne korrelasjonen mellom dyktigheten til en elev og sannsynligheten for å løse oppgaven.»

(Matematikksenteret, 2015, s. 38), der dyktighet er definert ut fra det som måles i en Item Response Theory analyse (IRT-analyse), slik det brukes i de aller fleste storskala prøvesystemer, nasjonale prøver og internasjonale komparative undersøkelser i dag.

Et resultat av Matematikksenteret sin «Vurdering av eksamen i matematikk» (Matematikksenteret, 2015) var at Utdanningsdirektoratet finansierte en evaluering av matematikkeksamen på 10. trinn for perioden 2017 – 2019. Evalueringen utføres av Fafo, i samarbeid med OsloMet, med overordnet perspektiv å evaluere om eksamen både er og oppleves å være rettferdig. Dette er belyst ved å se på hvordan eksamen er utformet (oppbygging, språk og design/layout), sammenhengen mellom læreplanen og opplæringen, vanskegrad, arbeidsmengde og vurdering. I tillegg har hver rapport sitt fokusområde. I 2017 var det utformingen av eksamen, for 2018 bruken av digitale hjelpemidler. Evalueringene har innhentet data i form av dokumentanalyse av bl.a. læreplan, eksamen- og sensorveiledninger, eksamensbesvarelser, vurderingsskjemaer, eksamen og lærebøker, spørreskjemaer til sensorer og lærere og intervju med et utvalg av lærere som fikk trukket ut elever til eksamen og elever som hadde avlagt eksamen i matematikk. Eksamensoppgavene ble, med utgangspunkt i rammeverket utarbeidet av Johan Lithner i artikkelen «A research framework for creative and imitative reasoning» (Lithner, 2008), klassifisert som enten å kreve algoritmisk løsning eller kreativ løsning. Klassifiseringen la til grunn Niss og Jensen (2002) og Mullis og Martin (2013) sine beskrivelser av resonneringskompetanse (Andersen, Fossum, Rogstad & Smestad, 2017; Bjørnset, Fossum, Rogstad, Smestad & Talberg, 2018).

I begge Fafo-rapportene oppgir nærmere 90 % av lærerne i spørreundersøkelsen at eksamen står i forhold til læreplanen/kompetansemålene i matematikk. Rapporten fra 2017 peker imidlertid på, ut fra poenggivingen av elevbesvarelsene, at det er ulikheter i vanskegraden mellom hvert hovedområde: Tall og algebra er det området som gir mest poengutdeling, mens geometri gir minst. I tillegg har hovedområdet måling minst spredning i vanskelighetsgrad, mens tall og algebra har størst spredning. Her pekes det også på at det er en sterk sammenheng mellom oppgavenes vanskelighetsgrad og kategoriseringen av oppgavene som henholdsvis algoritmiske eller kreative, der algoritmisk løsning henspiller på løsningsmetoder som elevene er kjent med fra eksempler og oppgaver fra læreboka, forlagsprøver eller tidligere eksamener og kreativ løsning innebærer at lignende oppgaver og eksempler ikke er å finne i lærebøkene, forlagsprøver eller tidligere eksamensoppgaver, slik at rutinepreget algoritmisk løsning ikke er mulig: Kun tre av de ti oppgavene som empirisk blir kategorisert som vanskeligst er kan betraktes som algoritmiske, mens alle de ti enkleste oppgavene faller inn under kategorien algoritmisk løsning (Andersen et al., 2017).

En alternativ tilnærming til Fafo-rapportenes analyseverktøy for kategorisering av hvorvidt oppgaver i matematikk vil oppleves vanskelig å løse for elevene, er rammeverket utviklet av *the PISA Mathematics Expert Group* (MEG). Ekspertgruppen har, siden 2003, arbeidet med å utvikle et rammeverk for kategorisering av oppgaver, slik at man på forhånd skal kunne si noe om vanskegraden til matematikkoppgaver som blir gitt til 15-åringer på PISA-testene. Gruppen satte i utgangspunktet opp to hovedkategorier, henholdsvis oppgavenes overflate-karakteristika (*surface features of items*) og oppgavenes kognitive krav (*cognitive demand characteristics of items*). De videre MEG-studiene viste imidlertid at overflate-karakteristika som matematisk emne, oppgaveutforming og kontekst var lite avgjørende for hvor vanskelig oppgavene var for elevene, og de satte derfor heller inn støtet for å utvikle et analyseverktøy for oppgavenes kognitive krav, ut fra seks definerte matematiske kompetanser. MEG-skjemaet har i dag, pr. siste revisjon i 2013, en detaljert beskrivelse av de seks kompetansene *communication*, *devising strategies*, *mathematising*, *representation, using symbols, operations and formal language* og *reasoning and argument*, med tilhørende nivåskalering for hver kompetanse fra 0 – 3. Utarbeidelsen av MEG-skjemaet bygger på en tilpasset versjon av de ulike matematiske kompetansene (*kompetanseblomsten*) slik det er beskrevet i arbeidet til Niss og Jensen (2002), og er nærmere gjort rede for i artikkelen «Using Competencies to Explain Mathematical Item Demand: A work in Progress» (Turner, Blum & Niss, 2015), en tråd jeg nøster videre opp i det påfølgende teorikapittelet.

I artikkelen *Identifying competency demands in mathematical tasks: Recognising what matters* redegjør Pettersen og Nortvedt (2018) for hvordan MEG-skjemaet kan anvendes for å kartlegge hvilke kompetanser som prøves for elevene som avla PISA-testen i matematikk i 2012 og eksamen i matematikk på 10. trinn i 2014. De setter sammen et kode-panel på fem personer, bestående av lærere og studenter på lærerutdanningen, og konkluderer med at denne gruppen i det store og hele klarer å bruke MEG-analyseverktøyet til å identifisere hvilke matematiske kompetanser som krever aktivisering for å løse de ulike oppgavene.

Kategorisering av eksamensoppgavene ut fra MEG-skjemaet har med andre ord potensiale til å si noe om hvilke matematiske kompetanser elevene testes i, og på hvilket nivå. Dette kan være interessant med tanke på om eksamen faktisk tester en så bred matematisk kompetanse av kunnskaper, ferdigheter og prosesser som rammeverket for eksamen legger opp til, og om man på forhånd, under utarbeidelse av eksamensoppgavene, kan bruke MEG-skjemaet som en indikator på om eksamenssettet tester ulike kompetanser på ulikt nivå. Dette vil i så tilfelle være i tråd med Niss og Jensen (2002, s. 126) sine anbefalinger, der de påpeker at «for at få et dækkende og righoldigt billede af en persons matematiske kompetencer, må [man]undersøge personens virksomhed inden for en bredere palet af matematiske aktiviteter».

1.2 Formålet med studien

Denne masteroppgaven har til hensikt å si noe om i hvilken grad eksamensoppgavene i matematikk på 10. trinn tester ulike aspekter ved elevenes matematiske kompetanse. Metoden som er brukt for å kaste lys over dette er hentet fra et analyseverktøy utviklet av *the PISA Mathematics Expert Group* (Turner et al., 2015), og er for norske forhold kun brukt en gang tidligere av Pettersen (2019) til å analysere PISA-testen fra 2012 og eksamen i matematikk på 10. trinn fra 2014. I tillegg har Kristianslund (2015) brukt tilsvarende analyseverktøy på eksamensoppgaver i matematikk for grunnskolelærerutdanningen på 5. – 10. trinn.

En av utfordringene med oppgaveanalyseskjemaet utviklet av MEG er at det, på tross av egne definisjoner og beskrivelser for hver kompetansekategori, gir rom for tolkning. I tråd med anbefalingene til utviklerne av MEG-skjemaet satte jeg derfor sammen en kodegruppe bestående av meg selv og tre andre lærere med lang fartstid (mellom 17 og 28 år) fra ungdomsskolen, for på den måten å få et diskusjonspanel og en «member check» av egen koding.

Målet med studien er todelt. På den ene siden ønsker jeg, på bakgrunn av anvendt teori og analysemetode, å få innsikt i hvilke sider og på hvilket nivå matematisk kompetanse testes på skriftlige eksamen, for på grunnlag av dette å kunne si noe om intensjonene i Kunnskapsløftet (LK06) gjenspeiles i eksamensoppgavene i matematikk. På den andre siden ønsker jeg også å undersøke hvor hensiktsmessig det er å anvende MEG-skjemaet på norske eksamensoppgaver. Denne studien innbefatter derfor også en evaluering av dette oppgaveanalyseskjemaet, der samsvarsindikatoren *intraclass correlation coefficient* (ICC) er benyttet for å undersøke kodesamsvaret.

Nå kan det for så vidt innvendes at spørsmålet om samsvaret mellom eksamens innhold og vanskegrad, undervisning og læreplan er belyst i rapportene fra Matematikksenteret (2015) og evalueringene i Fafo-rapportene (Andersen et al., 2017; Bjørnset et al., 2018). For å skape et bredere bilde av hva som testes på eksamen, kan det imidlertid være hensiktsmessig å bruke andre analyseverktøy på samme datasett, for på den måten å få fram andre nyanser i materialet enn det tidligere analyser og evalueringer har pekt på. Min tanke er at det å bruke rammeverktøyet og oppgaveanalyseskjemaet til MEG for å analysere eksamensoppgaver i matematikk på grunnskolenivå kan bidra til dette. Resultatene herfra kan være av interesse både for fagnemnden som lager eksamensoppgaver i matematikk, og som et mulig verktøy for lærere når de skal velge ut oppgaver som nører oppunder utviklingen av elevenes matematiske kompetanse.

Ifølge *Rammeverk for eksamen* skal eksamen blant annet «dekke sentrale deler av læreplanen og gjenspeile læreplanen og formålet med faget» og «gi kandidater på alle nivå mulighet til å vise sin kompetanse» (Utdanningsdirektoratet, 2017g, s. 5). Flere studier (Niss og Jensen, 2002; Wilson, 2007; Jensen, 2007; Boesen, 2006; Alseth, Breiteig & Brekke, 2003; Michelsen, 2001; Schoenfeld, 2007) har påvist sammenhenger mellom vurdering av elever og undervisningspraksis, og pekt på at tester med viktig utfall har en tilbakevirkende effekt på undervisningen. I tillegg påpeker Mogens Niss (1993) at selv om læreplanene i matematikk i mange land har gjennomgått store endringer med tanke på å beskrive matematisk kompetanse og arbeidsmåter i faget, er vurderingsformene ikke endret tilsvarende, med det til følge at vurderingen i liten grad evaluerer den kunnskap og kompetanse man ønsker skal prege undervisningen. Mine studier vil kunne bidra til å gi en dypere innsikt i hvilke sider av matematisk kompetanse som blir testet på eksamen, i tillegg til å si noe om hvilket nivå kompetansen blir testet på. Resultatene herfra vil kunne antyde om det er samsvar mellom de målene samfunnet har satt for grunnskoleopplæringen i matematikk, og hva som testes på skriftlig eksamen i dette faget.

I disse dager fases Kunnskapsløftet (LK06) ut, og Fagfornyelsen (LK20) toger inn. Nye kjerneelementer er fastsatt av Kunnskapsdepartementet for alle fag som omfattes av fagfornyelsen, og skal ligge til grunn for utforming av læreplaner for fag (Kunnskapsdepartementet, 2018b). Kjerneelementene innbefatter både det viktigste innholdet og det elevene må lære for å kunne mestre og å bruke faget, som f.eks. kunnskapsområder, metoder, begreper, tenkemåter og uttrykksformer (Utdanningsdirektoratet, 2017f). For faget matematikk er kjerneelementene *utforsking og problemløsing*, *modellering og anvendelser*, *resonnering og argumentasjon*, *representasjon og kommunikasjon*, *abstraksjon og generalisering* og *matematiske kunnskapsområder*, der «De fem første kjerneelementene beskriver arbeidsmåter, metoder og tenkemåter i matematikk», mens «Det sjette kjerneelementet beskriver de sentrale kunnskapsområdene i matematikk», som elevene skal møte gjennom de fem første kjerneelementene (Kunnskapsdepartementet, 2018b, s. 15).

Ut fra innholdsbeskrivelsen i de fem første kjerneelementene ser vi fortsatt fellestrekk med kompetansebeskrivelsene til Niss og Jensen (2002) og PISA-rammeverket i matematikk (OECD, 2017) som blir presentert i påfølgende teorikapittel. Selv om min analyse tar for seg eksamen fra 2017 – 2019 under Kunnskapsløftet, tenker jeg derfor at analyseverktøyet jeg bruker også kan ha relevans for påfølgende eksamener under LK20.

I september 2018 nedsatte Kunnskapsdepartementet en eksamensgruppe som skulle «utrede et helhetlig eksamensordningssystem for fagene som omfattes av fagfornyelsen» (Utdanningsdirektoratet, 2019f, s. 2). Målet med dette er å sikre en god sammenheng mellom læreplanene og vurderingsformene i faget, der eksamensgruppa har tatt utgangspunkt i kompetansebegrepet i LK20, med vekt på hvordan eksamen kan understøtte opplæringen og være en valid og reliabel prøving av kompetanse ut fra beskrivelsene i den nye læreplanen (Utdanningsdirektoratet, 2019f). I dette arbeidet kan det være behov for ulike analyseverktøy, der jeg tenker rammeverket til MEG-studien med tilhørende oppgaveanalyseskjema kan være et mulig verktøy for å sikre dette.

Både gjeldende læreplanen i matematikk (Kunnskapsløftet) og rammeverket for PISA-testene er, ifølge henholdsvis Valenta, Norset, Åsenhus og Wæge (2015) og Niss (2015), tuftet på kompetansebeskrivelsene fra den danske rapporten «Kompetenser og Matematikklæring» (KOM-prosjektet) til Niss & Jensen (2002). Begge prøvene tester noenlunde samme aldersgruppe (15-åringer), og ifølge Nortvedt og Pettersen (2016) måler PISA-testene matematisk kompetanse som er relevant i forhold til Kunnskapsløftet.

Til å analysere eksamensoppgavene i matematikk på 10. trinn har jeg derfor valgt et analyseverktøy utviklet av Program for International Student Assessment (PISA) (Turner et al., 2015). Dette analyseverktøyet ble i utgangspunktet utarbeidet for å kunne forutsi hvor vanskelig ulike oppgaver ville være for elevene som gjennomfører PISA-testene i matematikk. Analyseskjemaet er i utgangspunktet utviklet av og for eksperter, men Pettersen og Nortvedt (2018) viste at også fem personer, med ulik grad av undervisningserfaring og matematikkbakgrunn, som gruppe leverte en nokså konsistent analyse av hvilke kompetanser de ulike oppgavene testet. Forfatterne påpeker imidlertid at operasjonaliseringene av de ulike kompetansene gir rom for tolkning. Det er derfor hensiktsmessig å la flere personer individuelt kode samme oppgave, slik at man kommer fram til en gjennomsnittsverdi for hva løsningen av oppgaven krever for hver av de seks kompetansene i MEG-skjemaet. Pettersen og Nortvedt (2018) konkluderer med at dette analyseverktøyet kan være nyttig å bruke til å analysere og velge ut hensiktsmessige oppgaver for å utvikle elevenes matematiske kompetanse. Jeg tenker at så også kan være tilfelle med tanke på utforming og utvelgelse av oppgaver på skriftlig eksamen i matematikk, slik at man kan sikre at eksamen tester et bredest mulig spekter av matematisk kompetanse på ulike nivå.

Både «Vurderingen av eksamen i matematikk» (Matematikksenteret, 2015) og Fafo-rapportene som evaluerte skriftlig eksamen i matematikk på 10. trinn for årene 2017 og 2018 anbefaler pilotering av eksamensoppgaver. Dette har, så vidt jeg har klart å bringe på det rene, aldri blitt gjennomført. Pilotering av eksamensoppgaver kan være en utfordrende og arbeidskrevende prosess, men kanskje finnes det andre alternative ruter, som f.eks. å benytte samme analyseverktøyet som i denne masteroppgaven til å kartlegge og forutsi hvor kognitivt krevende oppgavene vil virke på elevene? Dette bør i så tilfelle testes empirisk mot faktiske elevresultater, noe som ligger utenfor rammen av min masteroppgave.

Ut fra formålet med oppgaven, den felles referanserammen for matematisk kompetanse mellom Kunnskapsløftet og PISA-rammeverket og potensialet i analyseverktøyet til MEG-studien har jeg formulert følgende problemstilling:

I hvilken grad tester oppgavene på skriftlig eksamen i matematikk for 10. trinn ulike aspekter ved elevenes matematiske kompetanse?

1.3 Oppgavens struktur

Jeg ønsker i denne studien å si noe om i hvilken grad eksamensoppgavene i matematikk på 10. trinn tester ulike sider av elevenes matematiske kompetanser. For å undersøke dette har jeg brukt et oppgaveanalyseskjema (MEG-skjema) utviklet av *Program for International Student Assessment (PISA)*, der jeg og tre ungdomsskolelærere individuelt kodet alle eksamensoppgavene for perioden 2017 – 2019. Kodesamsvaret i gruppa ble kartlagt ved å bruke samsvarsindikatoren *intraclass correlation coefficient* (ICC).

I teoridelen, kapittel 2, ser jeg på ulike tilnærminger til begrepet matematisk kompetanse, og tar for meg ulike rammeverk for matematisk kompetanse og kategorisering av oppgaver i matematikk. Kapitlet omhandler også læreplanteori, formålet med matematikkfaget i skolen og en sammenligning av rammeverket for vurdering av eksamen og det oppgaveanalyseskjemaet jeg brukte. Avslutningsvis tar jeg opp forskning på sammenhengen mellom vurdering og undervisning.

Kapittel 3 handler om metode, og tar for seg forskningsdesign, datamaterialet og analysemetode. Det gis en prosedyrebeskrivelse for kodingen av oppgavene med flere kodeeksempler. Til sist diskuteres aspekter ved forskningskvaliteten på mitt arbeid, der jeg blant annet redegjør for valg av samsvarsindikator og verdier for styrkesamsvar for bestemmelse av kodesamsvaret i min kodegruppe.

I kapittel 4 presenteres koderesultatene for alle eksamensoppgavene fordelt på kompetanser og nivå. Del 1 og del 2 for det enkelte året presenteres hver for seg, for så å vise samlet resultat pr. år. Til sist presenteres en oversikt for hele 3-års perioden fra 2017 – 2019.

I kapittel 5 diskuteres resultatene opp mot MEG-rammeverket og teorien som er presentert i kapittel 2. Jeg diskuterer også kodesamsvar og evaluerer bruken av MEG-skjemaet på eksamensoppgaver på 10. trinn. Avslutningsvis oppsummerer jeg hovedfunnene i denne studien.

2 Teori

I dette kapitlet ser jeg på ulike tilnærminger til begrepet matematisk kompetanse og tar for meg ulike rammeverk for matematisk kompetanse og kategorisering av oppgaver i matematikk. Her blir særlig KOM – prosjektet og PISA – rammeverket grundig presentert, i det disse rammeverkene danner utgangspunktet for mitt oppgaveanalyseverktøy. Deretter ser jeg på ulike perspektiver på læreplanen, tar for meg kompetansebegrepet i Kunnskapsløftet, LK20 og PISA, og sammenligner formålet med matematikkfaget i Kunnskapsløftet og rammeverket for PISA-testene. Etter dette følger en sammenligning av rammeverket for vurdering av eksamen og MEG-skjemaet, før jeg til sist tar fram forskning som har sett på sammenhenger mellom vurdering av elever og undervisningspraksis.

2.1 Matematisk kompetanse – et sammensatt begrep

I *Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook 1: Cognitive Domain* (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956) setter sensor-teamet bak Blooms Taksonomi opp følgende likning for vurdering av eksamenskandidatens helhetlige kompetanse: «Arts or skills + knowledge = abilities» (Bloom et al., 1956, s. 38), som på norsk vel helst kan oversettes med «ferdigheter + kunnskap = kompetanse». Likningen kan forstås som at det å vise kompetanse avhenger av eleven/studenten sine evner til å kombinere ferdigheter og kunnskaper, og at enten bare kunnskaper eller bare ferdigheter ikke dekker kompetansebegrepet.

På papiret ser denne likningen tilforlatelig ut, men ifølge redegjørelsen til Jeremy Kilpatrick i oppslagsverket *Encyclopedia of Mathematics Education* (Kilpatrick, 2014, s. 85) har det vist seg utfordrende å definere et enhetlig innhold i begrepet kompetanse:

> *The concept of competence is one of the most elusive in the educational literature. Writers often use the term competence or competency and assume they and their readers know what it means. But arriving at a simple definition is a challenging matter. […] Competence seems to possess a host of near synonyms: ability, capability, cognizance, effectuality, efficacy, efficiency, knowledge, mastery, proficiency, skill, and talent – the list goes on.*

Niss, Bruder, Planas, Turner og Villa-Ochoa (2017) velger å nærme seg begrepene kompetanse, viten og kunnskap i matematikk ved å stille spørsmålet «Hva betyr det å mestre matematikk?» De peker på at hvis man betrakter dette spørsmålet med tradisjonelle kunnskapsorienterte briller, vil svaret være at mestring i matematikk innebærer å kunne begreper, definisjoner, regler, teoremer, formler, metoder og fakta i faget, der pensumlister og prosedyrekunnskap står i sentrum. Velger man derimot en mer prosessorientert tilnærming til faget, vil svaret være hvilke karakteristiske kognitive prosesser man må beherske for å mestre matematikk, og at det å mestre faget innebærer å kunne anvende og ikke bare besitte kunnskap: «"Knowing" and "being able to do" are two different things» (Niss et al., 2017, s. 236). En tidlig representant som forfektet dette synet var George Polya, der han i forordet til boken *How to Solve It* sier det slik (Polya, 1957 (1945), s. v):

> *If [a teacher of mathematics] fills his allotted time with drilling his students in routine operations he kills their interest, hampers their intellectual development, and misuses his opportunity. But if he challenges the curiosity of his students by setting them problems proportionate to their knowledge and helps them to solve their problems with stimulating questions, he may give them a taste for, and some means of, independent thinking.*

Om enn man velger å fokusere på kunnskap og innhold, prosedyrer og teknikker, hva det innebærer å gjennomføre karakteristiske matematiske prosesser (the *enactment of mathematics*) eller at matematikk er et redskapsfag for menneskelig aktivitet og utvikling, representerer disse tilnærmingene ulike, men gjensidig avhengig tankegods. Idet ingen av fokusområdene står alene i et vakuum, vil det å forfekte ett av disse fenomenologiske ståstedene for å slå de andre i hjel, være lite hensiktsmessig. Snarere bør man heller ha i bakhodet at det finnes en kontekstavhengig balansegang mellom dem (Niss et al., 2017).

Det tradisjonelle absolutistiske synet på matematikk, der tilegnelse av matematisk kunnskap innebar å huske og bruke matematiske regler, fakta og prosedyrer på korrekt vis, møtte imidlertid kritikk, som vi ser av sitatet til Polya (1945). Opp gjennom 1960- og 1970-årene sluttet flere matematikere og filosofer seg til kritikken (Ernest, 1991): Det å mestre matematikk innebar mer enn et *kunnskapsprodukt* av innøvde regler og prosedyrer, man måtte også ta med i beregningen hvilke karakteristiske matematiske *prosesser* elevene måtte beherske for å løse oppgaver av ulik art (Niss et al., 2017). En følge av den prosess-orienterte tilnærmingen til matematikk ble fornyet fokus på problemløsing i undervisningen: Mens problemløsing tradisjonelt hadde dreid seg om elevene var i stand til å løse en matematikkoppgave eller ei, førte nå det prosess-orienterte perspektivet til fokus på de kognitive aktivitetene (f.eks. strategisk tankegang og metakognisjon) som kreves i for å løse et problem av matematisk art (Niss et al., 2017). I norske læreplaner kommer dette til syne i M87, der problemløsning var et hovedemne som skulle være en del av all matematikkopplæring (Kirke- og undervisningsdepartementet, 1987).

I følge Niss et al. (2017) tok *The National Council of Teachers of Mathematics* (NCTM) i USA en ledende rolle i å sette det prosess-orienterte perspektivet og problemløsning inn i skolematematikken. Allerede i publikasjonen *An Agenda for Action: Recommendations for School Mathematics of the 1980s* (NCTM, 1980) slår de fast at undervisningen i matematikk skal fokusere på problemløsning, at grunnleggende ferdigheter i matematikk er mer enn algoritmer, og at suksess i faget skal evalueres bredere enn med konvensjonelle prøver. I 1989 gikk NCTM enda et skritt videre, og konkretiserte fire overordnede prosess-standarder for skolematematikken på alle nivåer i USA: «Mathematics as problem solving», «mathematics as communication», «mathematics as reasoning» og «mathematical connections» (NCTM, 1989). Etter en del motstand og opposisjon kom så den reviderte utgaven *Principles and Standards for School Mathematics* (NCTM, 2000). Her beskrives fem «Content Standards» *(Number and Operations, Algebra, Geometry, Measurement* og *Data Analysis and Probability)* som elevene skal lære, i tillegg til fem «Process Standards» (*Problem Solving*, *Reasoning and Proof, Communication, Connections* og *Representations*) som setter søkelyset på måter elevene skal erverve og anvende fagkunnskap. Denne måten å beskrive og organisere de karakteristiske matematiske prosessene på, skal vi se har mange likhetstrekk med de senere omtalte rammeverkene til Kilpatrick, Swafford og Findell (2001), Niss og Jensen (2002), TIMSS (Mullis & Martin, 2013) og OECD (2017).

I engelsk faglitteratur brukes begreper som *mathematical literacy* (OECD, 2017), *numeracy* (Storbritannia og Australia), *mathematical proficiency* (Kilpatrick et al., 2001) og *mathematical competence/competencies* (Niss, 2003) for å beskrive ulike sider av hva det vil si å beherske matematikk. De engelskspråklige begrepene er ikke alltid uten videre lett å oversette direkte til norsk, men for innholdet i *mathematical literacy* og *mathematical proficiency* er begrepet matematisk kompetanse benyttet (Nortvedt & Pettersen, 2016; Valenta et al., 2015).

Ifølge Niss et al. (2017) er alle disse termene beslektede, dog ikke identiske begreper, men med det til felles at alle ser på matematikk som bestående av to essensielle aspekter, nemlig matematikk som produkt og matematikk som prosess. Matematikk som produkt er for eksempel faktakunnskaper og prosedyreferdigheter innenfor et område, mens prosessaspektet er knyttet til de kognitive aktivitetene som kreves for å løse oppgaven.

Kompetansebegrepet har med andre ord mange fasetter, men beskriver på ulike måter hvor godt rustet man er til å anvende kunnskap og ferdigheter i ulike situasjoner. En helhetlig matematisk kompetanse innebærer både å kunne anvende kunnskap om faginnholdet og å kunne aktivere kognitive ferdigheter, der de kognitive ferdighetene er egenskaper man ser på som nødvendig for å mestre matematisk tankegang i møte med matematikkoppgaver, uavhengig av temaet (Niss, et al., 2017).

Rammeverk i matematikk kan på samme måte ha ulike vinklinger, der noen tar for seg både organisering av faginnhold og kognitive ferdigheter, mens andre fokuserer på de kognitive prosessene/områdene man trenger for å mestre faget (Kilpatrick, 2014). Eksempler på førstnevnte kan være rammeverket i TIMSS og PISA, mens sistnevnte kan representeres ved *kompetanseblomsten* i KOM-prosjektet og *trådmodellen* til Kilpatrick, Swafford og Findell.

2.2 Rammeverk for matematisk kompetanse

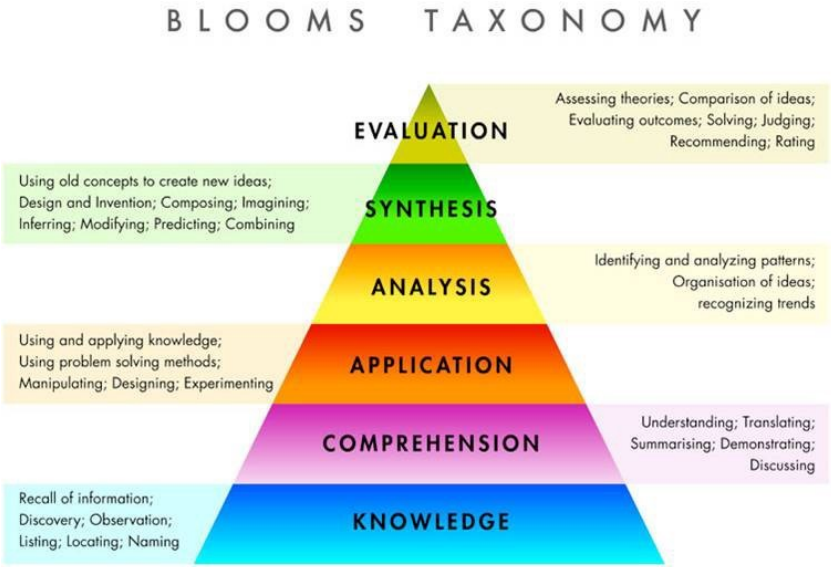
2.2.1 Blooms taksonomi

I følge Kilpatrick (2014) er Blooms *Taxonomy of Educational Objectives* (Bloom et al., 1956) selve stamfaren for utarbeidelsen av senere rammeverk for matematisk kompetanse i undervisningssammenheng, og Webb (2014, s. 63) hevder at «Bloom's Taxonomy has been applied and adapted by mathematics educators since its publication».

Rammeverket ble i utgangspunktet laget som en ressurs for å etablere et felles språk for evaluering av elever og studenters kognitive ferdigheter, på tvers av fag. Taksonomien tar derfor ikke for seg ulike fagspesifikke emner, men søker heller å identifisere de kognitive prosessene som ligger under det å besitte fagkunnskap. Boken til Bloom inkluderer flere eksempler på hvordan slike oppgaver kan utformes, dog med få eksempler fra matematikk (Webb, 2014).

Rammeverket beskriver seks hierarkisk ordnede nivåer/klasser for kognitive ferdigheter (Bloom et al., 1956), og kan kort beskrives slik:

Knowledge innebærer å kunne huske fakta og gjengi fagstoff. *Comprehension* handler om at eleven skal kunne plukke ut relevant fagstoff, beskrive, og i noen grad utdype, det man har lært med egne ord. Nivået *Application* innebærer at eleven skal kunne bruke og anvende kunnskap, f.eks. det å ta i bruk lærte prinsipper og regler i løsning av ulike typer problemer. I *Analysis* ligger det å kunne identifisere analysere og forstå sammenhengen, mens *Synthesis* referer til å kunne sette sammen deler til en helhet for på den måten å kunne trekke egne slutninger ut fra tilgjengelig informasjon. *Evaluation* anses som det mest avanserte nivået i modellen, og handler om å kunne gjennomføre ulike former for vurderinger. En generell illustrasjon av taksonomien for kognitive ferdigheter er gjengitt i figur 1.



Figur 1: Blooms Taksonomi (Webb, 2014, s. 64)

Den hierarkiske organiseringen av de kognitive læringsmålene i Blooms taksonomimodell, fra enkel til kompleks og fra konkret til abstrakt, innebærer at en oppgave som f.eks. ligger på nivået *Application* normalt vil oppfattes som mer kognitivt krevende enn for oppgaver som ligger på nivåene under.

I følge Webb (2014) fikk imidlertid teamet bak taksonomien kritikk fra matematikere som f.eks. Hans Freudenthal for at kategoriene var lite egnet for hva det vil si å beherske matematikk. Freudenthal var ikke uenig i selve kategoriene, men derimot motstander av hvordan de ble brukt for å angi et bestemt nivå. Mens Blooms taksonomi beskriver *kapasiteten til å løse et gitt problem* som indikator på et bestemt kognitivt nivå, ser Freudenthal heller på hvilke *metoder* eleven bruker for å løse problemet som avgjørende for å bestemme hvilket nivå eleven befinner seg på. I tillegg påpekte andre at distinksjonene mellom nabo-kategoriene var uklare, at taksonomien ikke er hierarkisk, snarere en oppramsing av kategorier og at taksonomien er endimensjonal i det den helt ser bort fra fagkunnskap til fordel for prosess. For som Bloom et al. (1956, s. 12) sier det: «We are not attempting to classify the particular subject matter or content».

Som svar på kritikken publiserte Anderson og Krathwohl (2001) en revidert todimensjonal utgave av Blooms taksonomi, der kunnskaps-kategorien ble delt inn i *factual*, *conceptual*, *procedural* og *metacognitive*, og skilt fra de kognitive prosess-kategoriene *remember, understand, apply, analyze, evaluate* og *create.* Revisjonen tok altså opp i seg kritikken om å inkludere fagkunnskap i rammeverket, men utelater viktige matematiske prosesser som handler om representasjoner, hypoteser/antagelser og bevis, i tillegg til at forståelse havner langt nede i hierarkiet (Kilpatrick, 2014).

Tankegodset fra Blooms Taksonomi har ikke gått upåaktet hen. I flere internasjonale tester i matematikk, som TIMSS og PISA, finner vi igjen kognitive ferdigheter med tilhørende underkategorier. Selv om referansen til Blooms Taksonomi ikke er direkte nevnt i TIMMS-rammeverket for matematikk, sier Webb (2014) at det er tydelig at denne taksonomien har påvirket organisasjonens strukturering av kognitive ferdigheter med tilhørende underkategorier. TIMSS opererer f.eks. med de tre kognitive kategoriene kunne (*knowing*), anvende (*applying*) og resonnere (*reasoning*), som ut fra beskrivelsene i rammeverket til TIMSS tydelig bærer preg av en hierarkisk kompleksitet fra det å kunne huske fakta, via det å kunne anvende fakta til å løse rutineoppgaver, til det å kunne løse oppgaver der man må resonnere seg fram til en hensiktsmessig løsningsmetode (Webb, 2014). Mullis et al. (2003, s. 25) advarer imidlertid mot en instrumentell hierarkisk forståelse av de kognitive kategoriene i TIMMS-testene: «cognitive complexity should not be confused with item difficulty. For nearly all of the cognitive skills listed, it is possible to create relatively easy items as well as very challenging items». I PISA-undersøkelsen er de tre kognitive kategoriene formulere (*formulate*) bruke (*employ*) og vurdere (*interpret/evaluate*) organisert horisontalt for å unngå en slik hierarkisk tolkning (Webb, 2014).

2.2.2 Trådmodellen til Kilpatrick, Swafford og Findell

Med utgangspunktet «All young Americans must learn to think mathematically, and they must think mathematically to learn» (Kilpatrick, Swafford & Findell, 2001, s. 1) satte The National Research Council (NRC) ned en bredt sammensatt forskergruppe for å se på hvordan man kunne forbedre matematikkundervisningen. Ved å definere hva matematisk kompetanse består av, tenkte forskergruppen at det ville være mulig å utforme undervisning for å stimulere til utvikling i disse kompetansene hos elevene.

I *Adding it up: Helping children learn mathematics* tilkjennegir forfatterne imidlertid at «no term captures completely all aspects of expertise, competence, knowledge and facility in mathematics, we have chosen *mathematical proficiency* to capture what we believe is necessary to learn mathematics successfully (Kilpatrick et al., 2001 s. 116). De beskriver videre et rammeverk for mathematical proficiency bestående av fem tråder (*strands*):

Conceptual understanding handler om å kunne bygge opp begrepsmessige strukturer og å se sammenhenger mellom ulike begreper, ideer og prosedyrer. Begrepet innbefatter også å tolke, forstå og benytte ulike representasjoner, og å kunne veksle mellom disse.

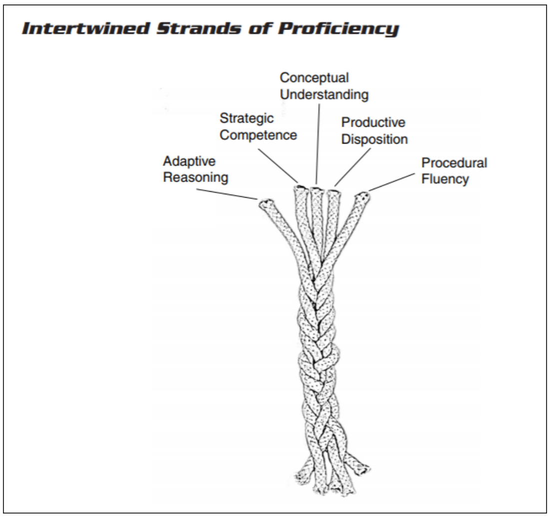
Procedural fluency handler om å kunne utføre ulike matematiske prosedyrer nøyaktig, fleksibelt og hensiktsmessig. Begrepet innebærer både en forståelse av hvordan en prosedyre skal/kan gjennomføres og hvorfor den er gyldig.

Strategic competence handler om å kunne gjenkjenne og formulere matematiske problem, representere dem på ulike vis, utvikle en løsningsstrategi og vurdere rimeligheten av løsningen.

Adaptive reasoning handler om å forklare hvordan man tenker, følge med i logiske resonnement og vurdere gyldigheten av resonnementet. Kompetansen innebærer også å kunne se og begrunne sammenhenger mellom ulike begreper, egenskaper, framgangsmåter og å kunne argumentere for gyldigheten av en hypotese ved å gå fra en kjent situasjon til det ukjente som skal undersøkes.

Productive disposition handler om å kunne se på matematikken som fornuftig, nyttig og verdifull. I tillegg innbefatter begrepet troen på at det er mulig bli kompetent i matematikk, og at innsats bidrar til læring.

Som figur 2 viser, legger forfatterne vekt på at de fem trådene i mathematical proficiency (matematisk kompetanse) må sees i sammenheng, og at de representerer ulike aspekter i en kompleks helhet: «The five strands are interwoven and interdependent in the development of proficiency in mathematics» (Kilpatrick et al., 2001, s. 116).



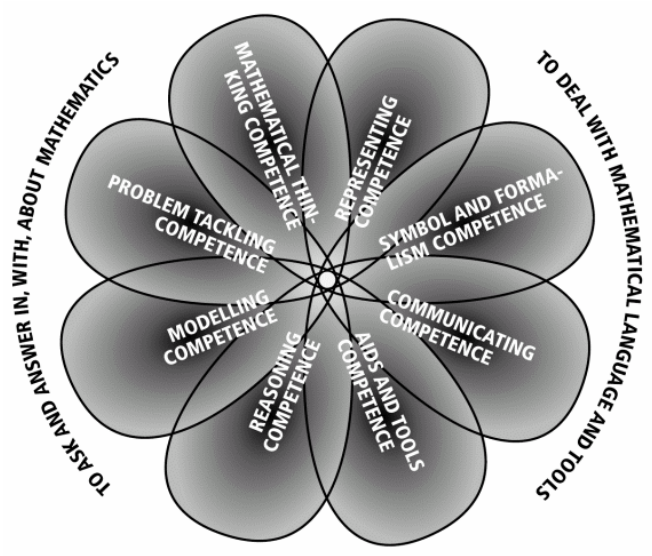
Figur 2: Five Strands of Proficiency in mathematics (Kilpatrick et al., 2001, s. 117)

Istedenfor å presisere hvilket matematisk innhold (*mathematical content knowledge*) opplæringen skal bestå av, fokuserer dette rammeverket primært på de kognitive prosessene/områdene som elevene må beherske for å kunne utøve matematikk (Kilpatrick, 2014).

2.2.3 Kompetanseblomsten til KOM-prosjektet

I 2002 utgav Undervisningsministeriet i Danmark rapporten *Kompetencer og Matematiklæring*. Rapporten var et resultat av arbeidet med å etablere en kompetansebasert systematisk forståelse og utvikling av matematikkfaget, med tilhørende forslag til fornyelse av matematikkundervisningen. Arbeidet ble ledet av Mogens Niss, professor ved Roskilde universitetssenter, og arbeidsgruppens anbefalinger var at læreplanene i matematikk bør fokusere på den kompetansen elevene skal oppnå i utdanningsløpet framfor den tradisjonelle vektleggingen av emnebaserte læreplaner og pensumlister (Niss & Jensen, 2002).

Prosjektet utviklet en pragmatisk beskrivelse av hva det vil si å mestre matematikk, der matematisk kompetanse består av «at have viden om, at forstå, udøve, anvende, og kunne tage stilling til matematik og matematikvirksomhed i en mangfoldighed af sammenhænge, hvori matematik indgår eller kan komme til at indgå» og at *en* matematisk kompetanse er «indsigtsfuld parathed til at handle hensigtsmæssigt i situationer, som rummer en bestemt slags matematiske udfordringer» (Niss & Jensen, 2002, s. 43). Rapporten påpeker at selv om de ulike matematiske kompetanser har sin egen selvstendige og rimelig avgrensede identitet, er det dog både forbindelseslinjer og overlapp mellom dem. Man kan tenke seg *en* matematisk kompetanse som et knutepunkt i en klynge av matematisk innhold, der konsentrasjonen er sterkest nær midten, for så gradvis å tynnes ut mot kanten, der de i sin tur overlapper med andre kompetanser i matematikk. Dette impliserer også at en matematisk kompetanse ikke kan erverves isolert sett, men må sees i sammenheng med sine «kompetanse-kompanjonger» (Niss & Jensen, 2002). Dette kan illustreres med *kompetanseblomsten*:



Figur 3: Kompetansene i KOM-prosjektet visualisert som kronblader i en blomst (Niss 2015, s. 41)

Som vi ser av figuren over, utarbeidet KOM-prosjektet en modell for hva som kjennetegner matematisk kompetanse, bestående av to hovedkompetanser («overkompetanser») med tilhørende åtte delkompetanser, der hver kompetanse innbefatter å være i stand til, på grunnlag av viten og konkrete ferdigheter, å kunne utøve bestemte typer av matematiske aktiviteter (Niss & Jensen, 2002).

Kompetansen *å kunne spørre og svare i, med og om matematikk* karakteriseres av fire matematiske delkompetanser (Niss & Jensen, 2002):

Tankegangskompetanse handler om å være klar over hvilke typer spørsmål som er karakteristisk for matematikk, å kunne stille slike spørsmål selv, og å ha blikk for hvilke typer av svar som kan forventes. Dette innbefatter også å kjenne, forstå og kunne bruke matematiske begreper, å kunne abstrahere og generalisere, og å kunne skille mellom påstander, antagelser og bevis.

Problembehandlingskompetanse handler om å kunne identifisere, formulere, avgrense og presisere ulike matematiske problemer og å kunne løse matematiske problemer, gjerne på forskjellige måter.

Modelleringskompetanse handler om å kunne analysere og avkode eksisterende matematiske modeller og å kunne utføre aktiv modellbygging fra en gitt situasjon. Elementer i aktiv modellbygging er å kunne identifisere og strukturere situasjonen som skal modelleres, matematisere situasjonen (oversette situasjonen til et matematisk språk, med matematiske problemstillinger, nødvendige symboler og matematiske uttrykk), behandle det matematiske uttrykket modellen har gitt oppgav til og til sist validere og analysere modellen kritisk. Kompetansen inneholder også det å kunne diskutere modellen med andre og å kunne vurdere ulike modeller opp mot hverandre.

Resonneringskompetanse handler om å kunne følge og bedømme matematisk argumentasjon man blir presentert for. I tillegg karakteriseres kompetansen av evnen til å tenke ut og kunne gjennomføre uformelle og formelle resonnementer, å kunne omforme resonnementer og antagelser til gyldige bevis, og å kunne følge og bedømme matematiske resonnementer og forstå hva et bevis er.

For kompetansen *å kunne håndtere språk og redskaper i matematikk* beskriver KOM-prosjektet fire spesifikke kompetanser (Niss & Jensen, 2002):

Representasjonskompetanse handler om å kunne forstå (i betydningen avkode, tolke og skille mellom) og bruke ulike representasjoner av matematiske objekter, fenomener, problemer eller situasjoner (det være seg symbolsk, visuelt, verbalt eller konkret). Til denne kompetansen ligger også det å kunne forstå forbindelsene mellom ulike representasjonsformer, og det å kunne velge blant og oversette mellom dem.

Symbol- og formalismekompetanse handler om det å kunne bruke og avkode symbol- og formelspråk, oversette mellom matematisk symbolspråk og dagligtale og å kunne håndtere «spillereglene» for formelle matematiske systemer.

Kommunikasjonskompetanse handler om å kunne kommunisere i, med og om matematikk. Det være seg å kunne sette seg inn i og tolke andres matematikkholdige utsagn (skriftlige, muntlige eller visuelt) eller det å kunne uttrykke seg om matematiske forhold på ulike måter på forskjellige nivåer med tanke på teoretisk eller teknisk presisjon.

Hjelpemiddelkompetanse handler om å kjenne til og kunne bruke ulike hjelpemidler som egner seg til matematisk virksomhet. Innunder dette hører også det å ha blikk for hjelpemidlenes muligheter og begrensninger og det å kunne bruke dem på en hensiktsmessig måte.

Som vi ser av de ulike kompetansebeskrivelsene innbefatter de både en undersøkende side (forstå, reflektere og analysere) og en produktiv side (evnen til å gjennomføre prosesser), der «The *radius of action* of a given competency refers to the *range of different kinds of contexts and situations* in which a person can successfully activate the competency.» (Niss, 2015, s. 44-45).

Tankegodset fra KOM-prosjektet har gitt kime til mye forskning på dette feltet (Kilpatrick, 2014), og har influert læreplanutvikling, oppgaveutforming og vurdering i flere land, f.eks. Kunnskapsløftet i Norge (Valenta et al., 2015) og PISA-undersøkelsene i regi av OECD (Niss, 2015; OECD, 2017).

Rammeverkene som er beskrevet til nå, med unntak av Anderson og Krathwohl sin revisjon av Blooms Taksonomi, fokuserer alle på de kognitive prosessene/områdene som må aktiviseres og utvikles for å beherske matematikk. Læreplaner og nasjonale og internasjonale tester er imidlertid ofte bygd opp i to dimensjoner, med beskrivelse av både faginnhold og kognitive prosesser som kjennetegn på hva det vil si å mestre faget. Både PISA- og TIMSS-undersøkelsene, som begge gjennomføres for et utvalg av norske skoleelever, er eksempler på internasjonale tester i sistnevnte kategori.

2.2.4 TIMSS (Trends in International Mathematics and Science Study)

Trends in International Mathematics and Science Study (TIMSS) er en internasjonal komparativ trendstudie i matematikk og naturfag for grunnskolen med røtter tilbake til 1960-tallet. Undersøkelsen gjennomføres hvert fjerde år, og ble i 2015 gjennomført i Norge for et utvalg av elever på 5. trinn og 9. trinn. Formålet med studien er «å kartlegge faktorer som fremmer læring, følge med på utvikling i eget land (trend) og kunne sammenlikne med andre lands utdanningssystemer og gi informasjon til læreplanutvikling» (Nilsen & Kaarstein, 2016, s. 179). TIMMS måler, i motsetning til PISA, kompetanse ut ifra et rammeverk basert på deltagerlandene sine læreplaner i matematikk og naturfag og «det er et mål at rammeverkene skal ligge så tett som mulig opp til læreplanene i deltakerlandene» (Nilsen & Kaarstein, 2016, s. 182). Ifølge Kjærnsli, Jensen, Bergem, Kaarstein og Nilsen (2019) er oppgavene i TIMSS nesten alltid enkeltstående, det er flere oppgaver uten kontekst og tekstene er ofte korte. TIMSS innbefatter også en analyse av faktorer som påvirker læringsutbytte gjennom spørreskjemaer til elever, lærere, rektorer og foreldre (Nilsen & Kaarstein, 2016).

For utforming av oppgaver i matematikk tar rammeverket til TIMSS utgangspunkt i to dimensjoner: Faglige emneområder og kognitive kategorier. Emneområdene er tall, algebra, geometri og statistikk, mens de kognitive kategoriene er beskrevet som hvilke kognitive ferdigheter det er forventet at elevene kan ta i bruk i en gitt matematisk kontekst (Mullis & Martin, 2013). Ifølge Pedersen (2013) bygger dette kognitive rammeverket på Anderson og Krathwohl (2001) sin revisjon av Blooms taksonomi.

I rammeverkene for matematikk og naturfag er det tre kognitive kategorier: å kunne (*knowing*), å anvende (*applying*) og å resonnere (*reasoning*). De tre kognitive kategoriene for matematikk, er ifølge Nilsen og Kaarstein (2016, s. 183), på norsk beskrevet slik:

Å kunne innebærer å huske fakta, gjenkjenne objekter og uttrykk, beherske de fire regningsartene for heltall, brøker og desimaltall, hente informasjon fra tabeller og diagrammer, måle og klassifisere.

Å anvende innebærer å bruke kunnskapene og ferdighetene sine til å velge metoder og strategier, representere informasjon, modellere situasjoner, følge instruksjoner og løse rutineproblemer.

Å resonnere innebærer å tenke logisk, analysere situasjoner og sammenhenger, generalisere resultater, kombinere informasjon, begrunne påstander og løse problemer som ikke er rutinepreget.

Ut fra beskrivelsene av de ulike kategoriene ser vi her en løst ordnet hierarkisk kognitiv kompleksitet fra det å kunne huske, gjenkjenne, klassifisere og utføre matematiske algoritmer, til det å anvende kunnskaper og ferdigheter til å velge, representere, modellere og løse rutineproblemer, for på høyeste nivå å resonnere gjennom å analysere, generalisere, kombinere, begrunne og løse ikke-standardiserte problemer.

2.2.5 Program for International Student Assessment (PISA)

PISA-undersøkelsen er en internasjonal sammenlignende trendstudie av 15-åringers kunnskaper og kompetanser innenfor fagområdene lesing (*reading literacy*), matematikk (*mathematical literacy*) og naturfag (*scientific literacy*) i regi av OECD. Undersøkelsen gjennomføres hvert tredje år, der Norge har deltatt siden år 2000, med vektlegging av et fagområde for hver treårssyklus. Sist gang matematikk var fokusområde var i 2015. Tanken bak å velge aldersgruppen 15-åringer, er at dette i det store og hele harmonerer med avsluttet obligatorisk skolegang for medlemslandene. Slik mener man å kunne vurdere hvor godt skolesystemet i ulike land forbereder elevene til videre utdanning, arbeidsliv og samfunnsdeltagelse (Kjærnsli & Jensen, 2016). Dette harmonerer godt med formålet for matematikkfaget slik det er beskrevet i den norske læreplanen i matematikk (Nortvedt & Pettersen, 2016).

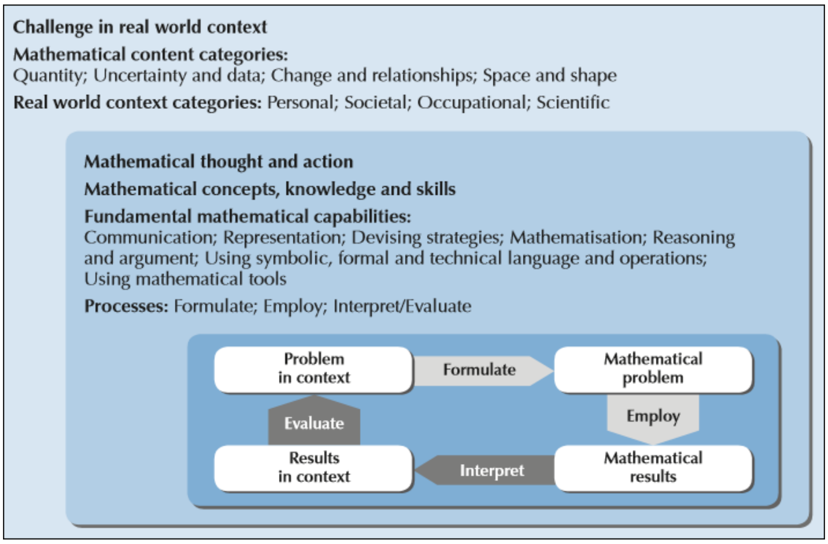
Rammeverket som ligger til grunn for utviklingen av PISA-oppgavene i matematikk, der et av utgangspunktene er å beskrive hva som antas å være viktige fagspesifikke kunnskaper og ferdigheter for en 15-åring, er utviklet av et internasjonalt team av forskere og fagdidaktikere. Ideen bak oppgavene er å få fram hvordan matematisk kompetanse (*mathematical literacy*) kan forstås og brukes i ulike sammenhenger, noe som medfører at alle oppgavene tar utgangspunkt i reelle, konkrete og kontekstuelle situasjoner. I motsetning til f.eks. TIMSS og eksamen i matematikk på 10.trinn er derfor ferdig oppstilte regnestykker fraværende i PISA-oppgavene (Nortvedt & Pettersen, 2016).

Selv om PISA-undersøkelsen har flere likhetstrekk med den norske læreplanen i matematikk, er den, i motsetning til f.eks. TIMSS-testen, uavhengig av de ulike landenes læreplan. PISA-rammeverket for matematikk bærer imidlertid mange felles argumenter både med formåls- og kompetansebeskrivelsene for den norske læreplanen i matematikk, og tidligere analyser har pekt på at PISA-testene i matematikk måler kompetanse som er relevant i forhold til den norske læreplanen i faget (Nortvedt & Pettersen, 2016)

PISA-undersøkelsen har som mål å teste deler av 15-åringer sin kompetanse i aktive problemløsing og modellering, og definerer, ifølge OECD (2017, s. 67), begrepet mathematical literacy som

> *an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens.*

Sammenhengene i definisjonen er illustrert i følgende modell, som også bærer likhetstrekk med NRC (2001) sin beskrivelse av hva det innebærer å beherske matematikk:



Figur 4: *A model of mathematical literacy in practice* (OECD, 2017, s. 68)

OECD (2017) framhever «to formulate, to employ, to interpret» som nyttige og meningsfulle begreper for å organisere de matematiske prosessene som beskriver hva en aktiv problemløser trenger for å kople konteksten i et problem med matematikk, for på den måten å løse problemet. De tre problemløsingsprosessene i denne syklusen er på norsk beskrevet av Nortvedt og Pettersen (2016) som *gjenkjenne og formulere*, *bearbeide og bruke* og *tolke og vurdere*, der førstnevnte prosess handler om kognitive aktiviteter som det å kunne (ut ifra en reell problem-situasjon) identifisere matematiske aspekter ved problemet. Dette innebærer f.eks. å skille relevant informasjon fra irrelevant informasjon, omforme (og eventuelt forenkle) problemet ved å bruke matematiske ord og uttrykk. *Bearbeide og bruke* innebærer at eleven må lage en strategi for å løse problemet og ta i bruk matematiske metoder og verktøy (begreper, fakta, prosedyrer og resonnementer) som trengs for å løse problemet. Neste steg blir så å *tolke og vurdere* løsningen opp mot den opprinnelige situasjonen, der det å oversette fra et matematisk språk til et hverdagsspråk, vurdere egnetheten til løsningen og å identifisere begrensninger i modellen som ligger til grunn for resultatet, er sentrale aspekter. Denne idealiserte og forenklede framstillingen for en problemløsningssyklus er illustrert nederst på figur 4.

For å mestre de tre prosessene i denne problemløsningssyklusen trekker OECD (2017) veksler på arbeidet til Mogens Niss og KOM-prosjektet, og identifiserer og beskriver sju fundamentale matematiske kapasiteter som ligger til grunn for en aktiv og kompetent problemløser. Ut ifra tradisjonen for begrepsbruk i Norden har Nortvedt (2013) valgt å bruke kompetanse framfor kapasitet i denne sammenhengen, og hun beskriver kort de sju kompetansene i PISA-rammeverket (Nortvedt, 2013, s. 48) som det å kunne:

- *kommunisere med, i og om matematikk*
- *matematisere og modellere både matematiske og virkelige situasjoner*
- *representere matematiske størrelser, velge, veksle mellom og bruke matematiske representasjoner i oppgaveløsing*
- *resonnere og argumentere matematisk*
- *planlegge, velge ut og bruke problemløsingsstrategier*
- *bruke symbol- og formalspråk, regler og formelle matematiske metoder*
- *velge ut og bruke matematiske verktøy og hjelpemidler*

Hun påpeker at de ulike kompetansene vil aktiviseres ulikt, alt etter problemet som skal løses, men at alle kompetansene tas i bruk i alle tre prosessene (Nortvedt, 2013). Forholdet mellom de tre matematiske prosessene (gjenkjenne og formulere, bearbeide og bruke og tolke og vurdere) og kompetansen *communication* er i rapporten for PISA 2015 illustrert på følgende måte (OECD, 2017, s. 71):

	Formulating **situations mathematically**	***Employing*** **mathematical concepts, facts, procedures and reasoning**	***Interpreting*****, applying and evaluating mathematical outcomes**
Communicating	Read, decode, and make sense of statements, questions, tasks, objects or images, in order to form a mental model of the situation	Articulate a solution, show the work involved in reaching a solution and/or summarise and present intermediate mathematical results	Construct and communicate explanations and arguments in the context of the problem

Figur 5: Forholdet mellom prosesser og kompetanse i matematikk.

Som vi ser av beskrivelsen over, knyttes en helhetlig matematisk kompetanse i PISA-rammeverket til det å kunne håndtere matematikk i ulike og ukjente situasjoner. Det å løse reelle oppgaver i konkrete situasjoner står sentralt i PISA-testen, og knyttes til modellerings- og problemløsningssyklusen (fig. 4). Elevene testes som aktive problemløsere som må ta i bruk matematikkfaglig kunnskap for å kjenne igjen og løse problemene som beskrives i oppgaven. Dette har, slik jeg ser det, klare paralleller til beskrivelsen av matematisk kompetanse i Kunnskapsløftet, der nettopp *problemløsning og modellering* blir beskrevet som sentrale aspekter under formålet med faget (Utdanningsdirektoratet, 2013).

2.3 Rammeverk for kategorisering av oppgaver i matematikk

Når forskere skal analysere oppgaver i matematikk kategoriserer de ofte ved å skille mellom hvordan oppgavene er presentert og utformet og hva som kreves for å løse oppgaven. Til førstnevnte kategori regnes egenskaper som matematisk innhold og tema, illustrasjoner og figurer, tekstutforming og kontekst, mens man i sistnevnte kategori ser på hva slags kunnskaper, operasjoner og kognitive aspekter som kreves for å løse oppgaven (Pettersen, 2019). Ifølge Turner, Dossey, Blum og Niss (2013) er karakteristika som matematisk innhold, kontekst og visuelle trekk å regne som overflate-karakteristika (*surface characteristics*) som er mer eller mindre lette å observere direkte.

I følge Niss et al. (2017) er det, på tross av alle studiene av utforming og krav til kunnskap for å løse matematikkoppgaver, få studier som tar for seg et komplett kompetansesystem der de ulike kompetansene som kreves for å løse ulike oppgaver analyseres og koples sammen. Boesen, Lithner og Palm (2018) peker på at dette kan skyldes at slike rammeverk fortrinnsvis er laget for å utvikle læreplaner og ikke for empirisk analyse av kompetanse knyttet til løsning av ulike matematikkoppgaver.

2.3.1 The Mathematical Competency Research Framework (MCRF)

Boesen et al. (2018) brukte imidlertid en slik vinkling med et komplett kompetansesystem på et empirisk materiale når de analyserte matematikkoppgaver gitt på nasjonale prøver i Sverige. Forskergruppen tok utgangspunkt i *the Mathematical Competency Research Framework* (MCRF) beskrevet i Lithner, Bergquist, Bergquist, Boesen, Palm og Palmberg (2010). Rammeverket tar utgangspunkt i Niss og Jensen (2002) sin beskrivelse av hva det innebærer å besitte matematisk kompetanse, og definerer tre kompetanse-relaterte aktiviteter (*interpret*, *do and use* og *judge*) som de mener er nødvendig evner (*ability*) for å beherske en matematisk kompetanse. MCRF henter fem av sine seks kompetanser fra NCTM sine *Process Standards* (NCTM, 2000), noe som ifølge Lithner et al. (2010) skyldes at dette er prosessmål som er godt spredt og kjent i fagmiljøene og i tillegg har vært gjenstand for mye forskning. På bakgrunn av dette definerer MCRF de seks kompetansene *Problem solving ability, Reasoning ability, Representation ability, Connection ability, Communication ability* og *Applying procedures ability*, der sistnevnte kompetanse ikke er med i NCTM (2000) sine *Process Standards*, men som ifølge Lithner et al. (2010) er nødvendig å ta med for å kunne beskrive eksisterende praksis. Boesen et al. (2018) fant at dette rammeverket var hensiktsmessig å bruke med tanke på å bestemme om matematikktester har en jevn fordeling på tvers av kompetanser, og finner at rammeverket favner et bredt spekter av hva det vil si å beherske matematikk og at det forteller i hvilken grad ulike tester kartlegger en helhetlig matematisk kompetanse.

2.3.2 The PISA Mathematics Expert Group (MEG)

En annen studie (Turner et al., 2013; Turner et al., 2015) som også berører hvilke matematiske kompetanser som kreves for å løse ulike matematikkoppgaver, er arbeidet utført av *the PISA Mathematics Expert Group*. De har, med utgangspunkt i matematikk-rammeverket til PISA og matematikkoppgavene gitt til 15-åringer på PISA-testene fra 2003 til 2012, utarbeidet et skjema for kartlegging av i hvilken grad oppgaveløsingene krever aktivisering av ulike matematiske kompetanser. MEG-skjemaet tar for seg de seks matematiske kompetansene *communication*, *devising strategies*, *mathematising*, *representation, using symbols, operations and formal language* og *reasoning and argument*, der kompetansene *reasoning* og *mathematical thinking* i KOM-beskrivelsen er stått sammen til *reasoning and argument* og kompetansen *mathematical aids and tools* er utelatt. Dette arbeidet er, som vi ser av kompetansebenevningene, også influert av KOM-prosjektet (Niss, 2015). Turner et al. (2013) har dessuten vist at MEG-skjemaet, når brukt av eksperter, er hensiktsmessig å bruke for å identifisere kompetansekravene til ulike matematikkoppgaver i PISA-testene.

Det er imidlertid en distinkt forskjell mellom KOM-rammeverket og MEG-skjemaet: Mens de åtte kompetansene i KOM blir beskrevet bestående av en kjerne som gradvis overlappes av andre kompetanser (jf. *kompetanseblomsten* til Niss og Jensen, 2012), er kompetansene i MEG-skjemaet forsøkt beskrevet og definert så presist som mulig for å minimalisere overlapping. Dette fordi MEG-skjemaet i utgangspunktet ble brukt til å forutsi empirisk vanskelighetsgrad til PISA-testene, og derfor måtte ha mest mulig adskilte kategorier for å oppnå størst mulig psykometrisk reliabilitet (Niss, 2015). I tillegg er KOM-kompetansene *reasoning* og *mathematical thinking* rekonstruert til *reasoning and argument* i MEG-rammeverket, og *mathematical aids and tools* er tatt ut (Turner et al., 2015). Det er i dag derfor ikke en en-til-en korrespondanse mellom kompetanse-kategoriene i de to rammeverkene: MEG-skjemaet, utarbeidet på bakgrunn av PISA-rammeverket i matematikk, står på egne ben, med mål om en så klar og tydelig operasjonalisering og nivådeling av de seks kompetansene som mulig (Niss, 2015). Turner et al. (2015, Appendix 2) beskriver de seks kompetansene som:

Communication har et tosidig aspekt, og handler både om hvor krevende det er å lese, tolke og forstå hva oppgaven handler om og hva som kreves av kommunikasjon for å formidle utregninger og svar.

Devising strategies dreier seg om det å velge/bruke/lage en plan for å løse oppgaven, samt det å overvåke sin egen problemløsningsprosess.

Mathematising handler om i hvilken grad elevene må bruke modellering for å oversette fra en situasjon utenfor matematikk til det matematiske domenet eller visa versa. Dette innebærer f.eks. å lage en matematisk modell ut fra en reell situasjon og tolke og vurdere hvor godt modellen passer med den opprinnelige beskrivelsen av problemet.

Representation dreier som om å tolke, lage, bruke, velge og oversette mellom ulike matematiske representasjoner i jakten på en løsning.

Using symbols, operations and formal language handler om å kunne forstå og bruke matematiske begreper, fakta, regler og prosedyrer, huske og bruke matematisk symbolspråk og å forstå og manipulere algebraiske uttrykk og formler.

Reasoning and argument handler om å kunne trekke logiske slutninger forankret i matematiske tankeprosesser for å kople sammen argumenter som munner ut i en løsning.

Rekkefølgen i MEG-skjemaet er heller ikke tilfeldig, men lagt opp slik man tenker seg en ideell problemløser vil nærme seg oppgavene i PISA-testene: Lese og forstå oppgaven, lage en plan for å løse problemet, matematisere situasjonen og ta i bruk relevante representasjoner som i sin tur må løses ved å bruke matematiske regler og prosedyrer. Det hele skal munne ut (kommuniseres) i en løsning forankret i matematiske argumenter og resonnement (Turner et al., 2015; Niss, 2015).

For å nyansere de kognitive kravene for hver av de seks delkompetansene satte forskerteamet nivåer fra 0 til 3, der nivå 0 beskriver fravær eller lave krav til anvendelse av en kompetanse, mens beskrivelsene på nivå 3 kjennetegnes av avanserte og komplekse krav til den aktuelle kompetansen. For anvendelse av kompetansen *devising strategies*, kan dette illustreres på følgende vis (Turner et al., 2015, s. 111 – 112):

> *0: Take direct actions, where the solution process needed is explicitly stated or obvious.*
> *1: Find a straight-forward strategy (usually of a single stage) to combine or use the given information.*
> *2: Devise a straight-forward multi-stage strategy, for example involving a linear sequence of stages, or repeatedly use an identified strategy that requires targeted and controlled processing.*
> *3: Devise a complex multi-stage strategy, for example that involves bringing together multiple sub-goals or where using the strategy involves substantial monitoring and control of the solution process; or evaluate or compare strategies.*

Vi ser her tydelig hvordan kravet til strategitenkning for å løse oppgaven øker fra det å kunne «gyve» direkte løs på oppgaven fordi løsningsprosessen er opplagt, via det å finne en enkel (ett-stegs) strategi som løser oppgaven, til det å måtte lage mer sammensatte strategier, som enten krever flere steg, eller er så krevende at de innbefatter komplekse strategier over flere trinn, der man kanskje også må vurdere og sammenligne ulike strategier.

2.3.3 Utprøving av MEG-skjemaet på eksamensoppgaver i Norge

For norske forhold er skjemaet testet ut på to ulike arenaer. Kristianslund (2015) brukte MEG-skjemaet for å analysere kognitive utfordringer på eksamensoppgaver i matematikk gitt ved tre norske grunnskolelærerutdanninger. Med utgangspunkt i beskrivelsene i MEG-rammeverket kodet hun 473 matematikkfaglige oppgaver og fant at lærerstudentene ble testet på lavt kognitivt nivå i de fleste kategoriene som MEG-rammeverket bruker for å beskrive matematisk kompetanse. Unntakene var fakta, prosedyrer og resonnement, der hun så et noe høyere nivå.

I artikkelen *Identifying Competency Demands in Mathematical Tasks: Recognising What Matters* (Pettersen & Nortvedt, 2018) beskriver og analyserer forskerne arbeidet til fem deltagere, som etter en felles skolering, individuelt brukte MEG-skjemaet for å kategorisere oppgavene gitt til eksamen i matematikk for 10. trinn i 2014 og oppgavene på PISA-testen for 2012. Deltagerne bestod av to lærere og tre lærerstudenter i sitt siste år av masterprogrammet i lærerutdanning. Resultatene fra analysearbeidet til Pettersen og Nortvedt indikerer at deltagerne klarte å bruke MEG-skjemaet til å identifisere de seks kompetansene som skjemaet beskriver i forbindelse med oppgaveløsing, men strevde mer når de, på skalaen fra 0 til 3, skulle avgjøre hvor kognitivt krevende oppgaver var. På tross av individuelle forskjeller i kategorisering av oppgavenes kognitive krav, leverte imidlertid deltagerne som gruppe et nokså konsistent resultat. Dette tar forfatterne til inntekt for at MEG-skjemaet kan være et nyttig verktøy for lærere når de skal diskutere hvorvidt matematikkoppgavene de presenterer for elevene legger til rette for aktivisering av ulike kompetanser og deres grad av kompleksitet. Dette er i tråd med Turner et al. (2015) sin påstand om at MEG-skjemaet kan være hensiktsmessig å bruke for lærere når de skal velge ut og tilrettelegge egnede oppgaver for vurdering av elevenes matematiske kompetanse.

Andreas Pettersen var imidlertid også nysgjerrig på om kategoriseringen av kompetanse-kompleksitet i MEG-skjemaet kunne brukes til å forutsi hvilke oppgaver som vil vise seg vanskelig å løse for elever. I artikkelen Mathematical *Competency Demands of Assessment Items: a Search for Empirical Evidence* (Pettersen og Braeken, 2019) tar forfatterne opp tråden fra analysearbeidet rundt bruken av MEG-skjemaet til Turner et al. (2013), som ved hjelp av lineære regresjonsanalyser konstruerte modeller for sammenhengen mellom teoretisk kodet nivå og faktisk elevprestasjon på PISA-tester. Turner et al. (2013) fant, ut fra gjennomsnittet av nivå-kategoriseringen til de åtte fagpersonene som utførte kodingen, at kompetansene *reasoning and argument, symbols and formalism,* og *problem solving* kunne forklare over 70 % av variansen i elevens resultater.

Pettersen og Braeken (2019) bygger i sitt arbeid videre på analysen til Pettersen og Nortvedt (2018), som lot «noviser» med liten erfaring i bruken av MEG-skjemaet kode både norsk eksamen i matematikk for 10. trinn anno 2014 og de frigitte oppgavene fra PISA 2012. I tillegg utvider de perspektivet til MEG-skjemaet ved å bruke koderesultatene fra en norsk eksamen for å undersøke sammenhengen mellom nivå-kategoriseringene og hvor vanskelig oppgavene i praksis viste seg å være for elevene. Resultatene herfra viste seg å være i tråd med hovedtrekkene fra Turner et al. (2013): For PISA-oppgavene kunne vanskelighetsgraden for kategoriene *symbols and formalism*, *reasoning and argument* og *devising strategies* forklare 55 % av variansen i de faktiske elevresultatene, mens for eksamensoppgavene var tilsvarende tall 48 %, der kun vektingen av de to førstnevnte kategoriene kunne brukes til å forutsi empirisk resultat. I tillegg fant forskerne ut at klassifiseringsnivået til *Representation* i PISA-oppgavene kan være en bestemmende faktor for elevresultat, men at denne som regel opptrer i følge med andre kategorier (f.eks. *symbols and formalism),* og derfor ikke er en sterk nok indikator til å stå på egne ben.

I det både PISA-testene og eksamen i matematikk for 10. trinn har til hensikt å vurdere elevenes helhetlige matematiske kompetanse, og begge bygger på kompetansebeskrivelsene i KOM-rammeverket, forventet Pettersen og Braeken at kompetansene beskrevet i MEG-rammeverket også ville være representert i de ovennevnte to testene, selv om de selvfølgelig ikke var konstruert direkte med tanke på kategoriene i MEG-skjemaet. Resultatene til Pettersen og Braeken (2019) viste imidlertid at MEG-rammeverket for matematisk kompetanse passet best med oppgavene i PISA-testene, og de mener det er et tankekors at det, ut ifra deres analyse, kun er *Symbols and formalism* og *Reasoning and argument* som kan relateres til elevenes faktiske resultater (Pettersen & Braeken, 2019, s. 421):

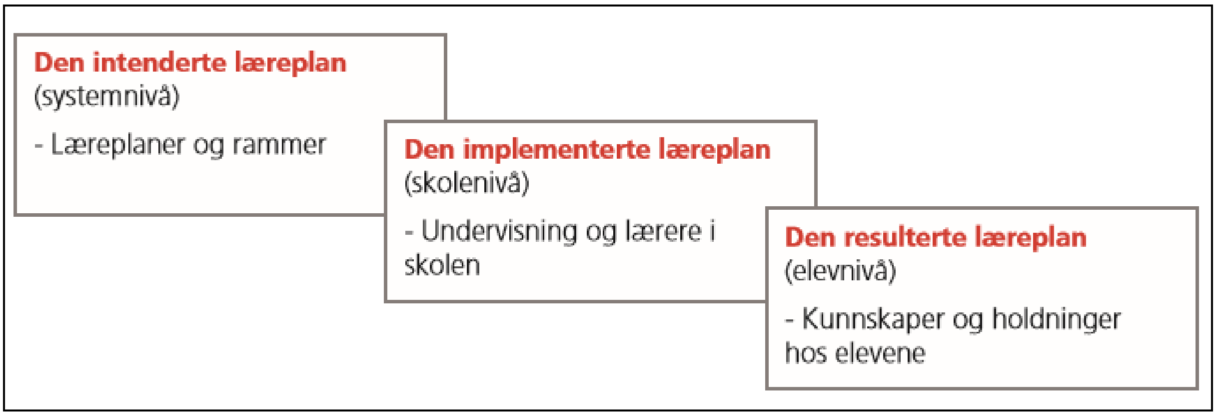
> *These results could indicate that the narrow focus on procedural skills for a rather large proportion of the exam items might be at the expense of a valid measure of more general mathematical competence as represented through the six MEG competencies.*

2.4 Læreplanteori

2.4.1 Tre perspektiver på læreplanen

Læreplaner er styringsdokumenter der myndighetene legger føringer for hva slags faginnhold og kompetanser de anser som viktige. Læreplaner kan dog oppleves ulikt, alt etter om man representerer myndigheter, lærer eller elev. En tradisjonell inndeling er, ifølge Pettersen (2019) og Grønmo (2010), som begge refererer Goodlad (1979), å se på læreplanene fra tre perspektiv: Den intenderte læreplanen, den implementerte læreplanen og den resulterte læreplanen.

Det øverste nivået, *systemnivået*, handler om tilrettelegging og organisering av skolen fra nasjonale myndigheter sin side. Innholdet i den intenderte læreplanen gjenspeiles i formelle dokumenter som lover, læreplandokumenter og andre uttalte mål og forventninger for skolen fra ansvarlige myndigheter. Andre rammefaktorer, som hvordan skolesystemet er organisert og hvilke muligheter elevene har for valg av skole og fag, er også en del av dette. I det eksamensordninger kan antas å være et viktig styringsredskap i skolen, er det vanlig å legge den til det intenderte nivået. Alt i alt kan man si at den intenderte læreplanen er den tilsiktede læreplanen, som beskriver hva slags utdanningstilbud myndighetene *ønsker* at elever skal få (Grønmo, 2010). Den implementerte læreplanen handler om hvordan intensjonene fra systemnivå blir oversatt til faktisk *praksis*, hvordan læreplan og rammer fra systemnivået nedfeller seg på den enkelte skole og i den enkelte klasse, og med selve undervisningen som en avgjørende faktor (Grønmo, 2010). Den resulterte læreplanen handler om *elevenes læringsresultater*: Altså hvilken kompetanse elevene har ervervet i form av både på kunnskaper, ferdigheter og holdninger (Grønmo, 2010).



Figur 6: Framstilling av læreplaner på tre nivåer (Grønmo, 2010, s. 28)

Som antydet i beskrivelsene av de ulike læreplannivåene over, er det ikke selvsagt at det går en sammenhengende ubrutt linje fra hva nasjonale myndigheter ønsker å oppnå med utdanningen til hva elevene faktisk lærer på skolen. Både Polikoff, Porter og Smithson (2011) og Boesen, Helenius, Bergqvist, Bergqvist, Lithner, Palm og Palmberg (2014) har pekt på uoverensstemmelser mellom intendert og resultert læreplan i matematikk, der matematisk kompetanse i læreplanen f.eks. innsnevres til hovedsakelig å dreie seg om innlæring av prosedyrekunnskap for elevene. Til det siste punktet kan kanskje innføringen av nye læreplaner i Norge (LK20) sees på som en reaksjon fra myndighetene sin side for å unngå denne tolkningen, idet det her eksplisitt påpekes at «I matematikk skal elevene jobbe mer med metoder og tenkemåter slik at de får større forståelse for faget» (Kunnskapsdepartementet, 2018a).

I denne oppgaven er det forholdet mellom den intenderte læreplanen i matematikk, slik den er beskrevet i Kunnskapsløftet, og den implementerte læreplanen for samme fag, slik den kommer til uttrykk på eksamen i matematikk for 10. trinn, som er gjenstand for analyse.

2.4.2 Matematisk kompetanse i Kunnskapsløftet og PISA

I innledningen til læreplanen i matematikk fellesfag (MAT1-04) beskrives matematisk kompetanse som viktig både for utviklingen av det norske samfunnet og for den enkelte borger. Matematisk kompetanse er nødvendig for å forstå og kunne påvirke prosesser i samfunnet, og «Kompetanse i matematikk er ein viktig reiskap for den einskilde, og faget kan leggje grunnlag for å ta vidare utdanning og for deltaking i yrkesliv og fritidsaktivitetar.» (Utdanningsdirektoratet, 2013, s. 2). I tillegg beskriver læreplanen fem grunnleggende ferdigheter som er nødvendige forutsetninger for læring og utvikling i skole, arbeid og samfunnsliv. Formålet med matematikkfaget blir dermed å utvikle den matematiske kompetansen som er nødvendig for både individet og samfunnet. Dette sammenfaller i stor grad med begrunnelsene for å undervise i matematikk som vi finner i matematikkrammeverket til PISA (Nortvedt & Pettersen, 2016).

Kompetanse er et gjennomgående begrep i Kunnskapsløftet, med beskrivelse av hovedområder og kompetansemål for hva elevene skal kunne etter endt opplæring. De grunnleggende ferdighetene er gjennomgående integrert i kompetansemålene på alle trinn og i alle fag (Utdanningsdirektoratet, 2017h). I læreplanverket for Kunnskapsløftet (Utdanningsdirektoratet, 2016, s. 1) er kompetanse forstått som *evnen til å løse oppgaver og mestre komplekse utfordringer. Elevene viser kompetanse i konkrete situasjoner ved å bruke kunnskaper og ferdigheter til å løse oppgaver. Det kan handle om å mestre utfordringer på konkrete områder innenfor utdanning, yrke- og samfunnsliv eller på det personlige plan.*

Overordnet del av LK20 (Utdanningsdirektoratet, 2018f, s. 11) definerer kompetanse som

> *Kompetanse er å kunne tilegne seg og anvende kunnskaper og ferdigheter til å mestre utfordringer og løse oppgaver i kjente og ukjente sammenhenger og situasjoner. Kompetanse innebærer forståelse og evne til refleksjon og kritisk tenkning*.

Beskrivelsen av matematisk kompetanse i Kunnskapsløftet bygger i stor grad på arbeidet til KOM-prosjektet (Dale, Engelsen & Karseth, 2011; Valenta et al., 2015), og er overlappende med begrepet *mathematical literacy* som brukes i rammeverket for matematikk i PISA-undersøkelsen (Nortvedt, Pettersen, Petterson & Sollerman, 2016). I begge beskrivelsene er problemløsning og modellering sentrale aspekter ved matematisk kompetanse, og det legges vekt på at elevene er aktive problemløsere som kan bruke et bredt spekter av matematisk kompetanse i ulike sammenhenger (Utdanningsdirektoratet, 2013; OECD, 2017).

Rammeverket for matematikktestene i PISA er utviklet med utgangspunkt i hva OECD anser som viktige kompetanser for 15-åringer med tanke på videre utdanning, yrkes- og samfunnsliv. PISA-testene tar utgangspunkt i matematisk handling i form av problemløsing og modellering, de sju kapasitetene/kompetansene som trengs for å mestre dette og de fire sentrale ideene/innholdsområdene *Forandring og sammenheng*, *Rom og form*, *Tall og mål* og *Usikkerhet* (Nortvedt, 2013). I motsetning til matematikkeksamen for 10. trinn er PISA-testene ikke lærplanbasert, men har som formål å måle elevenes kompetanse i å bruke matematikk for å løse virkelighetsnære problemer. Analyser av data fra PISA-testene viser at den matematiske kompetansen som uttrykkes i de mål som er satt for Kunnskapsløftet ikke er vesentlig annerledes enn den kompetansen som kommer til uttrykk i PISA-prøvene (Olsen, 2010; Nortvedt, 2013).

Kunnskapsløftet og rammeverket for matematikk i PISA-testene viser stor grad av sammenfallende begrunnelser for formålet med det å undervise matematikk. De har en felles plattform i KOM-prosjektet sin beskrivelse av matematisk kompetanse og legger en kompetansebasert vurdering til grunn. Jeg ser det derfor som relevant å bruke konkretiseringen av matematikkrammeverket fra PISA-undersøkelsen (MEG-studien) når jeg skal analysere eksamensoppgaver i matematikk på 10. trinn.

2.5 Sammenligning av rammeverk for eksamensvurdering og MEG-skjemaet

Ifølge *Rammeverket for eksamen*, som er retningsgivende for arbeidet med sentralt gitt skriftlig eksamen, skal det ved utforming av eksamensoppgaver tas utgangspunkt i kompetansemålene i læreplanen for faget, der de grunnleggende ferdighetene regnes som en integrert del av kompetansemålene (Utdanningsdirektoratet, 2017g).

Formålet med eksamen er å gi informasjon om elevens individuelle kompetanse i faget i form av en eksamenskarakter, men også at eleven skal få anledning til å vise sin kompetanse i samsvar med læreplanen. Selve eksamen-settet utarbeides, på oppdrag fra Utdanningsdirektoratet, av egne fagnemnder, der begge instanser har ansvar for at oppgavene er i tråd med læreplanverket for Kunnskapsløftet og relevante forskrifter (Utdanningsdirektoratet, 2017g).

Ut fra Kunnskapsløftet sin definisjon av kompetanse velger fagnemndene tema, hovedområder, kompetansemål og oppgavetyper som er relevante å prøve til eksamen. Ifølge *Eksamensveiledningen 2019* (MAT0010) «prøver eksamensoppgaven kandidatene bredt i kompetansemål fra alle hovedområdene i læreplanen. Elevene prøves ikke nødvendigvis i *alle* kompetansemålene i læreplanen.» (Utdanningsdirektoratet 2019a, s. 6).

På samme måte som MEG-rammeverket har eksamensveiledningen for MAT0010 utarbeidet en matrise for å beskrive kjennetegn på måloppnåelse (kvaliteten på) for hvor godt elevene mestrer faget. I MEG-rammeverket beskrives kompleksiteten på en skala fra 0 – 3, der nivå 0 angir at kompetanse-kategorien enten er fraværende eller kun krever anvendelse på et (svært) lavt nivå. For vurdering av mestring i faget opererer Utdanningsdirektoratet med matrisen *Kjennetegn på måloppnåelse* i form av beskrivelser for karakterene 1 og 2, og for intervallene 3 – 4 og 5 – 6, der «karakteren 1 uttrykker at eleven har svært lav kompetanse i faget» (Utdanningsdirektoratet, 2019a, s. 15).

Vurderingen av matematikkompetansen på eksamen i matematikk for 10. trinn er delt inn i tre kategorier: 1) *begreper, forståelse og ferdigheter*, 2) *problemløsing* og 3) *kommunikasjon*. Kategorien *begreper, forståelse og ferdigheter* handler om i hvilken grad elevene kjenner, forstår og bruker matematiske begreper, og hvordan elevene avkoder, oversetter og behandler symboler og formler (Utdanningsdirektoratet, 2019a). Dette sammenfaller, slik jeg forstår eksamensveiledningen til Utdanningsdirektoratet, godt med kategoriene *using symbols, operations and formal language* og *representation* i beskrivelsene til MEG (Turner et al., 2015)

Begrepet *problem* er i eksamensveiledningen til MAT0010 gitt en vid betydning, fra enkle rutinemessige oppgaver, til større og mer komplekse oppgaver. Kategorien problemløsning handler om hvordan elevene bruker kunnskaper og ferdigheter for å løse matematiske problemstillinger og deres evne til å se sammenhenger i faget, og er ifølge eksamensveiledningen for 2019 den mest sentrale kategorien for sensors vurderingsgrunnlag (Utdanningsdirektoratet, 2019a). Elevenes evne til å lage, bruke og vurdere modeller faller også inn under denne kategorien. Selv om beskrivelsene av problemløsning (strategisk tankegang) og matematisering er mye tydeligere formulert i MEG-rammeverket (f.eks. beskrivelsene av de ulike stegene i både problemløsning- og modelleringsprosessen), ser jeg klare paralleller til det Turner et al. (2015) beskriver som *devising strategies og mathematising.*

I hvilken grad elevene er kjent med og kan bruke ulike hjelpemidler er også lagt under kategoriene *problemløsning* i eksamensveiledningen for 2019 for MAT0010 (Utdanningsdirektoratet, 2019a). Denne kompetansen er utelatt fra MEG-rammeverket, så dette aspektet vil derfor ikke bli fanget opp i analyseskjemaet.

I hvilken grad elevene er i stand til å sette seg inn i en matematisk tekst og uttrykke seg skriftlig ved hjelp av matematisk symbolbruk, er i Utdanningsdirektoratet (2019a) beskrevet under kategorien *kommunikasjon*. Dette passer, slik jeg ser det, godt med *the communication competency* hos Turner et al. (2015), idet de også beskriver kommunikasjon som bestående av både en lese- og forstådel (*receptive component*) og en produserende side (*constructive component*).

I den ovenfor omtalte eksamensveiledningen (Utdanningsdirektoratet, 2019a) står det også å lese at de tre kategoriene ikke skal forståes adskilt, men at de beskriver matematikkompetansen på tvers av læreplanens hovedområder og kompetansemål, og som i KOM-prosjektet (Niss & Jensen, 2002) presiseres det at kompetanse-kategoriene har flytende overganger. Både beskrivelsene av kompetansemålene for matematikk i Kunnskapsløftet og analyseskjemaet til MEG for bestemmelse av kognitivt nivå på matematikkoppgaver bygger i stor grad på arbeidet til KOM-prosjektet. Som beskrevet over ser jeg derfor mange fellestrekk med tanke på kompetansen som må brukes for å vurdere og kategorisere matematikkoppgavenes kompleksitet. Dette understøttes også av Pettersen (2019, s. 16), som sier at «the MEG scheme and its six competencies are regarded as highly relevant to mathematics education in Norway and a suitable tool for analysing mathematical tasks in Norwegian secondary education».

2.6 Vurdering og undervisning

Flere forskningsresultater har påpekt sammenhenger mellom vurdering av elever og undervisningspraksis (Niss & Jensen, 2002; Wilson, 2007; Jensen, 2007; Boesen, 2006; Alseth, Breiteig & Brekke, 2003; Michelsen, 2001; Schoenfeld, 2007)

Niss og Jensen (2002) påpeker at kompetanse som ikke gjøres til gjenstand for vurdering fort kan bli utelatt fra undervisningen. Presisering av kompetansemål i læreplanen for å sette rammer for undervisningen er med andre ord ikke tilstrekkelig. De må også settes på dagsorden i vurderingssammenheng (Niss & Jensen, 2002, s. 168):

> *I al matematikundervisning er evalueringsspørgsmål af central betydning, hvad enten man tænker på forskellige former for afsluttende evaluering, herunder prøver og eksamener, eller på løbende evaluering knyttet til selve undervisningen. Der er overvældende forskningsmæssig evidens for, at uanset hvilke evalueringsformer der benyttes, udøver evalueringen en væsentlig tilbagevirkende indflydelse på undervisnings- og læreprocesser.*

Vurdering i form av f.eks. eksamen kan med andre ord påvirke hva skolen, lærere og elever anser som viktig i undervisningssammenheng. Hvis man fra sentralt hold ønsker å befeste, utvikle eller endre undervisningspraksis bør man lage vurderingsformer som understøtter læreplanmål og ønsket undervisningspraksis (Niss & Jensen, 2002).

Schoenfeld (2007) og Wilson (2007) argumenterer likeledes for at tester spiller en sentral rolle for undervisningspraksisen. Førstnevnte forfatter bruker begrepet *WYTIWYG – What You Test Is What You Get*, og poengterer at jo større konsekvenser testen har for elevene, desto sterkere effekt vil den ha på undervisningspraksisen (Schoenfeld, 2007, s. 12):

> *For example, if the high-stakes assessment in mathematics focuses on procedural skills, teachers may drill their students for procedural fluency — and conceptual understanding and problem solving skills may be left unaddressed as a consequence.*

I sine litteraturstudier av *high-stakes tests* (prøver med viktig utfall) fant Wilson (2007) at lærere bruker mye tid på oppgavetyper som ligger tett opptil slike tester, og at lærere dermed tilpasset sin undervisningspraksis til gjeldende prøver og eksamener hvor utfallet var viktig.

For norske forhold har Alseth, Breiteig og Brekke (2003) i sin evaluering av L97 tilsvarende påpekt at sentralt gitt eksamen i matematikk er sterkt styrende for innholdet i matematikkundervisningen. De ser at eksamen har en tilbakevirkende effekt på undervisningen, og at lærere på ungdomsskolen fokuserer på oppgaver som er i tråd med eksamen i matematikk.

Niss (1993) peker på at selv om læreplanene i matematikk har gjennomgått en stor utvikling siden 1970-tallet, og arbeidsformer som problemløsning, modellering, utforsking og hypotesetesting er kommet inn i mange lands læreplaner i matematikk, er ikke vurderingsformene endret i tilsvarende grad, med det til følge at det blir et sprik mellom intensjonene i læreplanen og eksisterende vurderingsformer. Gjeldende vurderingsformer klarer derfor i liten grad å evaluere den kunnskap og kompetanse man ønsker at undervisningen skal fremme. Tilsvarende sammenheng finner Jensen (2007) for danske forhold, der han viser til at eksamen i matematikk mer er et hinder enn et gavn for bruk av modellering og problemløsing i matematikkundervisningen.

Slik jeg vurderer utsagnene og resultatene til Niss (1993), Niss og Jensen (2002), Alseth et al. (2003), Wilson (2007) og Schoenfeld (2007) har vurdering, og da i særlig grad vurdering hvor utfallet er viktig, stor påvirkningskraft på undervisningen. Dersom myndighetene ønsker at undervisningen i matematikk skal gi elevene bred matematisk kompetanse og ivareta intensjonene i formålet med faget, der f.eks. problemløsning og modellering står eksplisitt nevnt, må dette også gjenspeiles i oppgavene på eksamen. Utdanningsdirektoratet sitt *Rammeverket for eksamen*, med tilhørende forskrifter, og *Eksamensveiledning* legger føringer for arbeidet med utforming av eksamen i matematikk. Praktisk sett vil det i løpet av en eksamen på fem timer ikke være mulig å lage oppgaver som berører alle kompetansemålene i læreplanen. Fagnemndene må derfor gjøre et utvalg av oppgaver de finner relevante i forhold til læreplanen i matematikk. Slik jeg forstår læreplannivåene til Grønmo (2010) vil dermed det som testes på eksamen være et uttrykk for den implementerte læreplanen slik den er tolket og forstått av fagnemnden i matematikk, og derfor ikke nødvendigvis samsvare med den intenderte læreplanen slik den er beskrevet i politisk bestemte læreplandokumenter og forskrifter fra Utdanningsdirektoratet.

2.7 Oppsummering av teorikapitlet

Jeg har i dette kapitlet gjort rede for ulike rammeverks beskrivelse av matematisk kompetanse. Blooms taksonomi beskriver et faguavhengig hierarkisk ordnet system med seks nivåer (figur 1), som senere ble revidert av Anderson og Krathwohl til også å inkludere fagkunnskap. På begynnelsen av 2000-tallet lanserte the National Research Council *trådmodellen* (figur 2), bestående av fem innbyrdes avhengige tråder som til sammen representerer ulike aspekter ved matematisk kompetanse. Omtrent samtidig la KOM-prosjektet, med Niss og Jensen i spissen, fram sin *kompetanseblomst* (figur 3) bestående av åtte kompetanser for hva det vil si å mestre matematikk. Med unntak av rammeverket til Anderson og Krathwohl, er alle de ovenfor nevnte kompetansebeskrivelsene endimensjonale, i den forstand at de beskriver hvilke kognitive prosesser som er nødvendig å beherske for å mestre matematisk tankegang, uavhengig av emnet.

TIMSS- og PISA-testene i matematikk har rammeverk som inkluderer både faglige emneområder og kognitive kategorier, der TIMMS, ifølge Pedersen (2013), bygger sitt rammeverk på Anderson og Krathwohl, mens PISA tar utgangspunkt i arbeidet til KOM-prosjektet (Nortvedt, 2013). Dette todimensjonale perspektivet finner vi også igjen i LK06 (Kunnskapsløftet), som både beskriver emner (hovedområder) innenfor matematikkfaget og kompetanser elevene må beherske for å mestre faget. Denne kompetansebeskrivelsen er ifølge Valenta et al. (2015) og Dale et al. (2011) sterkt influert av KOM-prosjektet sine kompetansebeskrivelser. På samme måte er også innholdet til LK20 (Fagfornyelsen) organisert, med fem kjerneelementer som beskriver sentrale arbeidsmåter, metoder og tenkemåter i matematikk, mens det sjette kjerneelementet beskriver sentrale kunnskapsområder i faget.

De presenterte rammeverkene for matematisk kompetanse beskriver med ulike, men dog beslektede, termer hva som kjennetegner det å beherske matematikk. Felles for de ulike rammeverkene er imidlertid at de alle beskriver matematisk kompetanse som en mosaikk, sammensatt av flere delkompetanser, som til sammen beskriver en helhetlig matematisk kompetanse.

På samme måte som det finnes ulike rammeverk for matematisk kompetanse, er det også utformet ulike rammeverk for kategorisering av hvilke kompetanser som kreves for å løse ulike oppgaver i matematikk. MEG-skjemaet, utviklet på bakgrunn av PISA-rammeverket i matematikk, er et slikt oppgaveanalyseverktøy. Skjemaet definerer i hvilken grad de seks matematiske kompetansene *communication*, *devising strategies*, *mathematising*, *representation, using symbols, operations and formal language* og *reasoning and argument* krever aktivisering fra nivå 0 til 3. Skjemaet er prøvd ut på norske eksamensoppgaver for 10. trinn av Pettersen (2019), som mener at MEG-skjemaet er et hensiktsmessig verktøy til å analysere norske eksamensoppgaver for denne aldersgruppen. Min sammenligning av matematisk kompetanse i rammeverket for eksamensvurdering og MEG-skjemaet viser også dette.

Veien fra den intenderte læreplanen på systemnivå, via den implementerte læreplanen på skolenivå, til den resulterte læreplanen på elevnivå kan være krokete og usammenhengende. Eksamen i matematikk utarbeides på bakgrunn av kompetansemålene i faget, der *Rammeverk for eksamen* er styrende for fagnemndene sitt arbeid med utforming av eksamen (Utdanningsdirektoratet, 2017g). Det som testes på eksamen er derfor et uttrykk for den implementerte læreplanen slik den er tolket og forstått av fagnemnden i matematikk. Matematikkeksamen er derfor ikke nødvendigvis i samsvar med den intenderte læreplanen, slik den er beskrevet i politisk bestemte læreplandokumenter og forskrifter fra Utdanningsdirektoratet.

Flere forskningsresultater har pekt på sammenhengen mellom vurdering av elever og undervisningspraksis, og poengtert at kompetanse som ikke gjøres til gjenstand for vurdering fort blir utelatt i undervisningen (Niss og Jensen, 2002), og at prøver med viktig utfall er sterkt styrende på undervisningen (Wilson, 2007; Schoenfeld, 2007). Dersom myndighetene ønsker at undervisningen i matematikk skal gi elevene bred matematisk kompetanse og ivareta intensjonene i formålet med faget, må derfor dette også gjenspeiles i oppgavene på eksamen.

3 Metode

For å svare på problemstillingen

> *I hvilken grad tester oppgavene på skriftlig eksamen i matematikk for 10. trinn ulike aspekter ved elevenes matematiske kompetanse*

tok jeg for meg læreplanen i matematikk fellesfag i Kunnskapsløftet (MAT1-04) og skriftlige eksamensoppgaver i matematikk (MAT0010) med tilhørende rammeverk og veiledningsmateriell. Jeg satte sammen en gruppe på 4 ungdomsskolelærere, meg selv inkludert, for å kategorisere matematikkoppgavene etter et bestemt rammeverk. Med bakgrunn i resultatene herfra har jeg sett på i hvilken grad det å løse eksamensoppgavene utfordrer elevenes matematiske kompetanse.

Dette kapitlet beskriver metodevalgene som er gjort for å gjennomføre denne studien. Jeg tar først for meg mitt forskningsdesign, og går deretter over til å beskrive utvalg og data. Deretter beskriver jeg analysemetoden og kodeprosedyren, før jeg avslutningsvis tar opp aspekter omkring forskningskvaliteten på dette arbeidet.

3.1 Forskningsdesign

3.1.1 Kvantitativ og kvalitativ metode

I samfunnsvitenskapelig forskning skilles det tradisjonelt mellom to hovedtyper av metodetilnærminger, henholdsvis kvantitativ og kvalitativ metode. Begrepene refererer til egenskapene til det datamaterialet som er samlet inn og skal analyseres, og hvilken av metodene som er mest hensiktsmessig avhenger av problemstillingen man ønsker å finne svar på (Robson & McCartan, 2016).

Kvantitativ forskning er tallbasert og egner seg godt for statistiske analyser av store mengder data, mens kvalitativ forskning egner seg best for analytiske beskrivelser der formålet er, ved hjelp av fortolkende analyser, å forstå og få dybdeinformasjon om et fenomen eller en situasjon (Husén, 1988).

Ifølge Robson og McCartan (2016) er kvalitative studier preget av både fleksibilitet og sensitivitet, der forskningen kan bli endret og tilpasset underveis, i tillegg til at forskeren selv arbeider nært med sine kilder. Kvantitative studier karakteriseres i sin tur av stringent struktur: Alle enhetene skal behandles likt, og man søker etter (lovmessig) korrelasjon, varians, verifikasjon og kausalitet (Husén, 1988). I kvantitative studier vil det dessuten være en viss distanse mellom undersøker og det undersøkende objekt, mens det i kvalitativ forskning kan hende at anskuelser forskeren bringer til torgs faktisk kan påvirke forskningsresultatene (Robson & McCartan, 2016)

I følge både Cohen, Manion og Morrison (2018) og Robson og McCartan (2016) kan imidlertid kvalitative og kvantitative metoder utfylle hverandre: Et fenomen kan ha både kvalitative og kvantitative aspekter, og kvalitative og kvantitative tilnærminger kan derfor stå i et komplementært forhold til hverandre, der den ene metodes svakhet kan oppveies av den andre, og vise versa.

3.1.2 Mixed methods

Ifølge Robson og McCartan (2016) kan valg av forskningsmetode sammenlignes med detektivarbeid: Forskeren kan *observere* mennesker og prøve å finne ut hva som skjer, og/eller forskeren kan *intervjue* (spørreskjemaer/tester) folk for å finne ut hva som skjer og/eller *look for fingerprints* dvs. se etter andre spor som er lagt igjen på åstedet, som f.eks. gjenstander, DNA-avtrykk, eller som i mitt tilfelle, analyse av dokumenter.

Utgangspunktet for valg av studieområde var et ønske om innsikt i hvilke sider og på hvilket nivå av matematisk kompetanse elevene prøves i på eksamen. Videre var jeg nysgjerrig på hvordan matematikkeksamen passet overens med det som i Kunnskapsløftet er beskrevet under formålet med faget.

Jeg har i denne studien gjennomført en kvalitativ undersøkelse av deler av læreplanen i matematikk for Kunnskapsløftet med tilhørende eksamensoppgaver og rammeverk for vurdering. Ut fra rammeverket til *the PISA Mathematics Expert Group* (Turner et al., 2015) er eksamensoppgavene kategorisert med hensyn på matematisk kompetanse. Innplasseringen i ulike kategorier og nivå er tolket ut fra et gitt rammeverk (MEG-rammeverket), og jeg ser derfor på dette som en kvalitativ vurdering.

For å få en oversikt over hvordan de ulike kategoriene fordeler seg, har jeg deretter telt opp og kvantifisert resultatene i form av tellinger, tabeller og diagrammer. Dette for å få data som kan hjelpe meg i jakten på å besvare mitt forskningsspørsmål. Min studie ligger dermed i skjæringspunktet mellom kvalitativ og kvantitativ analyse, og passer, slik jeg har forstått, godt med det Robson og McCartan (2016) og Cohen et al. (2018) beskriver under paraply-betegnelsen *Mixed methods*. Forfatterne nevner denne metoden som en nyttig strategi for å skaffe seg en mer fullstendig forståelse av et problem, for på den måten å oppnå kunnskapsutvikling.

Mixed methods er en metodisk tilnærming som kan opptre i mange varianter, men der en fellesnevner for gjennomføring av en slik metodikk kan være «data collection (both quantitative and qualitative), analysing and interpretation of studies that, singly or together, address a particular phenomenon.» (Cohen et al., 2018, s. 32).

I mine studier har jeg, etter innsamling og kvalitativ analyse av data, brukt en kvantitativ tilnærming for bedre å kunne si noe om datamaterialets beskaffenhet. En slik metodisk tilnærming har i Robson og McCartan (2016, s. 178) fått merkelappen *Sequential transformative design,* og blir her beskrevet som

> *One method precedes the other with either the qualitative or the quantitative method first. Priority may be given to either method. The results are integrated during interpretation. This design is guided primarily by a theoretical perspective (e.g. by the conceptual framework adopted)*

, noe jeg tenker passer godt overens med min metodiske tilnærming.

I det jeg er hovedinstrument for studien, og tolker, analyserer, drøfter og vurderer det endelige datamaterialet, vil derfor kvaliteten på analysen i stor grad avhenge av kvaliteten på meg som forsker (Robson & McCartan, 2016). Dette innbefatter også hvor godt jeg har klart å opparbeide en felles forståelse for anvendelsen av MEG-skjemaet i gruppa som kategoriserte eksamensoppgavene. Jeg forfekter imidlertid at min tolkning av datamaterialet ikke er den eneste rette, og vil derfor plassere mitt ståsted som forsker ut fra en hermeneutisk tilnærming. Jeg tenker at det ikke finnes én sannhet, og at de tolkningene jeg gjør av et fenomen, alltid vil være farget av konteksten og min forforståelse. I Cohen et al. (2018, s. 52) uttrykkes dette slik: «Hermeneutics […] seeks to understand situations through the eyes of the participants, […] and premised on the view that reality is socially constructed».

3.1.3 Dokumentanalyse

For å svare på problemstillingen i denne studien, har jeg brukt dokumentanalyse. Dokumentanalyse er ifølge Cohen et al. (2018) hensiktsmessig hvis man ønsker å foreta sammenligninger og/eller spore trender over tid. Dokumenter kan inneholde verdifull informasjon, men man må samtidig ta høyde for at de kan være ufullstendige, mangelfulle og er skrevet til et annet formål (og i en annen tid) enn det forskeren ønsker å fokusere på (Robson & McCartan, 2016).

Dokumentene jeg har analysert, er Kunnskapsløftet og sentralt gitt eksamen i matematikk for 10. trinn, med tilhørende eksamensveiledninger. Kunnskapsløftet er et offentlig dokument skrevet for å gi innhold og rammer for det 13-årige opplærings-løpet for elever i Norge. Sentralt gitt eksamen er offentlige dokumenter skrevet av fagpersoner i den hensikt å måle kunnskaper og ferdigheter elevene sitter igjen med etter 10-års skolegang. Rammeverktøyet for kategorisering av oppgaver er også offentlig tilgjengelig. Dokumentene og analyseverktøyet jeg har brukt, eksisterer i «permanent format», så det bør derfor ikke by på problemer hvis andre forskere skulle ønske å gjøre en lignende studie på nytt, eller for ev. å imøtegå mine resultater. Selv om dokumenter ikke endrer innhold etter at de er lest, kan det jeg har med meg i bagasjen av antagelser og forutinntatte meninger (forforståelsen), påvirke analysen av dokumentene. Jeg har derfor etterstrebet så grundig som mulig å redegjøre for de valg jeg har tatt underveis i tolkning- og analyseprosessen, slik at det skal være mulig å følge mine tankebaner og tilnærminger underveis i dette forskningsprosjektet. Det betyr at jeg ikke kan ta for gitt at min tolkning av datamaterialet er den eneste rette, og at det derfor hele tiden skal være mulig for leseren å følge den motsatte vei fra min tolkning tilbake til dataene. For å ivareta åpenheten omkring analyseprosessene og synliggjøre mine valg og tolkninger, vil jeg derfor presentere eksempler på hvordan kategoriseringer av oppgavene er utført. I tillegg er tolkninger av Kunnskapsløftet og eksamensveiledninger søkt synliggjort ved sitater og referanse til disse og andre relevante dokumenter.

3.2 Utvalg og data

Utvalget i denne studien består av læreplanen i matematikk for Kunnskapsløftet og tre skriftlige eksamenssett for 10. trinn, med tilhørende veiledninger. Cohen et al. (2018) nevner fem sentrale utvalgskriterier man må ta hensyn til for å få et hensiktsmessig utvalg, henholdsvis utvalgsstørrelse, representativitet, tilgjengelighet, utvalgsstrategi og hvilken type studium utvalget skal brukes til.

Utgangspunktet for denne studien var min nysgjerrighet omkring hvilke sider av matematisk kompetanse som blir testet på eksamen, og hvordan dette harmonerer med formålsbeskrivelsen for matematikk i Kunnskapsløftet. I det eksamen i matematikk på 10. trinn skal måle et ståsted for hva elevene behersker av matematikk etter endt grunnskoleutdanning, passet det godt med det jeg skulle undersøke. Jeg kunne også valgt å analysere muntlig eksamen i matematikk, men denne er lokalt gitt og ikke kvalitetskontrollert i samme grad som sentralt gitt skriftlig eksamen. I tillegg ville det ha blitt krevende å få samlet min «kode-gruppe» i en periode da de selv er opptatt med egne elevers eksamen og sensoroppdrag. Skriftlig eksamen i matematikk for de tre siste årene er dessuten åpent tilgjengelig på Matematikk.net, i tillegg til at man kan få tilgang til dette materialet på Utdanningsdirektoratets nettsider. Dette gjør mitt datamateriale lett tilgjengelig.

Jeg har valgt å ta for meg de tre siste eksamenssettene. Selv om Kunnskapsløftet som plandokument ikke er endret siden 2013, er utvalget av oppgaver som gis på eksamen likevel gjenstand for skjønn. Hva som prøves på eksamen, kan derfor variere noe fra år til år, og jeg ønsket derfor å se på de nyeste oppgavene for nærmest mulig å kunne si noe om nåsituasjonen, i tillegg til om jeg ev. kunne spore en endring eller konsistens over tid. Dette også med tanke på arbeidet med ny fagplan i matematikk som nå er inne i sin avsluttende fase.

Valg av antall eksamenssett ble også påvirket av tiden og omfanget man har til rådighet på en masteroppgave. Det er begrenset hvor stort datamateriale som er hensiktsmessig å bearbeide innenfor rammen av dette studiet, men der jeg tenker at mitt utvalg på 165 eksamensoppgaver er stort nok til å kunne antyde en trend om hvilke sider av matematisk kompetanse som elevene testes i på eksamen.

For ikke å stå alene i kategoriseringsarbeidet og for å få inn flere perspektiver på tolkningen av analyseverktøyet (MEG-skjemaet), valgte jeg å alliere meg med 3 ungdomsskolelærere. MEG-skjemaet er ikke et «plug and play»-verktøy, og både Turner et al. (2015) og Pettersen (2019) anbefaler at man, for å oppnå god kvalitet, lar en kodegruppe kategorisere de samme oppgavene. Ut fra dette studiet sin egenart valgte jeg derfor personer med både nærhet og erfaring fra undervisning og eksamen på ungdomsskolen.

Kodegruppa bestod av en kvinne med 28 års undervisningspraksis fra ungdomskolen, og to menn, med henholdsvis 22 og 17 års undervisningspraksis fra ungdomsskolen. Av sensorerfaring fra skriftlig eksamen på 10. trinn hadde kvinnen vært sensor en gang, mannen med lengst erfaring tre ganger, mens den andre mannen i gruppa aldri hadde vært skriftlig sensor. Selv har jeg 12 års undervisningserfaring fra ungdomstrinnet og vært skriftlig sensor på eksamen for 10. trinn to ganger.

3.3 Analyse

I denne delen presenterer jeg framgangsmåten for hvordan klassifiseringen og analysen ble gjennomført. Målet med analysen er å kunne si noe om hvilke sider av matematisk kompetanse elevene testes i på eksamen (og om dette er i tråd med gjeldende læreplan). Eksamensoppgavene er analysert for å kunne kategorisere og sammenligne innhold, for på den måten å få fram kvantitative sider ved materialet. Analyseverktøyet som blir benyttet legger imidlertid føringer for kategoriseringen, og kan ha utfordringer knyttet til subjektive tolkninger (Robson & McCartan, 2016) For å imøtegå dette valgte jeg et allerede eksisterende og utprøvd rammeverk for kategorisering av matematikkoppgaver, i tillegg til å utføre «member check», der en gruppe av ungdomsskolelærere kodet de samme oppgavene som meg.

Datamaterialet mitt bestod av 165 oppgaver fordelt på tre skriftlige eksamenssett. Til å klassifisere oppgavene valgte jeg et analyseverktøy utviklet av *the PISA Mathematics Expert Group* (MEG). MEG-skjemaet er en modifisert versjon av KOM-rammeverket, og består av operasjonelle definisjoner for seks matematiske kompetanser, beskrevet i kapittel 2.3.2.

3.3.1 Prosedyrebeskrivelse for koding av oppgaver

MEG-skjemaet som ble benyttet for å kategorisere eksamensoppgavene gir rom for tolkning, og for å sikre best mulig kvalitet i kodingen satte jeg meg først godt inn i utviklingen av MEG-rammeverket fram til dagens versjon. Jeg studerte deretter kode-eksemplene i Turner et al. (2015), Pettersen (2019) og Kristianslund (2015) og kodet deretter en tid etterpå disse eksemplene selv, for å se hvordan min tolkning av MEG-skjemaet harmonerte både med MEG-ekspertene og resultatene til Pettersen og Kristianslund. Jeg ønsket også å få tilgang til resultatene av kodingen av 2014-eksamenssettet som gruppen til Pettersen og Nortvedt (2018) hadde gjort, men det fikk jeg dessverre ikke.

Jeg kodet deretter 2014-eksamenssettet selv, og delte deretter kompetansedefinisjonene med nivåbeskrivelsene og kodeeksempler fra Turner et al. (2015) med min kodegruppe. Kodegruppen bestod av tre lærere fra 2 forskjellige ungdomsskoler, og alle fikk satt av tid på sin arbeidsplass til å sette seg inn i materialet og kode del 1 av eksamen fra 2014. Sammen med selve oppgavesettet fikk også kodegruppa tilgang på eksamensveiledning, sensorveiledning og forhåndssensur, i tillegg til at de underveis i kodingen løste alle oppgavene selv. Når vi så kodet oppgavene med hensyn på kompetansedefinisjoner med tilhørende kompetansenivå, la vi vekt på hvilke løsninger som kan forventes av elever i aldersgruppen 15 – 16 år i matematikk.

Kodegruppa hadde til sammen tre møter av to timer, der vi, for å oppnå en best mulig felles forståelse av MEG-skjemaet, diskuterte de ulike kompetansedefinisjonene og tolkningen av nivåbeskrivelsene. Vi kodet ferdig hele 2014-settet hver for oss, før vi til slutt drøftet og ble enige om en felles koding av hele eksamenssettet. Deretter kodet gruppa de eksamenssettene som skulle brukes i denne studien individuelt, og sendte resultatene over til meg. Dette i tråd med anbefalingene til Turner et al. (2013).

Etter at vi hadde prøve-kodet et stykke ut i 2014-settet, opplevde vi imidlertid et behov for å raffinere MEG-skjemaet. Flere av oppgavene på del 1 krevde bare ett svar uten å vise utregning, mens oppgaveteksten kunne være krevende å forstå for en 10. klassing. Lese og forstå oppgaveteksten (*the receptive component of communication)* kunne f.eks. havne på nivå 1 eller 2, mens nivået for den kommunikasjonen som krevdes for å formidle svaret (*the constructive component of communication*) lå på nivå 0, som tilsier at eleven kun trenger å svare med et enkeltord eller et tall. Det ble derfor vanskelig, og til tider uhensiktsmessig, å skulle operere med et snitt på nivået for kommunikasjon. For bedre å få fram bredden i kompetansen *communication*, bestemte jeg derfor å dele kommunikasjons-kategorien i vårt kode-skjema i to: En for hvilket nivå som kreves for å lese, forstå og tolke oppgaven, og en for hva som kreves for å formidle utregninger og svar.

Når det gjelder bruk av hjelpemidler for å løse oppgaver, er hjelpemiddelkompetanse ikke en del av kompetansekategoriene i MEG-skjemaet (Turner et al., 2015). På del 1 av eksamen er det bare tillatt med skrivesaker, passer, linjal og vinkelmåler, mens del 2 åpner opp for alle hjelpemidler med unntak av Internett og verktøy som tillater kommunikasjon (Utdanningsdirektoratet, 2018e). Under kodingen av oppgavene på del 2 har vi derfor tatt høyde for at bruk av regneark, graftegner og kalkulator kan påvirke på hvilket nivå de ulike kompetansene blir klassifisert. Jeg tenker at dette i første rekke gjelder kompetansen *using symbols, operations and formal language*, fordi enkelte oppgaver blir enklere å løse med tilgang på f.eks. regneark og lærebøker, noe som i sin tur gjør at koding av *using symbols, operations and formal language* blir kodet på et lavere nivå enn tilsvarende oppgave på del 1.

En annen faktor som kan påvirke hvilke kompetanser som aktiveres for å løse de ulike eksamensoppgavene, er valg av løsningsstrategi. Oppgavene kan ofte løses på ulikt vis, noe som gjør at kodingen av samme oppgave kan bli litt ulik, alt etter hvordan man tenker eleven har løst oppgaven. For eksempel kan en oppgave i statistikk på del 1 løses helt instrumentelt ved hjelp av en regel, eller man kan resonnere seg fram til svaret ved å se etter et system eller mønster for hvordan oppgaven kan løses. En variasjon i kodingen er derfor å forvente når man legger til grunn hvilke løsninger som er relevante for en 10. klassing å bruke.

3.3.2 Illustrasjon av klassifisering gjennom kodeeksempler

For å vise hvordan jeg og har gjennomført kodingen, vil jeg her vise noen eksempler på koding av ulike oppgaver. Jeg vil deretter vise eksempler på hvordan kodingen kan se ut når man ser på hele kodegruppen. De fullstendige resultater følger i kapittel 4.

3.3.2.1 Eksempel 1: Koding av oppgave 12a del 1, eksamen 2019

Gloria passer barna til naboen på lørdager. Hun får 80 kroner for å møte opp. I tillegg får hun en timelønn på 50 kroner.

En lørdag fikk Gloria til sammen 180 kroner.

a) Hvor mange timer passet hun barna denne lørdagen?

1 time	2 timer	3 timer	4 timer
○	○	○	○

Figur 7: Bilde av oppgave 12a på del 1, eksamen 2019 (Utdanningsdirektoratet, 2019d).

17 av totalt 30 deloppgaver på del 1 i eksamenssettet for 2019 er flervalgsoppgaver, der elevene skal krysse av for ett rett alternativ. Svar i form av avkryssing utgjør med andre ord over 55% av oppgavene på del 1, så jeg finner det derfor relevant å ta for meg en slik type oppgave i eksempelgjennomgangen.

Oppgaven er kodet på nivå 0 med hensyn på krav som stilles for å lese og forstå oppgaven (*receptive communication*). Setningene er korte og gir direkte all den informasjonen man trenger for å løse oppgaven.

Problemstillingen er presentert som tekst, uten hint om løsningsstrategi, der eleven selv må sette sammen relevant informasjon i oppgaven for å komme fram til en strategi for å løse oppgaven. Dette harmonerer med nivå 1 under kompetansen *devising strategy*, der eleven selv må lage en plan for å løse oppgaven, men at denne planen er enkel og ukomplisert (*a stright-forward strategy*).

Mathematising handler i denne oppgaven om å lage en matematisk modell ut fra en reell situasjon (inntekter for barnepass). All informasjonen man trenger for å konstruere en relevant matematisk modell er oppgitt i oppgaven, så da passer nivå 1 med beskrivelsen i Turner et al. (2015).

I det elevene her bare trenger å lese numeriske verdier rett fra teksten er *representation* kodet til nivå 0.

Selve utregningen er enkel: Man får oppgitt et konstant beløp for oppmøte, og må bestemme antall timer Gloria har arbeidet når hun har tjent 180 kr. Tallene er enkle (80 kr og 50 kr), så *using symbols, operations and formal language* er kodet til nivå 0.

Reasoning and argument er også kodet til nivå 0, fordi elevene her kan trekke slutninger direkte fra oppgitt informasjon.

For presentasjon av løsningen kreves bare et kryss i en av de fire rubrikkene. Dette tilsvarer nivå 0 til kompetansen *constructive communication*.

Denne oppgaven kodet alle i gruppa helt likt.

3.3.2.2 Eksempel 2: Koding av oppgave 4a del 2, eksamen 2019

Oppgave 4 (4 poeng) REGNEARK

Therese lager et budsjett for en ferietur til Tokyo i høstferien. Ferieturen skal vare i 7 dager.

a) Lag og fullfør Therese sitt budsjett som er vist nedenfor.
Vis også hvilke formler du har brukt. Alle beløp er gitt i norske kroner.

	A	B	C	D
1	**Budsjett for ferietur til Tokyo**			
2				
3	**Feriedager**	7		
4				
5	**Utgiftsposter**		Gjennomsnitt per dag	Utgifter
6	Flyreise tur/retur			7000,00
7	Kollektivtransport		100,00	
8	Hotell		800,00	
9	Mat		300,00	
10	Shopping		400,00	
11	Inngang severdigheter		100,00	
12	Sum utgifter			
13				
14	**Inntektsposter**	Timer	Timelønn	Inntekter
15	Sommerjobb	120	115,00	
16	Pengegaver bursdag			3000,00
17	Sum inntekter			

Figur 8: Bilde av oppgave 4a på del 2, eksamen 2019 (Utdanningsdirektoratet, 2019e).

Denne oppgaven var interessant, både fordi jeg og min kodegruppe ikke er enige med løsningsforslaget i sensorveiledningen, og fordi oppgaven forutsetter bruk av digitale hjelpemidler i form av regneark. Uenigheten med sensorveiledningen bestod i hvordan man teller antall overnattinger mot antall dager når man er på tur, og blir forklart under gjennomgangen av kompetansen *reasoning and argument* for denne oppgaven.

Oppgaveinformasjonen består av både tekst, foto og bilde av en tabell fra Excel, der eleven skal lage og fullføre den avbildede tabellen på egen PC. Her må elevene identifisere og lenke relevante informasjonen gitt i teksten med tilhørende tabell, noe som passer med nivå 1 under *receptive communication.*

Strategien eleven må bruke for å fullføre den oppgitte tabellen i Excel er ikke direkte beskrevet i oppgaveteksten, men jeg anser den, så sant eleven er kjent med regneark, for å være nokså rett fram, der riktig formel bare skal plasseres i rett rute. Devising *strategy* er derfor kodet til nivå 1.

Matematiseringen i denne oppgaven handler om å oversette et budsjett for en ferietur til oppsatte utgift- og inntektsposter uttrykt som tallstørrelser i Excel. All informasjon elevene trenger for å gjøre beregninger i den oppgitte modellen (regnearket) er gitt, så *mathematising* er satt på nivå 1.

I jakten på å løse oppgaven må eleven forholde seg til tabellen som er vist i oppgaven. Dette anser jeg som en enkel og standardisert representasjon, noe som sammenfaller med nivå 1 under beskrivelsen av *representation*.

For å løse oppgaven må elevene beherske symbolspråket, prosedyrer og formelbruk i et regneark. Jeg anser at kravene som stilles til dette i denne oppgaven er enkle (nivå 1), men ikke elementære (nivå 0), i det man her må beherske et sett av prosedyrer, formler og symboler. *Using symbols, operations and formal language* setter derfor til nivå 1.

For å unngå en liten kortslutning, eller kort slutning, om man vil, må man i denne oppgaven bruke litt logisk resonnement. Ferieturen er beskrevet å vare i 7 dager. Er man to dager på tur, gir dette en overnatting. Er man 7 dager på tur, gir dette 6 overnattinger. Oppgaven legger opp til en god kobling mellom dagligliv og matematikk, men denne tas det dessverre ikke høyde for i Sensorveiledningen (Utdanningsdirektoratet, 2019b). Den nødvendige logiske slutningen som trengs for å se dette, gjør at jeg setter *reasoning and argument* til nivå 1.

For å uttrykke en løsning må eleven her lage tabellen som vist i oppgaven i sitt regneark, og fylle inn riktig formel for å få rett verdi i cellene. Dette passer med nivåbeskrivelse 1 under *constructive communication*, i det elevene her bare skal fylle inn rett formel i rett celle uten nærmere forklaring enn en formelutskrift.

Som det framgår av tabellen under, der jeg er koder 1, var tre av fire kodere samstemt for kodingen av denne oppgaven.

Oppg. 4a, del 2, 2019	**Koder 1**	**Koder 2**	**Koder 3**	**Koder 4**
Receptive communication	1	1	1	1
Constructive communication	1	1	1	2
Devising strategies	1	1	1	1
Mathematising	1	1	1	1
Representation	1	1	1	1
Symbols, operations & formal lang.	1	1	1	0
Reasoning & argument	1	1	1	1

Tabell 1: Koderesultater fra oppgave 4a, del 2, eksamen 2019.

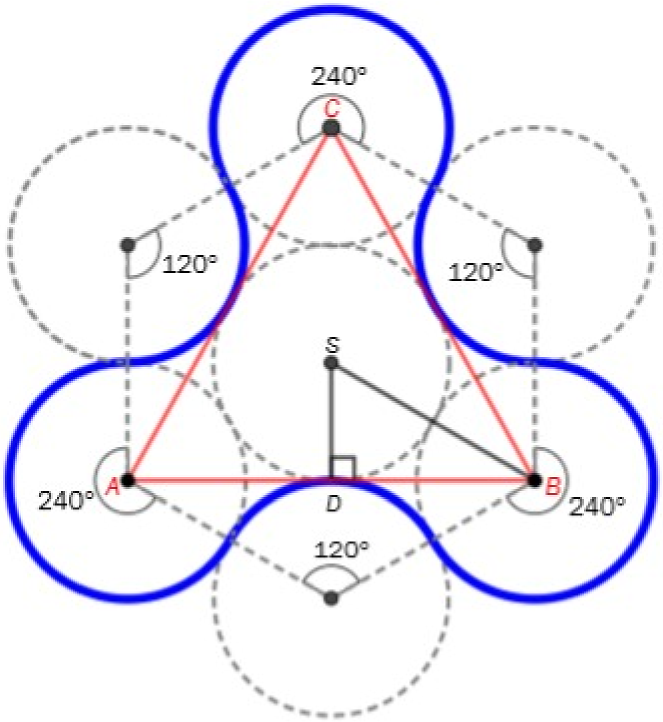
Koder 4 anså imidlertid oppgaven som enklere med tanke på bruk av matematisk faktakunnskap (*using symbols, operations and formal language*) enn resten av gruppa, men tolket tydeligvis *constructive communication* til å ligge et nivå over oss andre. I det beskrivelsene av de ulike nivåene i MEG-skjemaet inneholder begreper som bla. *simple*, *elementary* og *complex* gir dette rom for tolkning av hva som kan forventes av en elev på 10. trinn. Vi forsøkte å legge så klare som mulig premisser for dette under vår prøvekoding, men det er vanskelig å fange opp alle oppgavetyper i gjennomgangen av ett oppgavesett. Det overordnede bildet viser likevel at gruppa som helhet var godt samkjørt i kodingen av denne oppgaven.

3.3.2.3 Eksempel 3: Koding av oppgave 8a del 2, eksamen 2018

Oppgave 8 (4 poeng)

En fidget spinner er en leke som er laget med et sirkelformet kulelager i midten og med tre «armer».

Nedenfor ser du en del av en forenklet tegning av en fidget spinner.
$\triangle ABC$ (rød farge) er en likesidet trekant. Alle sirklene har radius lik 1,5 cm.



a) Vis at omkretsen av fidget spinneren, markert med blå sirkelbuer, er 9π cm.

Figur 9: Bilde av oppgave 8a på del 2, eksamen 2018 (Utdanningsdirektoratet, 2018e).

Oppgaven er valgt som eksempel fordi den viste seg å bli vurdert av kodegruppa til et klart høyere nivå enn flertallet av oppgavene i dette oppgavesettet. I tillegg er den, ifølge Fafo-rapporten som omhandler dette eksamenssettet, en av de oppgavene som empirisk viste seg vanskelig for elevene (lav poengskår). Et siste moment jeg har tatt hensyn til i utvelgelsen av denne oppgaven er poenget til Pettersen og Nortvedt (2018), der de sier at spriket i kodesamsvaret øker når kompetansekravene som stilles for å løse oppgaven øker. Det kan være interessant å se om dette også gjelder min kodegruppe.

Jeg har kodet *receptive communication* til nivå 2 fordi jeg tenker at elevene her må identifisere og velge ut elementer fra figuren, og knytte dette sammen med oppgaveteksten, for å forstå hva oppgaven spør etter. Oppgaveteksten inneholder begreper som *fidget spinner*, *radius*, *omkrets*, *sirkelbue* og at svaret skal bli *9π*, og jeg antar at elevene må gå fram og tilbake i den informasjonen som er presentert i oppgaven for å forstå oppgaven.

Devising strategies handler om å velge eller lage en strategi for å løse oppgaven, samt å overvåke og kontrollere gjennomføringen av den strategien man bruker. Det er her, som på de fleste av eksamensoppgavene, mulig å velge ulike framgangsmåter. En plan for å løse oppgaven vil imidlertid, slik jeg tenker at en elev på 10. trinn kan nærme seg oppgaven, fordre to trinn: Man må bruke informasjonen om størrelsene på sirkelbuene og radiusen til sirklene for så å bestemme omkretsen til spinneren. Dette passer med beskrivelsen på nivå 2 for *devising strategies*, der kjennetegnet er at man lager en lineær plan som består av flere trinn, men at dette steget kan brukes direkte som utgangspunkt for neste trinn.

Oppgaven handler om et virkelig objekt (en spinner) som både er avbildet og vist i en forenklet tegning. For å oversette mellom den ekte spinneren og den matematikken de må bruke for å løse oppgaven, kan de bruke tegningen (modellen) i oppgaven. Dette passer med nivå 1 under *mathematising*, som sier at eleven kan trekke konklusjoner om situasjonen han/hun skal løse direkte fra en gitt modell i oppgaven.

Denne oppgaven er den eneste i dette oppgavesettet jeg har skåret til nivå 2 på *representation*. Figuren man må bruke og dekode i jakten på en løsning består for så vidt stort sett av geometriske figurer som bør være kjent for en 10. klassing (sirkler, sirkelsektorer, likesidet trekant, rettvinklet trekant og regulær sekskant), men dette satt sammen i en figur, med omkretsen av spinneren både innenfor og utenfor de ulike geometriske figurene, gjør at jeg klassifiserer dette som en representasjon som er kompleks både å forstå og å bruke.

Når det gjelder kompetansen *using symbols, operations and formal language* har jeg, ut fra MEG-skjemaet, plassert den på nivå 1. Jeg anser tallene, definisjonene og regneprosedyrene elevene må bruke for å komme fram til 9 π som enkle, der de kan bruke faktakunnskap som vinkelsummen i sirkel og formel for omkrets av sirkel for å regne ut eller bruke som geometriske argumenter i svaret.

Selv om jeg kodet regelbruk og utregningen som enkel i denne oppgaven, betyr ikke dette at oppgaven også har lave krav til logisk resonnement (*reasoning and argument)*. Elevene kan ikke her bare basere seg på et argument fra informasjonen i oppgaven (nivå 0) for å rettferdiggjøre sin løsning, men må trekke flere slutninger ut fra opplysningene i figuren og oppgaveteksten (nivå 1), og knytte argumentene herfra sammen i en kjede av slutninger som bygger på hverandre (nivå 2). For å vise at omkretsen av spinneren i figuren er 9π må de med andre ord argumentere for at omkretsen tilsvarer tre sirkler, at det da blir relevant å bruke formel for omkrets av sirkel (selv om spinneren ikke er sirkulær) og at omkretsen da kan uttrykkes som 9π, noe som tilsvarer nivå 2 på kompetansen *reasoning and argument.*

For å løse denne oppgaven er det, ut fra min koding, mange kompetanser som må aktiviseres på nivå 1 eller 2. Når det kommer til *constructive communication* har jeg, ut ifra beskrivelsen i MEG-skjemaet, falt ned på nivå 2, i det eleven her enten må presentere svaret i form av flere steg med utregninger (legge sammen sirkelsektorene og bruke svaret herfra til å beregne den totale omkretsen) og/eller en kort beskrivelse eller forklaring som viser at omkretsen av spinneren er 9 π.

Ideelt sett skulle de tre lærerne og jeg ha vurdert kompetansekravene på denne oppgaven helt likt, men vi ser av tabellen under at så er ikke tilfelle.

Oppg. 8a, del 2, 2018	**Koder 1**	**Koder 2**	**Koder 3**	**Koder 4**
Receptive communication	2	2	2	3
Constructive communication	2	2	2	2
Devising strategies	2	1	2	2
Mathematising	1	1	1	1
Representation	2	2	1	2
Symbols, operations & formal lang.	1	1	2	2
Reasoning & argument	2	2	2	2

Tabell 2: Koderesultater fra oppgave 8a, del 2, eksamen 2018.

I tråd med observasjonene til Pettersen (2019) øker spriket noe når kompleksiteten i oppgavene øker. Vi ser likevel at nivåforskjellen aldri er større enn 1, og at det overordnede bildet dermed fortsatt er temmelig samstemt.

3.4 Forskningskvalitet

For at resultatene av en undersøkelse skal være pålitelige og gyldige, er det en forutsetning at datamaterialet og analysene er av tilfredsstillende kvalitet. Dette kapitlet tar for seg faktorer som kan påvirke forskningskvaliteten.

3.4.1 Reliabilitet

Reliabilitet handler om påliteligheten av resultatene i en undersøkelse, og er knyttet opp mot utvelgelsen og gjennomføringen av datainnsamlingen, måleinstrumentenes nøyaktighet og etterprøvbarhet. Likevel er innholdet i begrepet, ifølge Cohen et al. (2018), litt ulikt om man ser på det med kvantitative eller kvalitative briller. Mens man i kvantitativ forskning antar at samme metode på samme data i samme undersøkelse vil gi samme resultat, kan dette være en utfordring i kvalitativ forskning.

Dataene jeg har plukket ut, består av eksamensoppgaver for årene 2017 – 2019. Dette innbefatter alle de skriftlige eksamensoppgavene for de tre siste årene det er avholdt eksamen i matematikk på 10. trinn, til sammen 165 eksamensoppgaver. Med utgangspunkt i samtlige eksamensoppgaver tenker jeg da å ha et godt grunnlag for å kunne peke på trender i matematikkeksamen med tanke på matematisk kompetanse for denne 3-års perioden. Jeg valgte eksamensoppgaver for grunnskolen fordi dette i stor grad sammenfaller med den aldersgruppen MEG-skjemaet er utarbeidet for (15-åringer). Alle dataene jeg har samlet er i utgangspunktet «døde dokumenter», som på forespørsel er tilgjengelig på Utdanningsdirektoratet sine hjemmesider. De vil for så vidt være upåvirket av min lesing, men under analysen av dokumentene i denne oppgaven har jeg etterstrebet å understøtte min argumentasjon ved å vise til den aktuelle teksten og forklare min tolkning, slik at leseren kan se og vurdere hvordan jeg har forstått dokumentene.

Koding av oppgavene har vært en «krevende sport», der jeg særlig har vektlagt to faktorer for at kodingen skal bli mest mulig pålitelig, nemlig det som dreier seg om at man klarer å kode konsistent gjennom hele prosessen, og det som handler om i hvilken grad kodingen gjennomføres likt, uavhengig av person. Turner et al. (2013) benevner dette som henholdsvis intrakoderelabilitet og interkodereliabilitet. For å sikre intrakodereliabilitet satte jeg meg godt inn i MEG-rammeverket, delte materialet med resten av kodegruppa, prøvekodet deler av eksamenssettet for 2014 på egenhånd og deretter sammen med gruppa. Deretter kodet vi hver for oss, før vi igjen satte oss sammen og diskuterte resultatet. Dette var en verdifull prosess, i det alle lærerne var kjent med elevgruppen og oppgaveformen eksamen er utarbeidet for. Som gruppe kom vi fram til ulike løsningsmåter som var aktuelle for en 10. klassing, for på den måten bedre å kunne kartlegge hvilke kompetanser som ble aktivisert for å løse oppgaven. I tillegg sammenlignet vi vår koding med eksemplene i Turner et al. (2015), slik at vi fikk synkronisert vår koding opp mot en ekspert. Vi diskuterte også definisjoner og beskrivelsene for den enkelte kompetanse, slik at vi tydeligere klarte å skille kategoriseringene for hver enkelt kompetanse fra hverandre. Dette er tydeligvis et kjent problem for utviklerne av MEG-skjemaet, for de presiserer flere ganger i skjemaet hva aktivisering av en spesifikk kompetanse innebærer, og hva som skiller de ulike kompetansene fra hverandre.

3.4.2 Kodesamsvar

For å styrke reliabiliteten i koderesultatene satte jeg sammen, som anbefalt av Turner et al. (2013), en gruppe som individuelt kodet alle eksamensoppgavene i denne studien. Gruppa bestod av meg selv og tre ungdomsskolelærere, alle med lang fartstid fra undervisning i matematikk på ungdomstrinnet. Ideelt sett burde jeg da fått fatt i personer med erfaring og kunnskap om mitt valgte rammeverk, men det klarte jeg ikke å oppdrive. Pettersen (2019) viser imidlertid at også lærere og lærerstudenter som blir gitt opplæring i bruken av MEG-skjemaet kan håndtere dette, så med det i mente plukket jeg ut tre lærere med lang fartstid fra ungdomsskolen til individuelt å kode de samme oppgavene som meg. Disse ble gitt opplæring over tre økter av to timer, i tillegg til koding på egenhånd og korrespondanse via telefon og e-post i opplæringsperioden.

For å undersøke interkodereliabiliteten, dvs. i hvilken grad vi vurderte samme kompetanse i samme oppgave til likt nivå, valgte jeg å bruke *intraclass correlation coefficient* (ICC). Dette skyldes i hovedsak to faktorer. For det første er dette en samsvarsindikator som, ifølge begrunnelsen i neste avsnitt, egner seg for mitt materiale. Som student ved NTNU fikk jeg tilgang til analyseverktøyet *IBM SPSS Statistics* (SPSS), slik at jeg kunne utføre beregninger for mine ICC-verdier. For det andre benyttet Pettersen og Nortvedt (2018) tilsvarende samsvarsindikator når de analyserte kodesamsvaret for sin gruppe. Jeg tenker da at dette også er en relevant indikator for samsvaret i min kodegruppe. I tillegg gir dette meg muligheten til å sammenlikne mine resultater med deres forskning.

I mitt materiale er kodene heltall, med stigende vanskelighetsgrad fra 0 til 3, der avstanden mellom nivåene ikke er definert. Ifølge Fleiss og Cohen (1973) vil da en vektet Cohens kappa være et hensiktsmessig mål på samsvar i mitt datamateriale, fordi denne reliabilitetsmarkøren tar hensyn til hvor ulikt oppgavene er kodet når kodesamsvaret beregnes. En nivådifferanse mellom koderne på 3 vil da ha større negativ virkning på relabilitetsmarkøren enn hvis nivådifferansen mellom koderne er 1. I SPSS kan man imidlertid ikke beregne vektet Cohens kappa, men Fleiss og Cohen (1973) har vist at ICC kan betraktes som et spesialtilfelle av vektet Cohens kappa, slik at jeg kan bruke mine ICC-verdier som et mål for absolutt overenskomst.

Pettersen og Nortvedt (2018) bruker Landis og Koch (1977) sine retningslinjer når de beregner absolutt overenskomst (ICC) mellom sine kodere. Idet Pettersen og Nortvedt (2018), så vidt jeg har bragt på det rene, er de eneste som har publisert resultater av MEG-skjemaet anvendt på norske eksamensoppgaver, kan det være interessant å sammenligne mine resultater med deres forskning. Jeg valgte derfor å benytte meg av de samme retningslinjene for absolutt overenskomst som de norske forskerne. Et alternativ kunne ev. være å bruke Domenic Cicchetti (1994) sine verdier for styrken på ICC. Han legger imidlertid lista litt lavere for *excellent agreement* (0,75) enn Landis og Koch (1973), som setter grensen til 0,81 for *almost perfect agreement*. Tabell 3 viser verdiene jeg brukte for å vurdere kodesamsvaret.

Kappa Statistic	*Strength of Agreement*
<0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost Perfect

Tabell 3: Sammenhengen mellom kappaverdier og styrkesamsvar (Landis & Koch, 1973, s. 165).

Ved beregning av ICC, *two-way mixed model og absolute agreement*, i SPSS får man opp to mål for styrken på samsvaret, henholdsvis *single-measures* og *average-measures*. I mitt materiale refererer da *single-measure* til troverdigheten av resultatene hvis bare en koder i gruppa skulle ha stått som representant for kodingen av eksamensoppgavene, mens *average-measures* refererer til troverdigheten av resultatene når man ser på hele gruppen under ett, og bruker et gjennomsnitt av kodeverdiene for hver enkel kompetansene. I det vi var fire kodere, er derfor sistnevnte gjennomsnittsmål for hver enkelt kompetanse i MEG-skjemaet mest aktuell for meg. Jeg tar imidlertid også med *single-measure* verdiene, for på den måten å kunne si noe om den enkelte koder sin reliabilitet opp mot våre resultater som gruppe.

3.4.3 Validitet

Validitet handler om gyldigheten av resultatene: I hvilke grad er utvelgelse av informasjon, datamateriale og teorier relevant for de problemstillingene som skal belyses? Dekker de måleinstrumentene som brukes det som faktisk skal måles (intern validitet)? Er resultatene også gyldig i andre sammenhenger, utenfor denne undersøkelsen (ekstern validitet)? (Cohen et al., 2018).

For å imøtekomme kravet til intern validitet har jeg i denne studien søkt å koble relevant teori og rammeverk for matematisk kompetanse, læreplaner og eksamen opp mot mitt forskningsspørsmål, for så på bakgrunn av dette å plukke ut aktuelle data i form av eksamensoppgaver. Jeg bruker et kvalitativt måleinstrument (MEG-skjema) som er prøvd ut på ulike testregimer i matematikk, der resultatene er publisert (Turner et al., 2015;

Pettersen, 2019). Man kan derfor, med en viss styrke, hevde at dette måleinstrumentet faktisk måler det jeg undersøker i min studie. En trussel mot min forskning er likevel om jeg bruker MEG-skjemaet i tråd med utviklernes intensjoner. Jeg har derfor etterstrebet å tydelig presentere det teoretiske rammeverket som ligger bak, i tillegg til å beskrive og forklare kodeprosedyren med tilhørende valg og dilemmaer. Målet med dette er å klargjøre for leseren hvordan jeg har forstått og tolket det aktuelle rammeverket i lys av min problemstilling. På den måten ønsker jeg at andre forskere også skal ha anledning til å gå inn i mitt materiale og etterprøve mine resultater. For norske forhold kan min undersøkelse sees på som en oppfølging av forskningsarbeidet til Pettersen (2019) omkring studien i bruken av MEG-skjemaet for kategorisering av kompetansekrav på eksamensoppgaver i matematikk. Jeg tenker at mitt bidrag her vil være å kunne indikere noen tendenser til, framfor å generalisere om, hva som kjennetegner de siste års skriftlig eksamen i matematikk på 10. trinn, med tanke på matematisk kompetanse.

3.4.4 Etikk

Ifølge retningslinjene til Norsk senter for forskningsdata (NDS) skal man søke godkjenning dersom masteroppgaven inneholder personopplysninger (NDS, 2019). Identiteten til de som har kodet oppgavene sammen med meg er holdt anonym. De deltok på frivillig basis, og kunne når som helst trekke seg fra oppdraget. Vi var i utgangspunktet fem kodere, men en av personene valgte av ytre omstendigheter å trekke seg etter innledende runder. Gruppa har etter dette bestått av tre ungdomsskolelærere og meg selv. Resultatene av kodingen er også anonymisert i oppgaven. Oppgaven inneholder derfor ikke personopplysninger som krever godkjenning fra NSD.

I referansene til eksamensoppgavene med tilhørende veiledningsmateriell for perioden 2017 – 2019 har jeg valgt å oppgi lenkene til ansvarlig utgiver, Utdanningsdirektoratet (Udir). Lenken til dette materialet er imidlertid passordbeskyttet, et passord jeg fikk tilgang til fra Udir. Eksamenssettene jeg brukte er imidlertid også åpent tilgjengelig på *matematikk.net* via lenken https://matematikk.net/side/Eksamensoppgaver.

4 Resultater

I dette kapitlet presenterer jeg resultatene fra analysen av eksamensoppgavene på 10. trinn for årene 2017 – 2019. Resultatene danner grunnlag for diskusjonen i den neste delen av oppgaven.

Matematikkoppgavene er, med utgangspunkt i MEG-skjemaet, kodet med hensyn på hvilke matematiske kompetanser som testes, og på hvilket nivå. Oppgavene er altså kodet fra nivå 0 til nivå 3 i de syv kategoriene *receptive communication (RC)*, *constructive communication (CC),* d*evising strategies (DS), mathematising (M)*, *representation (R)*, *using symbols, operations and formal language (SF)* og *reasoning and argument (RA)*, der kompetansen *communication*, i forhold til MEG-skjemaet, er delt inn i de to kategoriene *receptive communication og constructive communication*. Dette ble gjort etter erfaringene fra prøvekodingene, der jeg så potensialet for å få fram ytterligere interessante sider ved datamaterialet ved å foreta denne todelingen.

Prosedyrebeskrivelsen for koding er forklart i foregående metodekapittel, der kodeeksempler fra tre ulike eksamensoppgaver viser hvordan oppgavene ble klassifisert med tanke på kompetansenivå. I likhet med Turner et al. (2013) og Pettersen og Nortvedt (2018) baserer jeg meg på et gjennomsnitt av koderesultatene for hver kompetanse til hver oppgave. Jeg har derfor valgt å vise tabeller og figurer som viser prosentvise nivå-klassifiseringer for hver enkelt kompetanse.

Totalt har fire personer klassifisert 165 oppgaver med hensyn på 7 ulike matematiske kompetanser. Til sammen utgjør dette 4620 individuelle kodinger. For å gjøre tabellene mer leservennlige, er det i alle tabeller kun benyttet heltall, slik at verdier under 0,5 % vil framstå som 0 % i tabellene. Dette betyr også at det ikke alltid summeres til 100 % i tabellene. Trendene i tallmaterialet vil likevel framstå som tydelige, der jeg ved behov for presisering kommenterer dette i teksten.

I kommentarene til resultatene tar jeg ofte utgangspunkt i laveste nivå (nivå 0). Dette skyldes at nivå 0 i MEG-skjemaet beskriver fravær eller lave krav til anvendelse av en kompetanse, mens nivåene over 0 inneholder beskrivelser av i hvilken grad, fra enkel til kompleks, den enkelte kompetansen krever anvendelse.

4.1 Fordeling av kodekategorier for eksamen 2019 – 2017

I denne delen har jeg valgt å presentere del 1 og del 2 hver for seg, før jeg til sist ser samlet på hele eksamenssettet. Bakgrunnen for dette er utformingen av eksamen, der størstedelen av oppgavene på del 1 kun krever et svar uten å vise framgangsmåte (tilsvarer kodenivå 0 for kompetansen *constructive communication)* og at elevene på del 1 har begrenset tilgang på hjelpemidler (kun skrivesaker, linjal, passer og vinkelmåler). I tillegg er vektleggingen av hva som testes litt ulik på del 1 og 2. Ifølge Utdanningsdirektoratet (2018a; 2019a) vektlegger del 1 begreps- og tallforståelse, regneferdigheter, evne til problemløsning og resonnement, mens man på del 2 vektlegger begreps- og tallforståelse, digital kompetanse og evne til problemløsning og resonnement.

Jeg begynner gjennomgangen med «ferskeste vare» (eksamen 2019), og beveger meg så suksessivt bakover til eksamenssettet for 2017, før jeg til sist tar et overblikk over hele treårsperioden.

4.1.1 Resultater eksamen 2019, del 1

Del 1 for eksamen 2019 inneholder 4 oppgaver med regnerute der elevene skal vise framgangsmåte. De resterende 26 oppgavene krever bare et enkelt svar (kortsvarsoppgaver) eller et kryss for riktig svaralternativ (flersvarsoppgaver).

Del 1 for eksamen 2019 inneholder totalt 30 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 30 x 7 x 4 = 840 koderesultater.

Eksamen 2019, del 1	**0**	**1**	**2**	**3**
Receptive communication	82 %	18 %	0 %	0 %
Constructive communication	87 %	6 %	8 %	0 %
Devising strategies	43 %	53 %	3 %	0 %
Mathematising	62 %	38 %	0 %	0 %
Representation	100 %	0 %	0 %	0 %
Using symbols, operations, formal language	18 %	63 %	20 %	0 %
Reasoning and argument	86 %	13 %	1 %	0 %
Samlet nivåfordeling	68 %	27 %	5 %	0 %

Tabell 4: Klassifisering av oppgaver på eksamen 2019, del 1.

Samlet sett ser vi at kode over nivå 0 er benyttet i 32 % av tilfellene, med ingen klassifiseringer på nivå 3, der nivå 0 beskriver fravær eller lave krav til anvendelse av en kompetanse, mens nivå 3 kjennetegnes av avansert og komplekse krav til anvendelse av kompetansen.

Når vi går inn på enkeltkategorier ser vi at for *receptive* og *constructive communication*, *representation* og *reasoning and argument* er nivå 0 sterkt dominerende. Dette betyr at oppgavene kodes på laveste nivå for de matematiske kompetansene som handler om å kunne lese og forstå oppgaven, kunne uttrykke et svar, tolke, bruke, velge og oversette mellom ulike representasjoner og kunne å trekke logiske slutninger forankret i matematiske tankeprosesser.

For kategorien *using symbols, operations and formal language* ligger imidlertid 83 % av kodingene på nivå 1 eller 2, hvilket betyr at det er aktivisering av denne kompetansen, som handler om å kunne huske, bruke og forstå matematiske begreper, fakta, regler og prosedyrer, som oppnår høyest score i dette datamaterialet. I den andre enden finner vi *representation*, der alle oppgavene er kodet til nivå 0. Et nivå som kjennetegnes av fravær av denne kompetansen, eller at man bare trenger å lese/plotte verdier ut fra en enkel representasjon eller lese tallverdier direkte fra oppgaveteksten for å løse oppgaven.

D*evising strategies* blir kodet til nivå 1 i over halvparten av tilfellene, mens *mathematising* har 38 % på nivå 1. Nivåbeskrivelsen tilsier da at elevene ofte legge en, om dog enkel, plan for å løse oppgaven, og at de i flere tilfeller, ut fra forutsetningene i oppgaven, må lage en matematisk modell eller bruke en oppgitt modell, for å løse oppgaven.

4.1.2 Resultater eksamen 2019, del 2

Del 2 for eksamen 2019 inneholder totalt 24 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 24 x 7 x 4 = 672 koderesultater.

Eksamen 2019, del 2	**0**	**1**	**2**	**3**
Receptive communication	26 %	70 %	4 %	0 %
Constructive communication	9 %	64 %	23 %	4 %
Devising strategies	29 %	52 %	19 %	0 %
Mathematising	36 %	64 %	0 %	0 %
Representation	51 %	46 %	3 %	0 %
Using symbols, operations, formal language	15 %	72 %	13 %	1 %
Reasoning and argument	63 %	30 %	7 %	0 %
Samlet nivåfordeling	33 %	57 %	10 %	1 %

Tabell 5: Klassifisering av oppgaver på eksamen 2019, del 2.

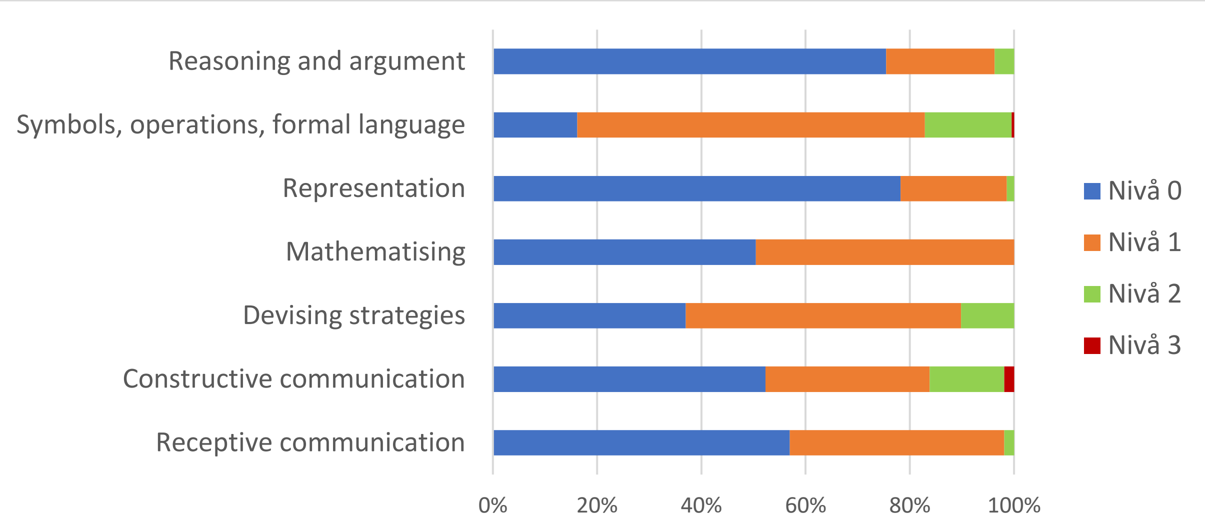
Samlet nivåfordeling for eksamen 2019 del 2 viser at koding over nivå 0 er benyttet i 68% av tilfellene. Med unntak av kategoriene *representation* og *reasoning and argument* ligger majoriteten av koderesultatene på nivå 1. *Representation* er imidlertid på det nærmeste kodet til nivå 0 like ofte som nivå 1, mens det for kategorien *reasoning and argument* sitt vedkommende er nivå 0 som dominerer.

Videre ser vi at det er kompetansene *constructive communication*, *devising strategies og using symbols, operations and formal language* som gjennomgående scorer utover nivå 1, mens det kun er noen få oppgaver som er kodet til nivå 3, og da bare innenfor kategoriene *constructive communication og using symbols, operations and formal language.* Disse to kategoriene skiller seg også ut ved at de har henholdsvis 91 % og 85 % av kodingene over nivå 0.

I forhold til del 1 ser vi nå at samlet nivåfordeling har tyngdepunkt på nivå 1, og at alle kompetanser har flere scoringer over laveste nivå på del 2 enn på del 1. Spesielt for kompetansene *receptive* og *constructive communication og representation* er forskjellene betydelige, mens vi i andre enden finner kompetansen *using symbols, operations and formal language*, som er på omtrent samme nivå på begge delene (18 % på nivå 0 på del 1 og 15 % på nivå 0 på del 2). Samlet sett indikerer dette at elevene må anvende flere kompetanser på et høyere nivå for å løse oppgavene på del 2 enn på del 1.

4.1.3 Samlet resultat for eksamen 2019

Eksamen 2019 inneholder totalt 54 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 54 x 7 x 4 = 1512 koderesultater.



Figur 10: Eksamen 2019: Fordeling på nivå 0 – 3 for de syv kodede kompetansene.

Samlet nivåfordeling på eksamen 2019 viser at koding på nivå 0 er benyttet omtrent like ofte som koding på nivå 1 og 2 til sammen (52 % versus 47 %). Samlet sett er kodenivå 3 benyttet på i underkant av 0,5 % av tilfellene. For alle kompetansene gjelder at koding utover nivå 1 er lite fremtredende.

Går vi inn på enkeltkompetansene ser vi imidlertid av *constructive communication*, *devising strategies og using symbols, operations and formal language* skiller seg ut ved å oppnå mellom 10% og 17 % scoring på nivå 2, der *constructive communication* har størst spennvidde med også en stor andel (52 %) på nivå 0, noe som nok skyldes oppgavene på del 1, som i liten grad krever at eleven viser framgangsmåter.

Av diagrammet ser vi også at både *representation* og *reasoning and argument* skiller seg ut ved å være kodet til nivå 0 i omtrent tre fjerdedeler av tilfellene. *Mathematising* er i sin tur kun kodet til å ligge på nivå 0 eller 1, med like stor andel i begge leire, mens *devising strategies* er kodet til over laveste nivå i 63 % av tilfellene, der oppgavene på del 2 trekker gjennomsnittet opp. For *receptive communication* er 57 % av oppgavene kodet til laveste nivå, mens det for kompetansen *using symbols, operations and formal language* kun er 16 % på nivå 0, med det til følge at denne kompetansen er den med klart høyest nivåscore totalt sett for hele 2019-settet.

4.1.4 Resultater eksamen 2018, del 1

Del 1 for eksamen 2018 inneholder seks oppgaver med regnerute, der elevene skal vise framgangsmåte, og en oppgave der de skal tegne en lineær graf. De resterende 23 oppgavene krever bare et enkelt svar (kortsvarsoppgave) eller et kryss for riktig svaralternativ (flersvarsoppgave).

Del 1 for eksamen 2018 inneholder totalt 30 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 30 x 7 x 4 = 840 koderesultater.

Eksamen 2018, del 1	**0**	**1**	**2**	**3**
Receptive communication	57 %	43 %	0 %	0 %
Constructive communication	77 %	12 %	12 %	0 %
Devising strategies	43 %	53 %	3 %	0 %
Mathematising	64 %	36 %	0 %	0 %
Representation	87 %	13 %	0 %	0 %
Using symbols, operations, formal language	13 %	75 %	13 %	0 %
Reasoning and argument	87 %	13 %	1 %	0 %
Samlet nivåfordeling	61 %	35 %	4 %	0 %

Tabell 6: Klassifisering av oppgaver på eksamen 2018, del 1.

Samlet for eksamen 2018, del 1, er kode på nivå 1 eller 2 benyttet i 39 % av tilfellene. Kategoriene *representation* og *reasoning and argument* er, som for del 1 på eksamen 2019, alt overveiende kodet til nivå 0. C*onstructive communication* er også kodet til nivå 0 i om lag tre fjerdedeler av tilfellene, men skiller seg ut, sammen med *using symbols, operations and formal language* ved å være de to kategoriene som har flest scoringer på nivå 2. Dette innebærer at elevene i gjennomsnitt for omkring hver åttende oppgavene må skrive en kort begrunnelse, eller vise en utregning som går over flere trinn og at de må bruke flere regler, definisjoner, fakta og regneprosedyrer i samme oppgave. *Using symbols, operations and formal language* er også den kategorien som skiller seg klart ut når det gjelder antall oppgaver kodet til nivå 1 eller 2. Syv av åtte oppgaver er kodet til ett av disse nivåene, og en frekvensanalyse viser at i 20 % av tilfellene er dette den eneste av de sju kategoriene som scorer over nivå 0, noe som indikerer at elevene i disse oppgavene bare testes i regler, fakta, begreper og prosedyrer. Tilsvarende frekvens for 2019 er 10%.

I det 23 av 30 oppgaver på del 1 i dette eksamenssettet ikke krever at elevene viser en framgangsmåte, er det heller ikke overraskende at om lag ¾ av oppgavene kodes til nivå 0 for kategorien *constructive communication*. Selv om flesteparten av oppgavene på del 1 stiller lite krav til å vise selve utregningen, er det imidlertid verdt å legge merke til at kategoriene *devising strategies og mathematising* blir kodet til nivå 1 for henholdsvis over halvparten og over en tredjedel av oppgavene. Dette impliserer at elevene ofte må legge en, om dog enkel, plan for å løse oppgaven, og at de i flere tilfeller, ut fra forutsetningene i oppgaven, må lage en matematisk modell eller bruke en oppgitt modell for å løse oppgaven.

4.1.5 Resultater eksamen 2018, del 2

Del 2 for eksamen 2018 inneholder totalt 23 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 23 x 7 x 4 = 644 koderesultater.

Eksamen 2018, del 2	**0**	**1**	**2**	**3**
Receptive communication	23 %	65 %	11 %	1 %
Constructive communication	9 %	59 %	32 %	1 %
Devising strategies	17 %	54 %	28 %	0 %
Mathematising	49 %	51 %	0 %	0 %
Representation	24 %	71 %	5 %	0 %
Using symbols, operations, formal language	17 %	58 %	24 %	1 %
Reasoning and argument	61 %	23 %	16 %	0 %
Samlet nivåfordeling	29 %	54 %	17 %	0 %

Tabell 7: Klassifisering av oppgaver på eksamen 2018, del 2.

Samlet nivåfordeling for del 2 på eksamen for 2018 viser at kodenivået over 0 er brukt i 71 % av tilfellene. Majoriteten av oppgavene krever derfor aktivisering av flere kompetanser for å løse oppgaven, men som det framgår av tabellen over, er det stort variasjon i kravet til graden av anvendelse for de ulike kompetansene.

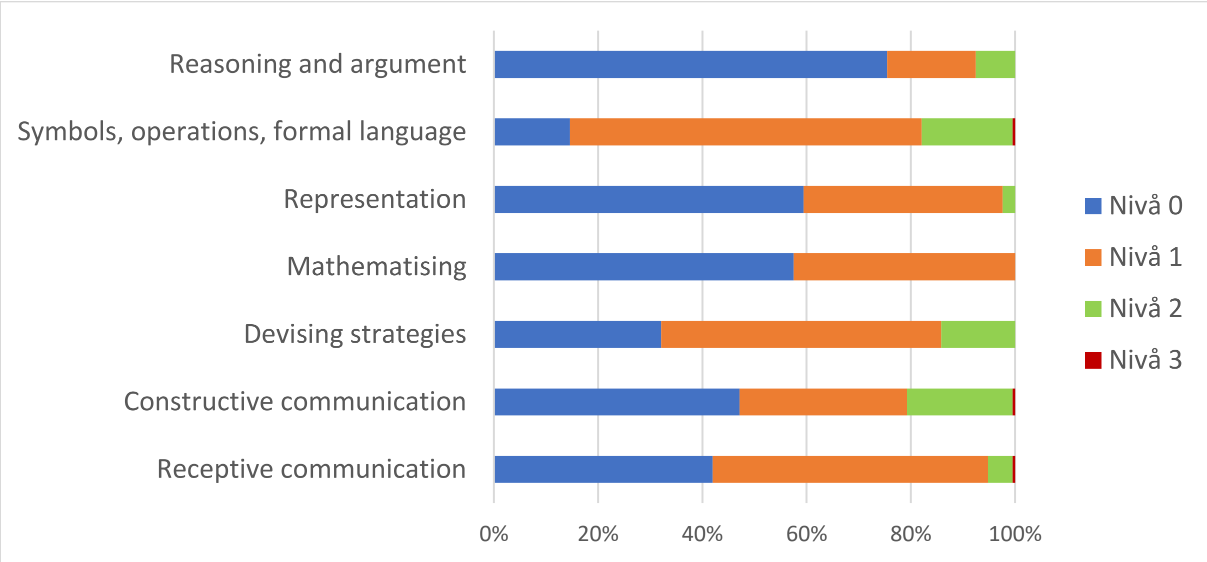
Av tabellen over ser vi at det er nivå 1 som dominerer i de fleste kategoriene. Unntakene er kompetansen *mathematising*, som er kodet til nivå 0 like ofte som nivå 1, og kompetansen *reasoning and argument*, der nivå 0 er dominerende. Kompetansene *constructive communication, devising strategies og using symbols, operations and formal language* utmerker seg imidlertid ved å ha mellom en fjerdedel og en tredjedel av kodingene på nivå 2, noe som tilsier at tyngdepunktet for kategoriseringene her heller mer mot nivå 2 enn nivå 0. En forsvinnende liten andel av kategoriene er kodet til nivå 3: Kun *receptive* og *constructive communication* og *using symbols, operations and formal language* er representert med en prosentandel hver her.

Totalt er nivå 0 brukt i 29 % av kodingene for de syv kompetansene. Variasjonen på tvers av kompetansene er stor, fra kun 9 % for *constructive communication* og til hele 61 % for *reasoning and argument*. I det nivå 0 beskriver fravær eller lav grad av anvendelse for en kompetanse, er det verdt å legge merke til at kategoriene *reasoning and argument* og *mathematising* har en stor andel på laveste nivå.

I forhold til del 1 ser vi nå at samlet nivåfordeling har tyngdepunkt på nivå 1, og at seks av syv kompetanser har flere scoringer over laveste nivå på del 2 enn på del 1. Unntaket er kompetansen *using symbols, operations and formal language* som er på omtrent samme score i begge delene, med 13 % på nivå 0 på del 1 og 17% på nivå 0 på del 2. Spesielt for kompetansene *constructive communication* er spranget stort, fra 77 % på nivå 0 på del 1 til 91 % over nivå 0 på del 2, noe som skyldes kravet til å vise framgangsmåter på del 2. Forskjellene i nivå er også betydelige mellom del 1 og del 2 for kompetansene *receptive* communication, *devising strategies*, *representation* og *reasoning and argument*.

4.1.6 Samlet resultat for eksamen 2018

Eksamen 2018 inneholder totalt 53 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 53 x 7 x 4 = 1484 koderesultater.



Figur 11: Eksamen 2018: Fordeling på nivå 0 – 3 for de syv kodede kompetansene.

Samlet nivåfordeling på eksamen 2018 viser, som for eksamen 2019, at koding over nivå 0 er benyttet omtrent like ofte som koding på nivå 0, men at det for 2018-settet sitt vedkommende, i motsetning til 2019-settet, er koding over nivå 0 som har en liten «ledelse» (53 % versus 47 %). 90 % av koderesultatene ligger på nivå 0 eller 1, og resultatene for enkeltkompetansene er nokså sammenfallende med 2019-settet.

På nivå 2 skiller kompetansene *constructive communication*, *devising strategies og using symbols, operations and formal language* seg ut ved å oppnå mellom 14 – 20 % scoring, der *constructive communication* har størst spennvidde med en stor andel (47 %) også på nivå 0, noe som skyldes oppgavene på del 1, der 23 av 30 oppgavene ikke krever at eleven viser framgangsmåter, og dermed blir kategorisert til nivå 0 for *constructive communication*.

Av tabellen ser vi at *reasoning and argument*, som for 2019-settet, skiller seg ut ved å være kodet til nivå 0 i tre fjerdedeler av tilfellene, mens kompetansen *representation* og *mathematising* har i underkant av 60 % av kodingene på dette nivået, der førstnevnte får nærmere 20 % flere scoringer over laveste nivå enn tilsvarende tall for 2019.

Devising strategies er kodet til over laveste nivå i 68 % av tilfellene, der mange kodinger på nivå 2 i del 2, som for 2019-settet, trekker gjennomsnittet opp. For *receptive communication* ligger 58 % av oppgavene over laveste nivå, noe som er økning i forhold til 2019-settet der tilsvarende prosentandel er 43. Kompetansen *using symbols, operations and formal language* skiller seg igjen ut, som i 2019, ved at hele 84 % av kodene ligger over nivå 0, med det til følge at denne kompetansen er den med klart høyest nivåscore totalt sett for hele 2018-settet.

4.1.7 Resultater eksamen 2017, del 1

Del 1 for eksamen 2017 inneholder ni oppgaver med regnerute, der elevene skal vise framgangsmåte, og en oppgave der de skal tegne perspektivlinjer for å finne forsvinningspunktene til en figur. De resterende 24 oppgavene krever bare et enkelt svar (kortsvarsoppgave) eller et kryss for riktig svaralternativ (flersvarsoppgave).

Del 1 for eksamen 2017 inneholder totalt 34 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 34 x 7 x 4 = 952 koderesultater.

Eksamen 2017, del 1	**0**	**1**	**2**	**3**
Receptive communication	57 %	43 %	0 %	0 %
Constructive communication	71 %	15 %	15 %	0 %
Devising strategies	45 %	49 %	7 %	0 %
Mathematising	69 %	26 %	4 %	1 %
Representation	82 %	16 %	2 %	0 %
Using symbols, operations, formal language	16 %	68 %	16 %	0 %
Reasoning and argument	87 %	10 %	3 %	0 %
Samlet nivåfordeling	61 %	32 %	7 %	0 %

Tabell 8: Klassifisering av oppgaver på eksamen 2017, del 1.

Nivåfordelingen for de syv kompetansene for eksamen 2017, del 1, er til forveksling lik tilsvarende tabell for 2018. Jeg vil derfor ikke gjenta alt som er sagt om denne fordelingen tidligere, men heller se på noen nyanser i resultatene.

Samlet nivåfordeling for kodinger over nivå 0 er den samme (39 %), men med en litt større andel (7 % versus 4 %) på nivå 2 for dette datamaterialet. Gjennomgående i denne delen har imidlertid alle kategoriene, med unntak av *receptive communication*, kodinger på nivå 2, men fortsatt tilnærmet 0 % (kun en koding) på nivå 3. På del 1 for 2018 er kun tre av syv kategorier representert med kodinger over 1 % på nivå 2.

Kategoriene *representation* og *reasoning and argument* er, som for både del 1 på eksamen 2019 og 2018, alt overveiende kodet til nivå 0. I tråd med at andelen oppgaver som bare krever et enkelt svar eller en avkryssing er mindre i del 1 for 2017 enn i del 1 for 2018, ligger også andelen på nivå 0 under kategorien *constructive communication* litt lavere (71 % versus 77 %) her. C*onstructive communication* og *using symbols, operations and formal language* skiller seg fortsatt ut, som i 2018, ved å være de to kategoriene som har flest scoringer (15 %) på nivå 2. *Using symbols, operations and formal language* er også den kategorien som skiller seg klart ut når det gjelder antall oppgaver kodet til nivå 1 eller 2. For denne kompetansen er 84 % av oppgavene kodet til nivå 1 eller 2, og en frekvensanalyse viser at i 17,6 % av tilfellene er dette den eneste av de sju kategoriene som koderne scorer over nivå 0. Ut fra mine resultater viser dette at elevene i disse oppgavene bare testes i regler, fakta, begreper og prosedyrer. Tilsvarende frekvens på del 1 for 2018 er 20 %, og for 2019 10 %.

4.1.8 Resultater eksamen 2017, del 2

Del 2 for eksamen 2017 inneholder totalt 24 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 24 x 7 x 4 = 672 koderesultater.

Eksamen 2017, del 2	**0**	**1**	**2**	**3**
Receptive communication	10 %	77 %	10 %	2 %
Constructive communication	4 %	47 %	49 %	0 %
Devising strategies	18 %	53 %	29 %	0 %
Mathematising	45 %	52 %	3 %	0 %
Representation	46 %	47 %	7 %	0 %
Using symbols, operations, formal language	20 %	51 %	28 %	1 %
Reasoning and argument	71 %	23 %	6 %	0 %
Samlet nivåfordeling	31 %	50 %	19 %	0 %

Tabell 9: Klassifisering av oppgaver på eksamen 2017, del 2.

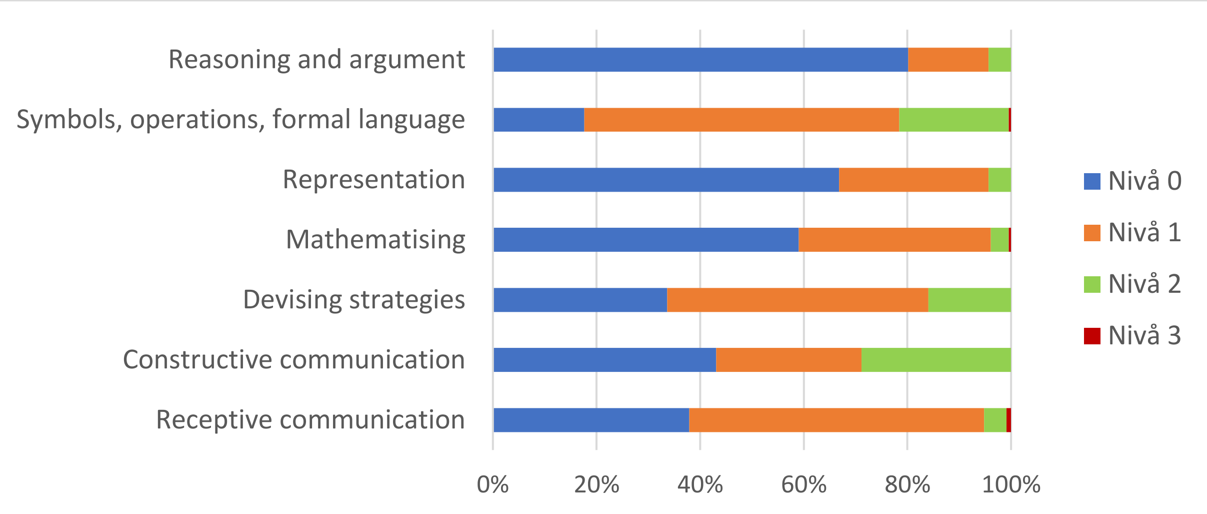
Samlet nivåfordeling for eksamen 2017, del 2, viser at koder over laveste nivå er benyttet i 69 % av tilfellene. Kodenivå 3 er benyttet i totalt 0,45 % av tilfellene og rundes derfor ned til 0 % i tabellen.

Kompetansen *reasoning and argument* peker seg ut ved å ha hele 71 % av kodingene på laveste nivå, mens det for kompetansene m*athematising* og *representation* er slik at de er kodet til nivå 0 omtrent like ofte som nivå 1. *Constructive communication* skiller seg ut ved å ha nesten alle kodinger omtrent likt fordelt mellom nivå 1 eller 2 (til sammen 96 %), mens kompetansene *devising strategies og using symbols, operations and formal language* begge har rundt 80 % over laveste nivå, og en ganske lik fordeling seg imellom på nivå 1 og 2. *Receptive communication* har tyngdepunkt på nivå 1 med om lag tre fjerdedeler av kodingene på dette nivået.

I forhold til del 1 ser vi nå at samlet nivåfordeling har tyngdepunkt på nivå 1, og at seks av syv kompetanser har flere scoringer over laveste nivå på del 2 enn på del 1. Unntaket er kompetansen *using symbols, operations and formal language* som er på omtrent samme nivå på begge delene, med 16 % på nivå 0 på del 1 og 20 % på nivå 0 på del 2. Spesielt for kompetansen *constructive communication* er spranget stort, fra 71 % på nivå 0 på del 1 til 96 % over nivå 0 på del 2, noe som skyldes kravet til å vise framgangsmåter på del 2. Forskjellene i nivå er også spesielt stort mellom del 1 og del 2 for kompetansen *receptive communication* (fra 57 % på nivå 0 på del 1 til 90 % over nivå 0 på del 2), hvilket indikerer at det kreves mer av elevene for å lese, tolke og forstå hva oppgaven handler om i del 2 enn hva som kreves i del 1.

4.1.9 Samlet resultat for eksamen 2017

Eksamen 2017 inneholder totalt 58 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 58 x 7 x 4 = 1624 koderesultater.



Figur 12: Eksamen 2017: Fordeling på nivå 0 – 3 for de syv kodede kompetansene.

Samlet nivåfordeling på eksamen 2017, som for eksamen 2019 og 2018, viser at koding over nivå 0 er benyttet omtrent like ofte som koding på nivå 0, med en liten favorisering for kodenivåene over nivå 0 (52 % versus 48 %). 88 % av koderesultatene ligger på nivå 0 eller 1. Tilsvarende tall for 2018-settet er 90 % og for 2019-settet 92 %.

Et i øyenfallende resultat i tabellen er for kompetansene *reasoning and argument* og *using symbols, operations and formal*, som har omtrent diametralt motsatt scoring: For førstnevnte havner 80 % av koderesultatene på nivå 0, mens for sistnevnte havner 82 % av koderesultatene over nivå 0. Den samme trenden finner vi også for de to andre eksamenssettene.

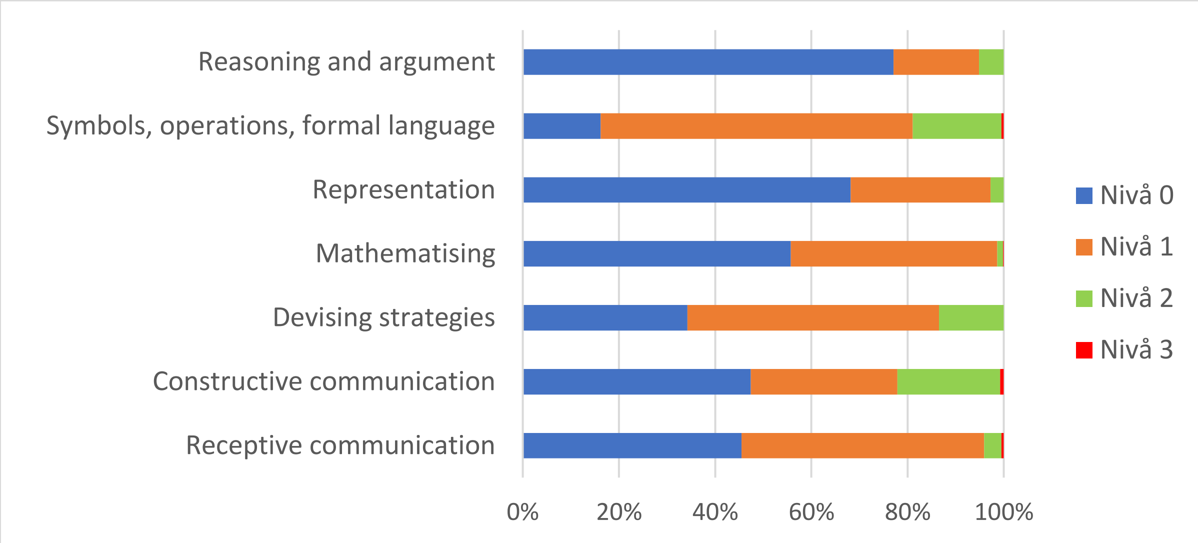
På nivå 2 skiller kompetansene *constructive communication*, *devising strategies* og *using symbols, operations and formal language* igjen seg ut ved å oppnå mellom 16 – 29 % scoring, der *constructive communication* gjør seg bemerket med både en stor andel (43 %) på nivå 0, men også hele 29 % på nivå 2, noe som er flest antall tilfeller på nivå 2 for denne kompetansen i samtlige analyserte eksamenssett.

Kompetansen *representation* og *mathematising* har begge, i dette eksamenssettet, godt over halvparten av kodingene på det laveste nivået, henholdsvis 67 % og 59 %. Denne trenden, med et stort antall oppgaver kategorisert til nivå 0 for disse to kompetansene, finner vi også igjen i de andre oppgavesettene.

Devising strategies er kodet til over laveste nivå i 66 % av tilfellene, der mange kodinger på nivå 2 i del 2, som for både 2018- og 2019-settet, trekker gjennomsnittet opp. For *receptive communication* ligger 62 % av oppgavene over laveste nivå. Detter er på nivå med 2018-settet (58 %), og økning i forhold til 2019-settet (43%).

4.1.10 Samlet resultat for eksamen 2017 – 2019

Eksamen 2017 – 2019 inneholder totalt 165 oppgaver. Oppgavene ble kodet med hensyn på syv kategorier av fire kodere. Totalt 165 x 7 x 4 = 4620 koderesultater.



Figur 13: Eksamen 2017 – 2019: Fordeling på nivå 0 – 3 for de syv kodede kompetansene.

Ut fra den samlede oversikten for årene 2017 – 2019 er koding på nivå 0 og 1 sterkt fremtredende. Samlet sett fordeler 90 % av kodingene for de syv kompetansene seg på disse to nivåene, med henholdsvis 49 % på nivå 0 og 41 % på nivå 1. Koding på nivå 2 er totalt sett benyttet i overkant av 9 % av tilfellene, mens kun 12 av 4620 kodinger har oppnådd nivåscore 3. Disse 12 fordeler seg med henholdsvis fem på *constructive communication*, tre på *receptive communication*, tre på *using symbols, operations and formal language* og en på *mathematising*, der kompetansen *constructive communication* i oppgave 8c på del 2 for 2019 er kodet til nivå 3 av samtlige kodere, mens en koder her også har gitt nivå 3 på *using symbols, operations and formal language*.

Mathematising, *representation* og *reasoning and argument* utmerker seg i stigende rekkefølge ved å ha henholdsvis mellom 56 %, 68 % og 77 % av kodingene på nivå 0, men der *reasoning and argument* er den kompetansen som oftest scorer på nivå 2 av disse tre (5,2 %). For kompetansen *using symbols, operations and formal language* er situasjonen ganske annerledes. Denne kompetansen skiller seg ut ved å ha den klart største andelen av kodingene over nivå 0, majoriteten av kodingene på nivå 1, og en betydelig andel (18,5 %) på nivå 2. Som en «god nummer to» i antall scoringer over nivå 0 kommer *devising strategies* med 66 %. Av disse utgjør koding på nivå 2 13 %, bare overgått av *constructive communication* (21 % på nivå 2) og *using symbols, operations and formal language* (18 % på nivå 2).

Constructive og receptive communication, som i utgangspunktet er kategorien *communication* i MEG-skjemaet, viser i mitt materiale en tilnærmet lik fordeling på nivå 0. Forskjellene blir imidlertid tydeligere på høyere nivå, der *constructive communication* har en mye større andel av sine scoringer på nivå 2 enn *receptive communication*, henholdsvis 21,4 % mot 3,6 %.

Trenden for eksamensoppgavene i den treårsperioden jeg studerte er at samtlige kompetanser som inngår i MEG-skjemaet alt overveiende kodes på nivå 0 og 1, og at koding på nivå 3 kun forekommer unntaksvis. En beregning av gjennomsnittlig vanskegrad for de syv kodede kompetansene (tabell 10) viser også dette, der kompetansen *using symbols, operations and formal language* er den eneste som så vidt bikker 1. I motsatt ende av skalaen finner vi *reasoning and argument* og *representation*, med en gjennomsnittlig nivåscore på henholdsvis 0,28 og 0,35.

Kompetanse	**Gjennomsnittlig nivåscore**
Receptive communication	0,59
Constructive communication	0,75
Devising strategies	0,79
Mathematising	0,46
Representation	0,35
Symbols, operations, formal language	1,03
Reasoning and argument	0,28

Tabell 10: Gjennomsnittlig nivåscore for de syv kodede kompetansene.

Utdypende kommentarer til trender i mitt materiale for eksamensoppgavene i perioden 2017 – 2019 følger i neste kapittel.

5 Diskusjon

Hensikten med denne masteroppgaven har vært å undersøke i hvilken grad eksamensoppgavene i matematikk for 10. trinn tester ulike aspekter ved elevenes matematiske kompetanse. For å kartlegge dette, har jeg brukt et oppgaveanalyseskjema (MEG-skjema) utviklet av PISA, der jeg og tre ungdomsskolelærere individuelt kodet samtlige eksamensoppgaver for årene 2017 – 2019 i matematikk for 10. trinn.

I denne delen av oppgaven diskuterer jeg resultatene fra forrige kapittel opp mot problemstillingen, i tillegg til å diskutere kodesamsvar og erfaringene med bruken av MEG-skjemaet på eksamensoppgaver, før jeg konkluderer.

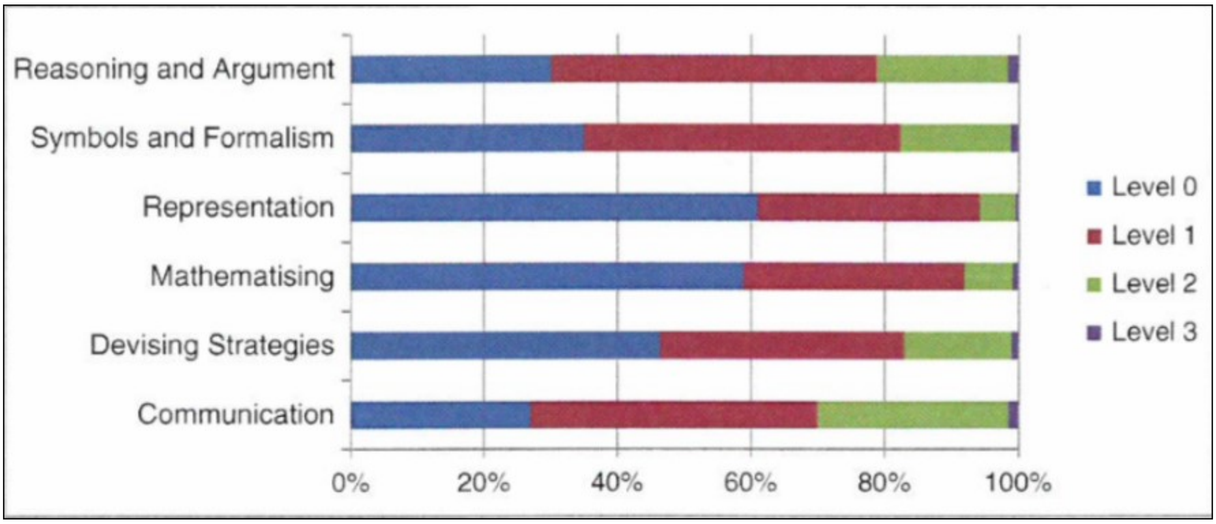
5.1 Kommentarer til koderesultatene

For eksamenssettene for perioden 2017 – 2019 er det, med utgangspunkt i MEG-rammeverket, jevnt over koding av kompetansene på nivå 0 og 1 som dominerer, der antall kodinger på disse to nivåene viser en svakt økende tendens fra 88 % i 2017, via 90 % i 2018, til hele 93 % i 2019. For alle eksamenssettene utgjør koding på nivå 0 om lag halvparten (2273/4620) av tilfellene, noe som innebærer at det for mange av de syv kompetansene jeg har sett på, ofte stilles lave/enkle krav til anvendelse, eller at eleven rett og slett ikke trenger å aktivisere den aktuelle kompetansen for å løse oppgaven. Koding på nivå 3 er kun unntaksvis benyttet (12 av 4620 koder), der kompetansen *constructive communication* står for fem av tilfellene. Ved beregningen av gjennomsnittlig vanskegrad for de syv kodede kompetansene (tabell 10) er det kompetansen *using symbols, operations and formal language* som scorer høyest (1,03), mens kompetansene *reasoning and argument* og *representation* i snitt scorer lavest og ligger på henholdsvis 0,28 og 0,35.

Kompetansen *reasoning and argument*, som i MEG-rammeverket handler om å kunne trekke logiske slutninger og bruke matematiske resonnement for å kople sammen gyldige slutninger på veien mot en løsning, peker seg ut i de samlede resultatene for alle årene ved å ha minst 75 % på nivå 0. Dette betyr ikke at anvendelse av denne kompetansen er totalt fraværende i flesteparten av eksamensoppgavene, men derimot at elevene, når de løser oppgaven, kan trekke slutninger direkte ut fra informasjonen/instruksjonene i oppgaven, og ikke trenger andre former for logiske resonnement for å komme opp med et gyldig svar. Gjennomgående øker kravet til aktivisering av denne kompetansen for oppgavene på del 2, der 37 % av oppgavene i 2019-settet og 49 % i 2018-settet blir klassifisert til nivå 1 eller 2. Samlet sett sier mine resultater at oppgavene på eksamen hovedsakelig legger opp til å teste elevene i kompetansen *reasoning and argument* på laveste nivå, og at elevene i liten grad må resonnere og trekke slutninger utover det som kreves i en problemløsningsprosess.

I rammeverkene til KOM-prosjektet (Niss & Jensen, 2002), TIMSS (Mullis & Martin, 2013) og Kilpatrick et al. (2001) finner vi igjen beskrivelse av tilsvarende kompetanse som henholdsvis *resonnerings- og tankegangskompetanse, reasoning* og *adaptive reasoning*. Alle rammeverkene anser dette som en (av flere) sentrale kompetanser i matematikk, og i eksamensveiledningen for 2019 (Utdanningsdirektoratet, 2019a) er evnen til å kunne gjennomføre logisk resonnement også en del av vurderingsgrunnlaget. Det er derfor overraskende, ut fra mine data, at oppgavene på eksamen i så liten grad utfordrer elevene innenfor dette kompetanse-områdetette. Dette kan imidlertid like fullt skyldes at innholdet i kompetansen *reasoning and argument* ikke er den samme i MEG-rammeverket som i eksamensveiledningen.

Sammenlignet med Pettersen og Nortvedt (2018) sine analyser av eksamenssettet for 2014 og PISA 2012-testen (figur 14), er mine resultater for nivåscore på kompetansen *reasoning and argument* litt overraskende. Analysen av deres materiale viser at bare om lag 30 % av kodingene for *reasoning and argument* er på nivå 0. Dette kan selvfølgelig skyldes at disse oppgavene samlet sett var mer krevende med tanke på denne kompetansen enn for mine tre oppgavesett. Det kan imidlertid også indikere at vi har lagt en annen nivå-tolkning til grunn for *reasoning and argument* enn det koderne til Pettersen og Nortvedt gjorde. I vår innledende runde, da vi kodet 2014-settet, kodet vi ofte kompetansen *reasoning and argument* til nivå 1. Etter å ha finlest beskrivelsen og kode-eksemplene i Turner et al. (2015), forsto vi det imidlertid slik at denne kompetansen dreier seg om evnen til å trekke valide slutninger for å rettferdiggjøre et svar, snarere enn generelle tankeprosesser og resonnement, som inngår i de andre kompetansene, som f. eks. *devising strategies* og *mathematising*. Når den enkelte koder anså at det trengtes resonnement **utover** hva som var innbakt i de andre kompetansene, ble *reasoning and argument* kodet til over nivå 0. Turner et al. (2015) nevner for øvrig at nivå-beskrivelsene for denne kompetansen har gjennomgått betydelige endringer siden førsteutkastet.



Figur 14: Samlet fordeling for 56 oppgaver fra eksamen 2014 og 85 oppgaver fra PISA 2012 på nivå 0 – 3 for de seks kompetansene i MEG-skjemaet (Pettersen & Nortvedt, 2018, s. 961).

Kompetansene *representation*, som handler om å tolke, lage, bruke, velge og oversette mellom ulike matematiske representasjoner, har for eksamenssettene fra årene 2017 – 2019 mellom 59 % - 78 % på laveste nivå, der del 1 for 2019 skiller seg ut med 100 % på nivå 0. At en så stor andel av oppgavene kodes til laveste nivå innebærer at elevene, for å løse oppgaven, sjelden må dekode, tenke ut eller manipulere matematiske representasjoner. I beste fall må de lese eller plotte tallverdier fra f.eks. et koordinatsystem, en tabell eller et enkelt diagram. I Kunnskapsløftet er representasjonskompetanse nevnt både under hovedområdene *tall og algebra* og *funksjoner* i kompetansemål etter 10. trinn (Utdanningsdirektoratet, 2013). Eksamensveiledningene nevner også at elevene skal ha kunnskap om/kunne behandle ulike representasjoner under kompetansen *begreper, forståelse og ferdigheter*

(Utdanningsdirektoratet, 2019a). Beskrivelsene på nivå 0 og 1 for kompetansen *representation* i MEG-skjemaet bruker ord som enkel (simpel) og standard representation for disse to nivåene, mens det først er på nivå 2 at elevene må kunne oversette mellom, og kunne bruke ulike enkle representasjoner. Denne beskrivelsen for nivå 2 i MEG-skjemaet kan, slik jeg ser det, passe godt med beskrivelsene for representasjonskompetanse slik den er beskrevet i kompetansemål etter 10. trinn i Kunnskapsløftet. Jeg finner det derfor overraskende at denne kompetansen, i mine resultater for årene 2017 – 2019, i alt overveiende grad testes på laveste nivå, og ytterst sjelden på nivå 2.

For kompetansen *mathematising* er mellom 50 % - 60 % av oppgavene i mitt datamateriale for årene 2017 – 2019 kodet til på laveste nivå, hvilket innebærer at elevene, når de skal løse oppgavene på eksamen, i liten grad trenger å matematisere (bruke modellering for å oversette fra en reell situasjon til matematikkens verden, eller vise versa). Høyest krav til anvendelse for denne kompetansen finner vi nok en gang på del 2, og andelen oppgaver over nivå 0 har økt fra 40 % i 2017 til 50 % i 2019.

Modellering er eksplisitt nevnt under formålet med faget i *Læreplan i matematikk fellesfag* (Utdanningsdirektoratet, 2013), og testes ifølge mine resultater for *mathematising* nærmest utelukkende på nivå 0 og 1. Dette innebærer at elevene på eksamen så godt som aldri blir utfordret til å lage modeller der de selv må definere variabler, sammenhenger og begrensninger, og validere/evaluere modeller eller sammenligne ulike modeller som beskriver virkelighetsnære situasjoner. Mot dette kan man hevde at slike krav kanskje vil være for tidkrevende for en eksamen på fem timer. På den annen side, skal man ta kompetansebegrepet i Kunnskapsløftet og Fagfornyelsen på alvor, er det kanskje nettopp oppgaver over nivå 1 for både *mathematising* og *devising strategies* i MEG-skjemaet som utfordrer elevene til å «mestre komplekse utfordringer» og erverve «forståelse og evne til refleksjon og kritisk tenkning».

Devising strategies dreier seg om det å velge/bruke/lage en plan for å løse oppgaven, samt det å overvåke sin egen problemløsningsprosess. Andelen oppgaver på nivå 1 eller 2 for anvendelse av denne kompetansen ligger for alle årene mellom 63 % – 68 %, men viser en litt fallende tendens for klassifiseringer på nivå 2, med 16 % i 2017, 14 % i 2018 til 10 % i 2019. Ifølge beskrivelsen for nivå 1 innebærer dette for vårt datamateriale at elevene ofte må lage en enkel strategi (*straight-forward strategy*) for å løse oppgaven, og at denne løsningsstrategien i noen tilfeller også må bestå av flere trinn for at elevene skal mestre oppgaven (nivå 2).

Problemløsning er den andre kompetansen, ved siden av modellering, som nevnes eksplisitt i «formålsparagrafen» i *Læreplan i matematikk fellesfag* (Utdanningsdirektoratet, 2013), og er ifølge eksamensveiledningen den mest sentrale kategorien for sensors vurderingsgrunnlag. Vurderingskategorien problemløsning inneholder for øvrig også kompetansen modellering (Utdanningsdirektoratet, 2019a). I mine resultater for *devising strategies* er det derfor positivt å se at denne kompetansen er en av tre (ved siden av *constructive communication* og *using symbols, operations and formal language*) som jevnt over har hovedvekten av nivåscore på over 0, og for gjennomsnittlig nivåscore ligger nest øverst med 0,79 (tabell 10).

Kompetansen som handler om å lese, tolke og forstå hva oppgaven handler om (*receptive communication*), viser i mitt materiale et fallende nivå fra 2017 til 2019. I 2017 lå 62 % av oppgavene over laveste nivå for denne kompetansen, mens tilsvarende tall for 2019 er 43 %. Ut fra det kodede materialet ser vi at det særlig er oppgavene på del 1 som er «rett fram» og enklest å lese og forstå, mens oppgavene på del 2 kjennetegnes av at elevene, for å løse oppgaven, i større grad må koble og knytte sammen den informasjon som er oppgitt i oppgaven. På del 2 for 2017 og 2018 er rundt 10 % av oppgavene klassifisert til nivå 2 og mellom 1 – 2 % til nivå 3, mens tilsvarende tall for 2019 er 4 % og 0 %.

Constructive communication dreier seg om hva som kreves av kommunikasjon for å formidle utregninger og svar, der nivå 0 refererer til oppgaver som bare krever et tall, et enkeltord eller et kryss i rett rubrikk som svar. Andelen oppgaver på del 1 som krever at elevene viser utregning har falt jevnt og trutt siden 2017, så ikke overraskende blir 82 % av oppgavene på del 1 for 2019 kodet til nivå 0, mens tilsvarende tall for 2017 er 71 %. På del 2 er imidlertid kravene til *constructive communication* betraktelig hevet for alle oppgavesettene, med over 90 % av kodingene over laveste nivå. 2017 del 2 har høyeste andel på nivå 2 med 49 %, 2018 har 32 % mens 2019 har 23 % på dette nivået. Til gjengjeld har 2019 4 % på nivå 3 for denne kompetansen. *Constructive communication* på nivå 2 innebærer at oppgaven krever en kort beskrivelse eller forklaring, eller at det må presenteres en utregningssekvens, mens oppgaver på nivå 3 i enda større grad krever at man må presentere en framgangsmåte som knytter sammen flere deler av problemet eller løsningen. For årene 2017 – 2019 kommer denne kompetansen ut med en gjennomsnittlig vanskegrad på 0,75 (tabell 10). Dette er på linje med *devising strategies* (0,79), og gir en tredjeplass med tanke på gjennomsnittlig vanskegrad.

Beskrivelsen av kompetansen *kommunikasjon* i eksamensveiledningen (Utdanningsdirektoratet, 2019a) passer godt overens med beskrivelsen av kompetansen *communication* i MEG-skjemaet (Turner et al., 2015), og alle eksamenssettene i mitt materiale viser tydelig økte krav til anvendelse av både *receptive* og *constructive communication* fra del 1 til del 2. Ved å separere *communication* i to kategorier fikk jeg også fram at selv om oppgavene kan være enkle å lese og forstå (*receptive communication* på nivå 0 eller 1), så kan det å formidle framgangsmåten kreve en utregningssekvens på nivå 2. Et eksempel på dette finner vi i på del 2 i 2019-settet, der kun 4 % av oppgavene er kodet til nivå 2 og 0 % til nivå 3 for *receptive communication*, mens tilsvarende tall for *constructive communication* er 23 % på nivå 2 og 4 % på nivå 3.

Ut fra erfaringene fra prøvekodingen av 2014-settet, der del 1 hadde mange oppgaver som ikke krevde visning av framgangsmåte, og vi på del 2 opplevde at oppgavene av og til kunne være lette å forstå, men kunne kreve omfattende kommunikasjon for å vise framgangsmåten, delte vi denne kompetansen i de to kategoriene *receptive* og *constructive communication*. MEG-skjemaet opererer med nivåbeskrivelser for begge disse sidene av *communication*, men har ikke delt den inn i separate kategorier slik som jeg valgte.

Using symbols, operations and formal language handler om å kunne huske, forstå og bruke matematiske begreper og symbolspråk, fakta, formler, regler og prosedyrer. Denne kompetansen skiller seg klart ut i datamaterialet ved å være den kompetansen som for alle oppgavesett, det være seg del 1 eller del 2, alltid har minst 80 % av kodingene over nivå 0. På del 1 for 2019 er hele 20 % av oppgavene kodet til nivå 2 for denne kompetansen, mens alle de andre kompetansene (med unntak av *devising strategies*) har majoriteten av kodingene på laveste nivå. En frekvensanalyse på del 1 for kategorien *Using symbols, operations and formal language,* der denne kategorien er den eneste av de sju kompetansene som alle koderne scorer til over nivå 0, viser i mitt materiale en fallende tendens fra 20 % i 2018 til 10 % i 2019. Dette kan indikerer at oppgaver som bare tester elevene i regler, fakta, begreper og prosedyrer kanskje har sett sine beste dager. Gjennomgående for alle oppgavesettene ligger mellom 61 – 67 % av kodene for *using symbols, operations and formal language,* på nivå 1, hvilket tilsier at man for å løse oppgaver på dette nivået må anvende aritmetiske beregninger med brøk eller desimaler, eller bruke en matematisk definisjon, fakta eller konvensjon, eller gjøre direkte bruk av enkle matematiske forhold som involverer variabler.

I eksamensveiledningen for 2019 (Utdanningsdirektoratet, 2019a) er tilsvarende kompetansen *using symbols, operations and formal language* beskrevet under *begreper, forståelse og ferdigheter*. I trådmodellen til Kilpatrick et al. (2001) innbefatter conceptual *understanding* og *procedural fluency* sider ved denne kompetansen, mens kompetanseblomsten til Niss og Jensen (2002), som MEG-skjemaet bygger på, opererer med *symbol- og formalismekompetanse*amme. Rammeverket for matematikk i PISA-testene bruker begrepet *using symbolic, formal and technical language and operations* (OECD, 2017). Alle rammeverkene ser med andre ord denne kompetansen som en del av det å mestre matematikk, men påpeker samtidig at en helhetlig matematisk kompetanse består av en pluralisme av kompetanser.

I mine resultater for årene 2017 – 2019 er *using symbols, operations and formal language* den kompetansen som har færrest kodinger på nivå 0 og majoriteten av kodene på nivå 1. I tillegg er 18,5 % av oppgavene kodet til nivå 2 for denne kompetansen. *Using symbols, operations and formal language* er med andre ord den kompetansen i mitt materiale som gjennomgående for alle tre årene i størst grad blir testet på ulikt nivå. Jeg tenker at skal elevene få vist en helhetlig matematisk kompetanse på eksamen i matematikk, i tråd med den intenderte læreplanen, burde så også være tilfelle for alle de syv kompetansene jeg har kartlagt

5.2 Kommentarer til kodesamsvar

Som beskrevet i metodedelen brukte jeg, for å kunne si noe om kodesamsvaret i gruppa, IBM SPSS Statistics, versjon 25, til å beregne *the intraclass correlation coefficient* (ICC). Tabellen under viser resultatene fra denne analysen i form av gjennomsnittsmål for gruppa (*average-measure*) og enkeltmål (*single-measure*) for hver av de syv kompetansene vi kodet for hver oppgave. *Average-measure* måler i hvilken grad vi som gruppe leverer pålitelige samsvarende kodinger, mens *single-measure* måler troverdigheten av resultatene hvis bare en av medlemmene i kodegruppa skulle stått som representant for all kodingen.

	Average measures	Single measures
Receptive communication	0,88	0,66
Constructive communication	0,96	0,85
Devising strategies	0,90	0,70
Mathematising	0,82	0,54
Representation	0,90	0,70
Symbols, operations, formal language	0,88	0,64
Reasoning and argument	0,88	0,64

Tabell 11: ICC-mål for kodesamsvar, eksamen 2017 – 2019, fordelt på kompetanse.

Med utgangspunkt i retningslinjene for samsvar til Landis og Koch (1977), tabell 3, ligger alle gjennomsnittsmålene for samsvar (*average-measures*) på *almost perfect agreement* (over 0,80). Dette viser at vi som gruppe var godt samkjørte i vår vurdering av nivået til de ulike kompetansene i MEG-skjemaet. Særlig når det gjelder kompetansen *constructive communication* var kodingen samkjørt, hvilket jeg ikke finner særlig overraskende, da en stor del av oppgavene på del 1 kun krever et svar uten begrunnelse, og derfor havner på nivå 0.

Kompetansen *mathematising* befinner seg med en ICC-verdi på 0,82 derimot helt i nedre grense for *almost perfect agreement*, så her var tydeligvis kode-panelet ikke så samstemt. Ut fra oversikten over koderesultatene for denne kompetansen i perioden 2017 – 2019 ser vi at kode 0 eller 1 er benyttet i nesten samtlige tilfeller. MEG-skjemaet beskriver nivå 0 for *mathematising* å gjelde oppgaver som tar for seg rene matematiske problemer (*the situation is purely intra-mathematical)* eller oppgaver, der koblingen mellom den reelle situasjonen som beskrives og modellen som trengs for å løse oppgaven ikke er relevant. Under prøvekodingen kom det fram ulike forståelser av dette nivået. Noen tolket reelle situasjoner beskrevet i oppgaver som irrelevant for å løse oppgavene: Her jobbet man med et reint matematisk problem. Andre mente derimot at eleven her trengte å bruke den reelle situasjonen for å matematisere problemet, og derfor satte nivået til 1. Resultatene fra kodesamsvaret i tabellen kan tyde på at ulikheter i tolkningen av *mathematising* også har vedvart inn i kodingen i mitt materiale.

ICC-målene for *single-measure agreement* viser, i tråd med Turner et al. (2015) sine anbefalinger om at flere bør kode samme oppgave, at vi som enkeltindivider ikke er like pålitelige som en gruppe. Hvis en av oss skulle stå som representant for kodingen av disse tre eksamenssettene, ville troverdigheten til kodingen av *mathematising* være *moderate*, *constructive communication* fortsatt være *almost perfect agreement*, mens troverdigheten av kodingene for de andre kompetansene ville være *substantial Agreement*. Dette tilsier at reliabiliteten i mitt materiale øker når oppgavene blir kodet av en gruppe framfor en enkelt person i denne gruppa.

Sammenlignet med Pettersen og Nortvedt (2018, s. 958) sine ICC-verdier for eksamen 2014, viser mitt kodesamsvar temmelig like nivåer for alle kompetansene på *average-measure*, med unntak av *representation* og *using symbols, operations and formal language*, der jeg har en høyere grad av kodesamsvar, henholdsvis 0,90 mot deres 0,79 og 0,88 mot deres 0,82. Det samme bildet tegner seg for kodesamsvaret for *single-measure*, men mine resultater viser sterkere reliabilitet enn Pettersen og Nortvedt for de to ovennevnte kompetansene. Her er Pettersen og Nortvedt nede på 0,43 og 0,47 for henholdsvis *representation* og *using symbols, operations and formal language*, mens tilsvarende tall for meg er 0,70 og 0,64. Dette tolker jeg som at hvis en av koderne i min gruppe skulle ha kodet hele eksamenssettet alene, så ville resultatene herfra ha vært mer troverdige enn om en i gruppa fra Pettersen & Nortvedt skulle ha gjort det samme. En mulig årsak til høyere kodesamsvar hos meg kan også ligge i utvalget av personene som kodet oppgavene: Min kodegruppe bestod av ungdomsskolelærere med minst 17 års fartstid fra ungdomsskolen, mens gruppa til Pettersen og Nortvedt hovedsakelig bestod av personer med liten praksis og erfaring fra undervisning og vurdering på ungdomsskolenivå.

5.3 Evaluering av MEG-skjemaet

MEG-skjemaet er influert av kompetansebeskrivelsene i KOM-prosjektet og er utarbeidet på bakgrunn av PISA-rammeverket i matematikk (Niss, 2015). Målet er å oppnå en tydelig operasjonalisering og nivådeling for de seks kompetansene *communication*, *devising strategies*, *mathematising*, *representation, using symbols, operations and formal language* og *reasoning and argument*, slik at MEG-skjemaet kan brukes som et verktøy til å forutsi vanskelighetsgraden på matematikkoppgavene i PISA-testen (Turner et al., 2015)

Utviklingen av MEG-rammeverket er ifølge Turner et al. (2015) en dynamisk prosess, og er i utgangspunktet utviklet for et annet test-regime enn norske eksamensoppgaver i matematikk på 10. trinn. MEG-skjemaet er imidlertid utviklet for tilsvarende aldersgruppe, og prøvd ut på 2014 – eksamenen i matematikk for 10. trinn av Pettersen og Nortvedt (2018). Turner et al. (2013) er også selv nysgjerrig på om dette oppgaveanalyseskjemaet kan ha et bredere anvendelsesområde enn bare PISA-tester i matematikk, og ser det som en interessant utvikling hvis MEG-skjemaet kan prøves ut også på andre arenaer. Det kan derfor være av interesse å bringe til torgs noen av mine erfaringer med bruken MEG-rammeverket.

Fokuset i min studie har vært, med utgangspunkt i MEG-skjemaet, å undersøke i hvilken grad eksamensoppgavene i matematikk for 10. trinn tester ulike aspekter ved elevenes matematiske kompetanse. Jeg har med andre ord vært ute etter å kategorisere eksamensoppgaver, framfor å si noe om faktisk vanskelighetsgrad, som er MEG-skjemaets opprinnelige funksjon. På bakgrunn av mine resultater og diskusjon, vil jeg si at MEG-skjemaet også har vært hensiktsmessig til mitt bruk. Jeg har f.eks. fått fram at kompetansene som er representert i MEG-skjemaet hovedsakelig testes på de to laveste nivåene på eksamen, at regel- og prosedyrekunnskap fortsatt står sterkt på eksamen og er den kompetansen der det stilles høyest krav, og at representasjonskompetansen gjennomgående testes på laveste nivå.

Diagrammene som viser nivå-fordelingen for hver enkelt kompetanse for årene 2017 – 2019 er også påfallende like. Totalt sett viser de en betydelig andel (2273 av 4620) kodinger på nivå 0, noe som indikerer at det på eksamen, ifølge vår bruk av MEG-skjemaet, ofte stilles lave krav til elevene når det gjelder anvendelse av kompetansene som inngår i MEG-skjemaet. I tillegg er det tankevekkende at vi kun unntaksvis har gitt nivåscore 3. Rammeverket er i utgangspunktet utviklet for tilsvarende aldersgruppe som norske 10. klassinger, med referanse til samme KOM-rammeverk, så jeg undres litt her.

En forklaring kan være at PISA-oppgavene faktisk har en bredere nivåfordeling enn norsk eksamen for 10. trinn. Et moment som understøtter dette, er Pettersen og Nortvedt (2018) sine resultater, som viser at spriket mellom koderne blir større desto mer krevende oppgavene er. De har i sitt materiale størst nivåsprik for PISA-oppgavene, og mindre for de norske eksamensoppgavene. Dette kan indikere at PISA-oppgavene faktisk er mer krevende enn eksamensoppgavene for 10. trinn.

Et annet moment kan være at vi rett og slett, som koderne til Pettersen og Nortvedt (2018), ikke klarte å bruke MEG-skjemaet etter intensjonene for oppgaver som lå på høyeste kompetansenivå. For kompetansen *constructive communication* blir f.eks. nivå 2 beskrevet som «Any constructive communication involves providing a brief description or explanation, or presenting a sequence of calculation steps, mens beskrivelsen for nivå 3 lyder «Any constructive communication would involve presenting argumentation that links multiple elements of the problem or solution» (Turner et al., 2015, s. 111). Det er her ikke opplagt hva som skiller «a sequence of calculation steps» fra «argumentation that links multiple elements of the solution», eller hvor omfattende en forklaring må være før den faller inn under nivå 3. Det er likevel påfallende at vi som gruppe var så konsistent i vår koding, noe som tyder på at vi som gruppe i hvert fall har tolket rammeverket noenlunde likt. Her tenker jeg at to faktorer har vært essensielle, henholdsvis det å sette seg godt inn i rammeverket og det å foreta prøvekodinger med påfølgende diskusjoner av resultatene.

MEG-rammeverket er langt ifra et «plug and play»-verktøy, og som Turner et al. (2015) anbefaler, bør man sette sammen en gruppe av kodere for å få de beste resultatene. I sin studie for å undersøke hvor egnet MEG-skjemaet er til å forutsi vanskelighetsgraden av oppgaver på PISA-testene i matematikk bruker Turner et al. (2013) åtte kodere. Min kodegruppe ble bare halvparten så stor, men til gjengjeld var jeg så heldig å få alliert meg med tre ungdomsskolelærere med mellom 17 og 28 års fartstid fra ungdomsskolen. Alle hadde utdanning på hovedfagsnivå, og to av tre hadde erfaring som sensor på skriftlig eksamen i matematikk på 10. trinn. Diskusjonene vi hadde omkring kompetansebeskrivelsene og nivåfordeling var til tider heftige, men resultatet må sies å være at vi som gruppe oppnådde en vesentlig omforent enighet, noe som vises i resultatene fra *the intraclass correlation coefficient*. Dette impliserer imidlertid ikke at vår omforente enighet er «rett forståelse» av MEG-skjemaet, men innenfor vår gruppes referanseramme var vi i hvert fall godt synkronisert.

Utfordringer med MEG-skjemaet var f.eks. ord som *simple* og *complex* i beskrivelsen av nivå: Hvor enkelt er *simple* og hvor krevende er *complex*? Her var *Level Descriptions* i kodeeksemplene til Turner et al. (2015) til god hjelp, men vi måtte dog, ut fra egne kodinger på 2014-settet, bestemme oss for «hvor landet lå». Fafo-rapporten har kanskje en enklere kriteriebasert kategorisering med tanke på om oppgavene krever algoritmisk løsning eller kreativ løsning, men jeg tenker at denne kategoriseringen ikke fanger opp kompetanse-variansen i oppgavematerialet på eksamen i samme grad som MEG-skjemaet.

Hjelpemiddelkompetanse er ikke med i MEG-skjemaet. I eksamensveiledningens *Kjennetegn på måloppnåelse* er denne kompetansen beskrevet under *problemløsning* (Utdanningsdirektorater, 2019a). I og med at alle ikke-kommuniserende hjelpemidler er tillatt på del 2 av eksamen, og flere av oppgavene her forutsetter bruk av digitale hjelpemidler, vil det nok for norske forhold være hensiktsmessig å inkorporere denne kompetansen i MEG-skjemaet. Jeg tenker da ikke primært som en indikator på vanskelighetsgrad, slik som Turner et al. (2013) var ute etter, men snarere slik jeg har brukt MEG-skjemaet i denne studien: For å kartlegge i hvilken grad elever må anvende/aktivisere en bestemt kompetanse for å løse en oppgave, for på den måten å sikre at eksamen tester et bredt spekter av ulike matematiske kompetanser. At elevene kunne bruke digitale hjelpemidler på del 2 gjorde også at vi stilte strengere krav til nivåene for *using symbols, operations and formal language*, slik at f.eks. det å bruke aritmetisk kalkulasjon med brøk og desimaler ikke automatisk havner på nivå 1, i motsetning til hva MEG-skjemaet beskriver.

I min bruk av MEG-skjemaet valgte jeg, etter innledende prøvekoding, å dele kategorien communication i to deler, henholdsvis *receptive communication* og *constructive communication,* for på den måten å få tydeligere fram sentrale karakteristikker ved mitt materiale. Det kom da tydelig fram at del 1 på eksamen stiller lave krav til *constructive* *communication*, men også at langt flere oppgaver på del 2 stiller høyere krav til *constructive communication* enn *receptive communication*. Oppgaver kunne med andre ord være ganske greie å forstå (nivå 1), men kreve at elevene viste en framgangsmåte der de måtte beskrive/forklare eller vise en sekvens av utregninger (nivå 2). Denne forskjellen tror jeg ikke ville kommet tydelig fram hvis vi her isteden skulle kodet en «middelverdi» for kategorien *communication*, uten å skjele til denne kategoriens tosidighet. MEG-skjemaet opererer forøvrig med nivåbeskrivelser for begge de ovennevnte sidene av *communication*, men uten å dele inn i to separate kategorier slik jeg valgte.

5.4 Konklusjon

Ut fra kompetanse- og nivåkategoriseringen utviklet av MEG-studien, viser mine studier at elevene på skriftlig eksamen i matematikk for 10. trinn i perioden 2017 – 2019 hovedsakelig blir testet på de to laveste nivåene for alle kompetansene som inngår i MEG-skjemaet. Kompetansene *representation* og *reasoning and argument* testes hovedsakelig på laveste nivå, der nivåbeskrivelsene kjennetegnes av fravær eller svært lave krav til anvendelse av nevnte kompetanse, mens kompetansen *mathematising* alt overveiende kodes like ofte til nivå 0 som nivå 1. Kompetansen *using symbols, operations and formal language* er i mine studier den kompetansen som har lavest andel på nivå 0, og samlet sett, for årene 2017 – 2019, er den kompetansen som blir testet på høyest nivå. Kompetansen *constructive communication* er i mine studier den kompetansen som viser størst «strekk i feltet» mellom del 1 og 2, med mellom 71 % og 87 % på nivå 0 i del 1 og mellom 27 % - 49 % over nivå 1 på del 2.

Nå skal man være forsiktig med å sammenlikne kompetanse-kategorier fra ulike rammeverk, men da både rammeverket til MEG-studien og eksamen for 10. trinn begge bygger på kompetansebeskrivelsene i KOM-prosjektet (Turner et al., 2015; Valenta et al., 2015) er det fristende å trekke noen tråder.

Eksamen i matematikk skal, ifølge Utdanningsdirektoratet (2019a), måle et bredt spekter av matematiske ferdigheter og prosesser som krever ulik kompetanse. Resultatene fra mine studier, med utgangspunkt i MEG-rammeverket, viser imidlertid at elevene, på en skala fra 0 til 3, stort sett testes på de laveste nivåene (nivå 0 og 1), der kategorisering på nivå 0 kan indikere fravær av anvendelse av den aktuelle kompetansen. En forsvinnende liten del av kompetansene testes på nivå 3: Av totalt 4620 kodinger har nivå 3 kun 12 scoringer, hvorav 5 på *constructive communication*, 3 på *Using symbols, operations and formal language*, 3 på *receptive communication* og en på *mathematising*. Dette indikerer, når man legger MEG-skjemaet til grunn, at potensialet i «bredt spekter av prosesser som krever ulik kompetanse» i eksamensveiledningen ikke utnyttes fullt ut for eksamensoppgavene. I mine studier viser særlig kompetansen *representation* og *reasoning and argument* høyt score på lavt nivå. For sistnevnte kategori er jeg tilbøyelig til å tenke at definisjonen av kompetansen er litt ulik i MEG-rammeverket og eksamensveiledningen (jf. tidligere drøfting under resultater), mens jeg for *representation* anser kategoriene som såpass like at jeg anser mine resultater som urovekkende: Aktivisering av denne kompetansen er enten fraværende eller, alt overveiende, på lavt nivå for alle eksamenssettene fra 2017 – 2019. Dette er også betenkelig når vi ser framover for hva som møter elevene med tanke på kjerneelementene i LK20 (Kunnskapsdepartementet, 2018b), der *representasjon* er nevnt som et sentralt element elevene må lære å mestre i matematikkfaget.

I formålet med faget for matematikk fellesfag (Utdanningsdirektoratet, 2013) og eksamensveiledningen for eksamen på 10. trinn (Utdanningsdirektoratet, 2019a) er problemløsning og modellering trukket fram som viktige kompetanser for hva det vil si å beherske matematikk. I mine analyser av eksamensoppgavene er *devising strategies* representert med majoriteten av nivåscorene på nivå 1 og med en nedadgående trend for nivå 2 fra 16 % i 2017 til 10 % i 2019, mens *mathematising* ligger omtrent jevnt på nivå 0 og 1. Ut fra mine resultater kan man da spørre seg om eksamen i matematikk på 10. trinn faktisk følger opp ordlyden i *Rammeverk for eksamen* (Utdanningsdirektoratet, 2017g), som både sier at eksamen skal gjenspeile formålet med faget og at kandidater på alle nivå skal få mulighet til å vise sin kompetanse på eksamen. Alle de kompetansene jeg har kartlagt, scorer riktig nok i større eller mindre grad over nivå 0, så man kan si at elevene i løpet av eksamen må aktivere alle syv kompetansene jeg har kartlagt, men med en såpass liten andel på nivå 2, og nesten ingen på nivå 3, tyder mine resultater på at elever på høyt nivå ikke får vist fram sitt kompetansespekter på eksamen. I mitt materiale er det kompetansene *constructive communication* og *using symbols, operations and formal language* som i størst grad dekker hele skalaen fra nivå 0 til 3.

Matematisk kompetanse er, som nevnt i innledningen til kapittel 2, et sammensatt begrep, der ulike rammeverk har ulike vinklinger. Alle rammeverkene omtalt i denne masteroppgaven, fra «stamfaren» Bloom via trådmodellen til NRC, kompetanseblomsten til KOM-prosjektet og test-regimene til TIMSS og PISA vektlegger imidlertid det å beherske ulike kognitive prosesser som en essensiell side av det å mestre matematikk. Selv om ulike rammeverk legger ulike termer til grunn for å beskrive og definere disse kognitive prosessene, påpeker Niss et al. (2017) at begrepsbruken er beslektet. Legger man mine resultater til grunn, samtidig som man skjeler til den beslektete begrepsbruken i rammeverkene fra teoridelen i denne oppgaven, indikerer mine studier at elevene på skriftlig eksamen i matematikk i liten grad utfordres til å anvende kognitive ferdigheter utover de to laveste nivåene i MEG-skjemaet. Dette på tross av at samtlige rammeverk jeg har beskrevet vektlegger dette, og at matematikkoppgaver som er kognitivt krevende fremmer større begrepsmessig forståelse hos elevene (Hiebert & Grouws, 2007). Det kan dermed tyde på at Niss (1993) fortsatt har sine ord i behold, når han sier at selv om læreplanene i matematikk i mange land nå også vektlegger kompetanse, metoder og tenkemåter fanges dette i liten grad opp av vurderingsformene. Følgen av dette blir dermed at vurderingen i liten grad evaluerer den kunnskap og kompetanse man ønsker skal prege undervisningen. Grønmo (2010), Pilikoff, Porter og Smithson (2011) og Boesen et al. (2014) har også i sine studier pekt på gapet mellom intensjonene i læreplanen og hva elevene faktisk lærer. I enkelte tilfeller kan det se ut som om matematisk kompetanse i enkelte klasserom innsnevres til hovedsakelig å handle om innlæring av fakta- og prosedyrekunnskap.

Ut fra mine resultater, der *using symbols, operations and formal language* er den kompetansen som gjennomgående for perioden 2017 – 2019 testes utover laveste nivå, kan man hevde at skriftlig eksamen i matematikk er med på å manifestere gapet mellom intensjonene for matematikk i Kunnskapsløftet og måten læreplanen i matematikk blir implementert på i form av eksamen: Fakta og prosedyrekunnskap er fortsatt det som testes på høyest nivå på eksamen, mens problemløsing og modellering (som vektlegges i formålsbeskrivelsen til faget) i større grad testes på lavere nivå.

En rekke forskere har sett på sammenhenger mellom vurdering av elever og undervisningspraksis, og pekt på at kompetanse som ikke gjøres til gjenstand for vurdering kan bli utelatt i undervisningen. Særlig tester med viktig utfall kan være med å sette rammer for lærernes undervisningspraksis (Schoenfeld, 2007; Wilson, 2007), og skriftlig eksamen i matematikk på 10. trinn må sies å være en slik viktig test. Fafo-rapportene som evaluerte eksamen i matematikk på 10. trinn for 2017 og 2018 (Andresen et al., 2017; Bjørnset et al., 2018) sier i sine hovedfunn for begge år at eksamen var svært god, og at opplæringen som er gitt i stor grad samsvarte med eksamensoppgavene. Ut fra den ovennevnte forskning på sammenhengen mellom vurdering og undervisningspraksis, finner jeg ikke dette overraskende: I det eksamen er en prøve med viktig utfall, vil trolig mange lærerne på 10. tinn innrette sin undervisning mot det som testes på eksamen. Faren med dette kan imidlertid være at kompetanser som ikke blir særlig vektlagt på eksamen heller ikke blir vektlagt i undervisningen, med det til følge at intensjonen i læreplanen om å gi elevene en helhetlig matematisk kompetanse ikke blir ivaretatt. For å sikre at eksamen dekker sentrale kompetanser på alle nivå, kan en vei å gå være å benytte MEG-rammeverket slik jeg har gjort i min studie. Man vil da få ut en kompetanse-profil for hele eksamenssettet, og kan ut fra dette vurdere om eksamen dekker et bredt kompetansespekter på ulikt nivå, og bruke dette som et supplement til kartleggingen av om eksamen også dekker relevante hovedområder/kjerneelementer og kompetansemål i den aktuelle læreplanen. Dette bygger jeg også på erfaringene i min kodegruppe, der deltagerne, på bakgrunn av studier av MEG-rammeverket, prøvekoding og diskusjon, for en stor del leverte samstemte kodinger av eksamensoppgavene i matematikk for 10. trinn for perioden 2017 – 2019.

Jeg har i min studie kun anvendt MEG-skjemaet på eksamensoppgaver. Jeg ser imidlertid, i tråd med Turner et al. (2013) og Pettersen (2019), potensiale for også å kunne bruke MEG-skjemaet på andre oppgaver matematikklærere gir til elevene. Ved å analysere oppgaver ut fra MEG-skjemaet, kan lærerne bli mer bevisst på hvilke kompetanser de ulike oppgavene aktiviserer. På den måten kan man sikre at undervisningen favner et bredt kompetansespekter, som inkluderer alle kompetansene elevene trenger å øve på for å opparbeide en helhetlig matematisk kompetanse. Lærerne som deltok i min kodegruppe kommenterte uoppfordret dette, og sa at de nå, etter denne intensive kode-seansen, var blitt mer bevisst på å velge oppgaver ut fra kompetansebeskrivelsene de hadde lært å kjenne gjennom MEG-skjemaet. Setter skoleeiere og rektorer dette i system, kan MEG-skjemaet også brukes som et verktøy for å drive utviklende undervisningspraksis på egen skole.

Referanser

Alseth, B., Breiteig, T., & Brekke, G. (2003). *Endringer og utvikling ved R97 som bakgrunn for videre planlegging og justering - matematikkfaget som kasus*. Notodden: Telemarksforskning.

Andresen, S., Fossum, A., Rogstad, J. & Smestad, B. (2017). *På prøve. Evaluering av matematikkeksamen på 10. trinn våren 2017*. Faforapport 2017:36. Oslo: Fafo

Anderson, L. W. (Red.), Krathwohl, D. R. (Red.), Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Bjørnset, M., Fossum, A., Rogstad, J., Smestad, B. & Talberg, N. (2018). *Digitale skillelinjer. Evaluering av matematikkeksamen på 10. trinn våren 2018*. Fafo-rapport 2018:36. Oslo: Fafo

Bloom, B.S. (Red.), Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: Longman.

Boesen, J. (2006). *Assessing mathematical creativity: Comparing national and teachermade tests, explaining differences and examining impact*. Doktoravhandling. Umeå (Sverige): Umeå universitet.

Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *The Journal of Mathematical Behavior, 33*, s. 72–87. https://doi.org/10.1016/j.jmathb.2013.10.001

Boesen, J., Lithner, J., & Palm, T. (2018). Assessing mathematical competencies: An analysis of Swedish national mathematics tests. *Scandinavian Journal of Educational Research, 62(1),* s. 109–124. https://doi.org/10.1080/00313831.2016.1212256

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), s. 284–290. http://dx.doi.org/10.1037/1040-3590.6.4.284

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8 utg.). London: Routledge.

Dale, E.L., Engelsen, B.U. & Karseth, B. (2011). *Kunnskapsløftets intensjoner, forutsetninger og operasjonaliseringer: En analyse av en læreplanreform*. Sluttrapport. Universitetet i Oslo. Pedagogisk forskningsinstitutt.

Ernest, P. (1991). *The philosophy of mathematics education*. London: Falmer Press

Fleiss, J. L., & Cohen, J. (1973). The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement, 33(3)*, s. 613–619. http://dx.doi.org/10.1177/001316447303300309

Goodlad, J. I. (1979). *Curriculum Inquiry. The Study of Curriculum Practice*. New York: McGraw-Hill.

Grønmo, L.S. (2010). TIMSS Advanced – et matematikkdidaktisk perspektiv. I Grønmo, L. S., Onstad, T., & Pedersen, I. F. (Red.). *Matematikk i motvind. TIMSS Advanced 2008 i videregående skole*, s. 27–43. Oslo: Unipub.

Grønmo, L.S., Onstad, T. & Pedersen, I.F. (2010). *Matematikk i motvind. TIMSS Advanced 2008 i videregående skole*. Oslo: Unipub.

Hiebert, J., & Grouws, D. A. (2007). The Effects of Classroom Mathematics Teaching on Students' Learning. I F. Lester (Red.). *Second Handbook of Research on Mathematics Teaching and Learning,* s. 371–404. Charlotte, NC: Information Age.

Husén, T. (1988). Research paradigms in education. *Interchange 19 (2),* s. 2–13. https://doi.org/10.1007/BF01815504

Jensen, T. H. (2007). *Udvikling af matematisk modelleringskompetence som matematikundervisningens omdrejningspunkt - hvorfor ikke?* Doktoravhandling. Roskilde: Roskilde universitet.

Kilpatrick, J. (2014). Competency Frameworks in Mathematics Education. I Lerman, S. (Red.), *Encyclopedia of Mathematics Education*, s. 85–87. Netherlands: Springer Reference.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). The Strands of Mathematical Proficiency. I National Research Council (Red.), *Adding it up. Helping Children Learn Mathematics*, s. 115–156. Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington DC: National Academy Press.

Kirke- og undervisningsdepartementet (1987). *Mønsterplan for grunnskolen: M87*. Oslo: Aschehoug

Kjærnsli, M. & Jensen, F. (2016). Gjennomføring og noen sentrale resultater. I Kjærnsli, M. & Jensen, F. (Red.). *Stø kurs. Norske elevers kompetanse i naturfag, matematikk og lesing i PISA 2015*, s. 11–31. Oslo: Universitetsforlaget.

Kjærnsli, M., Jensen, F., Bergem, O. K., Kaarstein, H. & Nilsen, T. (2019, 26. juli). *PISA og TIMSS – hva er forskjellen?* Hentet fra https://www.uv.uio.no/ils/om/aktuelt/aktuelle-saker/2016/pisa_og_timss_likheter_forskjeller.html

Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory into Practice, 41(4)*, s. 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kristianslund, T. K. (2015). *Eksamensoppgaver i matematikk i grunnskolelærerutdanningen 5.-10. trinn. En analyse av tematisk innhold og kognitive utfordringer*. Masteroppgave i matematikkdidaktikk. Institutt for lærerutdanning og skoleforskning. Utdanningsvitenskapelig fakultet. Universitetet i Oslo.

Kunnskapsdepartementet (2018a, 26. juni). *Fornyer innholdet i skolen*. [Pressemelding nr. 132-18]. Hentet fra https://www.regjeringen.no/no/aktuelt/fornyer-innholdet-i-skolen/id2606028/

Kunnskapsdepartementet (2018b). *Kjerneelementer i fag*. Hentet fra https://www.regjeringen.no/contentassets/3d659278ae55449f9d8373fff5de4f65/kjerneelementer-i-fag-for-utforming-av-lareplaner-for-fag-i-lk20-og-lk20s-fastsatt-av-kd.pdf

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics 33(1)*, s. 159–174.

Lithner, J. (2008). A Research framework for creative and imitative reasoning. *Educational Studies in Mathematics, 67(3),* s. 255–276

Lithner, J., Bergqvist, E., Bergqvist, T., Boesen, J., Palm, T. & Palmberg, B. (2010). *Mathematical competencies—A research framework.* Paper presented at the MADIF 7 Mathematics and mathematics education: Cultural and social dimensions, Stockholm, Sweden.

Matematikksenteret (2015). *Vurdering av eksamen i matematikk*. Hentet fra https://www.matematikksenteret.no/sites/default/files/attachments/resource/vurdering_av_eksamen_i_matematikk.pdf

Michelsen, C. (2001). *Begrebsdannelse ved domæneudvidelse: Elevers tilegnelse af funktionsbegrebet i et integreret undervisningsforløb mellem matematik og fysik*. Doktoravhandling. Odense: Syddansk universitet.

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowski, S.J. & O'Connor, K.M. (2003). *TIMSS Assessment Frameworks and Specifications, 2nd Edition*. Chestnut Hill: International Study Center, Boston College. Hentet fra https://timss.bc.edu/timss2003i/PDF/t03_af_book.pdf

Mullis, I.V.S. & Martin, M.O. (2013). *TIMSS 2015 Assessment Frameworks*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College og International Association for the Evaluation of Educational Achievement (IEA). Hentet fra https://timssandpirls.bc.edu/timss2015/downloads/T15_Frameworks_Full_Book.pdf

National Council of Teachers of Mathematics. (1980). *An agenda for action: Recommendations for school mathematics of the 1980s*. Reston: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston: National Council of Teachers of Mathematics.

National Research Council (2001). *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9822

Nilsen, T. & Kaarstein, H. (2016). TIMSS og statistiske metoder. I Bergem, O.K., Kaarstein, H. og Nilsen, T. (Red.). *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015*, s. 178–198. Oslo: Universitetsforlaget.

Niss, M. (1993). Assessment in mathematics education and its effects: An introduction. I M. Niss (Red.), *Investigations into assessment in mathematics education: An ICMI study*, s. 1–30. Dordrecht (Nederland): Kluwer academic publishers.

Niss. M. (2003). The Danish KOM project and possible consequences for teacher education. I R. Strässer, G. Brandell, B. Grevholm, & O. Helenius (Red.) (2003). *Educating for the future. Proceedings of an International Symposiumon on Mathematics Teacher Education*, s. 178–192. Gothenburg: Royal Swedish Academy of Science.

Niss, M. (2015). Mathematical competencies and PISA. I K. Stacey & R. Turner (Red.), *Assessing Mathematical Literacy*, s. 35–55. New York: Springer.

Niss M., Bruder R., Planas N., Turner R. & Villa-Ochoa J.A. (2017). Conceptualisation of the Role of Competencies, Knowing and Knowledge in Mathematics Education Research. I Kaiser G. (Red.) *Proceedings of the 13th International Congress on Mathematical Education,* s. 235–248. ICME-13 Monographs. Cham: Springer. https://doi.org/10.1007/978-3-319-62597-3_15

Niss, M., & Jensen, T. H. (Red.). (2002). *Kompetencer og matematiklæring—Idéer og inspiration til udvikling af matematikundervisning i Danmark. Uddannelsesstyrelsens temahæfteserie nr. 18*. Copenhagen: The Ministry of Education.

NOU 2015: 8. (2015). *Fremtidens skole. Fornyelse av fag og kompetanser*. Oslo: Kunnskapsdepartementet.

Norsk senter for forskningsdata (2019). *Må jeg melde prosjektet?* Hentet fra https://nsd.no/personvernombud/meld_prosjekt/index.html

Nortvedt, G.A. (2013). Matematikk i PISA – matematikkdidaktiske perspektiver. I M. Kjærnsli & R.V. Olsen (Red.). *Fortsatt en vei å gå. Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012*, s. 43–66. Oslo: Universitetsforlaget.

Nortvedt, G. A. & Pettersen, A. (2016). Matematikk. I Kjærnsli, M. og Jensen, F. (Red.). *Stø kurs. Norske elevers kompetanse i naturfag, matematikk og lesing i PISA 2015*, s. 107–135. Oslo: Universitetsforlaget.

Nortvedt, G.A., Pettersen, A., Petterson, A. & Sollerman, S. (2016). Is PISA 2012 relevant to mathematics education in Norway and Sweden? I M. Nordengen og Thorsen, H. (Red.) *Northern Lights on PISA and TALIS*, s. 27–58. København: Norden.

OECD (2017). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving. Revised edition*, PISA, Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264281820-en

Olsen, R.V. (2010). Matematikk i PISA. I Kjærnsli, M. & Roe, A. (Red.) *På rett spor. Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009*, s. 138–158. Oslo: Universitetsforlaget.

Pedersen, I. F. (2013). Is TIMSS Advanced an appropriate instrument for evaluating mathematical performance at the advanced level of Norwegian upper secondary school? An analysis of curriculum documents and assessment items. *Acta Didactica Norge, 7(1)*, s. 1–24. Hentet fra https://www.journals.uio.no/index.php/adno/article/view/1112

Pettersen, A. (2019). *Towards competency-oriented mathematics education. An investigation of task demands and teachers' knowledge of task demands from a competency perspective.* Doktoravhandling. Universitetet i Oslo. Institutt for lærerutdanning og skoleforskning. Det utdanningsvitenskaplige fakultet.

Pettersen, A. & Braeken, J. (2019). Mathematical Competency Demands of Assessment Items: a Search for Empirical Evidence*.* I *International Journal of Science and Mathematics Education*, 17(2), s. 405–425. http://dx.doi.org/10.1007/s10763-017-9870-y

Pettersen, A., & Nortvedt, G. A. (2018). Identifying competency demands in mathematical tasks: Recognising What Matters. *International Journal of Science and Mathematics Education, 16(5),* s. 949–965. https://doi.org/10.1007/s10763-017-9807-5.

Polikoff, M.S., Porter, A.C., & Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? I *American Educational Research Journal, 48(4)*, s. 965–995. https://doi.org/10.3102/0002831211410684

Pólya, G. (1945). *How to solve it?* Princeton: Princeton University Press. 2nd edition (1957).

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31(7),* s. 3–14.

Porter, A. C., & Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. I S. H. Fuhrman (Red.), *From the capitol to the classroom: Standards-based reform in the states—One hundredth yearbook of the National Society for the Study of Education, Part II*, s. 60 – 80. Chicago, IL: University of Chicago Press.

Robson, C. & McCartan, K. (2016). *Real world research. A Resource for Users of Social Research Methods in Applied Settings* (4. utg.) Oxford: Blackwell.

Schoenfeld, A. H. (2007). What Is Mathematical Proficiency and How Can It Be Assessed? I *Assessing Mathematical Proficiency, MSRI Publications, Vol. 53*, s. 59–73. Cambridge University press.

Stein, M., Smith, M., Henningsen, M., & Silver, E. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development.* New York: Teacher College Press.

Turner, R., Dossey, J., Blum, W., & Niss, M. (2013). Using mathematical competencies to predict item difficulty in PISA: A MEG study. I M. Prenzel, M. Kobarg, K. Schöps, & S. Rönnebeck (Red.). *Research on PISA: Research outcomes of the PISA Research Conference 2009*, s. 23–37. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-4458-5_3.

Turner, R., Blum, W., & Niss, M. (2015). Using competencies to explain mathematical item demand: A work in progress. I K. Stacey & R. Turner (Red.), *Assessing mathematical literacy: The PISA experience,* s. 85–115. New York: Springer. https://doi.org/10.1007/978-3-319-10121-7.

Utdanningsdirektoratet. (2013). *Læreplan i matematikk fellesfag. Læreplankode MAT1-04*. Hentet fra http://data.udir.no/kl06/MAT1-04.pdf.

Utdanningsdirektoratet (2016). *Å forstå kompetanse*. Hentet fra https://www.udir.no/laring-og-trivsel/lareplanverket/forsta-kompetanse/.

Utdanningsdirektoratet (2017a). *Eksamensveiledning – om vurdering av eksamensbesvarelser 2017, MAT0010 Matematikk, Sentralt gitt skriftlig eksamen*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2017b). *Sensorveiledning MAT0010 Matematikk, 16.05.2017*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2017c). *Forhåndssensurrapport 31.05.17. MAT0010 Matematikk.*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2017d). *Eksamen 16.05.2017. MAT0010 Matematikk. Del 1.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2017e). *Eksamen 16.05.2017. MAT0010 Matematikk. Del 2.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2017f). Kjerneelementer – fag i grunnskolen og gjennomgående fag i vgo. Hentet fra https://www.udir.no/laring-og-trivsel/lareplanverket/fagfornyelsen/kjerneelementer/

Utdanningsdirektoratet (2017g). *Rammeverk for eksamen.* Hentet fra https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-eksamen/

Utdanningsdirektoratet (2017h). *Rammeverk for grunnleggende ferdigheter.* Hentet fra https://www.udir.no/laring-og-trivsel/lareplanverket/grunnleggende-ferdigheter/rammeverk-for-grunnleggende-ferdigheter/

Utdanningsdirektoratet (2018a). *Eksamensveiledning – om vurdering av eksamensbesvarelser 2018, MAT0010 Matematikk, Sentralt gitt skriftlig eksamen*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2018b). *Sensorveiledning MAT0010 Matematikk, 16.05.2018*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2018c). *Forhåndssensurrapport 31.05.18. MAT0010 Matematikk. Forhåndssensur 30. – 31. mai 2018.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2018d). *Eksamen 16.05.2018. MAT0010 Matematikk. Del 1.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2018e). *Eksamen 16.05.2018. MAT0010 Matematikk. Del 2.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2018f). *Overordnet del av læreplanverket*. Hentet fra https://www.udir.no/laring-og-trivsel/lareplanverket/overordnet-del/

Utdanningsdirektoratet (2019a). *Eksamensveiledning – om vurdering av eksamensbesvarelser 2019, MAT0010 Matematikk, Sentralt gitt skriftlig eksamen*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2019b). *Sensorveiledning MAT0010 Matematikk, 16.05.2019.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2019c). *Forhåndssensurrapport 29.05.19. MAT0010 Matematikk. Forhåndssensur 28. – 29. mai 2019*. Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2019d). *Eksamen 16.05.2019*. *MAT0010 Matematikk. Del 1.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2019e). *Eksamen 16.05.2019*. *MAT0010 Matematikk. Del 2.* Hentet fra https://sokeresultat.udir.no/eksamensoppgaver.html

Utdanningsdirektoratet (2019f). *Kunnskapsgrunnlag for evaluering av eksamensordningen*. Hentet fra https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/Kunnskapsgrunnlag-for-evaluering-av-eksamensordningen/

Utdanningsdirektoratet (2019g). *Vurderinger og foreløpige anbefalinger fra eksamensgruppa.* Hentet fra https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/vurderinger-og-forelopige-anbefalinger-fra-eksamensgruppa/

Valenta, A., Nosrati, M., Åsenhus, R., & Wæge, K. (2015). *Skisse av den «ideelle læreplan i matematikk»*. Trondheim: Nasjonalt senter for matematikk i opplæringen. Hentet fra https://docplayer.me/1644531-Skisse-av-den-ideelle-laereplan-i-matematikk.html

Webb, N. L. (1997). Determining alignment of expectations and assessments in mathematics and science education. *NISE Brief, 1(2)*. National Institute for Science Education, University of Wisconsin-Madison.

Webb, D. C. (2014). Bloom's Taxonomy in Mathematics Education. I Lerman, S. (Red.), *Encyclopedia of Mathematics Education*, s. 63–68. Netherlands: Springer Reference.

Webb, N. L. (2002). *Alignment study in language arts, mathematics, science, and social studies of state standards and assessments for four states*. Washington, DC: Council of Chief State of School Officers

Wilson, L. D. (2007). High stakes test in mathematics. I Lester, F. K. (Red.), *Second handbook of research on mathematics teaching and learning*, s. 1099–1110. Charlotte (USA): Information age publishing inc.