Eksamensoppgaver i matematikk i grunnskolelærerutdanningen 5.-10. trinn

En analyse av tematisk innhold og kognitive utfordringer

Tiina Klami Kristianslund

Masteroppgave i matematikkdidaktikk

Institutt for lærerutdanning og skoleforskning
Utdanningsvitenskapelig fakultet

UNIVERSITETET I OSLO

Våren 2015

Eksamensoppgaver i matematikk i grunnskolelærerutdanningen 5.-10. trinn

En analyse av tematisk innhold og kognitive utfordringer

2015

Eksamensoppgaver i matematikk i grunnskolelærerutdanningen 5.-10. trinn: En analyse av tematisk innhold og kognitive utfordringer

Tiina Klami Kristianslund

http://www.duo.uio.no

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Denne studien har undersøkt innholdet i eksamensoppgaver i matematikk ved tre norske grunnskolelærerutdanninger. Forskning har vist at matematikkunnskap hos lærere er viktig for god undervisningskompetanse. Resultatene fra TEDS-M 2008 viser imidlertid at norske lærerstudenter skårer lavt sammenliknet med et internasjonalt gjennomsnitt, både når det kommer til matematikkfaglig og matematikkdidaktisk kunnskap. Gjennom «Lærerløftet» uttrykker Regjeringen et klart mål om å øke kvaliteten på lærerutdanningen. Forslag fra nåværende kunnskapsminister, Torbjørn Røe Isaksen, om innføring av nasjonale prøver i matematikk i grunnskolelærerutdanningene, illustrerer et ønske om kvalitetssikring av innholdet i matematikkfagene i disse utdanningene. Min studie undersøker noe av innholdet i GLU 5-10-utdanningen i matematikk ved å analysere hvordan lærerstudentene utfordres tematisk og kognitivt i skriftlige skoleeksamensoppgaver.

Datamaterialet mitt består av 530 eksamensoppgaver, hvorav 473 er matematikkfaglige oppgaver og 57 matematikkdidaktiske. Analysen er basert på nivårammeverket utviklet i MEG-studien. Resultatene fra analysen viser at lærerstudentene ved de utvalgte lærestedene blir testet på et lavt nivå i de fleste av de kognitive kategoriene som rammeverket bruker til å beskrive matematikkompetanse. Dette gjelder også for hvert av de inkluderte lærestedene separat. En typisk oppgave i datamaterialet er lite utfordrende kommunikativt, krever til en liten grad at studentene må matematisere, lage løsningsstrategier eller bruke matematiske representasjoner. De må imidlertid kunne en del fakta og prosedyrer for å løse oppgaven og de må kunne resonnere noe underveis. Lærestedene skiller seg fra hverandre når det kommer til hvor stor andel av oppgavene som tematisk ligger innenfor læreplanen i matematikk for 8.-10. trinn, som er det høyeste matematikknivået studentene kan undervise på som lærere. Andelen oppgaver med tema fra et høyere nivå enn det denne læreplanen beskriver, ligger mellom 51 % og 82 %.

Min oppgave viser altså at lærerstudentene ved tre utvalgte læresteder til en stor grad testes i oppgaver som ligger på et høyt tematisk nivå, det vil si oppgaver som strekker seg utenfor temaer studentene selv skal undervise når de kommer ut i skolen, men at studentene testes på et lavt kognitivt nivå. Det sistnevnte står i motsetning til forskning som viser at lærere trenger å beherske matematikk på et høyt kognitivt nivå for å gjennomføre god undervisning.

Forord

Denne masteroppgaven markerer slutten på en lærerik og spennende utdanning ved Lektorprogrammet ved Universitet i Oslo. Arbeidet med oppgaven har vært lærerikt, interessant og morsomt, samtidig som det har vært krevende.

Jeg vil først og fremst takke mine flotte veiledere, Hege Kaarstein og Torgeir Onstad. Jeg kom i kontakt med Hege allerede i juni 2014 og vi har hatt mange og gode samtaler fram mot innlevering av den endelige oppgaven. Takk for all god inspirasjon, for hjelp til å velge og konstruere et ønsket og realistisk mål for studien, for gode råd gjennom hele prosessen, for god faglig oversikt og for tålmodighet og oppmuntring. Torgeir ble inkludert som medveileder fra januar 2015. Tusen takk til deg for gode tilbakemeldinger, oppmuntringer og fine faglige diskusjoner gjennom veiledningene.

Videre vil jeg takke Andreas Pettersen for all hjelp jeg har fått av deg. For det første har det vært en enorm hjelp at du har kodet deler av datamaterialet parallelt med meg. I tillegg har jeg fått mange gode råd i forbindelse med valg av litteratur og rammeverk for koding.

Tusen takk til lærestedene som har gitt meg tilgang på eksamensoppgaver slik at studien ble mulig å gjennomføre.

Gode medstudenter gjør arbeidet med en slik oppgave mye lettere. Spesielt vil jeg takke Anne Line for gode råd og gode samtaler både om masteren og om alt annet. Tusen takk også til Ane Sofie for korrekturlesning og for at du er en så god venn.

Jeg har verdens fineste familie. Takk til pappa, Kristin og Truls for korrekturlesning. Takk til mamma for alt du gjør for at jeg skal ha det bra! Takk til min fantastiske mann, Eirik, som har stått på sammen med meg for at denne oppgaven skulle bli ferdig. Takk for oppmuntring og støtte, faglige diskusjoner og ikke minst for at du har passet på at jeg har det fint.

Oslo, 23. juni 2015

Tiina Klami Kristianslund

Innholdsfortegnelse

Figurer

Tabeller

1 Innledning

1.1 Bakgrunn for valg av tema

Norsk lærerutdanning får oppmerksomhet, både hos politikere og i media. Gjennom «Lærerløftet» uttrykker Regjeringen et klart mål om å øke kvaliteten på utdanningen av lærere. Forslag som trekkes frem er å kreve karakteren 4 eller bedre i norsk, matematikk og engelsk for opptak, og å gjøre om lærerutdanningen til en femårig mastergradsutdanning (Kunnskapsdepartementet, 2014). Lærerutdanningen er en av få utdanninger som Kunnskapsdepartementet har fastsatt rammeplaner for (Kunnskapsdepartementet, 2010a). I Nasjonale retningslinjer for grunnskolelærerutdanningen, konkretiseres dette nærmere med føringer for innhold og undervisningsform i utdanningen (Kunnskapsdepartementet, 2010b). For grunnskolelærerutdanningen 5.-10. trinn kan vi blant annet lese dette om matematikkfaget (Kunnskapsdepartementet, 2010b, s. 34):

> *Gjennom matematikkfaget for trinn 5-10 skal studentene utvikle undervisningskunnskap i matematikk. Dette innebærer at de må ha en solid og reflektert forståelse for den matematikken elevene skal lære og hvordan denne utvikles videre på de neste trinnene i utdanningssystemet. Videre kreves matematikkfaglig kunnskap som er særegen for lærerprofesjonen. Slik kunnskap omfatter, i tillegg til selv å kunne gjennomføre og forstå matematiske prosesser og argumenter, også å kunne analysere slike som foreslås av andre med tanke på å vurdere deres holdbarhet og eventuelle potensial*.

Nåværende kunnskapsminister Torbjørn Røe Isaksen har foreslått å innføre nasjonale prøver i matematikk ved grunnskolelærerutdannningene (Svarstad, 2014), noe som tyder på et ønske om kvalitetssikring av innholdet i utdanningen.

Noe av årsaken til fokuset på lærerutdanningen er trolig at norske elevers matematikkompetanse er dårligere enn det man kunne ønske. Programme for International Student Assessment (PISA) 2012 viste at norske 15 år gamle elever skåret lavere enn det internasjonale gjennomsnittet i matematikk. Spesielt skåret de norske elevene lavt i oppgaver som var knyttet til å bruke matematisk formalkompetanse (Kjærnsli & Olsen, 2013). Elevene skåret også noe lavere enn det internasjonale gjennomsnittet i Trends in International Mathematics and Science Study (TIMSS) 2011 (Grønmo et al., 2012). TIMSS er en annen stor internasjonal undersøkelse som tester elevprestasjoner på 4. og 8. trinn i matematikk og naturfag. I denne undersøkelsen så vi en positiv trend hvor resultatene var klart bedre i TIMSS 2011 enn i TIMSS 2003 og 2007. Denne positive utviklingen ble imidlertid ikke bekreftet gjennom PISA 2012, hvor norske elever skåret omtrent likt som i PISA 2003, som var forrige gang matematikk var det sentrale fagområdet for undersøkelsen (Nortvedt, 2013).

Forskning viser at matematikkunnskap hos lærere er viktig for god undervisningskompetanse i matematikk (Ball, Thames, & Phelps, 2008; Shulman, 1986). Teacher Education and Development Study – Mathematics (TEDS-M) 2008 viste imidlertid at også norske lærerstudenter skåret lavt i forhold til et internasjonalt gjennomsnitt, både når det kom til matematikkfaglig og matematikkdidaktisk kunnskap (Grønmo & Onstad, 2012). Undersøkelsen viste at norske lærerstudenter på alle trinn hadde store utfordringer med den faglige basisen i matematikk.

Anvendt matematikk og problemløsing i virkelighetsnære situasjoner er en viktig drivkraft i utviklingen av norske lærerplaner (Olsen & Grønmo, 2006). Rapporten fra TEDS-M 2008 påpekte imidlertid at anvendt matematikk ikke kan komme i stedet for ren matematikk. Skal man ha noe å anvende, forutsetter det at man har en god matematikkfaglig basis (Grønmo & Onstad, 2012). Det fremstår tydelig at det er viktig å styrke matematikkunnskapene til lærerstudenter i Norge. Resultatene i TEDS-M 2008 pekte mot et behov for en grundig gjennomgang og en debatt om innholdet i den norske lærerutdanningen (Grønmo & Onstad, 2012). TEDS-M 2008 ble gjennomført før overgangen fra allmennlærerutdanningen (ALU) til grunnskolelærerutdanningen (GLU), slik at resultatene herfra ikke nødvendigvis sier noe om nivået på norske lærerstudenter i dag. Likevel har resultatene fra denne undersøkelsen vært en viktig motivasjon og bakgrunn for mitt ønske om å kartlegge matematikkoppgaver gitt til skriftlig eksamen ved norske GLU-utdanninger.

Tonheim og Torkildsen (2010) har gjennomført en studie hvor de undersøkte hva som karakteriserte faget og fagopplæringen i matematikk i allmennlærerutdanningen. Studien gjennomgikk fagplaner, pensum og vurderingsmetoder og inkluderte en analyse av skriftlige eksamensoppgaver for det kullet som startet på allmennlærerutdanningen høsten 2006. Forfatterne så at det var stor enighet om intensjonene for hvilken kunnskap studentene skulle sitte igjen med etter endt studieløp. Analysen av eksamensoppgavene viste imidlertid at det var store variasjoner i hvilken kunnskap som faktisk ble vurdert. I tillegg viste undersøkelsen at matematikk som tematisk lå utenfor grunnskolepensum til en liten grad ble testet ved flertallet av lærestedene, noe som ble ansett som problematisk (Tonheim & Torkildsen, 2010). Undersøkelsen illustrer en mulig tilnærming til kartlegging av innhold i lærerutdanningen. Denne undersøkelsen er imidlertid, som TEDS-M 2008, gjennomført før overgangen til den nye GLU-utdanningen. Funnene fra undersøkelsen er derfor ikke nødvendigvis representative for innholdet i dagens lærerutdanning.

I 2010 gikk man over fra ALU til den todelte GLU med mer spesialisert utdanning rettet mot undervisning på barnetrinnet eller mellomtrinnet (Kunnskapsdepartementet, 2009). Overgangen ble innført som et forsøk på å ivareta behovet for spesialisert kompetanse innen ulike fag, årstrinn og funksjoner i skolen. Dagens grunnskolelærerutdanning består av to løp. GLU 1-7 omfatter minst fire undervisningsfag på 30 studiepoeng hver og pedagogikk, og kvalifiserer for undervisning fra 1.-7. trinn i barneskolen. GLU 5-10 omfatter tre undervisningsfag på 60 studiepoeng hver og 60 studiepoeng i pedagogikk, og kvalifiserer for undervisning fra 5.-10. trinn. Felles for begge løpene er at alle undervisningsfag skal omfatte fagdidaktikk og at de grunnleggende ferdighetene i læreplanverket for grunnskolen skal være integrert i fagene (Kunnskapsdepartementet, 2009).

Karakterundersøkelsen i matematikk i GLU-utdanningene, som ble gjennomført i 2014, omfattet gjennomgang og undersøkelse av emneplaner, eksamensoppgaver, eksamensbesvarelser og karakterlister i matematikkfagene (Christiansen, Enge, & Lode, 2015). Denne undersøkelsen var ment som et innlegg i en diskusjon om hvorvidt vurdering i lærerutdanningene i større grad bør styres av nasjonale veiledninger og en felles plattform (Universitets- og høgskolerådet, 2013). Gruppen som ledet arbeidet opplevde at fagmiljøene i GLU-utdanningene i matematikk ønsker et tettere samarbeid om utarbeidelse av eksamensoppgaver og vurdering av eksamensbesvarelser. Våren 2015 ble det utgitt en rapport som kan være et steg mot å etablere et slikt felles språk for vurderingspraksisen ved de ulike lærestedene (Christiansen et al., 2015).

Karakterundersøkelsen presenterte en oversikt over karakterfordelingen i noen utvalgte emner på GLU-utdanningene. For faget Matematikk 1 i GLU 5-10 første emne er snittkarakteren C eller D ved samtlige læresteder. Arbeidsgruppen ønsket å undersøke om det er en årsakssammenheng mellom institusjon og karakterfordeling. Årsaker til en slik sammenheng kan skyldes nivå på de aktuelle studentene, ulik vanskelighetsgrad i de gitte eksamenene, ulik vurderingsstrenghet og ulik eksamensform (Christiansen et al., 2015). Da rapporten ble skrevet forelå det imidlertid ikke beskrivelser av noe av dette og man kunne derfor ikke si noe om en slik årsakssammenheng (Christiansen et al., 2015). Dette viser et behov for å gjøre undersøkelser knyttet til vurdering i lærerutdanningen.

Gjennom karakterundersøkelsen i matematikk i GLU-utdanningene viste det seg at eksamensformen ved de ulike lærestedene varierer mellom muntlige eksamener, skriftlige skoleeksamener og skriftlige hjemmeeksamener. Rapporten fra undersøkelsen pekte på hvordan ulike eksamensformer tester ulike deler av lærerstudentenes kompetanse innen matematikk og matematikkdidaktikk og hvordan de ulike eksamensformene utfyller hverandre (Christiansen et al., 2015). En hjemmeeksamen kan egne seg godt til å måle fagdidaktisk kunnskap siden denne eksamensformen langt på vei kan minne om hvordan læreren kan jobbe med forberedelsene til timene. En skoleeksamen vil være godt egnet til både matematikkoppgaver og matematikkdidaktiske oppgaver. Den er spesielt godt egnet til å måle paratheten, evnen til umiddelbar respons, av den matematiske undervisningskunnskapen til studentene. Den muntlige eksamensformen trekkes frem som godt egnet til å måle sider ved kandidatens undervisningskunnskap som er tett knyttet til undervisningssituasjoner i klasserommet (Christiansen et al., 2015).

1.2 Formål, avgrensning og problemstilling

Jeg ønsker gjennom denne studien å gjøre en kartlegging av deler av vurderingspraksisen ved grunnskolelærerutdanninger i Norge. Spesielt ønsker jeg å se på hvordan lærerstudentene utfordres tematisk og kognitivt i matematikkfaglige oppgaver fordi forskning viser at lærerens matematikkforståelse er viktig for elevenes læringsutbytte (Ball et al., 2008; Shulman, 1987). Eksamensoppgaver er en lett tilgjengelig ressurs som gir god informasjon om hva lærerutdanningene velger å vektlegge. Det ville være mulig å undersøke GLU 1-7, GLU 5-10 eller begge. Det er naturlig for meg å se på de høyere trinnene da jeg selv har undervisningskompetanse for mellomtrinnet og derfor en bedre forståelse av hva matematikkundervisningen her inneholder. Jeg velger derfor å se på GLU 5-10. Det ble ikke vurdert som aktuelt for meg å inkludere muntlige eksamener i mitt datamateriale, da det ville vært svært arbeidskrevende og vanskelig å standardisere. Valget falt på å utelukkende analysere skriftlige skoleeksamensoppgaver.

På bakgrunn av formålet med oppgaven og de metodiske avgrensningene, har jeg formulert følgende problemstilling:

> *Hva kjennetegner skriftlige skoleeksamensoppgaver i matematikk ved GLU 5-10-utdanninger med tanke på matematisk tema og kognitiv kategori, og hvilke likheter og forskjeller finner vi mellom læresteder?*

1.3 Oppgavens struktur

Jeg ønsker å kunne si noe om hvordan de skriftlige skoleeksamensoppgavene tester den matematiske kompetansen hos lærerstudentene og hvordan dette sier noe om undervisningskunnskapen til de fremtidige lærerne. Kapittel 2 tar for seg ulike mulige kategoriseringsverktøy for matematisk kompetanse slik at jeg med utgangspunkt i dette kan velge et rammeverk for kategorisering av mine data. I tillegg beskrives kjennetegn på undervisningskunnskap i matematikk slik at jeg kan bruke dette til å diskutere i hvilken grad de inkluderte eksamensoppgavene vurderer dette. I kapittel 3 beskrives valg av studiedesign, datamateriale og analysemetode. Det gis også en beskrivelse av hvordan datainnsamling og koding av data har foregått og en oversikt over tidsrammene for studien. I tillegg diskuterer jeg hvilke validitetstrusler jeg må ta høyde for i min studie.

I kapittel 4 presenteres resultatet av kategoriseringen av eksamensoppgavene, både med tanke på hvilke matematiske temaer som testes i oppgavene, hvilket nivå de ligger på i forhold til læreplanen i matematikk for 8.-10. trinn og hvordan de utfordrer lærerstudentene kognitivt. I tillegg presenteres samsvarsdata for meg og en annen koder som har analysert deler av datamaterialet i studien. I kapittel 5 diskuteres funnene som presenteres i kapittel 4, opp mot teorien som er presentert i kapittel 2. Fordi valg av rammeverk for kategorisering av mine data har vært med på å forme studien, følger en diskusjon av hvordan jeg har oppfattet egnetheten til det valgte rammeverket. En vurdering av styrker og svakheter ved egen oppgave presenteres også.

I kapittel 6 avsluttes oppgaven med en kort oppsummering av hovedfunnene i studien etterfulgt av en refleksjon rundt mulige implikasjoner for videre studier.

2 Teori

Lærerstudentene som gjennomfører de eksamensoppgavene som jeg analyserer i denne studien er både studenter som gjennom sitt studieløp skal oppnå matematikkompetanse, og fremtidige lærere som må opparbeide undervisningskunnskap i matematikk. Matematikkompetanse og undervisningskunnskap i matematikk er derfor sentrale begreper i oppgaven.

I kapittel 2.1 definerer jeg hva som ligger i begrepene undervisningskompetanse, undervisningskunnskap, fagkunnskap, fagdidaktisk kunnskap, matematikkompetanse og kognitive kategorier fordi dette er begreper som vil gå igjen i teorien, og i analyse og diskusjon av mine resultater. Jeg ønsker å undersøke hvilken matematikkompetanse de inkluderte lærestedene tester hos lærerstudentene. For å få til dette trenger jeg et rammeverk som jeg kan kategorisere oppgavene i. I kapittel 2.2 presenteres et utvalg av mulige rammeverk for slik kategorisering. Det finnes tidligere studier som har gjort tilsvarende kategorisering av oppgaver som det jeg ønsker å gjøre. Jeg har i kapittel 2.3 sett på to studier som har brukt hvert sitt rammeverk som beskriver ulike komponenter av matematisk kompetanse til å analysere og sammenlikne matematikkoppgaver. I kapittel 2.4 presenteres ulike teoretiske tilnærminger til hva som kjennetegner undervisningskunnskap i matematikk.

2.1 Definisjoner av sentrale begreper

Undervisningskunnskap i matematikk bygger på både fagkunnskap og fagdidaktisk kunnskap (Grønmo & Onstad, 2012). Fagkunnskap handler om kunnskap om faginnhold som ikke er knyttet til undervisning eller elever. Fagdidaktisk kunnskap omfatter lærerplankunnskap, kunnskap om planlegging av undervisning og for læring og kunnskap om å gjennomføre undervisning og tilrettelegge for læring, på linje med definisjonene i rammeverket til TEDS-M 2008 (Grønmo & Onstad, 2012). Dette stemmer godt med Store norske leksikons definisjon av didaktikk: «Sammenhengen mellom undervisningens begrunnelse, innhold og gjennomføring» (Tjeldvoll & Skagen, 2014). Spesielt i forbindelse med beskrivelse av undervisningskompetanse, lærerkunnskap og ved kategorisering av oppgaver vil jeg bruke disse definisjonene som veiledning for hva som er matematikkdidaktikk.

Begrepet *kompetanse* er vanskelig å definere presist (Kilpatrick, 2014, s. 85):

> *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. Dictionaries give such definitions as "the state or quality of being adequately or well qualified"; "the ability to do something successfully or efficiently"; possession of required skill, knowledge, qualification, or capacity"; "a specific range of skill, knowledge, qualification, or capacity"; and "the scope of a person's or group's knowledge or ability." 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.*

Det er altså vanskelig å gi en presis definisjon av *kompetanse*, men vi ser at det handler om hvor godt rustet man er til å gjennomføre eller forstå det man skal utføre. God undervisningskompetanse kan ut fra definisjonene over bety at læreren har ferdighetene, kunnskapen, kvalifikasjonene og kapasiteten som skal til for å gjennomføre god undervisning. *Fagkunnskap* er altså en del av kompetansebegrepet. Å ha tilstrekkelig kunnskap om et tema i matematikken er en del av det å ha matematisk kompetanse, men kunnskap er imidlertid ikke tilstrekkelig for å ha god matematikkompetanse.

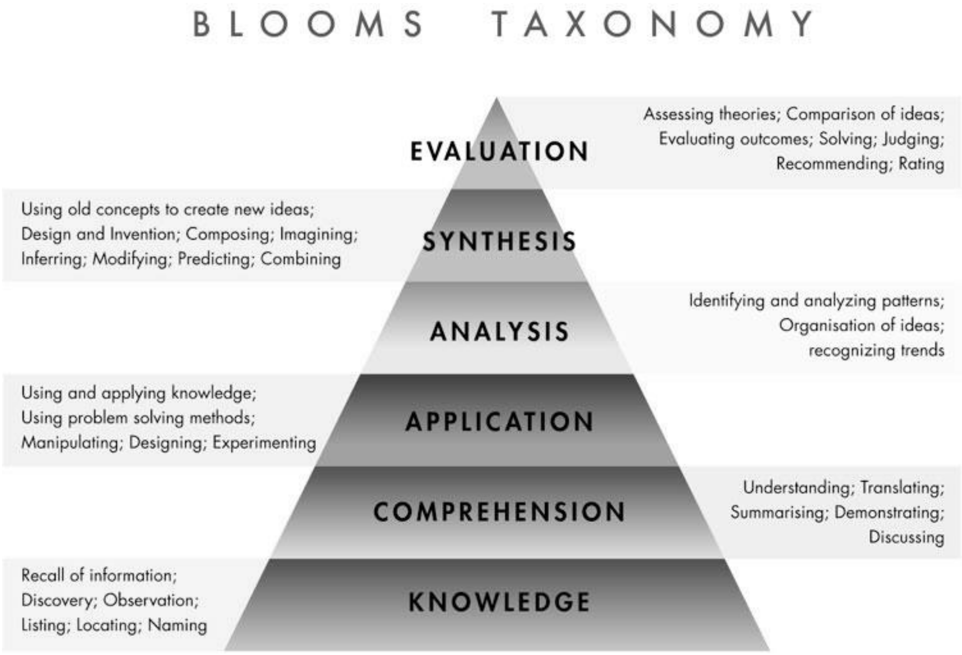
Ifølge Kilpatrick (2014) består en helhetlig matematikkompetanse av både kunnskap om faginnholdet og kognitive ferdigheter. De kognitive ferdighetene er uavhengige av hvilket tema som testes, men er egenskaper som anses som nødvendige for å kunne mestre matematisk tankegang. *Kognitive kategorier* i matematikk er beskrivelser av slike kognitive ferdigheter som elevene trenger i møtet med en matematikkoppgave (Grønmo et al., 2012).

2.2 Matematikkompetanse

2.2.1 Blooms taksonomi

Det første kompetanserammeverket for undervisning var Blooms taksonomi (Kilpatrick, 2014), publisert første gang i 1956 (Krathwohl, 2002). Rammeverket ble utviklet med utgangspunkt i et ønske om å lage en oppgavebank med grupper av oppgaver som testet samme ferdigheter hos elevene. For å få til dette trengtes et kategoriseringssystem som klassifiserte disse ferdighetene

(Krathwohl, 2002). Arbeidet ble initiert av Benjamin S. Bloom. Han satte ned en gruppe bestående av seg selv, Engelhart, Furst, Hill og Krathwohl som sammen arbeidet seg fram til taksonomien (Krathwohl, 2002). Blooms taksonomi så ikke på ulike emner innen bestemte fag, men utelukkende på de kognitive prosessene som ligger under det å inneha fagkunnskap (Kilpatrick, 2014). Blooms taksonomi består av seks hierarkisk ordnede klasser som hver beskriver en kognitiv ferdighet. De seks kategoriene er *Knowledge, Comprehension, Application*, *Analysis*, *Synthesis* og *Evaluation* (Krathwohl, 2002, s. 212). For illustrasjon av taksonomien og den hierarkiske ordningen, se figur 1.



Figur 1. Bloom's Taxonomy (Webb, 2014, s. 64)

I *Knowledge* ligger det å kunne gjengi innlært fagstoff. Med *Comprehension* menes at eleven kan plukke ut relevant fagstoff og gjengi dette med egne ord. I *Application* ligger å kunne bruke denne kunnskapen i adekvate situasjoner. *Analysis* handler om å kunne se sammenhenger og *Synthesis* om å kunne trekke egne slutninger. Den høyest rangerte kompetansen, *Evaluation*, handler om i hvilken grad eleven kan bedømme noe ut fra gitte kriterier (Krathwohl, 2002). Det som menes med at Blooms taksonomi er ordnet hierarkisk er at kompetansene er ordnet fra enkelt til komplekst og fra konkret til abstrakt (Krathwohl, 2002). Dette betyr at en oppgave som krever *Evaluation* for å løses, normalt vil oppfattes vanskeligere for elevene enn en oppgave som kun krever *Comprehension*.

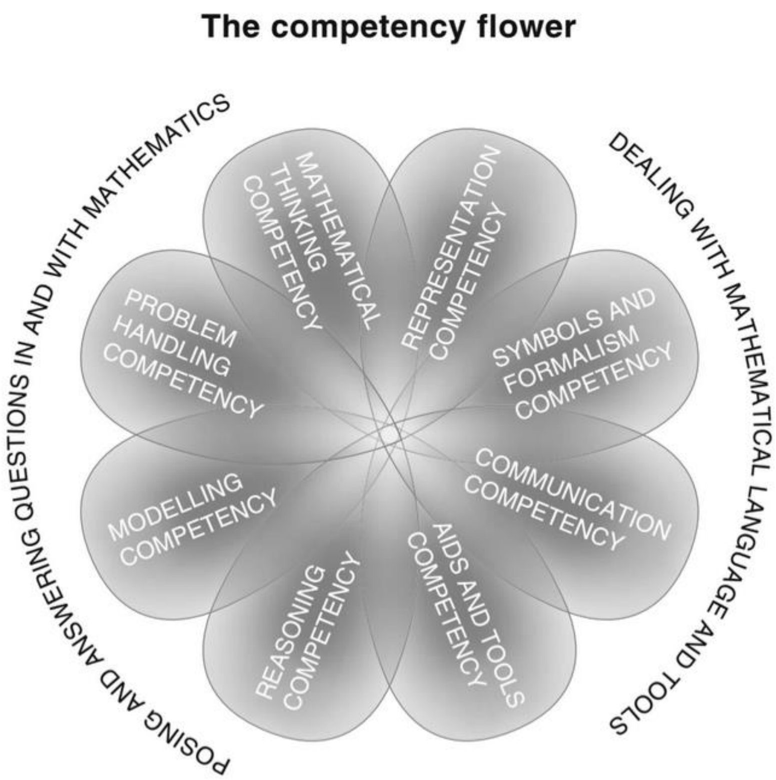
Blooms taksonomi ble tidlig kritisert, blant annet av Hans Freudenthal og Chris Ormell, for å være dårlig egnet til å beskrive matematikk (Krathwohl, 2002). Noe av kritikken rettet seg mot at taksonomien favoriserte prosess over innhold (Kilpatrick, 2014). Krathwohl (2002) svarte, gjennom en revidert versjon av Blooms taksonomi, på nettopp denne kritikken ved å separere kunnskap fra de kognitive prosessene i Blooms opprinnelige modell. På denne måten beveget han seg fra en endimensjonal til en todimensjonal modell hvor faglig innhold utgjorde den ene dimensjonen og prosessene som skal brukes til å behandle dette innholdet utgjorde den andre (Krathwohl, 2002).

2.2.2 KOM-prosjektet: Åtte kompetanser for å mestre matematikk

Gjennom arbeidet med å utvikle en kompetansebasert læreplan i Danmark, fremfor den tradisjonelle emnebaserte læreplanen, utviklet prosjektet Kompetencer og matematiklæring (KOM) en beskrivelse av hva som kjennetegner mestring av matematikk (Niss, 2003).

I arbeidet med utforming av ny læreplan tok arbeidsgruppen utgangspunkt i spørsmålet om hva det vil si å mestre matematikk (Kilpatrick, 2014). For å svare på dette ble det utarbeidet en modell for hva som kjennetegner matematisk kompetanse. Gruppen identifiserte følgende åtte kompetanser som de mente var nødvendige for å mestre matematikk: *Tankegangskompetanse, problembehandlingskompetanse, modelleringskompetanse, resonnementskompetanse, representasjonskompetanse, kompetanse i symbolbruk og formalisme, kommunikasjonskompetanse* og *hjelpemiddelkompetanse* (Niss & Jensen, 2002, s. 45). Jeg har gjennomgående brukt Matematikksenteret (2006) sine oversettelser av de begrepene KOM-prosjektet bruker i sin rapport. De åtte kompetansene er illustrert i figur 2.

Fire av kompetansene handler om å kunne stille og svare på spørsmål om og med matematikk (Niss, 2015, s. 41, min oversettelse). *Tankegangskompetanse* handler om å beherske matematisk tankegang, å ha bevissthet rundt hvilke spørsmål som er karakteristiske for matematikken og å kunne se for seg hva slags type svar som forventes på en matematisk problemstilling. Innen tankegangskompetansen ligger også begrepsforståelsen. Den handler om å kjenne, forstå og å kunne bruke ulike matematiske begrep, å kjenne omfang og grenser til disse begrepene og å kunne skille mellom ulike matematiske påstander, antagelser og bevis. *Problembehandlingskompetanse* handler om å kunne identifisere, presentere og spesifisere ulike matematiske problemstillinger og å kunne løse disse. *Modelleringskompetanse* handler om å kunne analysere og avkode eksisterende matematiske modeller og å aktivt kunne modellere ut fra en gitt situasjon. I dette ligger det at man kan identifisere og strukturere den situasjonen som skal modelleres, at man kan matematisere situasjonen og at man kan bearbeide, validere, analysere og kritisere modellen. Modelleringskompetanse handler også om å kunne kommunisere med andre om modellen. *Resonnementskompetanse* handler om å kunne følge og vurdere argumentasjoner som legges frem av andre. I denne kompetansen ligger også å kunne tenke ut og gjennomføre uformelle og formelle resonnementer, å kunne omforme resonnementer og antakelser til gyldige bevis og å kunne skille ulike typer matematiske resonnement (Niss & Jensen, 2002).



Figur 2. KOM-studiens åtte kompetanser for å mestre matematikk (Niss, 2015, s. 41)

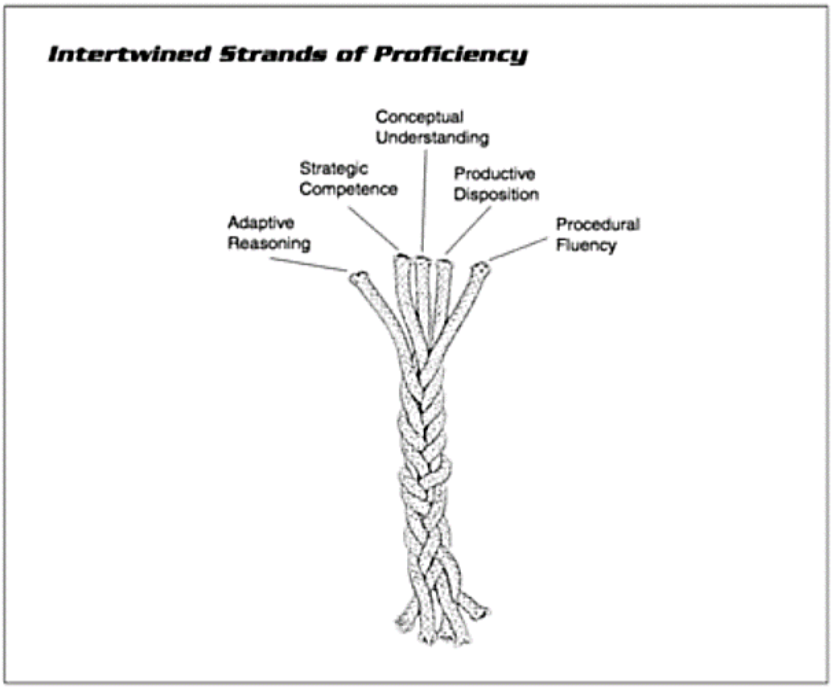
De siste fire kompetansene handler om å mestre det matematiske språk og om å kunne bruke hensiktsmessige verktøy (Niss, 2015, s. 41, min oversettelse). En elev som innehar *representasjonskompetanse* kan forstå, avkode, tolke og bruke ulike representasjoner av matematiske objekter, fenomener og situasjoner og se sammenhengen mellom ulike representasjoner. *Symbol- og formalismekompetanse* handler om å kunne bruke og avkode symbol- og formalismespråket i matematikk og å kunne oversette mellom dagligtale og matematisk symbolspråk. *Kommunikasjonskompetanse* handler om både å kunne sette seg inn i andres matematikkholdige tekster, skriftlige eller muntlige, og ikke minst å kunne uttrykke seg selv matematisk. *Hjelpemiddelskompetanse* handler om å kjenne til og kunne bruke varierte hjelpemidler, men også å kjenne rekkevidden og begrensningen til disse slik at man bruker dem på en hensiktsmessig måte (Niss & Jensen, 2002).

Dette rammeverket som KOM-prosjektet har utviklet for beskrivelse av matematisk kompetanse har vært av stor betydning for senere forskning på feltet, blant annet i arbeidet med utforming av oppgaver og vurdering av resultater i PISA-undersøkelsen (OECD, 2013). Det har også helt klart hatt konsekvenser for norsk skole, gjennom endring av læreplanen LK06 og gjennom utforming av nasjonale prøver (Matematikksenteret, 2006).

2.2.3 Kilpatrick: Fem tråder som tilsammen beskriver matematisk kompetanse

Amerikaneren Jeremy Kilpatrick og hans forskningsgruppe har hevdet at det er fem tråder som til sammen utgjør matematisk kompetanse: «The Strands of Mathematical Proficency» (Kilpatrick, Swafford, & Findell, 2001). Arbeidet med å definere hva som kjennetegner matematisk kompetanse var, som for andre forskningsgrupper, motivert av et ønske om å bedre matematikkundervisningen. Ved å forstå hva matematisk kompetanse består av, ville det være mulig å utforme undervisning med utgangspunkt i å stimulere endring i disse kompetanse-elementene hos elevene (Kilpatrick et al., 2001). Mens Niss og Jensen (2002) sine kompetanse-beskrivelser er prosessorienterte, er Kilpatrick sine fem tråder mer kognitive områder som kjennetegner den matematiske kompetansen. Disse fem trådene er illustrert i figur 3.

Conceptual understanding handler om at elevene må forstå matematiske begreper, operasjoner og relasjoner. *Procedural fluency* handler om at elevene skal være i stand til effektivt å utføre passende matematiske prosedyrer. *Strategic competence* handler om i hvor stor grad eleven kan formulere, representere og løse matematiske problemer. *Adaptive reasoning* handler om kapasiteten for logisk tenkning, refleksjon, forklaring og begrunnelse. *Productive disposition* handler om at elevene skal se på matematikken som nyttig, betydningsfull og aktuell for seg selv (Kilpatrick et al., 2001). Et viktig poeng med denne femdelte modellen er et fokus på hvordan disse trådene ikke kan forstås enkeltvis, men må ses i sammenheng med de andre: «The five strands are interwoven and interdependent in the development of proficiency in mathematics» (Kilpatrick et al., 2001, s. 116).



Figur 3. Five strands of Profiency (Kilpatrick et al., 2001, s. 117)

2.2.4 Programme for International Student Assessment (PISA)

De kompetansebeskrivelsene som hittil er beskrevet er endimensjonale med tanke på at de bare beskriver kognitiv kategori, med unntak av Krathwohl (2002) sin revisjon av Blooms taksonomi (se 2.2.1). Mange kompetanserammeverk i matematikk er imidlertid todimensjonale systemer med tematiske innholdsbeskrivelser i den ene dimensjonen og beskrivelser av kognitive prosesser i den andre. TIMSS og PISA; to store internasjonale undersøkelser som undersøker elevkompetanse, har rammeverk som organiseres rundt to slike dimensjoner; en innholdsmessig dimensjon og en kognitiv dimensjon (Kilpatrick, 2014). TEDS-M 2008, den internasjonale undersøkelsen av lærerstudenters kompetanse, brukte det samme matematiske rammeverket som TIMSS for å måle lærerstudentenes matematikkunnskap og er således også en representant for en slik todimensjonal kategorisering (Tatto et al., 2008).

PISA er en internasjonal undersøkelse som sammenlikner elever i ulike OECD-land innen tre sentrale fagområder: lesing, matematikk og naturfag. PISA-undersøkelsen gjennomføres hvert tredje år (Nortvedt, 2013). PISA ønsker å måle hvordan elevene bruker sin matematiske kompetanse til å løse anvendte, det vil si virkelighetsnære, problemer. Undersøkelsen er uavhengig av læreplaner, men tar sikte på å teste en form for *Mathematical literacy* som elevene vil trenge i hverdagen nå og i fremtiden (Nortvedt, 2013). Uttrykket *Mathematical literacy* har egentlig ingen god norsk oversettelse (Nortvedt, 2013), men den norske rapporten fra PISA 2012 oversetter dette med matematisk kompetanse. Jeg kommer til å bruke denne oversettelsen videre i oppgaven. PISA er avhengig av å definere hva som ligger i den matematiske kompetansen som skal testes for å utforme sine oppgavesett. En ekspertgruppe utnevnt av OECD har derfor utviklet et analyseverktøy ut fra en konsensus om hvilken kunnskap og hvilke ferdigheter som vil være nødvendige å inneha for å kunne delta som fullverdige borgere i fremtidens samfunn. Det undersøkelsen ønsker å teste er i hvilken grad skolene forbereder elevene på de utfordringene som de trolig vil møte i fremtiden (Nortvedt, 2013). Rammeverket bygger blant annet på NCTM sine beskrivelser av hva som ligger i matematisk forståelse (OECD, 2013).

PISA definerer matematisk kompetanse på følgende måte (OECD, 2013, s. 25):

> *Mathematical literacy is 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 recognize the role that the mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens*.

PISA trekker frem elevene som aktive problemløsere og kompetansene knyttes derfor til det å kunne bruke matematikk (OECD, 2013). De tre prosessene *formulere*, *bruke* og *vurdere* defineres som de tre prosessene som elevene må mestre for å håndtere problemløsings-situasjonen. I rammeverket til PISA defineres hva som ligger i disse prosessene. Å *formulere* innebærer at eleven skal gjenkjenne og identifisere muligheter til å bruke matematikk ved å lage en matematisk problemstilling med utgangspunkt i en virkelig situasjon. *Å bruke* innebærer at elevene skal bruke matematiske begreper, fakta og prosedyrer og resonnere rundt dette slik at de kan løse et problem som allerede har fått en matematisk form. *Å vurdere* handler om elevenes evne til å reflektere over matematiske løsninger og til å tolke og vurdere løsningene opp mot hverandre og opp mot den konteksten som problemet oppstod i (Nortvedt, 2013).

PISA definerer videre hvilke kompetanseområder elever trenger for å kunne mestre de tre prosessene for problemløsing. PISA har tatt utgangspunkt i arbeidene til Niss (se 2.2.2) og hans åtte kompetanser og har utformet syv kompetanseområder (OECD, 2013). De syv kompetanseområdene i PISA er (Nortvedt, 2013, s. 48):

- *Kommunisere med, i og om matematikk*
- *Matematisere og modellere matematiske og virkelige situasjoner*
- *Representere matematiske størrelser og bruke hensiktsmessige matematiske representasjoner i oppgaveløsing*
- *Resonnere og argumentere matematisk*
- *Planlegge, velge ut og bruke problemløsingsstrategier*
- *Bruke symbol- og formelspråk*
- *Velge ut og bruke hensiktsmessige matematiske verktøy og hjelpemidler*

For beskrivelse av hva som ligger i disse kompetanseområdene, henviser jeg til beskrivelsen av Niss og Jensen (2002) sine åtte kompetanser (se 2.2.2). De ulike kompetansene tas i bruk i større eller mindre grad i hver av de tre prosessene *å formulere, bruke* og *vurdere*. Disse kompetanseområdene beskriver dermed hva som må til for at en elev kan bli en matematisk problemløser og få matematisk kompetanse (Nortvedt, 2013).

2.2.5 Trends in International Mathematics and Science Study (TIMSS)

TIMSS er en internasjonal, sammenlignende studie av realfagundervisning i skolen (Grønmo et al., 2012). I motsetning til PISA er TIMSS læreplanbasert (Mullis, 2014). Hensikten i denne studien er å måle en form for skolekunnskap fremfor en allmenndannende matematikkunnskap som PISA ønsker å måle. Rammeverket til TIMSS Matematikk organiseres rundt to dimensjoner; en innholdsmessig dimensjon og en kognitiv dimensjon. Den kognitive dimensjonen beskriver hvilke kognitive ferdigheter det er forventet at elevene kan ta i bruk ved en gitt matematisk kontekst (Grønmo, Lindquist, & Arora, 2014). Ifølge Pedersen (2013, s. 5) bygger TIMSS den kognitive delen av rammeverket på Krathwohl (2002) sin revisjon av Blooms taksonomi som er beskrevet i 2.2.1.

TIMSS deler inn i tre kognitive kategorier; *å kunne, å anvende* og *å resonnere* (Onstad, 2010). De definerer videre hva som ligger i disse kategoriene. *Å kunne* betyr blant annet å huske definisjoner, kjenne terminologi og gjenkjenne matematiske objekter og uttrykk, å kunne utføre matematiske algoritmer og å kunne hente informasjon fra grafer, tabeller, matematiske tekster og andre kilder. *Å anvende* betyr å bruke kunnskapene sine til å velge hensiktsmessige metoder, strategier og verktøy, til å representere matematisk informasjon på en hensiktsmessig måte og til å modellere situasjoner for rutineoppgaver. Rutineoppgaver er oppgaver hvor innhold og løsningsmetoder er kjente for elevene. *Å resonnere* innebærer logisk og systematisk tenkning fra elevene. Dette innebærer at elevene kan avdekke innholdet i et matematisk problem, velge nødvendige strategier for å løse oppgaven og bruke kunnskap, representasjoner og prosedyrer til å løse ikke rutinepregede oppgaver. Kompetansen innebærer i tillegg at elevene skal kunne vurdere hvor godt egnet egne løsningsstrategier er, validitet til egne løsninger og muligheter for generalisering (Onstad, 2010).

2.3 Tidligere studier: Bruk av rammeverk til kategorisering av oppgaver

2.3.1 Mathematics Expert Group (MEG): Konkretisering av rammeverk fra PISA

Turner, Dossey, Blum, og Niss (2013) gir i sin artikkel «Using Mathematical Competencies to Predict Item Difficulty in PISA: A MEG Study» en mulig konkretisering av rammeverket til PISA. Formålet med denne studien var å lage et kategoriseringssystem for å forutsi vanskelighetsgraden til matematikkoppgaver som blir gitt i PISA (Turner et al., 2013). Med PISA sitt kategoriseringssystem som utgangspunkt delte denne forskergruppen matematikkompetanse inn i seks delkompetanser. Delkompetansene bygger på PISA sine, men hjelpemiddelkompetansen er fjernet. De seks delkompetansene som MEG-studien kategoriserte dataene i er (Turner et al., 2013, s. 24):

- *Reasoning and argument*
- *Communication*
- *Modelling*
- *Representation*
- *Solving problems mathematically (referred to as problem solving)*
- *Using symbolic, formal and technical language and operations (referred to as Symbols and formalism)*

Reasoning and argument handler om evnen til logisk tenkning, til å trekke logiske slutninger og resonnere matematisk. *Communication* handler om hvor vanskelig oppgaven er å lese og hvor vanskelig det er å tolke hva oppgaven spør om. I tillegg handler denne kompetansen om graden av kommunikasjon som kreves for å formidle svar og utregninger. *Modelling* handler om i hvilken grad eleven må oversette mellom matematikk og noe utenfor matematikken, for eksempel ved å lage en matematisk modell fra en praktisk problemstilling. *Representation* handler om å tolke, bruke og oversette mellom ulike matematiske representasjoner. *Solving problems mathematically* handler om å velge og bruke en overordnet matematisk strategi for å løse oppgaven. *Using symbolic, formal and technical language and operations* handler om å kunne bruke, manipulere og forstå et matematisk symbolspråk, å kjenne matematiske regler og å kunne bruke disse.

Forskergruppen delte inn hver kategori i fire nivåer fra 0 til 3 og ga nøyaktige beskrivelser for hvilken kompetanse som kreves for å være innen et gitt nivå. Et eksempel på en slik beskrivelse er kompetansebeskrivelsen for hva som kreves for å være på nivå 0 innen *Communication* (Turner et al., 2013, s. 27):

> *Understand a short sentence or phrase relating to a single familiar concept that gives immediate access to the context, where it is clear what information is relevant, and where the order of information matches the required steps of thought.*

Kjennetegn på å være på nivå 1 innen samme kategori ble beskrevet som følger (Turner et al., 2013, s. 27):

> *Identify and extract relevant information. Use links or connections within the text that are needed to understand the context and task, or cycle within the text or between the text and other related representations. Any constructive communication required is simple but beyond the presentation of a single numeric result.*

Selv om begge disse beskrivelsene er kjennetegn på kommunikasjon i matematikk, har kategori 1 større grad av både abstraksjon og kompleksitet enn kategori 0. Denne videre nivådelingen nyanserer derfor rammeverket betraktelig. Det er viktig å merke seg at nivå 0 ikke nødvendigvis betyr fravær av kategorien. Det kan også bety at oppgaven krever kompetanse innen kategorien, men på det laveste nivået.

Turner et al. (2013) prøvde ut sitt kategoriseringssystem på 8 personer som hver kodet 48 oppgaver fra PISA 2003 og PISA 2006. Deretter undersøkte de sammenhengen mellom kategoriseringen som forskerne gjorde i de ulike nivåene og hvor vanskelig oppgavene slo ut for elevene i de faktiske PISA-undersøkelsene. Det viste seg at forskergruppen, ved hjelp av den nivåinndelte konkretiseringen av rammeverket, kunne forutse hvilke oppgaver som ville slå vanskelig ut i PISA. Gjennom regresjonsanalyser konstruerte forskerne modeller for disse sammenhengene. Klassifiseringene kunne forklare over 70 % av variansen i elevenes resultater. Det viste seg også at koding med utgangspunkt i dette rammeverket var konsistent innad i forskningsgruppen, noe som ga studien høy reliabilitet (Turner et al., 2013).

2.3.2 Pedersen (2013): Eksempel på bruk av rammeverk fra TIMSS

Pedersen (2013) har gjennomført en studie hvor hun gjør en innholdsanalyse av matematikk-oppgavene som ble gitt i TIMSS Advanced i 1998 og 2008 og læreplanene for R1/R2 fra LK06 og 2MX/3MX fra R94. Målet med studien var å vurdere samsvaret mellom oppgavene som gis i TIMSS Advanced og læreplanene til de elevene som gjennomfører disse testene (Pedersen, 2013). Fordi dette var en sammenliknende studie som analyserte innhold i eksamensoppgaver, så jeg klare likhetstrekk til min oppgave og dermed muligheter for en tilsvarende tilnærming. Pedersen (2013) tok utgangspunkt i rammeverket til TIMSS Advanced 2008 (Grønmo, Onstad, & Pedersen, 2010) som er beskrevet over, men reviderte dette noe ved å inkludere Blooms taksonomi. Resultatrammeverket bestod av de seks kategoriene *å kunne, å forstå, å bruke, å analysere, å evaluere* og *å skape.* Kategoriene var løst hierarkisk ordnet (Pedersen, 2013). Revisjonen ga smalere og mer nyanserte kategorier enn TIMSS sitt rammeverk alene og den hierarkiske ordningen ga muligheter til å sammenlikne nivå for oppgaver som er kategorisert med utgangspunkt i dette rammeverket.

Pedersen kategoriserte både TIMSS-oppgavene og læreplanmålene med hensyn på matematisk tema og kognitiv kategori. Målene og oppgavene ble dermed kategorisert i en todimensjonal matrise hvor matematisk tema utgjorde den ene dimensjonen og kognitiv kategori den andre (Pedersen, 2013). Metodeverktøyet for dataanalysen bygget på Porter (2002) sitt analyse-verktøy for å måle innhold i undervisning og undervisningsmateriell (Pedersen, 2013). Pedersen brukte disse verktøyene til å finne prosentvis fordeling av oppgaver i de ulike kategoriene. Deretter sammenliknet hun innholdet i de ulike kategoriene i fordelingsmatrisene for læreplanene og TIMSS Advanced-oppgavene. Studien viste et moderat samsvar mellom oppgavene i TIMSS Advanced og innholdet i læreplanene, noe som impliserte at vekten i de norske læreplanene ikke nødvendigvis gjenspeiles i TIMSS Advanced-oppgavene (Pedersen, 2013).

2.4 Undervisningskunnskap i matematikk

2.4.1 Shulman: Konseptualisering av lærerkunnskap

Lee S. Shulman var en av de første som konseptualiserte lærerkunnskap (Ball et al., 2008). Professoren ved Universitetet i Stanford har forsket på undervisning og lærerutdanning i en årrekke og er en ledende skikkelse innen didaktisk forskning (Grossman & Wineburg, 2001). Shulman (1986) viste at god undervisning avhenger av mer enn gode fagkunnskap. Han hevder at det finnes en nødvendig kunnskapsbase for lærere og foreslår å dele denne inn i tre kategorier: *Content Knowledge*, *Pedagogical Content Knowledge* og *Curricular Knowledge* (Shulman, 1986).

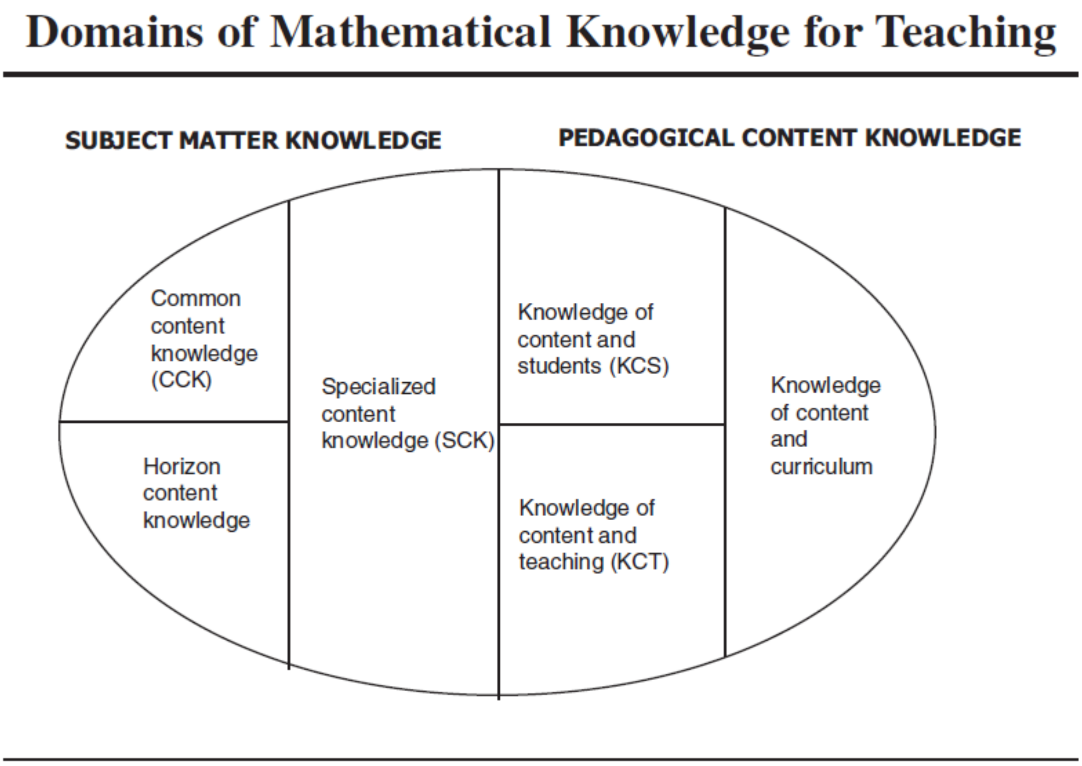
Fauskanger, Mosvold, og Bjuland (2010) oversetter *Content Knowledge* med fagkunnskap. Fagkunnskap innebærer kunnskap om fakta og begreper, men også kunnskap om hvordan disse er organisert og henger sammen. Denne typen kunnskap er fundamentet for lærerens forståelse og legger grunnlaget for elevenes forståelse: "The teacher has special responsibilities in relation to content knowledge, serving as the primary source of student understanding of subject matter" (Shulman, 1987, s. 9). *Pedagogical Content Knowledge* oversettes av Fauskanger et al. (2010) med fagdidaktisk kunnskap. Dette er kunnskap om hvordan fagkunnskapen kan brukes i en undervisningskontekst og omfatter blant annet kunnskap om hva slags forkunnskap elevene har, hvilke misoppfatninger som er vanlige, eksempler og forklaringer som er spesielt godt egnet til å hjelpe elevene og kunnskap om hvordan undervisningen kan tilpasses slik at den enkelte eleven lærer best (Shulman, 1986). *Curricular Knowledge* er kunnskap om relevant undervisningsmateriell, blant annet lærebøker og læreplaner. I denne kunnskapskategorien ligger også kunnskap om hva elevene kan fra før og hva de skal lære senere. Dette omtaler Shulman (1986) som *Vertical knowledge*, jeg kommer til å bruke horisontkunnskap når jeg bruker dette begrepet i diskusjonen senere i oppgaven.

2.4.2 Ball og medarbeidere: Videre konseptualisering av lærer-kunnskap

Det hadde fra før vært udiskutabelt at lærerens fagkompetanse var viktig, men betydningen av kunnskap om det pedagogiske innholdet og hvordan elever lærer, en kunnskap som var profesjonsspesifikk for lærerne, var nytt ved Shulmans forskning (Ball et al., 2008). Deborah

Ball og hennes medarbeidere videreførte dette arbeidet ved å konseptualisere lærerkunnskap ytterligere og ved å knytte denne spesifikt til matematikk.

Mathematical knowledge for teaching (MKT) er et overordnet begrep for å beskrive den fagkunnskapen som er nødvendig for å undervise i matematikk (Enge & Valenta, 2010). MKT oversettes i de norske rammeplanene for lærerutdanningen med *undervisningskunnskap i matematikk* (Mosvold & Fauskanger, 2012). Ball et al. (2008) presenterer i sin artikkel «Content Knowledge for Teaching: What Makes it Special?» arbeidet med å definere denne profesjonsspesifikke kunnskapen. Undervisningskunnskap i matematikk ble delt inn i to hovedområder tilsvarende Shulman (1986) sin inndeling i *Content Knowledge* og *Pedagogical Content Knowledge*. De utdypet dette videre ved å dele opp hvert av de to hovedområdene i tre mindre kunnskapsområder (Ball et al., 2008). Se figur 4 for en illustrasjon av disse.



Figur 4. Undervisningskunnskap i matematikk (Ball et al., 2008, s. 403)

Common Content Knowledge er fagkunnskap som er felles på tvers av profesjoner, mens *Specialized Content Knowledge* er kunnskap som er nødvendig spesielt for lærere. Det er mulig for eksempel å tenke seg at lærere i matematikk trenger å vite mer om sammenhenger, bevis og forklaringer enn en som ikke skal forklare dette til andre, men bare anvende matematikken. *Horizon Content Knowledge* handler om en bevissthet rundt hvordan matematiske temaer henger sammen gjennom den matematikken som kommer videre i de ulike læreplanene. For eksempel vil det være nødvendig at lærere i første klasse kjenner hvilken matematikk som kommer i tredje klasse for å kunne gi elevene det nødvendige fundamentet. Dette gjenkjenner vi som *Vertical knowledge* hos Shulman (1986).

Knowledge of Content and Students handler om kunnskap om hvordan elevene tenker. Læreren må kunne forutse hva elevene finner motiverende, hva de synes er vanskelig og hva de synes er lett. I dette ligger også kunnskap om elevenes forkunnskap. *Knowledge of Content and Teaching* handler om innsikt i ulike pedagogiske tilnærminger og hvordan læreren best kan legge frem et matematisk tema eller en oppgave for elevene. Dette kunnskapsområdet kombinerer kunnskap om undervisning med kunnskap om faget, matematikken. Kunnskap om hvilke eksempler på tavla som vil gi elevene en dypere forståelse er et eksempel på kunnskap som ligger innunder dette kunnskapsområdet. *Knowledge of Content and Curriculum* handler om at lærerne må kjenne innhold i fag og læreplaner for å vite hva som skal vektlegges i undervisningen. Behovet for horisontkunnskap som er beskrevet over, gjør at læreren også må kjenne innholdet i læreplaner i matematikkfag som elevene har hatt og som de skal ha senere.

Ball et al. (2008) beskriver altså seks kunnskapsområder som til sammen utgjør undervisningskunnskap i matematikk. De konkluderer med at lærerne først og fremst må kunne faget de underviser. «Teachers must know the subject they teach. Indeed there may be nothing more fundamental to teacher competency. The reason is simple: Teachers who do not themselves know a subject well are not likely to help students learn this content» (Ball et al., 2008, s. 404). Denne kunnskapen er imidlertid ikke tilstrekkelig. Lærerne må ha kunnskap som gjør at matematikkunnskapene kan formidles til elevene på en best mulig måte. De må altså ha kunnskap som går utenfor å kunne løse avanserte matematikkoppgaver. «It seems unlikely that just knowing more advanced math will satisfy all of the content demands of teaching» (Ball et al., 2008, s. 404)

2.4.3 Principles and Standards for School Mathematics

Den amerikanske organisasjonen National Council of Teachers of Mathematics (NCTM) har arbeidet med å lage standarder for bedre matematikkundervisning og læring av matematikk. De utarbeidet først tre slike standarder – *Curriculum and Evaluation Standards for School Mathematics* (1989), *Professional Standards for Teaching Mathematics* (1991) og *Assessment Standards for School Mathematics* (1995). Disse var det første kjente forsøket av en profesjonell organisasjon på å utvikle og formulere eksplisitte mål for lærere og politikere og har derfor vært av stor politisk betydning (National Council of Teachers of Mathematics, 2000). I 1996 startet NCTM arbeidet med å revidere standardene, et arbeid som ledet frem til *Principles and Standards for School Mathematics* i 2000 (National Council of Teachers of Mathematics, 2000).

Et av områdene som diskuteres er *The Teaching Principle* – hva som kjennetegner god matematikkundervisning. God undervisningskunnskap defineres som en viktig del av dette, sammen med blant annet god relasjon til elevene og et læringsmiljø som utfordrer elevene faglig (National Council of Teachers of Mathematics, 2000). Den gode undervisningskunnskapen er en syntese av fagkunnskap og didaktiske kunnskap hvor begge er helt sentrale for å skape undervisning med god læringseffekt for elevene (National Council of Teachers of Mathematics, 2000, s. 17):

> *To be effective, teachers must know and understand deeply the mathematics they are teaching and be able to draw on that knowledge with flexibility in their teaching tasks. They need to understand and be committed to their students as learners of mathematics and as human beings and be skillful in choosing from and using a variety of pedagogical and assessment strategies*.

The Teaching Principle gir ingen klare definisjoner av hva god undervisningskunnskap er, men går gjennom kjennetegn på dette. Det er sentralt at lærerne har dyp forståelse for stoffet som skal undervises, noe som inkluderer å kjenne ulike representasjoner av det samme fenomenet. I tillegg fremstår det sentralt at de har god kjennskap til læreplaner og at de kjenner til elevforutsetninger, vanlige misoppfatninger og hvilke tilnærminger som vil kunne fungere for ulike elever (National Council of Teachers of Mathematics, 2000).

2.4.4 COACTIV-studien: Evaluering av betydningen av fagkunnskap og fagdidaktisk kunnskap i undervisningen

Den tyske studien «Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress» (COACTIV) har arbeidet med å konseptualisere undervisningskunnskap blant annet gjennom å kartlegge hvilken betydning fagkunnskap og pedagogisk kunnskap hos lærere har for elevenes læring (Baumert et al., 2010). Det teoretiske rammeverket til studien tar utgangspunkt blant annet i Shulman (1986) sine arbeider med å definere lærerkunnskap som kunnskap satt sammen av kunnskap om faginnhold, *Content Knowledge* (CK), og pedagogisk kunnskap knyttet til det å undervise elever, *Pedagogical Content Knowledge* (PCK).

COACTIV-studien bygger på den tyske utvidelsen av PISA-undersøkelsen fra 2003/2004 og fulgte de elevene som ble testet her og deres lærere. Fordi gruppen ønsket å vurdere lærerkompetanse var det nødvendig å utarbeide kunnskapstester for å vurdere lærerkunnskap, både CK og PCK. Arbeidet med dette medførte en videre konseptualisering av hva som ligger i CK og PCK for å kunne teste disse separat. Arbeidet tok utgangspunkt i mange tidligere studier, blant andre arbeidene til Shulman og Ball (Baumert et al., 2010). I tillegg tok den utgangspunkt i Krauss, Baumert, og Blum (2008) sin hierarkiske klassifisering av matematikkunnskap (Baumert et al., 2010, s. 142):

(a) The academic research knowledge generated at institutes of higher education
(b) A profound mathematical understanding of the mathematics taught at school
(c) A command of the school mathematics covered at the level taught
(d) The mathematical everyday knowledge that adults retain after leaving school

I COACTIV-studien ble CK for matematikklærere konseptualisert som den andre typen i oversikten over: En dyptgående forståelse av den matematikken som undervises i skolen. PCK ble konseptualisert som matematikkunnskap som er direkte knyttet til undervisning og elevenes kunnskap og som dermed gjør matematikken tilgjengelig for elevene (Baumert et al., 2010).

Resultatene av kunnskapstestene fra PISA ble satt opp mot data som handlet om klasserominstruksjoner og elevresultater. På denne måten ble det mulig å si noe om betydningen av de ulike komponentene av lærerkunnskap. Studien konkluderte med at det så ut som at didaktisk kunnskap (PCK) hadde større betydning enn lærerens fagkunnskap (CK) for elevenes læring. Lærerens didaktiske kunnskap (PCK) påvirket blant annet den individuelle læringsstøtten og kvaliteten på undervisningen (Baumert et al., 2010). Betydningen av fagkunnskap i matematikk ble imidlertid trukket frem som viktig, selv om dette ikke så ut til å være en signifikant sammenheng med elevprestasjoner når man satte testparametere fra lærertestene og elevresultatene opp mot hverandre. Betydningen av tilstrekkelig fagkompetanse ble heller anerkjent som en nødvendig forutsetning for å kunne utvikle god didaktisk kunnskap:

Our theoretical assumption is that PCK is inconceivable without a substantial level of CK but that CK alone is not a sufficient basis for teachers to deliver cognitively activating instruction that, at the same time, provides individual support for students' learning. (Baumert et al., 2010, s. 163).

2.4.5 KOM-prosjektets beskrivelse av lærerkompetanse i matematikk

Som beskrevet i 2.2.2 munnet KOM-prosjektet ut i en beskrivelse av åtte kompetanser som kjennetegner mestring av matematikk. I tillegg ble det utarbeidet en beskrivelse av hva som ligger i lærerkompetanse i matematikk (Niss, 2006). For det første må matematikklæreren selv inneha alle de åtte kompetansene som KOM-prosjektet definerer som nødvendige for å mestre matematikk. For det andre må en kompetent matematikklærer mestre følgende didaktiske og pedagogiske kompetanser (Niss, 2006, s. 46-47, mine oversettelser):

- *Curriculum competency:* Å kunne analysere, relatere til og implementere læreplanen i undervisningen
- *Teaching competency:* Å kunne utvikle, planlegge, organisere og gjennomføre undervisningen. Dette inkluderer å skape et rikt spekter av læringssituasjoner, finne passende undervisningsmateriale, inspirere og motivere elevene, diskutere pensum med elevene og tilpasse undervisningen til elevgruppen
- *Uncovering of learning competency:* Å kunne avdekke, tolke og analysere hvordan elever lærer matematikk. Kompetansen inkluderer også evnen til å tilpasse undervisningen til den enkelte student
- *Assessment competency:* Å kunne identifisere, karakterisere og kommunisere om elevenes læringsutbytte. Dette inkluderer å velge ut varierte og passende vurderingssituasjoner og å gi fruktbare tilbakemeldinger på dette til elevene
- *Collaboration competency:* Evnen til samarbeid med kollegaer, foreldre, ledelse og andre om matematikkundervisning
- *Professional development competency:* Utvikling av sin egen kompetanse som matematikklærer (en metakompetanse). Dette inkluderer å delta og engasjere seg på ulike arenaer for selvutvikling, som etterutdanningskurs og konferanser.

Niss (2006) understreket altså betydningen av at lærerne selv mestrer den matematikken som de skal undervise i tillegg til å trekke frem hvilke fagdidaktiske og pedagogiske kompetanser som en god matematikklærer må mestre. At lærerne selv må mestre matematikken handler om at de har tilstrekkelig fagkunnskap. *Curriculum competency*, *Teaching competency*, *Uncovering of learning competency* og *Assessment competency* avhenger av god fagdidaktisk kunnskap. Undervisningskunnskap, bestående av fagkunnskap og fagdidaktisk kunnskap, står dermed også i Niss sin beskrivelse som helt sentralt for å gjennomføre god undervisning.

2.4.6 TEDS-M 2008: Vurdering av undervisningskunnskap i matematikk hos lærerstudenter

TEDS-M er en internasjonal undersøkelse som har vurdert kunnskapene til lærerstudenter. TEDS-M sitt rammeverk for beskrivelse av lærerkunnskap i matematikk bygger blant annet på arbeidene til Ball og hennes forskningsgruppe samt det tyske COACTIV-prosjektet (Tatto et al., 2008). Undersøkelsen testet både matematikkunnskap (MCK) og matematikkdidaktisk kunnskap (MPCK). Studien delte MPCK videre inn i kunnskap om læreplaner i matematikk, kunnskap om undervisningsplanlegging og kunnskap om tilrettelegging og gjennomføring av matematikkundervisning (Grønmo & Onstad, 2012). En lærer med god undervisningskunnskap i algebra er dermed ut fra rammeverket i TEDS-M en lærer som har gode fagkunnskap innen fagområdet algebra og som kan tolke de læreplanmålene som dekker området algebra i læreplanen. Læreren kan videre bruke dette til å planlegge, tilrettelegge og gjennomføre undervisning slik at det legges til rette for at også elevene kan få innsikt i fagkunnskapen.

I rapporten «Mange og store utfordringer» (Grønmo & Onstad, 2012) oppsummeres noen hovedfunn fra TEDS-M-undersøkelsen i 2008. Lærerstudenter i alle typer utdanningsprogram i Norge presterte lavt på oppgaver i matematikk i TEDS-M 2008 når man sammenlikner med et internasjonalt gjennomsnitt. Spesielt for lærere som skal undervise på ungdomstrinnet lå det norske gjennomsnittet langt under det internasjonale. Fra undersøkelsen fremstod det som spesielt viktig å styrke norske lærerstudenters rene matematikkunnskap i aritmetikk og algebra (Grønmo & Onstad, 2012). De norske lærerstudentene skåret noe høyere enn det internasjonale snittet på matematikkdidaktiske oppgaver. Undersøkelsen viste samtidig hvordan lærerstudentene skåret lavt på noen slike oppgaver fordi den matematikkdidaktiske kunnskapen er så tett sammenvevd med den matematiske. Spesielt gjaldt dette lærerstudentenes resultater på oppgaver for ungdomstrinnet (Grønmo & Onstad, 2012). Resultatene fra undersøkelsene antydet dermed en mulig sammenheng hvor man ser at kunnskap i matematikkdidaktikk forutsetter en tilstrekkelig faglig basis i matematikk.

2.5 Oppsummering av teorikapittelet

Som vi har sett i dette kapittelet, finnes det mange ulike muligheter for beskrivelser av hva som kjennetegner matematikkompetanse. Bloom beskriver et hierarkisk ordnet system av kategorier som kjennetegner en faguavhengig kompetanse. Gjennom KOM-prosjektet har Niss og hans forskningsgruppe laget en åttedelt modell, hvor de beskriver åtte komponenter som kjennetegner mestring av matematikk (Niss & Jensen, 2002). PISA bygger sitt rammeverk nettopp på arbeidene til KOM-prosjektet og trekker frem syv av komponentene fra KOM-prosjektet som tilstrekkelige for at elevene skal kunne bli problemløsere (Nortvedt, 2013). Kilpatrick et al. (2001) snakker på sin side om fem tråder som ikke kan sees separat, men som sett i sammenheng kjennetegner kompetanse i matematikk. Kompetansebeskrivelsene til Bloom, Niss og Kilpatrick er endimensjonale med tanke på at de kun beskriver kognitive ferdigheter og er uavhengige av det tematiske innholdet. Rammeverkene til PISA og TIMSS er på sin side todimensjonale ved at de beskriver både det matematiske temaet og kognitive kategorier.

Jeg ønsker å kategorisere mine oppgaver med hensyn på både tematisk innhold og kognitiv dimensjon, og jeg ønsker å se på både hvilke matematiske temaer som testes og hvilke kognitive utfordringer lærerstudentene møter. Jeg bestemte meg for å ta utgangspunkt i rammeverket til enten PISA eller TIMSS fordi disse er utviklet nettopp med sikte på å kategorisere oppgaver og fordi de har både en innholdsmessig og en kognitiv dimensjon. TIMSS har færre, men bredere og mer omfattende kompetansebeskrivelser enn PISA. Noe annet som skiller disse to fra hverandre, er at TIMSS sine kategorier i større grad er ordnet hierarkisk. Dette betyr at dersom jeg ønsker å vurdere vanskelighetsgraden til en oppgave, vil det muligens være enklere å kategorisere med utgangspunkt i TIMSS sitt rammeverk. På den annen side vil et rammeverk med bare tre kategorier sannsynligvis gi meg et litt lite nyansert bilde.

Pedersen (2013) viser at en utvidet versjon av TIMSS sitt rammeverk, hvor Bloom er inkludert, egner seg godt til å sammenlikne læreplanmål og oppgaver. Dette gir meg en indikasjon på at et tilsvarende rammeverk ville kunne egne seg for meg i mitt analysearbeid. Revisjonen gir smalere og mer nyanserte kategorier enn TIMSS sitt rammeverk alene og den hierarkiske ordningen gir muligheter til å sammenlikne nivå for oppgaver som er kategorisert med utgangspunkt i dette rammeverket, noe som vil være viktig også i mitt arbeid. Gjennom MEG-studien har vi empiri som støtter opp under at kompetansebeskrivelsene til PISA er relevante for hvordan matematikken faktisk oppleves for elevene (Turner et al., 2013). Fordi rammeverket er nyansert og fordi det er utviklet nettopp med sikte på å forutsi vanskelighetsgrad til oppgaver har jeg vurdert det som sannsynlig at også dette vil kunne egne seg godt til å gjennomføre min studie.

De ulike rammeverkene for beskrivelse av matematisk kompetanse er sammenfallende med tanke på at matematikkompetanse hos alle blir beskrevet som sammensatt av flere mindre kompetanser. Dessuten er de sammenfallende med tanke på at de ulike rammeverkene har delkompetanser som beskriver den samme ferdigheten. For eksempel inneholder *Evaluation* i Blooms taksonomi mye av den samme kompetansen som *Resonnementskompetanse* i KOM-prosjektet, *Adaptive reasoning* hos Kilpatrick, *Resonnere og argumentere* i PISA og *Å resonnere* i rammeverket til TIMSS (Kilpatrick et al., 2001; Krathwohl, 2002; Niss & Jensen, 2002; Nortvedt, 2013; Onstad, 2010). Dette betyr at uavhengig av hvilket rammeverk jeg velger for kategorisering av mine data, vil bildet som tegnes av om oppgavene tester en helhetlig matematisk kompetanse eller ikke være representativt også med utgangspunkt i den teorien som er presentert i dette kapittelet, men som ikke brukes som utgangspunkt for kategoriseringen.

Når forskere konseptualiserer lærerkunnskap, er det to kunnskapsområder som trekkes frem – fagkunnskap og fagdidaktisk kunnskap. De to henger sammen og læreren må inneha begge disse kunnskapene for å kunne gjennomføre god matematikkundervisning (Ball et al., 2008; Baumert et al., 2010; Grønmo & Onstad, 2012; National Council of Teachers of Mathematics, 2000; Shulman, 1986). COACTIV-studien finner at fagkunnskap er en parameter som ikke påvirker elevenes læringsutbytte i like stor grad som den didaktiske kunnskapen. Forskerne i denne studien antyder imidlertid at tilstrekkelige fagkunnskap er nødvendig for å kunne ha god fagdidaktisk kunnskap (Baumert et al., 2010). Andre forskere trekker fagkunnskap frem som helt avgjørende for lærerens undervisningskompetanse og som noe som må ligge som et fundament for å kunne undervise matematikk (Ball et al., 2008; Grønmo & Onstad, 2012; Shulman, 1987).

3 Metode

For å svare på problemstillingen presentert i kapittel 1 har jeg samlet og analysert matematikkoppgaver som gis til eksamen ved ulike GLU 5-10-utdanninger i Norge. Jeg har delt mellom didaktiske og matematikkfaglige oppgaver, telt opp antallet av disse, og jeg har kategorisert matematikkoppgavene med hensyn på matematisk tema og kognitiv kategori etter bestemte rammeverk. Med bakgrunn i dette har jeg sett på forskjeller og likheter mellom lærestedenes valg av oppgaver, hvilke temaer som inkluderes og i hvilken grad oppgavene utfordrer studentene innen de valgte kognitive kategoriene. I dette kapittelet beskrives metodevalgene som er gjort for å gjennomføre denne studien. I kapittel 3.1 beskrives valg av forskningsdesign. Kapittel 3.2 beskriver utvalget for studien, mens kapittel 3.3 beskriver datamaterialet inkludert i studien og kapittel 3.4 hva slags analysemetode som har blitt brukt. Kodeprosedyren presenteres i kapittel 3.5. I kapittel 3.6 presenteres en oversikt over hva som er gjort for å ivareta validiteten til studien.

3.1 Forskningsdesign

Utgangspunktet for valg av studie var et ønske om å kunne si noe om innholdet i norsk grunnskolelærerutdanning i matematikk. Jeg ønsket meg helst data som ga mulighet for generalisering, slik at mine funn kunne brukes til å si noe generelt om alle grunnskolelærerutdanningene i Norge. En tverrsnittstudie ville derfor vært ideelt for meg. Tverrsnittstudien kjennetegnes av at den prøver å tegne et bredt bilde av den sitasjonen som ønskes beskrevet ved et gitt tidspunkt (L. Cohen, Manion, & Morrison, 2011). Studien krever at det velges et representativt utvalg fra den populasjonen man ønsker å undersøke, utvalget må representere bredden og variasjonen i utvalget(L. Cohen et al., 2011). Jeg kunne gjennomført en slik studie ved å plukke ut noen eksamensoppgaver fra hvert lærested. Ideelt ville jeg kodet et fullt sett med eksamensoppgaver fra ett kull fra samtlige grunnskolelærerutdanninger i Norge for å gjøre en slik tverrsnittstudie.

Det er to begrensninger som har styrt mengden av datamateriale og dermed påvirket forskningsdesignet til denne studien. For det første gjorde tidsomfanget til denne oppgaven det umulig å kode så mange oppgaver som det ville tilsvare å dekke oppgaver fra alle læresteder i Norge. For det andre har et flertall av de norske grunnskolelærerutdanningene en vesentlig del av sin vurdering som muntlig eksamen. Dette kan ikke kodes på samme måte som de skriftlige eksamensoppgavene og disse lærestedene ble derfor utelukket fra studien.

Valget falt derfor på å gjennomføre en komparativ casestudie hvor jeg har analysert og sammenliknet innholdet i skriftlige eksamensoppgaver ved noen GLU 5-10-utdanninger i Norge. Den komparative studien kjennetegnes nettopp av at flere tilfeller eller populasjoner blir studert og sammenliknet (L. Cohen et al., 2011). Casestudien studerer en eller noen få enheter inngående for å trekke deskriptive slutninger, eller eventuelt gi innsikt i hvorvidt et fenomen fører til noe annet (Sterri & Wæhle, 2014). Den gir dermed mulighet til å gå i dybden med analyse, kategorisering og drøfting slik som jeg har gjort i denne studien. Ulempen med å velge en slik studie er at det er utfordrende å generalisere eventuelle funn til hele populasjonen, i mitt tilfelle matematikkundervisningen ved alle grunnskolelærerutdanninger i Norge. En slik komparativ casestudie vil imidlertid kunne indikere noen tendenser, spesielt dersom man inkluderer studien i en rekke av flere studier som belyser den samme problemstillingen (L. Cohen et al., 2011). Selv om casestudien tar for seg en eller noen få enheter og studerer disse grundig, brukes metoden ofte nettopp for å kaste lys over en hel klasse av viktige fenomener ut fra den grundige og helhetlige beskrivelsen av det enkelte tilfellet (Sterri & Wæhle, 2014).

Casestudier har en rekke kjennetegn. For det første er en rik beskrivelse av casene i fokus og den presenterer gjerne en rekke hendelser som er relevante for casen eller casene. Videre blander den gjerne en beskrivelse av casene med en analyse av dem. Forskeren er gjerne på en eller annen måte integrert i casen og casen kan være linket til noe som kjennetegner personligheten til forskeren. På denne måten har casestudien ofte et personlig preg (L. Cohen et al., 2011). Min studie blir en litt utypisk casestudie hvor mange av kjennetegnene på en slik studie ikke ivaretas.

For det første ønsker jeg ikke en veldig rik casebeskrivelse som utgangspunkt for analysen. Årsaken til dette er at jeg ønsker å ivareta størst mulig grad av anonymitet. Jo mer detaljerte beskrivelser jeg hadde gitt av de valgte lærestedene, jo lettere ville det vært å avdekke hvilke læresteder jeg har inkludert i studien. Jeg mener også at casebeskrivelsen er mindre relevant hos meg fordi jeg bare ønsker å gi en beskrivelse av innholdet i eksamensoppgaver i matematikk ved noen lærerutdanninger og lete etter likheter og forskjeller mellom disse. Jeg har ikke hatt som mål å se på om ulik profil på læresteder gir ulikt resultat, jeg har altså ikke vært ute etter årsakssammenhenger. Eksempler på ulike profiler kunne vært universitet kontra høyskole, et sentralt lærested kontra et desentralt lærested eller eventuelt lærested med få studenter mot lærested med mange studenter. Mitt datagrunnlag ville uansett vært for lite til å diskutere slike årsakssammenhenger. Grunnet tilgang til data og tidsbegrensning ville jeg maksimalt kunne bruke ett lærested med hver av disse profilene og det ville ikke vært mulig å trekke generelle slutninger på grunnlag av dette. De slutningene som jeg, og eventuelt leseren, ville trukket hadde vært for svakt fundert.

Noe annet som skiller seg fra den typiske casestudien er at jeg ikke har vært integrert i casen eller hatt et personlig forhold til noen av casene. Min subjektivitet har kommet frem i analysen gjennom min tolkning av rammeverket, men ellers står jeg som en objektiv forsker på utsiden av alle casene.

Det skilles ofte mellom kvantitative og kvalitative tilnærminger i den pedagogiske forskningen (Kleven, Tveit, & Hjardemaal, 2014). Den kvalitative forskningsmetoden kjennetegnes av en nær relasjon mellom forsker og informant, et lite datamateriale og forskerens subjektivitet i analyseprosessen. Styrken til den kvalitative metoden er muligheten til helhetlig å vurdere enkelttilfellene. Den kvantitative metoden kjennetegnes blant annet av et større datamateriale og en objektiv forskerrolle og har sin styrke i eksplorerende eller hypotesedannende undersøkelser (Kleven et al., 2014). Mitt datamateriale består utelukkende av tekst. Tekstanalysen hører tradisjonelt til den kvalitative forskningstradisjonen (Ryghaug, 2002). Jeg skal imidlertid ikke analysere tekst med fokus primært på språket, men skal gjøre en kvantitativ innholdsanalyse hvor jeg kategoriserer i et bestemt rammeverk og gjør en opptelling. Målet har vært å sitte inne med data som det er mulig å gjøre statistiske beregninger med. Dette har vært med på å gi studien et mer kvantitativt preg. Min studie ligger dermed i skjæringspunktet mellom de kvalitative og kvantitative studiene.

3.2 Utvalg

Jeg har valgt en komparativ casestudie der jeg ser på eksamensoppgaver fra utvalgte læresteder. Det er 16 GLU 5-10-utdanninger i Norge som jeg har kunnet velge ut fra (Samordna opptak, 2015). For å velge læresteder har jeg fulgt noen utvalgskriterier. Ifølge L. Cohen et al. (2011, s. 143) må det tas stilling til fem nøkkelpunkter når man foretar et slikt utvalg: Utvalgsstørrelsen, representativiteten til utvalget, tilgjengeligheten, hvilken utvalgsstrategi som skal brukes og hva slags studie utvalget skal gjøres til.

Utvalget mitt har vært lite. Som nevnt over fjerner et lite utvalg mulighetene for generaliseringer, men det gir samtidig muligheter for meg til å gå i dybden i analyse av oppgavene. For at utvalget skulle bli så representativt som mulig ønsket jeg meg at det skulle være variasjon i om lærestedene var fra universitet eller høyskole, variasjon i størrelse på lærested og variasjon i hvor sentralt lærestedet var. Likevel er det viktig å påpeke at utvalget mitt har vært for lite til at det har vært representativt. Tilgjengeligheten til data har vært et viktig utvalgskriterium i min studie, ved mange av lærestedene er en stor andel av vurderingen muntlige eksamener. Disse lærestedene er ekskludert fra studien. Videre har jeg forsøkt å velge læresteder som har eksamensoppgaver tilgjengelig på sine nettsider fordi dette har gjort datainnsamlingsprosessen enklere. Jeg har også vurdert responsen når jeg har tatt kontakt med lærestedene og latt dette være med å påvirke utvalget mitt. Jeg ser derfor at tilgjengelighet av analyserbare data har vært en styrende faktor i datainnsamlingen.

Jeg endte opp med tre GLU 5-10-utdanninger hvor jeg samlet de skriftlige skoleeksamensoppgavene fra to fullstendige kull fra hvert lærested. Lærestedene har blitt anonymisert. En beskrivelse av datamaterialet kommer i kapittel 3.3.

3.3 Data

Mitt datamateriale består av skriftlige skoleeksamensoppgaver fra GLU 5-10-utdanningen i matematikk fra tre ulike læresteder. Disse er samlet inn fra lærestedenes nettsider og ved direkte kontakt med lærestedene. GLU 5-10-utdanningen består av 60 studiepoeng pedagogikk og tre fag à 60 studiepoeng. Jeg har inkludert to kull fra hvert lærested i studien og har samlet alle eksamensoppgavene som studenter fra disse kullene møtte i sine matematikkurs. Dette har resultert i et samlet datamateriale bestående av 530 eksamensoppgaver. Oppgavene er hentet fra eksamener som er gjennomført i årene 2011, 2012, 2013 og 2014. Eksamenene består av både matematikkfaglige oppgaver og matematikkdidaktiske oppgaver.

I min studie har jeg valgt å fokusere på de rene matematikkfaglige oppgavene, altså på hvordan de utvalgte lærestedene vurderer fagkunnskapen til sine studenter. Jeg anerkjenner at studentene må mestre det matematikkfaglige for å besvare de didaktiske oppgavene, slik at også didaktiske oppgaver vil være med på å teste fagkunnskapen hos studentene. I tillegg anerkjenner jeg at de matematikkdidaktiske oppgavene er en svært viktig del av undervisningskunnskapen som skal opparbeides hos lærerstudentene. Det beste ville være å finne egnede rammeverk og kategorisere begge typer oppgaver. Grunnet studiens omfang så jeg at jeg måtte velge å kategorisere enten matematikkoppgavene eller de matematikkdidaktiske oppgavene. Forskning viser at fagkunnskap må ligge til grunn for den didaktiske kunnskapen (Ball et al., 2008; Baumert et al., 2010; Shulman, 1986). Tonheim og Torkildsen (2010, s. 213) illustrerer dette på en fin måte:

> *Gode matematikklærarar må altså ha det faglege på plass, då vert det enklare å sjå samanhengar og kunne vise elevane fleire innfallsvinklar. Ein matematikklærer med god fagdidaktisk kompetanse, men som ikkje kan det matematikkfaglege, er utenkjeleg. Ein lærar som er fagleg sterk i matematikk, kan gripe tak i dei didaktiske utfordringane*.

Dessuten avdekker TEDS-M 2008 store utfordringer knyttet til fagkunnskapen til norske lærerstudenter i matematikk (Grønmo & Onstad, 2012). Jeg valgte derfor kun å kategorisere de matematikkfaglige oppgavene. Jeg talte opp, men kategoriserte ikke videre, de oppgavene som helt eller delvis hadde en matematikkdidaktisk formulering.

Mine data består utelukkende av eksamener som er gjennomført som skoleeksamener ved de tre inkluderte lærestedene. Med skoleeksamen mener jeg at de er gjennomført ved lærestedet og uten mulighet til samarbeid eller annen type kommunikasjon mellom studentene. Ingen av disse lærestedene har muntlige skoleeksamener, men to av dem har mappevurdering som en del av vurderingen. Mappevurderingene består av mest skriftlige oppgaver, både matematikkfaglige og matematikkdidaktiske, som studentene kan gjennomføre hjemme med det de har tilgjengelig av hjelpemidler. Ved det ene lærestedet inngår også en muntlig komponent i mappevurderingen. Lærestedet uten mappevurdering har en stor didaktikkoppgave som studentene får som hjemmearbeid. Oppgavene fra mappevurderingen og den store didaktiske oppgaven er ikke inkludert i mitt datamateriale.

3.4 Analyse

Dataene mine består utelukkende av tekstmateriale, nemlig eksamensoppgaver. Det finnes flere tekstanalytiske metoder: Innholdsanalyse, hermeneutikk, diskursanalyse og strukturalisme er noen slike (Ryghaug, 2002). Jeg skal analysere tekstene for å kunne kategorisere og sammenlikne innholdet i disse, så i mitt tilfelle er en kvantitativ innholdsanalyse det aktuelle. Ett av den kvantitative innholdsanalysens hovedmål er nettopp å kunne bearbeide og sammenlikne et stort tekstmateriale (Ryghaug, 2002). Den kvantitative innholdsanalysen forsøker å beskrive innholdet på en nøytral og objektiv måte. For å oppnå dette må man dele opp innholdet i mindre, håndterbare deler eller analyseenheter. Det må finnes et rammeverk av kategorier som analyseenhetene kan plasseres i. Kategoriene må konstrueres før selve kodingen og må være slik at uavhengige personer skal kunne plassere innholdet likt i disse (Ryghaug, 2002).

En svakhet med tekstanalyse som metode er at kodeskjemaet splitter opp innholdet i kategorier som analyseres hver for seg slik at teksten som helhet ikke oppdages (Ryghaug, 2002). Dette er imidlertid ikke et stort problem for meg i denne oppgaven fordi det nettopp er enkeltelementene, de enkelte oppgavene, som er forskningsobjektene. Innholdsanalysen har en utfordring knyttet til subjektivitet i tolkningen. Selv om man kaller metoden for objektiv, må teksten tolkes og forstås. Allerede når forskeren lager kodeskjema gjøres det tolkninger av innholdet og når forskeren skal registrere innholdet gjøres ytterlige tolkninger (Ryghaug, 2002). Jeg har valgt et eksisterende rammeverk for å kategorisere mine data, noe som har redusert subjektiviteten i kodingen. Valg av rammeverk for kategoriseringen er beskrevet i de neste avsnittene.

3.4.1 Valg av rammeverk for koding med hensyn på matematisk tema

For å gjøre en tematisk kategorisering har jeg brukt norske læreplaner i matematikk fra ungdomstrinnet og videregående skole.

Jeg har først tatt utgangspunkt i læreplanen i matematikk for 8.-10. trinn fra Kunnskapsløftet 2006 (LK06). Denne læreplanen deler inn i følgende emneområder : *Tall og algebra*, *Geometri*, *Måling*, *Sannsynlighet, statistikk og kombinatorikk* og *Funksjoner* (Utdanningsdirektoratet, 2013). Der dette var mulig kategoriserte jeg oppgavene i henhold til beskrivelsene i denne læreplanen. I noen oppgaver har jeg vært nødt til å gå utenfor denne læreplanen fordi den ikke har dekket temaene i oppgavene jeg har kodet. Dette gjaldt for eksempel oppgaver som omhandler grenser, derivasjon, integraler, komplekse tall og statistiske fordelinger. I disse tilfellene har jeg gått inn i andre læreplaner fra LK06 og sjekket om jeg har funnet temaet der. Jeg har i dette arbeidet brukt læreplanen for Matematikk 1T (Utdanningsdirektoratet, 2013), Matematikk R1 (Utdanningsdirektoratet, 2006a), Matematikk R2 (Utdanningsdirektoratet, 2006a), Matematikk S2 (Utdanningsdirektoratet, 2006b) og Matematikk X (Utdanningsdirektoratet, 2006c). Navnene på emneområdene i disse læreplanene er imidlertid ikke helt samsvarende. Det som heter *Tall og algebra* i læreplanen for 8.-10. trinn har tilsvarende områder i andre læreplaner med andre navn, for eksempel *Algebra* i læreplanen for R1 og *Tallteori* i læreplanen for Matematikk X. Det er imidlertid ikke større forskjeller enn at det har vært mulig for meg å bruke kategoriene for tematisk inndeling fra læreplanen i matematikk for 8.-10. trinn.

I arbeidet med å kode oppgaver har jeg kodet hvor mange av oppgavene som ligger innenfor læreplanen i matematikk for 8.-10. trinn og hvor mange av oppgavene som går utover det som dekkes i læreplanen for ungdomstrinnet. Dette er interessant fordi lærerstudentene som får oppgavene som er inkludert i min studie har kompetanse til å undervise opp til og med 10. trinn, det vil si alle emner som ligger i læreplanen i matematikk for 8.-10. trinn, men ikke på videregående nivå.

I noen oppgaver må lærerstudentene bruke ferdigheter fra flere av emneområdene, for eksempel vil de i en derivasjonsoppgave bruke kompetanse som hører innunder både *Tall og algebra* og *Funksjoner*. Dette gjelder stort sett oppgaver som dekker *Tall og algebra* i tillegg til et annet emneområde. I disse tilfellene har jeg gjort en vurdering av hvilken eller hvilke kompetanser som er mest fremtredende i løsning av oppgaven og kategorisert oppgaven under dette emneområdet. Dette har medført at det er flere oppgaver som tester kompetanse som ligger under kategorien *Tall og algebra* enn det som kommer frem i mine resultater.

3.4.2 Valg av rammeverk for koding i kognitive kategorier

I teorikapittel 2.2 er det presentert ulike rammeverk for å beskrive kognitive kategorier i matematikk. Valget mitt stod mellom å bruke rammeverket utviklet i MEG-studien (Turner et al., 2013) og et rammeverk som bygger på TIMSS Advanced og Blooms taksonomi tilsvarende Pedersen (2013) fordi begge har vist seg godt egnet til å gjennomføre studier med kategorisering av oppgaver slik som jeg har ønsket. Jeg fikk opplæring i MEG-studien av en forsker ved Institutt for lærerutdanning og skoleforskning (ILS) som arbeider med dette i sin doktoravhandling. Min veileder hjalp meg med å sette meg inn i rammeverket og analysemetoden som Pedersen har brukt i sitt arbeid. En del av arbeidet med å velge rammeverk, var å prøvekode oppgaver som ikke ble brukt i studien. Dette gjorde jeg fordi jeg ønsket å se hvor lett det var å kode i de ulike rammeverkene, og fordi jeg ønsket å se hvordan de kodede dataene så ut, hvilke analyser jeg kunne gjøre på disse og hva jeg kunne lese ut fra analysene.

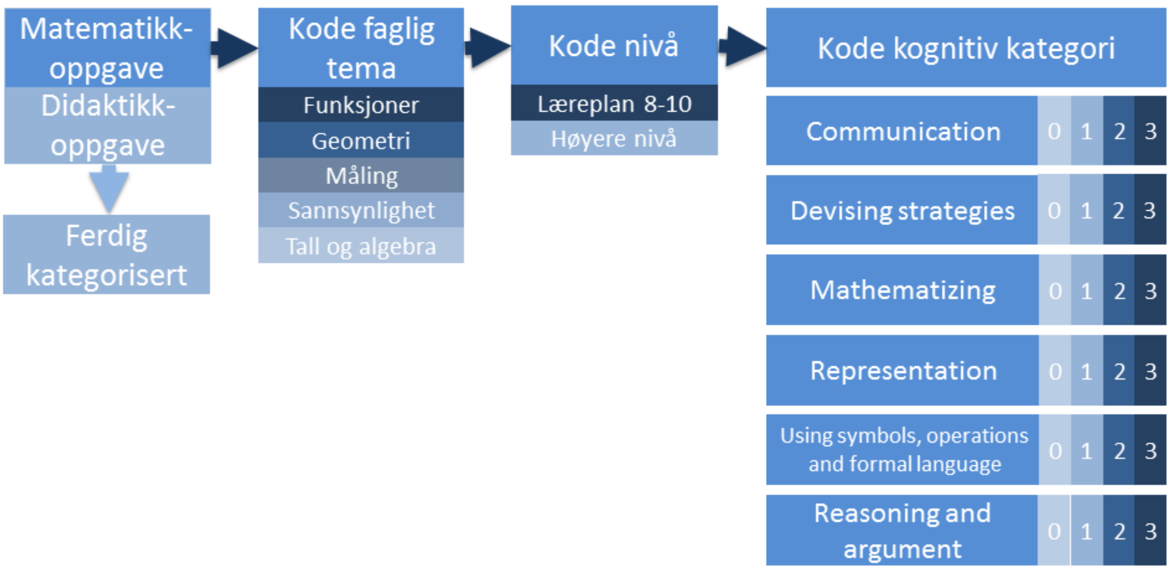
Ved det endelige valget vektla jeg rammeverkets tydelighet, tilgjengelig kompetanse ved eventuelle spørsmål til rammeverket, formatet til de ferdig kategoriserte dataene og hvilken informasjon jeg kunne hente. Kategorisering med utgangspunkt i rammeverket fra MEG-studien ga en mer detaljert beskrivelse av de ulike kategoriene enn Pedersen (2013) sin konkretisering av TIMSS Advanced sitt rammeverk ga. Spesielt var den detaljerte nivåbeskrivelsen utviklet i MEG-studien nyttig. Rammeverket i MEG-studien er i utgangspunktet utviklet med tanke på å forutsi vanskelighetsgraden til PISA-oppgaver (Turner et al., 2013). I følge Turner et al. (2013) kunne bruk av dette rammeverket faktisk forklare over 70 % av variansen i elevenes resultater. Det var derfor rimelig å anta at rammeverket ville egne seg godt til å si noe også om vanskelighetsgraden til de oppgavene som jeg skulle kategorisere. Dette var også mitt inntrykk da jeg prøvekodet oppgaver. En annen faktor som talte for å velge dette rammeverket var muligheten til å bruke den andre personen ved ILS som jobber med dette rammeverket som sparrepartner, informant og dobbeltkoder.

Det endelige valget endte derfor på MEG-studien sitt rammeverk for kategorisering av matematikkompetanse (Turner et al., 2013). Etter at jeg hadde bestemt meg for å bruke dette rammeverket, ble jeg gjort oppmerksom på at det er utarbeidet en ny versjon av dette. Dette er publisert våren 2015 (Turner, Blum, & Niss, 2015). Forskeren ved ILS som arbeider med dette rammeverket fikk, gjennom mailkorrespondanse med forskergruppen, den nye versjonen av rammeverket tidligere. Jeg fikk også tilgang til denne versjonen før jeg begynte kodearbeidet slik at jeg kunne bruke den i min studie.

Navnene på de seks kategoriene er endret noe i den nye versjonen, men innholdet er til en stor grad sammenfallende med MEG-studien fra 2013 (Turner, Blum, & Niss, 2015). Beskrivelsene av hva som ligger i hver kategori er imidlertid mer detaljerte, det samme gjelder nivåbeskrivelsene innen hver kategori. Dette betyr at den reviderte versjonen gir en ytterligere nyansering av rammeverket. Den nyeste kodebeskrivelsen, med oppdaterte navn og beskrivelser av hva som ligger i de ulike nivåene, er vedlagt oppgaven (vedlegg 1). De nye navnene på kategoriene er *Communication (C)*, *Devising strategies (DS)*, *Mathematising (M)*, *Representation (R)*, *Using symbols, operations and formal language (USOFL)* og *Reasoning and argument (R&A)* (Turner et al., 2015, forkortelsene er mine egne). Jeg har kategorisert alle matematikkoppgavene til ett av fire nivåer innen hver av disse seks kategoriene.

3.5 Koding: Prosedyrebeskrivelser

For å illustrere hvordan arbeidet med å kode oppgaver har foregått, har jeg laget figur 5 som viser prosessen med å samle oppgaver og kategorisere disse.



Figur 5. Prosedyrebeskrivelse

3.5.1 Gjennomføring av kodingen

For å sikre kvalitet i kodingen, satte jeg meg godt inn i rammeverket i forkant. Som en del av dette forarbeidet gjennomførte en forsker ved ILS og jeg prøvekoding hvor vi satt sammen og kodet oppgaver. Vi kodet oppgaver hver for oss, sammenliknet kodingen og diskuterte oss frem til enighet der vi var uenige. Rett i forkant av at jeg kodet datamaterialet i denne studien, jobbet jeg også som koder i et forskningsprosjekt som den samme forskeren leder. Her fikk vi også en grundig innføring i rammeverket og vi fikk trene på å kode PISA-oppgaver og eksamensoppgaver fra 10. trinn i det aktuelle rammeverket. Dette ga god trening i bruk av rammeverket, slik at kodingen av datamaterialet for denne studien ble lettere.

Jeg løste alle matematikkoppgavene i datamaterialet mitt før jeg gjennomførte kodingen av disse. En del av nivåbeskrivelsene i rammeverket skiller blant annet på antall regneoperasjoner, eller antall trinn i utvikling av en matematisk strategi. Fullstendige løsningsforslag var derfor nødvendig for samtlige eksamensoppgaver. I noen tilfeller inkluderte dette alternative løsninger for å vurdere om oppgaven skulle kodes annerledes med en annen løsningsstrategi.

I etterkant av dette kodet jeg alle oppgavene med hensyn på tema og kognitiv kategori som illustrert i figur 5. Hver matematikkoppgave ble kodet til ett matematisk tema ut fra inndelingen i læreplanen i matematikk for 8.-10. trinn (se kapittel 3.4.1), det ble kodet om oppgaven tematisk lå i eller utenfor denne læreplanen og til slutt ble oppgaven kodet til nivå 0-3 innen hver av de seks kognitive kategoriene for matematisk kompetanse (se kapittel 3.4.2).

3.5.2 Illustrasjon av koding gjennom kodeeksempler

For å illustrere hvordan jeg har arbeidet med kodingen, følger to eksempler på koding av oppgaver.

Eksempelkoding 1:

Løs likningen. Vis framgangsmåten:

$$\frac{4x}{3} + 2 = x + \frac{13}{6}$$

Denne oppgaven er hentet fra et eksamenssett fra et av lærestedene og er en av oppgavene som er inkludert i mitt datamateriale. Oppgaven er kodet som en matematikkfaglig oppgave som hører innunder emnekategorien *Tall og algebra* og oppgaven er kodet til å ligge innenfor læreplanen i matematikk for 8.-10. trinn. Oppgaven byr ikke på vesentlige utfordringer når det kommer til å lese oppgaven og forstå hva den spør om. Oppgaven er derfor kodet til nivå 0 innen *Communication.*[1] I nivåbeskrivelsen for *Devising strategies* står det at dersom løsningsstrategien er gitt eller åpenbar, skal oppgaven kodes til nivå 0 innen denne kategorien (Turner et al., 2015). Jeg har kodet oppgaven til nivå 0 i *Devising strategies*. Den eneste løsningsstrategien her, slik jeg har tolket rammeverket, er å finne ut at man skal få x alene. Dette mener jeg er gitt når det står at oppgaven skal løses som en likning. Jeg anerkjenner at det er mulig å tolke det som at studenten må lage en strategi for hvordan likningen skal løses og at den dermed skal kodes til nivå 1 i *Devising strategies*. Jeg har derimot tolket denne typen oppgaver til å handle om å manipulere likningen algebraisk slik at man får x alene og dette er kompetanse som hører innunder *Using symbols, operations and formal language*. Denne kategorien er kodet til nivå 2 fordi studentene må jobbe med en ukjent, fordi de må kombinere flere «regler» og fordi det krever flere regnetrinn for å fullføre oppgaven. I tillegg har jeg

[1] For beskrivelse av nivåene 0-3 innen hver av de kognitive kategoriene henvises til vedlegg 1

vurdert at studentene trenger kompetanse som hører innunder *Reasoning and argument*; evnen til å utføre enkle resonnementer rundt hvordan likningen kan løses. Jeg har kodet denne kategorien til nivå 1. *Mathematising* er satt til nivå 0 fordi studentene ikke må oversette mellom matematikk og noe utenfor matematikken. I det valgte rammeverket skal begge retninger, både å oversette noe fra dagliglivet over til matematikk, men også å oversette et tallsvar til noe utenfor matematikken, inkluderes i koding av denne kategorien. Studentene skal ikke bruke noen annen form for representasjon enn likningen de skal løse. *Representation* er derfor kodet til nivå 0, da jeg anser dette som en «simple representation».

Eksempelkoding 2:

Snorre, en elev i 10. klasse, fikk oppgaven:

Prisen på en bil vil øke med 3 % hver år de neste 4 årene. Prisen er i år 250 000 kroner. Hva blir prisen om 4 år?

Han løste oppgaven slik:

$$\frac{250\,000 \cdot 3}{100} = 7500 \cdot 4 = 30\,000$$

$$250\,000 + 30\,000 = \underline{\underline{280\,000}}$$

Om fire år koster bilen 280000 kr

(a) Gi Snorre tilbakemelding på løsningen.

(b) Løs oppgaven over på to ulike måter.

Denne oppgaven er hentet fra et annet eksamenssett fra et annet lærested og er også en av oppgavene som er inkludert i studien. Oppgave (a) er kodet som en matematikkdidaktisk oppgave, mens oppgave (b) er kodet som en matematikkfaglig oppgave. Det er derfor bare oppgave (b) som er kodet med hensyn på tema og kognitiv kategori. Denne oppgaven er kodet til det matematiske temaet *Tall og algebra* og den er kodet til å ligge utenfor læreplanen i matematikk for 8.-10. trinn. Det er mulig å løse denne typen oppgave uten å bruke vekstfaktor og da vil oppgaven kunne kategoriseres innenfor læreplanen 8-10. Imidlertid ber oppgaveteksten om to løsningsmetoder der jeg regner med at vekstfaktor er den ene. Begrepet *vekstfaktor* nevnes første gang i læreplanen i 1P/1T (Utdanningsdirektoratet, 2013). Jeg har også undersøkt et av de vanlige læreverkene fra ungdomsskolen, «Grunntall 10» (Bakke & Bakke, 2011), og jeg finner heller ikke bruk av vekstfaktor her.

Jeg har kodet oppgaven til nivå 1 innen *Communication* fordi det er presentert informasjon som ikke er relevant for å løse denne delen av oppgaven, nemlig løsningen til Snorre. Dette kan virke forvirrende for studentene som skal løse oppgaven. Oppgaven krever at det utvikles to ulike løsningsstrategier, og det er ikke sagt noe i oppgaveteksten om hvordan dette skal gjøres. Derfor kodet jeg oppgaven til nivå 2 innen *Devising strategies*. Jeg har kodet kategorien *Mathematising* til nivå 1 fordi studentene må gå fra informasjon om at noe øker med 3 % til å lage en matematisk modell hvor de bruker vekstfaktor og se at denne må være 1,03. *Representation* er kodet til nivå 0 fordi studentene kun skal lese numeriske verdier fra en tekst. Studentene må kunne anvende prosent og vekstfaktor og utføre beregninger over flere trinn. Jeg har derfor kodet oppgaven til nivå 2 innen *Using symbols, operations and formal language*. Jeg mener at studentene må utføre enkle resonnementer for å løse oppgaven og har kodet til nivå 1 innen *Reasoning and argument*.

Når jeg har kodet, har jeg ført inn i kodeskjemaer. I tabell 1 vises kodeskjemaer for de to eksempeloppgavene over.

Tabell 1. Eksempel på føring av kodedata

	Tema	C	DS	M	R	USOFL	RA
Eksempeloppgave 1	*Tall og algebra*	0	0	0	0	2	1
Eksempeloppgave 2	*Funksjoner*	1	2	1	0	2	1

3.5.3 Valg jeg har tatt i forbindelse med koding i kognitiv kategori

I kodingen har jeg foretatt noen valg når det kommer til tolkning av oppgaver og av rammeverk. Jeg har forsøkt å holde meg konsistent til disse gjennom hele kodingsprosessen.

Hjelpemiddelkompetanse måles ikke i det valgte rammeverket (Turner et al., 2015). Jeg har derfor tatt hensyn til om studentene har hjelpemidler, for eksempel kalkulator og formelverk, tilgjengelig når jeg har kodet i de andre kategoriene. Spesielt har dette påvirket hvilket nivå oppgaven ender på i *Using symbols, operations and formal language* fordi kalkulator påvirker hvor mye regning som må gjennomføres og formelverk påvirker hvor mye studentene må kunne når de møter oppgaven.

Oppgaver med sannsynlighetsregning har vært vanskelige å kode fordi det er vanskelig å vite om studentene bruker et sett med regler som de kjenner eller om de resonnerer seg frem til svaret. Her har jeg valgt å anta at studentene til en stor grad benytter seg av regler, spesielt i tilfellene hvor disse er tilgjengelige i et formelhefte. Dette har medført at disse oppgavene blir kodet høyere i *Using symbols, operations and formal language* og lavere i *Reasoning and argument.* Konstruksjonsoppgaver er også en oppgavetype hvor det er vanskelig å vurdere om studentene har pugget løsningsstrategier eller om de resonnerer seg frem til løsningen. Jeg har valgt å kode denne typen oppgaver som at studentene kombinerer dette, men kanskje i større grad enn i sannsynlighetsoppgavene resonnerer seg frem til en løsning på oppgaven.

Et siste gjennomgående kodevalg har vært bare å se på den mottagende delen av *Communication*, ikke den konstruerende delen. Dette betyr at jeg ikke har vurdert i hvilken grad studentene må formulere svar og vise utregninger, men bare har vurdert hvor vanskelig det er å forstå innholdet i oppgaveteksten.[2] Rammeverket er i utgangspunktet utviklet for å kode PISA-oppgaver. Veldig mange av disse er flervalgsoppgaver hvor studentene ikke trenger å vise utregning. I mitt datamateriale krever alle oppgavene at studentene viser utregning. Dette medfører at alle oppgavene ville blitt kodet til nivå 2 eller 3 dersom jeg skulle tatt hensyn til den konstruerende delen. Dersom jeg skulle kodet også den konstruerende delen, ville jeg for eksempel ikke fått frem ulik kommunikativ utfordring i eksempelkoding 1 og eksempelkoding 2 i avsnitt 3.5.2. Jeg så det derfor som mest fruktbart å kode bare den mottagende delen.

3.6 Forskningskvalitet

Dette kapittelet vil handle om validitet, reliabilitet og etikk som alle er faktorer som påvirker forskningskvaliteten.

Validitet handler om i hvilken grad de slutningene vi trekker er gyldige (Kleven et al., 2014). I pedagogisk forskning deler vi ofte validiteten inn i tre deler: Begrepsvaliditet, indre validitet og ytre validitet (Kleven et al., 2014). Begrepsvaliditet handler om samsvaret mellom begreper slik de er definert teoretisk og begrepene slik vi lykkes med å operasjonalisere dem. Indre

[2] Se vedlegg 1 for beskrivelse av hva som ligger i *Communication*

validitet handler om i hvilken grad slutninger som gjøres om relasjoner mellom variabler er sanne. Indre validitet er altså knyttet til tolkningen av relasjoner mellom variabler slik de er operasjonalisert i undersøkelsen, mens om operasjonaliseringen er godt gjennomført, handler om begrepsvaliditet. Ytre validitet handler om gyldigheten til resultatene utenfor deres gitte kontekst, altså om muligheten for generalisering av funn (Kleven et al., 2014). L. Cohen et al. (2011) beskriver validitet og reliabilitet knyttet til casestudier spesielt. Begrepsvaliditet, indre validitet og ytre validitet beskrives som de viktigste typene validitet også når de knyttes konkret til casestudier.

Mine oppfatninger og min forståelse av læreplaner, rammeverk og oppgaver er eksempler på mulige trusler mot begrepsvaliditeten dersom mine operasjonaliseringer gjennom kodearbeidet ikke gjenspeiler den teorien som jeg ønsker. Den indre validiteten må vurderes dersom jeg trekker slutninger om sammenhenger i datamaterialet, for eksempel om forskjeller mellom lærestedene. Ytre validitet er aktuelt å vurdere dersom jeg ønsker å si noe om nivået på norske lærerstudenter ut fra nivået på oppgavene som gis til eksamen eller dersom jeg ønsker å generalisere mine funn til å si noe om norsk grunnskolelærerutdanning generelt. Som diskutert i kapittel 3.1 vil jeg i utgangspunktet ikke kunne gå inn på årsakssammenhenger med mitt datamateriale eller forsøke å generalisere til andre læresteder. Intensjonen er heller å beskrive situasjonen og eventuelle forskjeller og bruke dette som utgangspunkt for en diskusjon rundt problemstillinger knyttet til testing av matematikkfaglige kunnskaper hos lærerstudenter. Dette har bakgrunn i de utfordringene jeg møter for den indre og ytre validiteten med et begrenset datamateriale med kun tre læresteder. Begrepsvaliditeten blir derimot svært viktig i min studie for at funnene jeg beskriver skal kunne forstås i sammenheng med relevant teori.

Reliabilitet handler om påliteligheten til forskningen, om hvorvidt dataene påvirkes av tilfeldige og systematiske målefeil (Kleven et al., 2014). I mitt tilfelle vil dette handle om i hvilken grad kodingen har vært pålitelig. Det er to typer reliabilitet som er aktuelt å diskutere i forbindelse med kodearbeid tilsvarende min studie, interkoderreliabilitet og intrakoderreliabilitet (Turner et al., 2013). Interkoderreliabilitet handler om i hvilken grad kodingen gjennomføres likt av uavhengige personer (Turner et al., 2013). For å vurdere denne fikk jeg en annen person til å kode noen av oppgavene parallelt med meg, som beskrevet i kapittel 3.6.1. Intrakoderreliabilitet handler om i hvilken grad kodingen som én koder gjennomfører er konsistent gjennom hele kodeprosessen (Turner et al., 2013). For å sikre dette har jeg brukt mye tid på å sette meg inn i rammeverket som jeg har valgt å bruke, og jeg har øvd på koding slik som beskrevet i 3.5.1.

Turner et al. (2013) antyder at erfaring med koding og kunnskap om rammeverket er avgjørende for reliabiliteten.

En diskusjon av forskningskvaliteten til denne studien presenteres i kapittel 5.6.

3.6.1 Undersøkelse av samsvar med annen koder

For å styrke reliabiliteten til studien fikk jeg en annen koder til å kode en del av eksamensoppgavene parallelt med meg. Jeg kom i kontakt med en stipendiat som bruker det samme rammeverket i sitt arbeid med doktorgraden og han sa seg villig til å kode oppgaver for meg. Dette er samme forsker som jeg har gjennomført kodetrening sammen med. I alt 50 matematikkoppgaver, omtrent 10 % av datamaterialet, ble kodet av denne andre koderen. Oppgavene ble valgt ut med tanke på at de skulle representere alle læresteder, men oppgavene fra hvert av lærestedene ble trukket tilfeldig.

For å vurdere samsvaret i kodingen ønsket jeg å bruke statistiske analyser. Ifølge Altman (1991) er Cohens kappa den foretrukne statistiske analysen i interkoderreliabilitetsstudier. Cohens kappa, κ, blir brukt som et mål på hvor godt samsvar det er mellom to eller flere forskeres skårer på en gitt variabel, korrigert for eventuelt tilfeldig samsvar (J. Cohen, 1960). Tilfeldig samsvar vil si at to kodere velger samme kategori bare ut fra tilfeldighet. Skalaen til Cohens kappa varierer fra -1 til 1. En κ-verdi på 1 indikerer perfekt samsvar, mens en κ-verdi på 0 forteller at samsvaret er helt tilfeldig. Negative κ-verdier indikerer at samsvaret mellom koderne er mindre enn det enn kunne forvente ved helt tilfeldige svar (Fleiss & Cohen, 1973) . I realiteten bruker vi stort sett bare de positive κ-verdiene, det er derfor disse jeg kommer til å diskutere videre i oppgaven. Det er utviklet flere retningslinjer for sammenheng mellom κ-verdier og styrken til samsvaret. Altman (1991, s. 404) har foreslått en skala for sammenhengen mellom de positive κ-verdiene og samsvaret i en reliabilitetsstudie. Altmans skala er gjengitt i tabell 2.

Tabell 2. Sammenheng mellom kappaverdi og styrke på samsvaret (Altman, 1991, s. 404)

Kappaverdi	Styrke på samsvar
<0.20	Dårlig
0.21-0.40	Svak
0.41-0.60	Moderat
0.61-0.80	God
0.81-1.00	Svært god

Mine oppgaver er kodet i nivå 0-3 innen de seks kognitive kategoriene, hvor nivåene har stigende vanskelighetsgrad. Mine kategorier er derfor på ordinalnivå, hvilket innebærer at de kan ordnes i rekkefølge, men at avstanden mellom kategoriene ikke er definert (Kleven et al., 2014). For å vurdere samsvar i studier hvor det er benyttet ordinale kategoriske variabler vil en vektet Cohens kappa være et bedre mål på reliabiliteten enn en ordinær Cohens kappa (Fleiss & Cohen, 1973). Denne reliabilitetsmarkøren tar hensyn til at dataene er ordinale, ved ikke bare å ta hensyn til om oppgavene er kodet likt eller ikke, men også til hvor forskjellig oppgavene er kodet når samsvaret beregnes (Fleiss & Cohen, 1973). Det vil for eksempel si at dersom en koder har kodet en 0 og den andre 3, så vil det ha større negativ virkning på reliabilitetsmarkøren enn dersom den ene er kodet 1 og den andre 2. Dersom man hadde brukt en vanlig Cohens kappa uten vekting, ville disse to situasjonene hatt samme virkning på reliabilitetsmarkøren.

Jeg har ønsket å kjøre mine statistiske analyser i SPSS. Ifølge Hallgren (2012) er det ikke mulig å kjøre analyse av vektet Cohens kappa i dette programmet. Fleiss og Cohen (1973) har imidlertid vist at en vektet kappa med kvadratiske vekter tilsvarer en ICC(2,1), det vil si en "intra-class correlation coefficient" med tilfeldig utvalgt objekt og tilfeldig utvalgt koder, der man ser på absolutt overenskomst. Dette betyr at dersom jeg kjører en statistisk analyse av ICC(2,1), der jeg ser på absolutt overenskomst, vil jeg få et mål på en kvadratisk vektet kappa for mine data.

Med tilfeldig utvalgte kodere menes at de to koderne er tilfeldig utvalgt fra en større populasjon av kodere, og at vi er interessert i overenskomsten mellom disse to som et estimat for overenskomsten i populasjonen av kodere. Dette gjør at målet vi får presentert sier noe om hvor godt ulike forskere generelt koder med dette rammeverket, ikke bare disse to koderne sett i forhold til hverandre (Hallgren, 2012). Jeg hadde bare en koder som var aktuell å velge, og kan derfor ikke si at han ble valgt tilfeldig. Ifølge (Hallgren, 2012) har dette valget imidlertid ikke noe å si for ICC-målet, det har kun betydning når vi skal tolke dataene. Vi antar videre at oppgavene som er dobbeltkodet er tilfeldig utvalgt fra en større populasjon av oppgaver, på denne måten vil målet si noe om reliabiliteten til hele datamaterialet, ikke bare det som er kodet dobbelt. I mitt tilfelle ble oppgavene valgt strategisk med tanke på å representere alle lærestedene, men oppgavene innen hvert lærested ble plukket ved tilfeldig utvalg. Å se på absolutt overenskomst vil si at vi er interessert i hvorvidt koderne koder akkurat det samme, ikke bare om de er konsistente (Kim, 2013). Dersom vi bare så på konsistens ville vi fått en høy verdi dersom koderne kodet de samme oppgavene høyt og lavt i poengsum, selv om den ene jevnt over kodet høyere verdier enn den andre, men absolutt overenskomst gir oss høyere verdier jo likere koderne koder.

Når man kjører en ICC(2,1) der man ser på absolutt overenskomst, får man i SPSS opp to verdier: *Single-measures* og *Average-measures*. Jeg har oversatt dette til enkeltmål og gjennomsnittsmål. Forskjellen mellom disse er at gjennomsnittsmål måler hvor reliabelt et gjennomsnitt av koderne er, for å bruke dette gjennomsnittet som mål for samsvaret også i resten av datamaterialet som er inkludert i studien. Ved enkeltmål vurderes heller reliabiliteten til den enkelte koderen (Hallgren, 2012). Det er nettopp reliabiliteten til meg som enkeltkoder som jeg ønsker å undersøke, derfor er enkeltmål det mest aktuelle for min studie.

Jeg har tidligere presentert Altman (1991) sin tabell over κ-verdier og styrken til samsvaret i dataene som er kodet (tabell 2). Cicchetti (1994) beskriver hvordan man kan bruke ICC-verdier til å vurdere samsvar. Han vurderer ICC-verdier under 0.40 for «poor», verdier mellom 0.40 og 0.59 som «fair», verdier mellom 0.60 og 0.74 som «good» og Verdier over 0.75 som «excellent». Vi ser at det er stort samsvar mellom dette og Altman (1991) sitt forslag over sammenhengen mellom κ-verdier og styrken til samsvaret.

MEG-studien som jeg har tatt utgangspunkt i når jeg har valgt rammeverk for å kode i kognitive kategorier, har også vurdert samsvaret mellom åtte kodere som har kodet PISA-oppgaver med utgangspunkt i rammeverket. De har imidlertid brukt Cronbachs alfa til å vurdere dette (Turner et al., 2013). Cronbachs alfa brukes for å sjekke om flere elementer i en test tester samme konstrukt og gir korrelasjonen mellom disse (Sim & Wright, 2005). Turner et al. (2013) har ikke begrunnet hvorfor de har valgt å vurdere samsvaret ut fra denne parameteren. Jeg har, ut fra mine vurderinger, konkludert med at jeg ønsker å bruke en vektet Cohens kappa til å vurdere mitt samsvar. Jeg har imidlertid også beregnet Cronbach alfa-verdier for å kunne sammenlikne samsvaret i min oppgave med samsvaret i MEG-studien.

3.6.2 Etikk

Retningslinjene fra Norsk samfunnsvitenskapelig datatjeneste (NSD) sier tydelig at godkjenning bare må søkes dersom oppgaven inneholder personopplysninger (Norsk samfunnsvitenskapelig datatjeneste, 2015). Dette er ikke tilfellet i denne studien.

Denne studien har i utgangspunktet vært gjennomført som en anonym studie ved at jeg har holdt identiteten til de tre lærestedene skjult. Oppgaven er skrevet med en målsetning om så stor anonymisering som mulig. Det vil imidlertid være slik at dersom man finleser alle beskrivelser av læresteder med tanke på andel didaktiske oppgaver, andel mappevurderinger og andre karakteristika vil det kanskje være mulig å nøste seg frem til hvilke læresteder dette må være. Det anses at dette ikke er noe problem for de aktuelle lærestedene, spesielt siden de til en stor grad har valgt å ha eksamensoppgavene tilgjengelig på nett.

3.7 Praktisk gjennomføring

For andre som eventuelt skal gjennomføre tilsvarende studier senere, har jeg valgt å legge ved en tidsplan som beskriver tidsomfang for ulike faser av prosessen. Jeg har også lagt til noen kommentarer som utdyper dette.

Måned	**Hva som er gjort**
Oktober 2014 - februar 2015	Datainnsamling
Oktober 2014 – februar 2015	Arbeid med å finne egnet rammeverk og trening på å bruke det valgte rammeverket
Februar 2015 – april 2015	Koding av datamaterialet
April 2015 – mai 2015	Analyse av data
Januar 2015 – juni 2015	Skriving av oppgave

Kodingen av eksamensoppgavene har foregått over en lang tidsperiode, fra februar til april. Omfanget har vært stort med tanke på antall oppgaver, men også med tanke på å sette seg inn i læreplaner og å lære å bruke MEG-rammeverket til koding av denne typen oppgaver. Fordi dette arbeidet har krevd svært høy tilstedeværelse og konsentrasjon har jeg forsøkt å ikke kode i for lange strekk uten pause. Jeg har heller strukket kodingen over en lengre tidsperiode og kombinert dette med andre arbeidsoppgaver slik at det ikke har blitt hele dager med kun koding.

Datainnsamlingen har også vært en omstendelig og forholdsvis tidkrevende prosess. Dette arbeidet har foregått over en periode på mange måneder, med oppstart tidlig høsten 2014 og avslutning vinteren 2015.

4 Resultater

I det følgende presenteres resultatene fra analyse av eksamensoppgavene fra tre GLU 5-10-utdanninger som beskrevet i metodekapittelet. Matematikkoppgavene er kodet med hensyn på matematisk tema og hvilken kognitiv kategori som testes. En diskusjon om betydningen av funnene med eventuelle implikasjoner finnes i kapittel 5.

4.1 Fordeling mellom matematikkoppgaver og didaktiske oppgaver

Fra de tre lærestedene samlet jeg til sammen 530 oppgaver, hvorav 473 var matematikkoppgaver mens 57 var fagdidaktiske oppgaver. Tabell 3 fremstiller antall oppgaver og fordelingen mellom matematikkoppgaver og fagdidaktiske oppgaver ved hvert av lærestedene.

Tabell 3. Antall matematikkoppgaver og didaktiske oppgaver for lærestedene samlet og for hvert lærested separat

	Totalt	Matematikk	Didaktikk	Andel didaktikk
Lærested A	169	130	39	23 %
Lærested B	179	173	6	3 %
Lærested C	182	170	12	7 %
Totalt	530	473	57	11 %

Vi ser at de totale oppgaveantallene fra de tre lærestedene er veldig like. En forskjell som imidlertid peker seg tydelig ut er andelen oppgaver som er didaktiske, med et spenn fra 3 % ved lærested B til 23 % ved lærested A. Som beskrevet i kapittel 3.3 er det kun skriftlige skoleeksamensoppgaver som er inkludert i studien, og det er mulig at fordelingen ville sett annerledes ut dersom oppgavene fra mappevurderingene var inkludert.

4.2 Kategorisering med hensyn på matematisk tema

Jeg har kodet alle matematikkoppgavene med hensyn på matematisk tema slik som beskrevet i kapittel 3.4.1. De er altså kategorisert med utgangspunkt i læreplanen i matematikk for 8.-10. trinn sin inndeling i de fem kategoriene *Tall og algebra*; *Geometri*; *Måling*; *Sannsynlighet, statistikk og kombinatorikk;* og *Funksjoner* (Utdanningsdirektoratet, 2013). I tabell 4 har jeg presentert den prosentvise fordelingen av temaer for hvert av lærestedene og for alle lærestedene samlet.

Tabell 4. Prosentvis fordeling av oppgaver innen de ulike matematiske temaene

	Lærested A	Lærested B	Lærested C	Samlet
Funksjoner	32 %	27 %	38 %	32 %
Geometri	12 %	23 %	3 %	13 %
Måling	2 %	2 %	4 %	3 %
Sannsynlighet, statistikk og kombinatorikk	17 %	28 %	15 %	21 %
Tall og algebra	36 %	21 %	39 %	32 %

Vi ser at det er store forskjeller på hvor stor andel av oppgavene som er kodet innen de ulike emneområdene. For lærestedene samlet, er *Måling* den klart minste kategorien med 3 % av oppgavene, mens *Funksjoner* og *Tall og algebra* er de største med 32 % av oppgavene hver. Andelen oppgaver i de ulike emneområdene skiller seg noe mellom lærestedene. Kategoriene *Geometri*, *Sannsynlighet, statistikk og kombinatorikk* og *Tall og algebra* er kategoriene hvor variasjonen er størst.

Som beskrevet i metodekapitlet måtte jeg i en del tilfeller bruke andre læreplaner enn læreplanen i matematikk for 8.-10. trinn for å finne temaet som ble dekket i eksamensoppgaven (se 3.4.1). Kompetansemål fra disse læreplanene har til sammen dekket nesten alle temaene som har dukket opp i koding av eksamensoppgavene fra de tre lærerutdanningene. Unntaket er oppgaver som handler om tallsystemer. Det går igjen at lærestedene har mange oppgaver med dette, men jeg finner det ikke igjen i noen læreplan som brukes i dag. Det lå imidlertid i læreplanen for Matematikk 2P frem til 2013 (Cappelen Damm, 2013). Jeg har kodet oppgavene som omhandler tallsystemer i kategorien *Tall og algebra.*

Jeg har registrert hvor mange oppgaver som ligger innenfor og utenfor læreplanen i matematikk for 8.-10. trinn. I tabell 5 er denne fordelingen presentert for de ulike lærestedene. Hovedvekten av oppgaver dekker temaer som ligger utenfor pensum i ungdomsskolen når vi ser på alle lærestedene samlet. Vi ser imidlertid at lærested A skiller seg ut og har en lavere andel av sine oppgaver med temaer svarende til pensum utenfor ungdomstrinnet enn de to andre lærestedene. Fordelingen av oppgaver i de to kategoriene er forholdsvis lik for lærested B og lærested C.

Tabell 5. Fordeling av andel oppgaver i og utenfor læreplanen 8-10

	Andel innenfor læreplanen 8-10	**Andel utenfor læreplanen 8-10**
Lærested A	49 %	51 %
Lærested B	22 %	78 %
Lærested C	18 %	82 %
Samlet	30 %	70 %

4.3 Kategorisering med hensyn på kognitiv kategori

Som beskrevet i 3.4.2 har jeg kategorisert kognitiv kompetanse i rammeverket som ble utviklet gjennom MEG-studien (Turner et al., 2013). Jeg har altså kategorisert alle oppgavene i nivå 0-3 i de seks kategoriene *Communication (C)*, *Devising strategies (DS)*, *Mathematising (M)*, *Representation (R)*, *Using symbols, operations and formal language (USOFL)* og *Reasoning and argument (R&A)*. I det følgende presenteres funnene fra koding i disse kategoriene for alle de matematikkfaglige oppgavene.

4.3.1 Samlet fordeling

Tabell 6 gir en oversikt over andelen matematikkoppgaver som er kodet i hvert nivå i de seks kategoriene for alle lærestedene sett samlet.[3]

Tabell 6. Andel oppgaver i hvert nivå innen hver kognitive kategori for alle lærestedene samlet

	0	**1**	**2**	**3**
Communication	39 %	46 %	13 %	2 %
Devising strategies	67 %	27 %	5 %	0 %
Mathematising	62 %	28 %	10 %	0 %
Representation	72 %	21 %	7 %	0 %
Using symbols, operations and formal language	7 %	40 %	43 %	10 %
Reasoning and argument	17 %	54 %	25 %	4 %

[3] I tabellene som viser prosentvis andel, har jeg rundet av til nærmeste heltall. Det vil si at 0 % representerer verdier under 0,5 %. Dette tilsvarer maksimalt to av de 473 matematikkfaglige oppgavene som er inkludert i studien.

Det er nivå 0 og 1 som dominerer i de fleste kategoriene. Unntakene er kategoriene *Using symbols, operations and formal language* og *Reasoning and argument* hvor nivå 1 og 2 er de to største. For tre av de kognitive kategoriene, *Devising strategies, Mathematising* og *Representation*, er nivå 0 alene det helt klart dominerende nivået. En stor andel av oppgavene i datamaterialet kodes altså til det laveste nivået i kompetansene som handler om å lage en overordnet løsningsstrategi, om å oversette mellom matematikk og noe utenfor matematikken og om å jobbe fleksibelt med ulike matematiske representasjoner.

For å undersøke hvor stor andel av oppgavene som er kodet til det høyeste nivået, utførte jeg en frekvensanalyse. Den viste at 16 % av oppgavene skårer nivå 3 i minst én kategori. Tabell 6 viser at kategorien *Using symbols, operations and formal language* er den med størst andel oppgaver kodet til nivå 3. Dette er oppgaver som krever at lærerstudentene må kunne bruke flere regneoperasjoner og regler i samme oppgave. I *Devising strategies, Mathematising* og *Representation* er 0 % av oppgavene kodet til nivå 3, det er også en liten andel som kodes til dette nivået i *Communication* og i *Reasoning and argument*.

Vi har sett at nivå 0 er det største for halvparten av de kognitive kategoriene. En frekvensanalyse viser at 9,1 % av oppgavene er på nivå 0 i alle andre kategorier enn *Using symbols, operations and formal language.* 9,1 % av oppgavene har altså følgende kodeprofil: C:0, DS:0, M:0, R:0 USOFL:1,2 eller 3 og R&A:0. Dette er oppgaver hvor den eneste kompetansen som testes hos studentene er regelkunnskap og regneferdigheter. Dersom vi ser bort fra *Use of symbols, operations and formal language* og analyserer kun med hensyn på de andre kategoriene, skårer bare 48 % av oppgavene nivå 2 eller 3 i minst én kategori. Det betyr at mer enn halvparten av oppgavene skårer 0 eller 1 i alle andre kategorier enn den som har med regelkunnskap og regneferdigheter å gjøre.

For å kunne si noe om hvorvidt det er forskjell på den kognitive vanskelighetsgraden innen oppgavene som tematisk ligger innenfor ungdomsskolepensum og de som ligger utenfor, har jeg funnet gjennomsnittsnivået innen hver kognitive kategori for disse to gruppene av oppgaver. Resultatene er presentert i tabell 7.

Tabell 7. Gjennomsnittlig vanskelighetsgrad innen hver kognitive kategori for oppgaver i og utenfor læreplanen i matematikk for 8.-10. trinn

	Innen læreplanen 8-10	**Utenfor læreplanen 8-10**
Communication	0.70	0.80
Devising Strategies	0.39	0.37
Mathematising	0.51	0.48
Representation	0.40	0.35
Using symbols, operations and formal language	1.22	1.69
Reasoning and argument	0.95	1.24

Den gjennomsnittlige vanskelighetsgraden til oppgavene som er kodet i og utenfor læreplanen i matematikk 8-10 er ganske lik. De kategoriene som skiller seg mest fra hverandre er *Using symbols, operations and formal language* og *Reasoning and argument* hvor oppgavene utenfor læreplanen 8-10 skårer noe høyere enn oppgavene som ligger innenfor denne.

4.3.2 Sammenlikning mellom lærestedene

Videre har jeg sett på nivåfordelingen innen de seks kognitive kategoriene for hvert av de tre lærestedene separat. Tabell 8 viser fordelingen for de tre lærestedene.

Nivå 0 og 1 er de to dominerende nivåene i alle kategoriene unntatt *Using symbols, operations and formal language* og *Reasoning and argument*, akkurat som for den samlede fordelingen. I kategoriene *Communication* og *Reasoning and argument* er fordelingen forholdsvis lik for de tre lærestedene. I de fire andre kategoriene er det noen forskjeller, men det er ikke ett lærested som skiller seg klart fra de andre med hensyn på å ha høyere eller lavere nivåer i alle disse fire kategoriene.

Tabell 8. Nivåfordeling innen hver kognitive kategori for hvert lærested separat

Communication

	0	1	2	3
Lærested A	34 %	48 %	15 %	4 %
Lærested B	44 %	43 %	12 %	2 %
Lærested C	39 %	48 %	14 %	0 %

Representation

	0	1	2	3
Lærested A	53 %	32 %	13 %	2 %
Lærested B	76 %	19 %	5 %	0 %
Lærested C	82 %	13 %	5 %	0 %

Devising strategies

	0	1	2	3
Lærested A	74 %	20 %	5 %	1 %
Lærested B	77 %	19 %	4 %	0 %
Lærested C	53 %	41 %	7 %	0 %

Using symbols, operations and formal language

	0	1	2	3
Lærested A	15 %	45 %	34 %	5 %
Lærested B	4 %	45 %	41 %	10 %
Lærested C	4 %	31 %	51 %	14 %

Mathematising

	0	1	2	3
Lærested A	52 %	29 %	19 %	0 %
Lærested B	61 %	30 %	9 %	1 %
Lærested C	70 %	25 %	5 %	0 %

Reasoning and argument

	0	1	2	3
Lærested A	18 %	47 %	34 %	2 %
Lærested B	19 %	52 %	25 %	5 %
Lærested C	16 %	60 %	19 %	5 %

4.4 Kodesamsvar

For å vurdere reliabiliteten i kodingen har en annen forsker, som er kjent med rammeverket, kodet omtrent 10 % av datamaterialet parallelt med meg. Jeg har brukt Cohens kappa og vektet Cohens kappa som mål på interkoderreliabiliteten. Dette er begrunnet i metodekapittelet (se 3.6.1). I tabell 9 er samsvaret for hver av de kognitive kategoriene *Communication*, *Devising strategies*, *Mathematising*, *Representation*, *Using symbols, operations and formal language* og *Reasoning and argument* presentert.

Tabell 9. Mål for samsvar i koding

	Cohens kappa	Vektet Cohens kappa gjennomsnittsmål	Vektet Cohens kappa enkeltmål
Communication	0,43	0,87	**0,77**
Devising strategies	0,31	0,67	**0,51**
Mathematising	0,59	0,82	**0,70**
Representation	0,28	0,68	**0,52**
Using symbols, operations and formal language	0,03	0,57	**0,40**
Reasoning and argument	0,28	0,75	**0,60**

Ut fra begrunnelse gitt i metodekapittelet er det en vektet Cohens kappa med enkeltmål som blir det riktigste å bruke som samsvarsmål i min studie. Gjennomgående gir vektet Cohens kappa et vesentlig høyere samsvar enn Cohens kappa. At normal Cohens kappa gir forholdsvis lavt samsvar, betyr at vi ofte er uenige om nivået innenfor en kategori, men at samsvaret er høyere for vektet Cohens kappa, betyr at vi sjeldnere er uenige med mer enn ett nivå. Kategorien *Using symbols, operations and formal language* er den med lavest samsvar, her har Cohens kappa verdien 0,032. Dette sier at vi var uenige omtrent like ofte som vi var enige når vi kodet. Vektet kappa er imidlertid 0,40, noe som tilsier at det sjeldnere er uenighet med mer enn ett nivå. At en vektet Cohens kappa er 0,40, betyr at samsvaret ifølge Cicchetti (1994) er helt i nedre grense av hva som klassifiseres som moderat, mens det ifølge Altman (1991) er helt i øvre grense for hva som klassifiseres som dårlig (se 3.6.1). Jeg kommer til å omtale samsvaret for kategorien *Using symbols, operations and formal language* som moderat i resten av oppgaven, men leseren bør ha i bakhodet at dette ligger helt i grensen mot dårlig samsvar.

De andre kategoriene har høyere verdier både for Cohens kappa og vektet Cohens kappa, noe som vitner om bedre kodesamsvar. Dersom vi ser på vektet kappa er samsvaret ifølge Altman (1991) moderat, godt eller svært godt for alle de andre kategoriene. Dette betyr at den andre koderen og jeg har hatt en stor grad av enighet i hvordan vi tolker og bruker rammeverket i møte med datamaterialet.

Vektet Cohens kappa gjennomsnittsmål gir høyere samsvarverdier enn vektet Cohens kappa enkeltmål. Vektet Cohens kappa gjennomsnittsmål viser reliabiliteten til et gjennomsnitt av koderne og er dermed et mål på hvor reliabel studien min ville vært dersom vi hadde vært to kodere til å kode hele datamaterialet. Resultatene viser at dette ville gitt klart høyere reliabilitet.

Vi ser en klar skjevfordeling mot de lave nivåene i kodedataene ved at et stort flertall av oppgavene er kodet til nivå 0 eller 1 i de fleste kognitive kategoriene. Ifølge Sim og Wright (2005) vil det i de tilfellene hvor det er asymmetri i klassifiseringen, det vil si i de tilfellene hvor ikke alle utfallene er like sannsynlige, bli en lavere kappaverdi. Dette betyr at mine samsvarsverdier muligens er lavere enn det reelle samsvaret i kodingen. I så fall vil den reelle reliabiliteten være sterkere enn det som kommer frem av tallene i tabellen.

For å kunne sammenlikne med kodesamsvaret i MEG-studien, beregnet jeg Cronbachs alfa for mine data. Disse er presentert i tabell 10, sammen med verdiene som de fant i MEG-studien.

Tabell 10. Sammenlikning av Cronbachs alfa-verdier for mine data og for data fra MEG-studien (Turner et al., 2013, s. 30)

	Cronbachs alfa i MEG-studien	**Cronbachs alfa for min studie**
Communication	0,95	0,91
Devising Strategies	0,90	0,67
Mathematising	0,81	0,85
Representation	0,83	0,68
Using symbols, operations and formal language	0,89	0,68
Reasoning and argument	0,62	0,78

Vi ser at samsvarsverdiene er noe høyere for de åtte koderne i MEG-studien enn for den andre koderen og meg i min studie i fire av seks kategorier. I kategoriene *Mathematising* og *Reasoning and argument* er imidlertid samsvaret høyere i min studie.

5 Diskusjon

I dette kapittelet diskuterer jeg hva mine resultater indikerer om hvor godt eksamensoppgavene tester ulike aspekter ved matematisk kompetanse, og i hvor stor grad det ser ut som at oppgavene tester god undervisningskunnskap i matematikk. Kapittelet inneholder også en diskusjon av kodesamsvaret, samt en kort diskusjon av styrker og svakheter ved egen studie.

5.1 Fordeling mellom fagoppgaver og fagdidaktiske oppgaver

I mitt datamateriale varierer andelen fagdidaktiske oppgaver mellom de ulike lærestedene. Totalt er 11 % av datamaterialet kodet som matematikkdidaktiske oppgaver. Lærested A er det lærestedet som skiller seg ut i denne kodingen, med 23 % av oppgavene vurdert som matematikkdidaktiske. Ved lærested B og C gjelder dette kun 3 % og 7 % av oppgavene. Som begrunnet i teorikapittelet er både matematikkfaglig og matematikkdidaktisk kunnskap viktig for god undervisningskompetanse i matematikk (Ball et al., 2008; Shulman, 1986). Baumert et al. (2010) undersøkte gjennom COACTIV-studien betydningen disse to kunnskapsformene har for elevenes læring og konkluderer med at det er den fagdidaktiske kunnskapen som er av størst betydning.

At fagdidaktisk kunnskap har stor betydning for elevenes læring gjør det interessant å diskutere den lave andelen fagdidaktiske oppgaver som gis til skoleeksamen ved mine inkluderte lærestedер. Det må imidlertid tas med i betraktningen at fordelingen av oppgaver i mappevurderingene ikke er inkludert i studien. To av lærestedene har mappevurdering som del av den totale vurderingen, og det er mulig at studentene får mange didaktiske oppgaver der. Det siste lærestedet har en stor didaktisk hjemmeeksamen som ikke er inkludert i kodingen. Det sistnevnte lærestedet er det lærestedet med lavest prosentvis andel av didaktiske oppgaver gitt til skoleeksamen. Denne store didaktiske hjemmeeksamenen er imidlertid av en annen karakter enn de matematikkdidaktiske oppgavene i de skriftlige skoleeksamenene, fordi den handler om et stort forskningsbasert arbeid som det skal gjøres rede for. De matematikkdidaktiske oppgavene som er inkludert i mitt datamateriale er kortere eksamensoppgaver, gjerne knyttet til en matematikkoppgave i det samme oppgavesettet.

Oppgaver gitt som hjemmeeksamen og skoleeksamensoppgaver vil ikke nødvendigvis teste den samme typen lærerkunnskap (Christiansen et al., 2015). Hjemmeeksamener vil være godt egnet til å teste kunnskap og kompetanser som vil tilsvare det lærerne vil komme til å bruke under forberedelser og etterarbeid knyttet til undervisningen. Fagdidaktiske oppgaver som gis til en skoleeksamen vil derimot kunne teste paratheten som kreves når læreren står i undervisnings-situasjonen (Christiansen et al., 2015). Dette er et argument for at det bør gis en viss andel fagdidaktiske oppgaver både til hjemmeeksamener, hvor studentene kan bruke hjelpemidler og samarbeide, og til skoleeksamener, gjerne i direkte tilknytning til en matematikkoppgave som de har løst.

De fagdidaktiske oppgavene knyttet til matematikkoppgaver, er blant annet viktige fordi de krever at studentene gjennomfører didaktiske refleksjoner knyttet til de matematiske temaene som de skal ut i skolen og undervise. Slike refleksjoner kan for eksempel dreie seg om hvilken forkunnskap elevene har, hvilke eksempler knyttet til et tema som kan være motiverende for elevene og om ulike eksempler og tilnærminger til et tema som til sammen kan gi en dypere forståelse for elevene (Ball et al., 2008; Niss, 2006; Shulman, 1986). De oppgavene som er kodet som didaktiske i mitt datamateriale handler gjerne om denne typen problemstillinger.

Didaktiske refleksjoner direkte knyttet til matematikkoppgavene vil også kunne være med å vise lærerstudentene hvorfor oppgavene som gis til skoleeksamen er relevante for den jobben de skal ut og gjøre. Ball et al. (2008, s. 404) peker nettopp på hvordan matematikkurs i lærerutdanningen for sjelden vinkler oppgavene slik at de blir relevante og profesjonsspesifikke for lærere: «Unfortunately, subject matter courses in teacher preparation programs tend to be academic in both the best and worst sense of the word, scholarly and irrelevant either way remote from classroom teaching». For mange oppgaver som kun tester ren matematikkunnskap til fordel for matematikkdidaktisk kunnskap, vil kanskje gjøre at lærerstudentene ikke ser relevansen av det de lærer. Resultatene i TEDS-M 2008 konkluderer også med et behov for bedre integrering mellom matematisk kunnskap og fagdidaktisk kunnskap i lærerutdanningen (Grønmo & Onstad, 2012). Jeg ser mange gode argumenter for at andelen didaktiske oppgaver som knyttes til matematikkoppgavene som studentene skal løse til eksamen, spesielt ved to av lærestedene, burde vært større.

5.2 Kategorisering med hensyn på matematisk tema

Jeg har gjennomført kategoriseringen med hensyn på matematisk tema ut fra inndelingen som ligger i læreplanen i matematikk for 8.-10. trinn (Utdanningsdirektoratet, 2013). For to av lærestedene separat, og for alle lærestedene sett samlet, er *Funksjoner* og *Tall og algebra* de største emneområdene. Ut fra læreplanen i matematikk for 8.-10. trinn ser det ut som at *Tall og algebra* er et stort emneområde også her (Utdanningsdirektoratet, 2013). Det blir imidlertid feil å trekke klare slutninger om hvor stor plass et tema får i undervisning, lærebøker og vurdering i skolen ut fra antall kompetansemål i læreplanen. Et enkelt kompetansemål kan dekke et stort fagområde og flere kompetansemål kan brukes for å beskrive et lite fagområde. I og med at jeg har brukt læreplaner også fra videregående skole, ville en slik kopling av tematisk fordeling mot fordeling i læreplanen dessuten måttet ta høyde for fordelingen i samtlige læreplaner, ikke bare læreplanen for ungdomstrinnet. *Funksjoner* er et eksempel på et område som det ser ut som får mindre plass på ungdomstrinnet. Området får imidlertid større plass i læreplanene på høyere trinn, for eksempel i læreplanen for Matematikk R1 (Utdanningsdirektoratet, 2006a), og mange av oppgavene som jeg har kodet har tematisk innhold som svarer til kompetansemål fra dette trinnet.

Da jeg ikke har tilgang til noen annen oversikt over emneområdenes plass i skolen enn antall kompetansemål i læreplanen, er det vanskelig å vurdere om fordelingen av matematisk tema i lærerstudentenes eksamensoppgaver gjenspeiler vektleggingen av disse temaene i undervisningen i grunnskolen. Nasjonale retningslinjer for grunnskolelærerutdanningen gir imidlertid klare mål for lærerstudentenes kunnskap etter endt utdanning: «Studenten har inngående undervisningskunnskap i matematikken elevene arbeider med på trinn 5-10, særlig tallforståelse og regning, geometri og måling, overgangen fra aritmetikk til algebra, algebra og funksjoner.» (Kunnskapsdepartementet, 2010b, s. 35). Temaene *Geometri* og *Måling* står til sammen for 15,4 % av oppgavene som jeg har kodet, *Måling* alene står for bare 3 % av oppgavene. Dette står i tydelig kontrast til målene i retningslinjene. Dersom det er mulig å indikere noe ut fra antall kompetansemål i læreplanen, kan det også se ut som at emneområdene *Geometri* og *Måling* har større plass i undervisningen på ungdomstrinnet enn i lærerstudentenes eksamensoppgaver.

Siden min studie viser at de tre inkluderte lærestedene har valgt å vektlegge de ulike emneområdene ulikt, synes jeg det er interessant å stille spørsmål ved hva som bestemmer fordelingen av de matematiske temaene ved de forskjellige lærestedene. I emneområdet *Sannsynlighet, statistikk og kombinatorikk* ser vi forholdsvis store forskjeller med 15 % av oppgavene ved lærested C, og 28 % av oppgavene ved lærested B kodet i denne kategorien. Tilsvarende forskjeller ser vi i emneområdet *Tall og algebra.* 21 % av oppgavene er kodet i denne kategorien ved lærested B, mens dette gjelder 39 % av oppgavene ved lærested C. Den største forskjellen finner vi i emneområdet *Geometri.* Her er det et spenn fra 3 % til 23 % mellom henholdsvis lærested B og lærested C. Dette emneområdet utgjør altså en minimal del av eksamensoppgavene ved det ene lærestedet, mens det utgjør en betydelig andel ved et annet. Det er påfallende at lærestedene velger å vektlegge de ulike emneområdene såpass ulikt, særlig sett i lys av målet, som vi finner i de nasjonale retningslinjene for lærerutdanningen, om at alle lærerstudentene skal ha inngående undervisningskunnskap i all matematikken som de kommer til å møte gjennom arbeidet i skolen.

Resultatene fra TEDS-M 2008 peker på at det fremstår som spesielt viktig å styrke norske lærerstudenters matematikkunnskap innen aritmetikk og algebra (Grønmo & Onstad, 2012, s. 195). Sett i dette perspektivet virker det fornuftig at en stor andel av oppgavene som gis til skoleeksamen ved de inkluderte lærestedene kodes nettopp til kategorien *Tall og algebra.* Denne andelen er i realiteten enda større enn det som kommer frem i tabell 4, jamfør diskusjonen av hvordan oppgaver som tematisk dekker flere områder er kodet (se 3.4.1). Vi kan imidlertid ikke si noe om lærerstudentene mestrer oppgaver innen emneområdet *Tall og algebra* ettersom vi ikke vet om studentene klarte å løse disse oppgavene til eksamen. Det eneste vi vet ut fra mine data er at de inkluderte lærestedene vektlegger dette hovedområdet i sine vurderinger.

Det er betydelige forskjeller mellom lærestedene i andelen oppgaver som tematisk ligger innenfor læreplanen for ungdomstrinnet. Ved lærested A ligger omtrent halvparten av oppgavene på et tematisk nivå som svarer til ungdomsskolepensum, mens det gjelder kun om lag en femdel av oppgavene ved lærested B og lærested C. Forskning peker på at både en dyptgående forståelse av det som skal undervises, og kunnskap om hva studentene skal lære senere, er viktig (Shulman, 1986). Lærerstudentene bør altså på den ene siden få en dyp forståelse av de temaene de skal undervise når de skal ut i jobb og på den andre siden ha kunnskaper om matematiske temaer som strekker seg utover dette. Det sistnevnte betegnes av Shulman (1986) og Ball et al. (2008) som horisontkunnskap. Dette er viktig fordi læreren må vite hva elevene har lært før og hva de skal lære senere for at undervisningen skal kunne legges opp etter dette. Kunnskap om hvilken rolle regning med brøk får for de som velger realfag i videregående skole, kan for eksempel være med på å gjøre at læreren på ungdomstrinnet oppfordrer til bruk av brøk fremfor avrunding til desimaltall. På samme måte kan kunnskap om hva slags funksjonsdrøfting elevene skal kunne utføre på videregående, være med på å påvirke hvordan funksjonsbegrepet introduseres på ungdomstrinnet.

Kunnskap om temaer som strekker seg utover temaene lærerstudentene skal undervise i skolen, kan sannsynligvis også bidra til å øke deres egen forståelse av matematikken. For eksempel kan forståelsen av funksjoner og funksjonsdrøfting bedres av at studentene lærer om derivasjon, kontinuitet og grenseberegninger, selv om dette ikke er pensum i ungdomsskolen. Slike temaer testes til en stor eller middels stor grad ved alle lærestedene som er inkludert i denne studien. Også de nasjonale retningslinjene for grunnskolelærerutdanningen 5.-10. trinn vektlegger hvordan studentene skal ha gode kunnskaper innen fagområder som ligger utenfor læreplanen 8-10. De sier følgende om mål for elevenes kunnskap etter endt utdanning: «Studenten har god kunnskap i matematisk analyse, inkludert derivasjon, integrasjon, differensiallikninger og enkle matematiske modeller, og kan relatere disse begrepene til det matematikkfaglige innholdet i trinn 5-10.» (Kunnskapsdepartementet, 2010b, s. 36).

Det er altså klare retningslinjer som tilsier at studentene bør lære om temaer som går utover ungdomsskolepensum. Det gjør at det er liten tvil om at det er riktig av lærestedene å inkludere oppgaver som strekker seg til et høyt tematisk nivå. Jeg merker meg imidlertid at studentene er forventet å kunne relatere disse temaene til undervisningen i grunnskolen. En nærmere diskusjon av dette, sett i sammenheng med oppgavenes kognitive nivå, kommer i kapittel 5.3.1.

5.3 Kategorisering med hensyn på kognitiv kategori

5.3.1 Samlet fordeling

Gjennom å kode innen de seks kognitive kategoriene gitt ved rammeverket utviklet i MEG-studien, kan jeg si noe om i hvilken grad oppgavene i datamaterialet mitt er kognitivt utfordrende for lærerstudentene. Rammeverket er bygget på PISA sin beskrivelse av hva som kjennetegner mestring i matematikk (Turner et al., 2013). Hjelpemiddelkompetansen er tatt ut i rammeverket til MEG-studien (Turner et al., 2013), men bortsett fra dette er det rimelig å anta at oppgaver som er kognitivt utfordrende innen alle de seks kategoriene, vil være oppgaver som utfordrer en helhetlig matematisk kompetanse. Niss (2006) peker på hvordan mestring innen de åtte delkompetansene i KOM-studien, som PISA bygger sitt rammeverk på, er en forutsetning for god undervisningskompetanse i matematikk. Således kan mine analyser sannsynligvis også antyde noe om hvordan kunnskapsdelen av undervisningskompetansen til lærerstudentene vurderes. Dette forutsetter selvsagt at oppgavene også spenner over et vidt tematisk område. Mine data sier ikke noe om lærerstudentene har klart å løse disse eksamensoppgavene og derfor ingenting om deres faktiske matematikkompetanse og undervisningskunnskap. Vi kan likevel si noe om hvilke deler av matematikkompetansen de inkluderte lærestedene vektlegger i sine eksamener.

Jevnt over er det de lave nivåene som dominerer innen de seks kognitive kategoriene. Unntakene er kategoriene *Using symbols, operations and formal language* hvor en betydelig andel av oppgavene også er kodet til nivå 2 og noen til nivå 3, og *Reasoning and argument* hvor størsteparten av oppgavene er kodet til nivå 1 og noen av oppgavene til nivå 2. Dette betyr at gjennomsnittsoppgaven når jeg ser på hele datamaterialet mitt samlet, er lite utfordrende kommunikativt, studentene trenger til en liten grad å matematisere, lage løsningsstrategier eller bruke representasjoner. De må imidlertid kunne en del fakta og prosedyrer for å løse oppgaven og de må kunne resonnere noe underveis. Kun 16 % av oppgavene skårer nivå 3 i minst én kategori. Dette er ikke forsvinnende lite, men vi må ha i bakhodet at rammeverket med nivåinndelingen er laget for oppgaver på ungdomsskolenivå, mens oppgavene som jeg har analysert er laget for å teste de som skal undervise på dette nivået.

Using symbols, operations and formal language er den kategorien som har størst andel av oppgavene kodet til nivå 3, men det er likevel bare 10 % av oppgavene som kodes til det høyeste nivået også i denne kategorien. Et høyt nivå innen *Using symbols, operations and formal language* tilsier at studentene må kunne bruke flere regler eller en rekke med regneoperasjoner, for eksempel ved å løse en likning som krever flere løsningssteg og bruk av ulike regler. Dette er altså oppgaver som tester regelkunnskap og regneferdigheter hos lærerstudentene.

Using symbols, operations and formal language tilsvarer delkompetanser som vi gjenfinner også i de andre rammeverkene for kjennetegn på matematisk kompetanse som jeg har beskrevet i teorikapittelet, men som jeg valgte å ikke bruke som utgangspunkt for kodingen (se 2.2). I Blooms taksonomi er *Knowledge* den lavest rangerte kognitive prosessen av de seks hierarkisk ordnede kategoriene som beskriver de kognitive ferdighetene til en elev (Krathwohl, 2002). Fra KOM-prosjektet kjenner vi igjen *Symbol- og formalismekompetansen* som den tilsvarende kompetansen (Niss & Jensen, 2002). MEG-studien bygger på KOM-prosjektet, så beskrivelsen av disse er svært samsvarende (Turner et al., 2013). Kilpatrick et al. (2001) sitt rammeverk har to tråder som sammen gir tilsvarende kompetanse, *Conceptual understanding* og *Procedural fluency.* PISA bruker i sitt rammeverk *Bruke symbol- og formelspråk* som vi kjenner igjen fra Niss og MEG-studien (Nortvedt, 2013), mens TIMSS bruker *Å kunne* (Onstad, 2010). Det som er felles for alle de teoretiske rammeverkene er at kompetanser tilsvarende *Using symbols, operations and formal language* er med i kompetansebeskrivelsen, men at det ikke er nok alene til å beskrive matematisk kompetanse. Den matematiske kompetansen er satt sammen av flere elementer, blant annet evnen til å finne løsningsstrategier og resonnere matematisk. Kilpatrick et al. (2001) tydeliggjør dette gjennom å fremstille matematisk kompetanse som fem tett sammenvevde tråder.

Ved så mye som 9,1 % av oppgavene i mitt datamateriale er det likevel kun *Using symbols, operations and formal language* som er kodet til nivå høyere enn 0. At mer enn halvparten av oppgavene skårer 0 eller 1 i alle andre kategorier enn i denne som har med regelkunnskap og regneferdigheter å gjøre, viser også tydelig hvordan denne ene delen av det som kjennetegner matematisk kompetanse blir vektlagt mer enn de andre. Et eksempel på en oppgave som bare skårer høyt i *Using symbols, operations and formal language* kan være «Deriver uttrykket $f(x) = 2x^2$». Denne oppgaven har ingen kommunikative utfordringer knyttet til seg. Studentene får oppgitt akkurat hva de skal gjøre i oppgaveteksten, og må følgelig ikke utvikle en løsningsstrategi. De trenger heller ikke å oversette til noe utenfor matematikken. Det er ingen representasjoner involvert annet enn uttrykket for f(x), som jeg har tolket som en enkel representasjon med tanke på hvilke operasjoner studentene skal utføre med uttrykket. Studentene trenger ikke å resonnere, de følger bare en enkel derivasjonsregel. I denne oppgaven er det dermed kun *Using symbols, operations and formal language* som skårer høyere enn nivå 0. Jeg mener at det er problematisk at lærerstudentene ikke utfordres mer innen de andre delene. Det er også problematisk dersom lærerstudentene selv opplever at matematikk handler om regler og algebraiske manipulasjoner når de skal ut i en undervisningsjobb i skolen.

Oppgaver som kun tester regneferdigheter og regler, vil i liten grad vurdere lærerstudentenes forståelse. Shulman (1987, s. 14) understreker nettopp behovet for vurdering av lærer-studentenes dyptgående forståelse av fagstoffet som de skal ut og undervise:

> *To teach is first to understand. We ask that the teacher comprehend critically a set of ideas to be taught. We expect teachers to understand what they teach and, when possible, to understand it in several ways. They should understand how a given idea relates to other ideas within the same subject area and to ideas in other subjects as well.*

For at læreren skal jobbe fleksibelt slik som Shulman beskriver i sitatet over, må forståelsen være så god at forklaringer, eksempler og metodevalg kan tilpasses den enkelte eleven. Det er lett å tenke at behovet for tilpasning i undervisningen i hovedsak dreier seg om elever med faglige utfordringer. Grønmo et al. (2012) peker imidlertid på hvordan det er viktig at lærere kan nok abstrakt, vurderende og resonnerende matematikk til å gi tilpasset opplæring også til de sterke elevene. De hevder at norsk skole i liten grad lykkes med å gi tilpasset opplæring til denne gruppen elever. Kanskje kan en større andel oppgaver som utfordrer lærerstudentene på høye nivåer innen de kognitive kategoriene som er undersøkt gjennom det valgte rammeverket, føre til at studentene utvikler høyere kompetanse i områder som handler om evnen til abstrakt tenkning, vurdering og resonnement. Lærere med høy kompetanse innen disse områdene kan igjen kanskje være med på å styrke den tilpassede opplæringen i skolen.

I beskrivelsen av hvilke grunnleggende ferdigheter elevene skal mestre ut fra læreplanen i matematikk fellesfag står det at elevene i grunnskolen blant annet skal kunne *reflektere*, *vurdere og kommunisere* matematikk (Utdanningsdirektoratet, 2013). Dette knytter jeg til kategoriene *Reasoning and argument* og C*ommunication* i rammeverket som jeg har brukt. Kun 15 % av oppgavene i mitt datamateriale utfordrer lærerstudentene på et høyt nivå (2 eller 3) innen *Communication*, mens dette gjelder for 29 % av oppgavene i kategorien *Reasoning and argument* (se tabell 6). Jeg mener at det er problematisk at de fremtidige lærerne testet såpass lite innen kompetanseområder som i så stor grad vektlegges i skolen.

De norske læreplanene har også et stort fokus på anvendt matematikk med målsetning om å utdanne elever som kan bruke matematikken i praktiske problemstillinger (Nortvedt, 2013). Funn fra TEDS-M 2008 antyder også et behov for at lærerstudentene blir gode på anvendt matematikk spesielt fordi norsk skole oppfordrer til utforsking og utvikling av egne metoder og strategier (Grønmo & Onstad, 2012). Vi ser imidlertid at eksamensoppgavene gjennomgående er kodet til lave nivåer i *Devising strategies* og *Mathematising* som er kategorier som har med anvendt matematikk og problemløsing å gjøre (Turner et al., 2015). Vi ser av tabell 6 at i kategorien *Devising strategies* er 0 % av oppgavene kodet til nivå 3 og 5 % av oppgavene til nivå 2, mens i kategorien *Mathematising* er de tilsvarende andelene 0 % og 10 %. Mine analyser viser dermed at det gis få eksamensoppgaver som tester studentenes evner til problemløsning til eksamen ved lærerutdanningene som er inkludert i denne studien. Dette er et viktig funn fordi høy kompetanse i disse kognitive kategoriene vil kunne styrke lærerstudentenes kunnskap om å løse praktiske problemstillinger, noe som anses som viktig kompetanse for elevene som lærerstudentene senere skal undervise.

At 0 % av oppgavene kodes til nivå 3 i kategoriene *Devising strategies, Mathematising* og *Representation* er bemerkelsesverdig. Nivå 3 i kategorien *Devising strategies* innebærer å utarbeide en kompleks løsningsstrategi over flere trinn. At løsningsstrategien er kompleks betyr at den inneholder flere delmål som studentene må kontrollere, evaluere og justere underveis i løsningsprosessen. Dette høyeste nivået kan også innebære at studentene må sammenlikne ulike løsningsstrategier. Nivå 3 innen *Mathematising* innebærer at studenten må lage modeller hvor de selv definerer antakelser, variabler, sammenhenger og begrensninger, at de må validere og evaluere modeller eller at de må sammenlikne ulike modeller som beskriver virkelighetsnære situasjoner. Det høyeste nivået i kategorien *Representation* tilsier at studentene skal forstå, bruke og trekke tråder mellom flere komplekse matematiske representasjoner, at de kan sammenlikne og evaluere disse, eller at de kan lage en representasjon som beskriver en kompleks matematisk sammenheng (Turner et al., 2015).

At det ikke vurderes om studentene kan utarbeide komplekse løsningsstrategier over flere trinn, eller om de kan lage modeller hvor de selv må definere antakelser og variabler, er tankevekkende, spesielt sett opp mot det uttalte målet om at norske skoleelever skal bli gode i anvendt matematikk (Grønmo & Onstad, 2012; Nortvedt, 2013; Utdanningsdirektoratet, 2013). Dersom man ser på læreplanen i matematikk for 8.-10. trinn, kan man lese at det forventes at norske ungdomsskoleelever skal mestre ulike representasjonsformer og skal kunne bruke disse til å fremstille matematiske sammenhenger (Utdanningsdirektoratet, 2013). Sett i sammenheng med dette, mener jeg at fremtidige ungdomsskolelærere også bør vurderes på det høyeste nivået innen kategorien som handler om matematiske representasjoner.

Som beskrevet i kapittel 5.2, har alle lærestedene i studien lagt det tematiske innholdet til et nivå som ligger utenfor ungdomsskolepensum på en forholdsvis stor andel av eksamensoppgavene. Denne typen oppgaver vil kunne være med på å teste horisontkunnskap, men forhåpentligvis også dyptgående kunnskap og forståelse. Dette er imidlertid avhengig av typen oppgaver som gis. Jeg vil hevde at enkle oppgaver med derivasjon, som i eksempelet

«Deriver funksjonen $f(x) = 2x^2$», ikke nødvendigvis vurderer forståelsen av funksjoner hos lærerstudentene. Dersom man derimot ber studentene om å drøfte betydningen av den deriverte eller kanskje enda bedre – om å tegne den deriverte ut fra grafen til en funksjon, tror jeg at dette til en større grad tester en dypere og mer grunnleggende forståelse av funksjoner. En slik forståelse kan ha stor nytteverdi for undervisning på ungdomstrinnet. Den sistnevnte typen oppgaver ville skåret høyt innen *Representation* og *Reasoning and argument*, mens hovedvekten av derivasjonsoppgavene i mitt datamateriale har vært av samme type som eksempelet over og tester bare delkompetansen *Using symbols, operations and formal language.*

Jeg mener at det er viktig å diskutere om det er mest nyttig å gi lærerstudenter oppgaver som er på et høyt nivå tematisk, men med lave nivåer i de kognitive kategoriene, eller om oppgaver som ligger på et lavere tematisk nivå og til gjengjeld har større kognitive utfordringer knyttet til seg kan være vel så fruktbare. Det ser ut som at lærestedene inkludert i denne studien har valgt å gi oppgaver som heller utfordrer studentene med et tematisk høyt nivå enn med høye nivåer i de kognitive kategoriene. Som jeg argumenterer for i avsnittet over, vil noe av effekten av å inkludere oppgaver som tematisk strekker seg til videregående nivå, bli borte dersom de samtidig skårer på et lavt nivå i alle andre kognitive kategorier enn *Using symbols, operations and formal language*. Jeg hevder at det er mulig å sette spørsmålstegn ved om det da er snakk om faktisk horisontkunnskap.

For å vurdere i hvilken grad lærestedene utfordret elevene med et høyt kognitivt nivå i oppgaver som lå innenfor ungdomsskolepensum, sammenliknet jeg nivåene innen de kognitive kategoriene for oppgavene som tematisk ligger innenfor og utenfor læreplanen 8-10. Tabell 7 viser at oppgavene som ligger innenfor ungdomsskolepensum, og som dermed skal være enklere tematisk, skårer omtrent samme gjennomsnittsverdi for alle de kognitive kategoriene som oppgaver som ligger på et høyere tematisk nivå. Når vi ser på oppgavene som tematisk ligger innenfor ungdomsskolepensum alene, er kategorien *Using symbols, operations and formal language* den eneste hvor gjennomsnittsnivået er over én. Med andre ord ser vi at det også for disse oppgavene er slik at de ikke er utfordrende i andre kognitive kategorier enn den som har med regel- og regne-ferdigheter å gjøre. Det kan altså se ut til at de inkluderte lærestedene til en liten grad har eksamensoppgaver som tester den dyptgående forståelsen av det lærerstudentene skal ut i skolen og undervise.

Hjelpemiddelkompetanse blir ikke kategorisert i rammeverket jeg har valgt for kodingen. Dette har gjort at jeg har vurdert tilgjengelige hjelpemidler før jeg har kodet *Use of symbols, operations and formal language* slik at nivået blir høyere eller lavere ut fra om det er kalkulator og/eller formelverktøy tilgjengelig. Det jeg imidlertid har merket meg, og som ikke kommer frem i noen av de presenterte dataene, er at studentene ikke har hatt tilgang til datamaskin ved noen av eksamenene som er inkludert i studien. Våren 2015 innførte Utdanningsdirektoratet krav til bruk av digitale verktøy for alle eksamener i matematikk både i grunnskolen og i videregående skole (Matematikksenteret, 2015). Det påpekes at det ikke lenger er tilstrekkelig med en kalkulator alene for å besvare en slik eksamen. Eksempler på hjelpemidler som kreves er Computer Algebra System (CAS), graftegnere og regneark.

Digitale ferdigheter er også en av de fem grunnleggende ferdighetene som er integrert i læreplanmålene gjennom hele utdanningsløpet i Norge (Matematikksenteret, 2015). Niss (1999) peker på hvordan bruk av IKT i undervisningen er en av de største og mest aktuelle utfordringene som lærere i skolen møter. Han påpeker videre hvordan dette stiller store krav til lærerne, de må ha kompetanse nok til å kunne bruke dette på en tilfredsstillende måte i undervisningen. Dette er like aktuelt i dag (Voogt, Erstad, Dede, & Mishra, 2013). Eksamensoppgavene som er inkludert i mitt datamateriale er imidlertid fra 2011-2014 og derfor eldre enn innføringen av det overnevnte kravet til bruk av digitale verktøy til eksamen i grunnskolen og i videregående skole. Dette kan kanskje forklare noe av hvorfor lærerstudentene som gjennomførte disse eksamenene ikke ble vurdert i bruk av slike verktøy. Jeg håper imidlertid at bruk av digitale hjelpemidler er inkludert i nyere eksamener ved lærerutdanningene.

5.3.2 Sammenlikning mellom lærestedene

Fordelingen innen de kognitive kategoriene skiller seg lite mellom lærestedene og dermed naturligvis heller ikke mye fra den samlede oversikten. Jevnt over kodes få oppgaver til nivå 3, mens mange oppgaver kodes til nivå 0 og 1. Diskusjonen og refleksjonene knyttet til nivåene innen de kognitive kategoriene for alle lærestedene samlet (kapittel 5.3.1), vil gjelde også for hvert av lærestedene separat. Det følger derfor ingen omfattende diskusjon av lærestedene satt opp mot hverandre når det kommer til koding i kognitiv kategori.

Det er imidlertid mulig å trekke fram noen forskjeller. Lærested A har færre oppgaver kodet til nivå 3 i kategoriene *Using symbols, operations and formal language* og *Reasoning and* *argument* enn de to andre lærestedene. Vi ser imidlertid at lærested A er det lærestedet som har oppgaver kodet til nivå 3 i flest kategorier, her er det bare kategorien *Mathematising* der 0 % av oppgavene er kodet til nivå 3. Ved lærested C er det hele fire kategorier hvor 0 % av oppgavene er kodet til nivå 3, mens dette gjelder to av kategoriene ved lærested B. Dette betyr at studentene ved lærested B og C går gjennom samtlige skriftlige skoleeksamener uten å bli testet i det som karakteriseres som det høyeste nivået innen flere av de kognitive kategoriene. Jamfør diskusjonen i kapittel 5.3.1 mener jeg at dette er problematisk, spesielt med tanke på at rammeverket er utviklet for å forutsi vanskelighetsgrad for ungdomstrinnet.

Vi så i kapittel 4.2 at det tematiske nivået ligger utenfor læreplanen 8-10 ved en betydelig lavere andel av oppgavene ved lærested A enn ved de to andre lærestedene. Dette kan nok være med på å forklare at lærested A skårer noe lavere i *Using symbols, operations and formal language* enn lærested B og C. Oppgavene fra lærested B og lærested C strakk seg i større grad opp på videregående nivå. Temaer som volum av omdreiningslegemer, komplekse tall og grenseverdiutregninger er eksempler på hva som ble inkludert ved disse lærestedene. Fordi disse oppgavene ofte inkluderer mange regneoperasjoner og regler som skal brukes sammen, vil de automatisk kodes til et høyt nivå innen kategorien *Using symbols, operations and formal language.*

Det ser ut som at lærestedene har lagt seg på en linje hvor de inkluderer mange oppgaver med høyt nivå på det matematiske temaet, men lavere nivå i de kognitive kategoriene. At lærested A har en kombinasjon av lavere andel oppgaver med høyt tematisk nivå enn de to andre lærestedene samtidig som nivåene innen de kognitive kategoriene ikke er høyere, indikerer at eksamensoppgavene fra lærested A sannsynligvis er lettere enn oppgavene fra lærested B og lærested C. Dette stemmer også godt med hvordan oppgavene ble oppfattet av meg og den andre koderen som løste disse oppgavene.

Som beskrevet i metodekapitlet, har jeg kodet skriftlige eksamensoppgaver fra to kull fra hvert lærested. Ved to av lærestedene, lærested B og lærested C, er eksamensoppgavene fra de to etterfølgende kullene veldig like. Fra det ene lærestedet ligger alle oppgaver tilgjengelig på nett, ved det andre lærestedet vet jeg ikke om studentene har tilgang til tidligere eksamensoppgaver. Uansett er nok muligheten for å få tak i tidligere oppgaver sannsynlig tilstede. Selv om jeg ikke har kodet med hensyn på hvor like eksamenssettene er, mener jeg at min oppfatning av dette er verdt å trekke frem. At eksamensoppgavene likner på hverandre er potensielt problematisk, spesielt fordi en så stor andel av oppgavene ikke er innen et høyt nivå i noen annen kognitiv kategori enn i *Use of symbols, operations and formal language*. Dette er oppgaver som det etter min erfaring vil være spesielt lett å pugge seg til løsningen på av flere grunner. De er ikke vanskelige å lese, det må ikke utformes noen løsningsstrategi, det kreves ikke oversettelser til noe utenfor matematikken og det må til en liten grad resonneres.

5.4 Kodesamsvar

Kodesamsvaret mellom den andre koderen og meg varierer fra moderat til godt. *Using symbols, operations and formal language*, *Representation* og *Devising strategies* er de kategoriene hvor vi er mest uenige. I disse kategoriene er samsvaret moderat. Jeg har forsøkt å analysere og muligens oppklare uenighetene i kodingen ved å se på kodeskjemaene våre og sammenlikne. Jeg har vært spesielt interessert i å undersøke om det ser ut som vi konsekvent har kodet skjevt i noen kategorier. Gjennom samtaler med den andre koderen i etterkant av kodearbeidet har jeg også forsøkt å oppklare hva disse skjevhetene i koding kan skyldes.

Når det gjelder *Representation* ser det ut som at forskjeller i koding kan skyldes at vi er uenige om hva som skal oppfattes som en representasjon. I kodeinstruksen fra rammeverket beskrives hva som skal oppfattes som en representasjon på følgende måte (Turner et al., 2015, s. 4):

> *By «representation of a mathematical entity» we understand a concrete expression (mapping) of a mathematical concept, object, relationship, process or action. It can be physical, verbal, symbolic, graphical, tabular, diagrammatic or figurative.*

Det er store muligheter for skjønn i tolkningen av hva som er en fysisk, verbal eller symbolsk representasjon. Vil det for eksempel være slik at enhver formel skal oppfattes som en symbolsk representasjon? Jeg har ikke kodet oppgaven til et høyere nivå innen *Representation* for alle oppgaver som inneholder en formel, men det ville kanskje være mulig å tolke rammeverket slik. For *Representation* er beskrivelsen av hva som kjennetegner nivå 0 følgende (Turner et al., 2015, s. 4):

> *Either no representation is involved, or read isolated values from a simple representation, for example from a coordinate system, table or bar chart; or plot such values; or read isolated numeric values directly from text.*

Hva betyr det at en representasjon er enkel? Her er det også rom for tolkningsforskjeller og det er sannsynlig at tolkningsforskjeller av denne typen er med på å forklare et moderat samsvar mellom den andre koderen og meg.

Samsvaret er moderat også for *Devising strategies*. Kategorien beskriver i hvilken grad den som løser oppgaven må lage, konstruere eller aktivere løsningsstrategier; i hvilken grad det må utvikles en overordnet plan for å løse problemet. Det ligger muligheter for tolkningsforskjeller også her. Som beskrevet i eksempelkoding 1 (s. 38) er det mulig å tolke ulikt om en oppgaves utfordring er knyttet til å lage en løsningsstrategi eller å resonnere, eventuelt om det handler om å kunne algebraiske manipulasjoner. Jeg ser at den andre koderen i noen tilfeller har kodet ett nivå høyere enn meg under *Devising strategies*, men lavere enn meg i *Using symbols, operations and formal language* og *Reasoning and argument*. Det ser altså ut som vi er uenige om hvorvidt oppgavene utfordrer studentene ved at de må finne løsningsstrategier, om de må resonnere seg frem til en løsning elle rom de rett og slett må kunne regler og bruk av algebraiske manipulasjoner. Slike tolkningsforskjeller kan forklare hvorfor vi i noen tilfeller er uenige med ett nivå og kan sannsynligvis være med på å forklare noe av uenighetene i koding i disse tre kategoriene.

I kategorien *Communication* er det godt eller svært godt samsvar i kodingen avhengig av om man velger å bruke Altman eller Hallgren sin tabell over sammenheng mellom samsvar og κ-verdi. Vi er altså i stor grad enige om nivået innen denne kategorien. Det er mulig å anta at dette har å gjøre med at nivåbeskrivelsene for denne kategorien er enkle å tolke. Når jeg har gått inn og sett på kodingene ser det ut som at det som er av forskjeller i kodingen i denne kategorien, skyldes ulik tolkning av om man skal ta hensyn til informasjon som er presentert i tidligere oppgaver dersom denne er nødvendig for å løse den aktuelle oppgaven. Jeg har tolket det som at informasjonen må medberegnes for hver eneste oppgave der den skal brukes, mens det virker som den andre koderen til en mindre grad har hensyntatt tidligere gitt oppgaveinformasjon.

Mathematising var den kategorien som var vanskeligst å kode, men samsvaret mellom oss er likevel godt med en vektet kappa på 0,7. At samsvaret ble så godt kan til en viss grad skyldes at en stor andel av oppgavene var opplagt «purely intra-mathematical» som skal kodes til nivå 0.[4] Fra tabell 6 ser vi at 62 % av oppgavene er kodet nettopp til nivå 0. Oppgaver av typen «Integrer uttrykket f(x) = 2x» eller «Regn ut $\frac{4x^2}{3x} + \frac{4}{6x}$» er eksempler på denne typen oppgaver

[4] Se vedlegg 1 for beskrivelse av nivå 0 innen kategorien Mathematising.

hvor det er opplagt at studentene ikke måtte oversette mellom noe utenfor matematikken og et matematisk språk. De oppgavene som var vanskelige å kode i denne kategorien, var de hvor man skulle vurdere om de var innen nivå 1, 2 eller 3. Det er mulig at samsvaret ville blitt dårligere dersom en større del av datamaterialet svarte til et høyere nivå innen *Mathematising*.

Etter å ha diskutert med den andre koderen, ser jeg at noe av uenigheten i koding skyldes konsekvente tolkningsforskjeller. Jeg har beskrevet de tydeligste av disse i avsnittene over. I tillegg ser det ut som at den andre koderen jevnt over har tolket rammeverket litt strengere enn meg. Unntaket fra dette er kategorien *Devising strategies*. Med en strengere tolkning, mener jeg at han har hatt en tendens til å kode de kognitive kategoriene i litt lavere nivåer enn meg. Fordi det delvis moderate samsvaret i koding ser ut til å delvis kunne skyldes konsekvente tolkningsforskjeller, er det rimelig å anta at min koding likevel har vært forholdsvis intrakonsistent. Det at variasjonen i tolkning ikke ser ut til kun å være tilfeldig, gir meg grunn til å anta at min tolkning har vært konsekvent gjennom kodingen av alle eksamensoppgavene.

Dersom den andre koderens tolkning av rammeverket er riktigere enn min, har jeg gjennomgående kodet i litt for høye nivåer i de kognitive kategoriene. I så fall burde funnene mine knyttet til kognitiv kategori vist at en enda lavere andel av oppgavene skårer innen de høyere nivåene i de kognitive kategoriene (tabell 6 og tabell 8). Dette betyr at argumentene som er trukket frem i diskusjonen over som går på at en liten andel av eksamensoppgavene er i de høye nivåene i de seks kognitive kategoriene, sannsynligvis ville vært like aktuelle også med en annen koder enn meg til å analysere datamaterialet.

At samsvarsverdiene er høyere for vektet Cohens kappa gjennomsnittsmål enn for vektet Cohens kappa enkeltmål, viser at reliabiliteten til studien ville blitt høyere dersom vi hadde vært to kodere til å kategorisere hele datamaterialet. Et forlenget samarbeid med den andre koderen i forkant av kodearbeidet ville sannsynligvis også redusert uoverensstemmelsene i kodingen, spesielt dersom vi hadde kodet et stort volum av oppgaver parallelt og diskutert uoverensstemmelsene i forkant av kodingen av datamaterialet som er inkludert. Det er tydelig at slike diskusjoner i ettertid av koding av et stort volum av oppgaver er fruktbart med tanke på å utvikle en god forståelse av rammeverket.

5.5 En evaluering av bruk av MEG-studiens rammeverk i min studie

Fordi valg av rammeverk, og bruk av dette, har vært med på å forme studien min, synes jeg at det er naturlig å inkludere en diskusjon av hvor godt det valgte rammeverket har fungert. Rammeverket jeg har brukt er i utgangspunktet utviklet for å kunne forutsi vanskelighetsgraden til PISA-oppgaver som skal teste ungdomsskoleelever. Det har derfor vært interessant å vurdere i hvilken grad rammeverket har egnet seg til å si noe om karakteristikker av, og forskjeller mellom, oppgaver på et annet faglig nivå.

Kategorisering av oppgaver ved hjelp av dette rammeverket har egnet seg godt til å trekke frem noen kjennetegn og karakteristikker ved lærestedene sett samlet. Det har egnet seg til å analysere hvordan lærerstudentene utfordres kognitivt, ikke bare hvilke regler og regneferdigheter de må beherske til eksamen. Mitt inntrykk er at viktige karakteristikker ved datamaterialet har kommet godt frem ved hjelp av rammeverket. At få oppgaver tester studentenes evne til problemløsing og at størsteparten av oppgavene stort sett bare tester lærerstudentenes regneferdigheter og regelkunnskap (jamfør kapittel 5.3.1), er eksempler på slike karakteristikker.

Samsvaret i kodingen mellom den andre koderen og meg var godt nok til at jeg vil si at tolkningen av kodeinstruksen var mulig å overføre til denne typen oppgaver. I MEG-studien analyserte de samsvaret mellom åtte kodere. De vurderte samsvaret ved hjelp av Cronbachs alfa (Turner et al., 2013). For å sammenlikne samsvaret i min koding med samsvaret i MEG-studien beregnet jeg Cronbachs alfa også for mine resultater (se tabell 10). Vi ser at i snitt er samsvaret i MEG-studien noe bedre, men vet samtidig at dette er samsvaret for åtte kodere som har vært med på å utvikle rammeverket. Dermed tenker jeg at det er naturlig at samsvaret er noe bedre enn i min studie. At differansen mellom samsvarsverdiene ikke er større enn den er, indikerer at rammeverket kan egne seg for å brukes utenfor den typen oppgaver som det opprinnelig er utviklet for.

I MEG-studien var målet å utvikle et rammeverk som kunne forutse vanskelighetsgraden til PISA-oppgaver, og forskergruppen konkluderte med at rammeverket var godt egnet til dette (Turner et al., 2013). Fokuset mitt i denne studien har ikke på samme måte vært å forutse vanskelighetsgraden til oppgavene, men heller å evaluere hva de tester, spesielt hvilke kompetanser de tester hos lærerstudentene. Dersom jeg skulle ønsket å evaluere hvordan rammeverket fungerer til å forutsi vanskelighetsgraden til mine data, måtte jeg hatt tilgang til eksamensresultater slik at jeg kunne vurdert om de oppgavene som kodes til høye nivåer i de kognitive kategoriene, også slår vanskelig ut på eksamen. I MEG-studien hadde de et objektivt mål på om det fungerte å forutse vanskelighetsgraden til PISA-oppgaver, nemlig resultatet på PISA-undersøkelsen for disse oppgavene (Turner et al., 2013). Å gjennomføre en liknende studie med oppgaver fra GLU 5-10-utdaningen kan være en mulighet for et senere prosjekt.

Jeg har ikke hatt noe objektivt mål på hvor vanskelig de ulike oppgavene ble oppfattet av lærerstudentene, men jeg har likevel kunnet vurdere om mitt inntrykk av vanskelighetsgrad stemmer overens med det som kommer frem etter kategoriseringen. Inntrykket til den andre koderen og jeg er at forskjellen i vanskelighetsgrad på oppgavene mellom de ulike lærestedene er enda større enn det som kom frem i mine undersøkelser. Den kategorien hvor dette har fått størst utslag, er *Using symbols, operations and formal language*. Nivåbeskrivelsene for denne kategorien skiller mellom antall regneoperasjoner og antall formler og regler som skal tas i bruk, i alle fall delvis fremfor vanskelighetsgraden til det som skal brukes (Turner et al., 2015). Jeg synes derfor at det har vært vanskelig å få fram variasjonen i vanskelighetsgrad i temaet til oppgaven. Det har derfor blitt slik at en oppgave som innebærer mange trinn med omformulering av et algebraisk uttrykk kanskje skårer like høyt i dette rammeverket, som en oppgave som inkluderer integrasjon, selv om det nok ikke er slik at disse oppfattes like vanskelig for studenten. Dette er et argument for at rammeverket muligens egner seg dårligere til å kode oppgaver som ligger på et høyere nivå enn de som brukes i PISA-undersøkelsen og som rammeverket er utviklet for. Jeg har imidlertid dekket noe av dette i min studie ved å inkludere koding i og utenfor læreplanen 8-10.

En annen svakhet ved rammeverket er at det ikke har et reelt nivå 0. Det laveste nivået betyr ikke at denne kompetansen nødvendigvis er fraværende. For eksempel i *Devising strategies* er det nivå 0 dersom studentene ikke trenger å lage noen strategi, men også dersom strategien er opplagt. Dersom det hadde vært slik at rammeverket brukte nivå 0-4 i stedet for 0-3, der nivå 0 betydde fravær av kompetansen, ville det kunnet nyansert kodingen og analysen ytterligere.

5.6 Styrker og svakheter ved egen oppgave

Denne studien er gjennomført med fokus på at den skal være reliabel og valid. Jeg har til enhver tid etterstrebet å sette meg inn i litteratur og rammeverk slik at begrepsvaliditeten ikke skulle trues. Det at jeg har brukt et eksisterende rammeverk som er knyttet til begrepet matematikk-kompetanse, har sikret operasjonaliseringen slik at det faktisk er matematikkompetansen som måles når jeg tester mine eksamensoppgaver.

Jeg ser imidlertid at jeg med fordel til en enda større grad enn jeg gjorde, kunne trent på bruk av og diskutert tolkninger av rammeverket utviklet i MEG-studien. Dette ville gjort meg enda bedre skikket til å gjennomføre kodingen i de kognitive kategoriene. Jeg lærte mye gjennom diskusjonene som den andre koderen og jeg hadde i etterkant av kodearbeidet, og ser at kodesamsvaret sannsynligvis ville blitt bedre dersom vi hadde hatt mulighet til å gjennomføre en like omfattende prøvekoding før jeg gjennomførte kodingen av datamaterialet som ble inkludert i studien. At det er to som har kodet deler av datamaterialet slik at det har vært mulig å undersøke samsvar i koding i kognitiv kategori, har styrket troverdigheten til studien ved å gi et mål på reliabiliteten. Kodesamsvaret er moderat, godt eller svært godt i de ulike kategoriene. Jeg vil si at dette er greit, men funnene i studien ville selvsagt fått styrket troverdighet dersom samsvaret var enda bedre.

Studien er gjennomført slik at det skal være mulig å etterprøve den, det skal være mulig å reprodusere det jeg har gjort. Etterprøvbarhet og gjennomsiktighet er viktig i studier med et kvantitativt preg som denne og er med på å løfte forskningskvaliteten (L. Cohen et al., 2011). Slutninger om og diskusjoner rundt resultatene er fundamentert i teori om hva som kjennetegner god matematikkompetanse og undervisningskunnskap i matematikk. Dette er gjort for at de slutningene som jeg har trukket skal være valide.

Det er en svakhet ved min oppgave at utvalget har vært for lite til å kunne si noe om lærerutdanningen i Norge generelt og til å påpeke eventuelle årsakssammenhenger. En tverrsnittstudie hvor jeg kunne kartlagt alle eksamensoppgaver fra samtlige GLU 5-10-utdanninger i Norge ville hatt en klar styrke ved muligheten for en slik generalisering. Som tidligere angitt måtte jeg da hatt metoder for å vurdere muntlige eksamener og mappevurderinger. Likevel mener jeg at en komparativ casestudie, slik som den jeg har gjennomført er nyttig med tanke på å at den kan avdekke tendenser og karakteristikker og vise frem nye mulige problemstillinger.

6 Oppsummering og implikasjoner

6.1 Oppsummering

Min analyse av skriftlige skoleeksamensoppgaver i matematikk fra tre GLU 5-10-utdanninger viser at lærerstudentene ved disse lærestedene har blitt testet på et lavt kognitivt nivå ut fra en kategorisering i nivårammeverket utviklet i MEG-studien. Dette står i motsetning til forskning som viser at lærerne trenger å beherske matematikk på et høyt kognitivt nivå for å gjennomføre god undervisning (Niss, 2006). Lærestedene som er inkludert i studien skiller seg lite fra hverandre når det kommer til nivåfordelingen i de kognitive kategoriene, de laveste nivåene dominerer i alle kategorier unntatt *Using symbols, operations and formal language* og *Reasoning and argument* ved alle lærestedene. Karakterundersøkelsen i matematikk i GLU-utdanningene i 2014 etterlyser etablering av et felles språk for vurderingspraksisen ved de ulike lærerutdanningene i Norge (Christiansen et al., 2015). Ut fra en kategorisering i rammeverket utviklet i MEG-studien, ser det ut som at det er stor enighet i hvilket nivå skoleeksamensoppgavene vurderer, men jeg er usikker på om det er dette nivået vi ønsker å være enige om.

I innledningen trekker jeg frem jeg følgende sitat fra de nasjonale retningslinjene for grunnskolelærerutdanningen 5.-10.trinn (Kunnskapsdepartementet, 2010b, s. 34):

> *Gjennom matematikkfaget for trinn 5-10 skal studentene utvikle undervisningskunnskap i matematikk. Dette innebærer at de må ha en solid og reflektert forståelse for den matematikken elevene skal lære og hvordan denne utvikles videre på de neste trinnene i utdanningssystemet*.

At en stor del av eksamensoppgavene som er inkludert i denne studien gjennomgående skårer på de lave nivåene i alle kategorier unntatt *Using symbols, operations and formal language* og *Reasoning and argument* vitner ikke om vurdering av en solid og reflektert forståelse slik som retningslinjene for utdanningen etterspør.

Andelen matematikkoppgaver og fagdidaktiske oppgaver gitt til skriftlig skoleeksamen varierer fra 3 % til 23 % mellom de ulike lærestedene. Hjemmeeksamen og mappevurderinger er ikke inkludert. Samtidig er det ulike deler av undervisningskompetansen som kan testes ved en hjemmeeksamen og en skoleeksamen. Det er ved en skoleeksamen at man til størst grad får testet paratheten og evnene til å finne spontane løsninger knyttet til klasseroms- og elev-konteksten, slik mye av lærerhverdagen kommer til å være (Christiansen et al., 2015). Jeg anser det derfor som viktig at norske lærerstudenter testes i fagdidaktiske oppgaver, gjerne direkte knyttet til en matematisk oppgave som de nettopp har løst, også til skoleeksamen.

Lærestedene skiller seg noe fra hverandre når det kommer til den tematiske fordelingen innen de fem emneområdene *Tall og algebra*; *Geometri*; *Måling*; *Sannsynlighet, statistikk og kombinatorikk;* og *Funksjoner*. Det ser også ut som at noen av disse områdene, spesielt *Geometri* og *Måling*, til en liten grad dekkes i eksamensoppgavene ved samtlige læresteder. Lærestedene skiller seg fra hverandre når det kommer til hvor stor andel av oppgavene som tematisk ligger innenfor læreplanen i matematikk for 8.-10. trinn. Alle lærestedene har oppgaver som tilsvarer pensum i ungdomsskolen og i videregående skole, men andelene oppgaver innenfor læreplanen 8-10 varierer mellom 18 % og 49 %.

Det er viktig at lærere har en kombinasjon av horisontkunnskap og dyptgående kunnskap om det de skal undervise (Ball et al., 2008; Niss, 2006; Shulman, 1987). Det er ikke noe opplagt svar på hvor stor andel av oppgavene som bør ligge innenfor og utenfor ungdomsskolepensum. At Lærested A har både høy andel av sine oppgaver hentet fra ungdomsskolepensum og oppgaver på et lavt kognitivt nivå, vitner imidlertid kanskje om en for stor andel oppgaver som verken vurderer horisontkunnskap eller dyptgående kunnskap om det studentene skal undervise. At en stor andel av oppgavene er kodet utenfor læreplanen 8-10 ved Lærested B og Lærested C viser at mange av oppgavene herfra er vanskelige tematisk. Jeg hevder imidlertid at en del av oppgavene med fordel kunne vært byttet ut med tematiske enklere oppgaver som utfordrer studentene mer kognitivt fordi dette er med på å vurdere viktig dyptgående kunnskap og forståelse. Dersom norske lærerstudenter generelt testes på et så lavt nivå innen de kognitive kategoriene som min studie kan antyde, mener jeg at dette er noe som bør diskuteres videre også på andre arenaer enn i denne oppgaven.

6.2 Implikasjoner og forslag til nye studier

Funnene i denne studien er interessante som del av en kartlegging av hva som undervises og vurderes ved ulike grunnskolelærerutdanninger i Norge. Studien gir imidlertid bare innblikk i en deler av vurderingspraksisen ved tre utvalgte læresteder. Jeg tror at en kartlegging av både hva som undervises og hva som vurderes ved lærerutdanningene vil være viktige innspill i en diskusjon om nivået på norske lærerstudenter, hvor blant annet resultater fra TEDS-M 2008 trekkes frem. I vurderingen av om det bør innføres nasjonale prøver i matematikk i lærerutdanningen, slik nåværende kunnskapsminister har foreslått (Svarstad, 2014), kan det også være nyttig å ha gjennomført en kartlegging av eksisterende vurderingspraksis. Kartleggende studier for en mulig ensretting innen lærerutdanningene blir etterspurt også i rapporten fra Karakterundersøkelsen 2014 (Christiansen et al., 2015). Min studie er et eksempel på hvordan slike kartleggende studier kan gjennomføres. Dersom man skulle ønske å gjennomføre en liknende studie for samtlige grunnskolelærerutdanninger i Norge, vil det være nødvendig å utvikle en metode og et rammeverk for å analysere muntlige eksamener, fordi mange læresteder inkluderer dette i studieløpet. Dersom man ønsker å kartlegge undervisningspraksis i tillegg til vurderingspraksis, må det settes inn tid og ressurser til observasjonsstudier, kartlegging av hva som finnes på digitale læringsplattformer og i lærebøker. Jeg mener imidlertid at dette er nyttig og viktig forskning å gjennomføre. Det vil for eksempel være mulig å hente inspirasjon fra den kartleggende studien som ble gjennomført av Tonheim og Torkildsen (2010) før overgangen fra ALU til GLU.

Det kan være nyttig å gjennomføre en tilsvarende studie som denne med et annet rammeverk for koding i kognitiv kategori. Det vil for eksempel være mulig å bruke en kombinasjon av Blooms taksonomi og TIMSS sitt rammeverk for beskrivelse av matematisk kompetanse inspirert av Pedersen (2013). Det ville vært spennende å se om et annet rammeverk ville kunne finne andre karakteristikker og tendenser enn bruk av rammeverket fra MEG-studien.

Jeg har ikke hatt noe objektivt mål på hvor vanskelig de eksamensoppgavene som er inkludert i denne studien ble oppfattet av lærerstudentene som gjennomførte eksamenene. Dersom det er mulig å innhente informasjon om hvordan lærerstudenter skårer på hver enkelt eksamensoppgave, kan det sannsynligvis være spennende å sette dette opp mot den kognitive vanskelighetsgraden til de respektive oppgavene. Da blir det mulig å si noe om hvilke kognitive kategorier, altså hvilke deler av den matematiske kompetansen, som studentene oppfatter som spesielt utfordrende. Dette kan og bør i så fall få implikasjoner for undervisningen og vurderingen ved de aktuelle lærestedene med tanke på å bygge opp den delen av den matematiske kompetansen som studentene ikke mestrer.

Litteraturliste

Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

Bakke, B., & Bakke, I. N. (2011). *Grunntall 10. Matematikk for ungdomstrinnet*. Drammen: Elektronisk undervisningsforlag AS.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of teacher education, 59*(5), 389-407. doi: 10.1177/0022487108324554

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133-180. doi: 10.3102/0002831209345157

Cappelen Damm. (2013). Sinus 2P. Endringer i læreplanen.   Lastet ned 10. april, 2015, fra http://sinus2p.cappelendamm.no/

Christiansen, A., Enge, O., & Lode, B. (2015). Rapport fra karakterundersøkelsene i matematikk i GLU-utdanningene i 2014. http://www.uhr.no/documents/Rapport_17032015_ENDELIG.pdf

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

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(1), 37-46. doi: 10.1177/001316446002000104

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education*: London: Routledge.

Enge, O., & Valenta, A. (2010). Utvikling av matematikklærerkompetansen hos studenter i allmennlærerutdanning. *Forskning og utvikling i praksis, 4*(3), 61-77. http://tapir.pdc.no/pdf/FOU/2010/2010-03-5.pdf

Fauskanger, J., Mosvold, R., & Bjuland, R. (2010). Hva må læreren kunne? *Tangenten 21*(4), 35-38. http://www.caspar.no/artikkel_pdf/35c_t2010-4.pdf

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), 613-619. doi: 10.1177/001316447303300309

Grossman, P., & Wineburg, S. (2001). Lee S. Shulman. I J. A. Palmer (Red.), *Fifty Modern Thinkers on Education: From Piaget to the Present*. London - New York: Routledge.

Grønmo, L. S., Lindquist, M., & Arora, A. (2014). TIMSS Advanced 2015 Mathematics Framework. I I.V.S. Mullis & M.O. Martin (Red.), *TIMSS Advanced 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Grønmo, L. S., & Onstad, T. (2012). *Mange og store utfordringer. Et nasjonalt og internasjonalt perspektiv på utdanning av lærere i matematikk basert på data fra TEDS-M*. Oslo: Unipub.

Grønmo, L. S., Onstad, T., Nilsen, T., Hole, A., Aslaksen, H., & Borge, I. C. (2012). *Framgang, men langt fram : norske elevers prestasjoner i matematikk og naturfag i TIMSS 2011*. Oslo: Akademika forlag.

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

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, *8(1)*, 23-34. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/

Kilpatrick, J. (2014). Competency Frameworks in Mathematics Education. I S. Lerman (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.

Kim, H. (2013). Statistical notes for clinical researchers: Evaluation of measurement error 1: using intraclass correlation coefficients. *Restorative Dentistry & Endodontics, 38*(2), 98-102. doi: 10.5395/rde.2013.38.2.98

Kjærnsli, M., & Olsen, R. V. (2013). PISA 2012 - sentrale funn. I M. Kjærnsli & R. V. Olsen (Red.), *Fortsatt en vei å gå: Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012* (s. 13-40). Oslo: Universitetsforlaget.

Kleven, T. A., Tveit, K., & Hjardemaal, F. (2014). *Innføring i pedagogisk forskningsmetode: En hjelp til kritisk tolking og vurdering* (2 utg. Vol. 2). Bergen: Fagbokforlaget.

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

Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *The International Journal on Mathematics Education, 40*(5), 873-892. doi: 10.1007/s11858-008-0141-9

Kunnskapsdepartementet. (2009). *St. melding nr. 11 (2008-2009). Læreren. Rollen og utdanningen*. Lastet ned fra https://www.regjeringen.no/nb/dokumenter/stmeld-nr-11-2008-2009-/id544920/.

Kunnskapsdepartementet. (2010a). *Forskrift om rammeplan for grunnskolelærerutdanningene for 1.–7. trinn og 5.–10. trinn*. Lastet ned fra https://www.regjeringen.no/no/dokumenter/forskrift-om-rammeplan-for-grunnskolelar/id594357/

Kunnskapsdepartementet. (2010b). *Nasjonale retningslinjer for grunnskolelærerutdanningen 5. – 10. trinn*. Lastet ned fra https://www.regjeringen.no/globalassets/upload/kd/rundskriv/2010/retningslinjer_grunnskolelaererutdanningen_5_10_trinn.pdf.

Kunnskapsdepartementet. (2014). *Lærerløftet*. Lastet ned fra https://www.regjeringen.no/globalassets/upload/kd/vedlegg/planer/kd_strategiskole_web.pdf.

Matematikksenteret. (2006). Kompetanser i matematikk (Niss 2002). Lastet ned 15. januar, 2015, fra http://matematikksenteret.no/content/307/Kompetanser-i-matematikk-Niss-2002

Matematikksenteret. (2015). Digitale verktøy til eksamen i matematikk. Lastet ned 1. juni, 2015, fra http://www.matematikksenteret.no/ny-eksamen-v2015/

Mosvold, R., & Fauskanger, J. (2012). Testing av matematikklærere - Nei takk, men ja til faglige diskusjoner. *Bedre skole*, *2*, 52-55. https://www.utdanningsforbundet.no/upload/Tidsskrifter/Bedre%20Skole/BS_nr_2-2012/5502-BS-2-12-web-ny_Mosvold_Fauskanger.pdf

Mullis, I. V. S. (2014). Introduction. I I. V. S. Mullis & M. O. Martin (Red.), *TIMSS 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics* (Vol. 1). Reston, Virginia: NCTM.

Niss, M. (1999). Aspects of the Nature and State of Research in Mathematics Education. *Educational Studies in Mathematics, 40*(1), 1-24. doi: 10.1023/A:1003715913784

Niss, M. (2003). *Mathematical competencies and the learning of mathematics: The Danish KOM project.* Paper presentert ved 3rd Mediterranean conference on mathematical education. Hentet fra http://www.math.chalmers.se/Math/Grundutb/CTH/mve375/1213/docs/KOMkompetenser.pdf.

Niss, M. (2006). What does it mean to be a competent mathematics teacher? A general problem illustrated by examples from Denmark *Praktika, 23 o Panellenio Synedrio Mathematikis Paideias. Patra 24-26 Noembriou 2006* (s. 39-47). Elleniki Mathematiki Etaireia, Patra.

Niss, M. (2015). Mathematical Competencies and PISA. I K. Stacey & R. Turner (Red.), *Assessing Mathematical Literacy:The PISA Experience* (s. 35-55). Switzerland: Springer International Publishing.

Niss, M., & Jensen, T. H. (2002). *Kompetencer og matematiklæring: Ideer og inspiration til udvikling af matematikundervisning i Danmark* København: Undervisningsministeriet.

Norsk samfunnsvitenskapelig datatjeneste. (2015). Må prosjektet meldes? Lastet ned 2. februar, 2015, fra http://www.nsd.uib.no/personvern/meldeplikt/

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. 43-62). Oslo: Universitetsforlaget.

OECD. (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*: OECD Publishing.

Olsen, R. V., & Grønmo, L. S. (2006). What are the Characteristics of the Nordic Profile in Mathematical Literacy? I J. Mejding & A. Roe (Red.), *Northern Lights on PISA 2003: A Reflection from the Nordic Countries* (s. 47-57). København: Nordisk ministerråd.

Onstad, T. (2010). Rammeverk og metoder. I L. S. Grønmo, T. Onstad & I. F. Pedersen (Red.), *Matematikk i motvind. TIMSS Advanced 2008 i videregående skole.* (s. 235-258). Oslo: Unipub.

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), 1-24. https://www.journals.uio.no/index.php/adno/article/view/1112

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational researcher, 31*(7), 3-14. doi: 10.3102/0013189X031007003

Ryghaug, M. (2002). Å bringe tekster i tale–mulige metodiske innfallsvinkler til tekstanalyse i statsvitenskap. *Norsk statsvitenskapelig tidsskrift*, *18*(4), 303-327.

Samordna opptak. (2015). Lærested og studier i Samordna opptak 2015. Lastet ned 15. mai, 2015, fra https://sok.samordnaopptak.no/studier?ord=&laerestedkode=&utdomrkode=L%C6RER&stikkordnr=3

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, *15*(2), 4-14. doi: 10.3102/0013189X015002004

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review, 57*(1), 1-23. doi: http://dx.doi.org/10.17763/haer.57.1.j463w79r56455411

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257-268. http://ptjournal.apta.org/content/85/3/257.long

Sterri, A. B., & Wæhle, E. (2014). Case studie. I Store norske leksikon. Lastet ned 03.03.2015, fra https://snl.no/case_studie.

Svarstad, J. (2014, 8. september). Røe Isaksen skal gi studentene nasjonale prøver, *Aftenposten*. Lastet ned fra http://www.aftenposten.no/nyheter/iriks/Roe-Isaksen-skal-gi-studentene-nasjonale-prover--7695199.html

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (Red.). (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics. Conceptual Framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Tjeldvoll, A., & Skagen, K. (2014). Didaktikk. I Store norske leksikon. Lastet ned 25.05.2015, fra https://snl.no/didaktikk.

Tonheim, O. H. M., & Torkildsen, O. E. (2010). Matematikk I i lærarutdanninga – kvalifiserande? I P. Haug (Red.), *Kvalifisering til læraryrket* (s. 209-230). Oslo: Abstrakt forlag.

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* (s. 85-115). Switzerland: Springer International Publishing.

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). Netherlands: Springer.

Universitets- og høgskolerådet. (2013). Karakterundersøkelsen i matematikk. Notat fra arbeidsgruppen. http://www.uhr.no/documents/Karakterundersokelse_i_matematikk___notat_fra_arbeidsgruppen_1.pdf

Utdanningsdirektoratet. (2006a). *Læreplan i matematikk for realfag - programfag i studiespesialiserende utdanningsprogram. Læreplankode MAT3-01*. Lastet ned fra http://data.udir.no/kl06/MAT3-01.pdf?lang=nob.

Utdanningsdirektoratet. (2006b). *Læreplan i matematikk for samfunnsfag - programfag i studiespesialiserende program. Læreplankode MAT4-01*. Lastet ned fra http://data.udir.no/kl06/MAT4-01.pdf?lang=nob.

Utdanningsdirektoratet. (2006c). *Læreplan i matematikk X - programfag i studiespesialiserende utdanningsprogram. Læreplankode MAT2-01*. Lastet ned fra http://www.udir.no/kl06/MAT2-01/.

Utdanningsdirektoratet. (2013). *Læreplan i matematikk fellesfag. Læreplankode MAT1-04*. Lastet ned fra http://data.udir.no/kl06/MAT1-04.pdf?lang=nno.

Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning*, *29*(5), 403-413. doi: 10.1111/jcal.12029

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

Vedlegg

Vedlegg 1: Rammeverk for koding (Turner et al., 2015)

Competency definitions and their level descriptions

Communication

The communication competency has both 'receptive' and 'constructive' components. The receptive component includes understanding what is being stated and shown related to the mathematical objectives of the task, including the mathematical language used, what information is relevant, and what is the nature of the response requested. The constructive component consists of presenting the response that may include solution steps, description of the reasoning used and justification of the answer provided.

In written and computer-based items, receptive communication relates to understanding text and images, still and moving. Text includes verbally presented mathematical expressions and may also be found in mathematical representations (for example titles, labels and legends in graphs and diagrams).

Communication does not include knowing how to approach or solve the problem, how to make use of particular information provided, or how to reason about or justify the answer obtained, rather it is the understanding or presenting of relevant information. It also does not apply to extracting or processing mathematical information from representations. In computer-based items, the instructions about navigation and other issues related to the computer environment may add to the general task demand, but is not part of the communication competency.

Demand for the receptive aspect of this competency increases according to the complexity of material to be interpreted in understanding the task; the need to link multiple information sources or to move backwards and forwards (to cycle) between information elements. The constructive aspect increases with the need to provide a detailed written solution or explanation.

Definition: Reading and ***interpreting*** statements, questions, instructions, tasks, images and objects; ***imagining*** and ***understanding*** the situation presented and ***making sense*** of the information provided including the mathematical terms referred to; ***presenting*** and ***explaining*** one's mathematical work or reasoning.

Level 0: Understand short sentences or phrases relating to concepts that give immediate access to the context, where all information is directly relevant to the task, and where the order of information matches the steps of thought required to understand what the task requests. Constructive communication involves only presentation of a single word or numeric result.

Level 1: Identify and link relevant elements of the information provided in the text and other related representation/s, where the material presented is more complex or extensive than short sentences and phrases or where some extraneous information may be present. Any constructive communication required is simple, for example it may involve writing a short statement or calculation, or expressing an interval or a range of values.

Level 2: Identify and select elements to be linked, where repeated cycling within the material presented is needed to understand the task; or understand multiple elements of the context or task or their links. Any constructive communication involves providing a brief description or explanation, or presenting a sequence of calculation steps.

Level 3: Identify, select and understand multiple context or task elements and links between them, involving logically complex relations (such as conditional or nested statements). Any constructive communication would involve presenting argumentation that links multiple elements of the problem or solution.

Devising strategies

The focus of this competency is on the strategic aspects of mathematical problem solving: selecting, constructing or activating a solution strategy and monitoring and controlling the implementation of the processes involved. 'Strategy' is used to mean a set of stages that together form the overall plan needed to solve the problem. Each stage comprises a sub-goal and related steps. For example a plan to gather data, to transform them and to represent them in a different way would normally constitute three separate stages.

The knowledge, technical procedures, mathematising and reasoning needed to actually carry out the solution process are taken to belong to those other competencies.

Demand for this competency increases with the degree of creativity and invention involved in identifying a suitable strategy, with increased complexity of the solution process (for example the number, range and complexity of the stages needed in a strategy), and with the consequential need for greater metacognitive control in the implementation of the strategy towards a solution.

Definition: ***Selecting*** or ***devising*** a mathematical strategy to solve a problem as well as ***monitoring*** and ***controlling*** implementation of the strategy.

Level 0: Take direct actions, where the solution process needed is explicitly stated or obvious.

Level 1: Find a straight-forward strategy (usually of a single stage) to combine or use the given information.

Level 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.

Level 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.

Mathematatising

The focus of this competency is on those aspects of the modelling cycle that link an extra-mathematical context with some mathematical domain. Accordingly, the mathematising *competency has two components. A situation outside mathematics may require translation into a form amenable to mathematical treatment. This includes making simplifying assumptions, identifying variables present in the context and relationships between them, and expressing those variables in a mathematical form. This translation is sometimes referred to as mathematising. Conversely, a mathematical entity or outcome may need to be interpreted in relation to an extra-mathematical situation or context. This includes translating mathematical results in relation to specific elements of the context and validating the adequacy of the solution found with respect to the context. This process is sometimes referred to as de-mathematising.*

The intra-mathematical treatment of ensuing issues and problems within the mathematical domain is dealt with under other competencies. Hence, while the mathematising *competency deals with representing extra-mathematical contexts by means of mathematical entities, the representation of mathematical entities is dealt with under the* representation *competency*.

Demand for activation of this competency increases with the degree of creativity, insight and knowledge needed to translate between the context elements and the mathematical structures of the problem.

Definition: ***Translating*** an extra-mathematical situation into a mathematical model, ***interpreting*** outcomes from using a model in relation to the problem situation, or ***validating*** the adequacy of the model in relation to the problem situation.

Level 0: Either the situation is purely intra-mathematical, or the relationship between the extra-mathematical situation and the model is not relevant to solving the problem.

Level 1: Construct a model where the required assumptions, variables, relationships and constraints are given; or draw conclusions about the situation directly from a given model or from the mathematical results.

Level 2: Construct a model where the required assumptions, variables, relationships and constraints can be readily identified; or modify a given model to satisfy changed conditions; or interpret a model or mathematical results where consideration of the problem situation is essential.

Level 3: Construct a model in a situation where the assumptions, variables, relationships and constraints need to be defined; or validate or evaluate models in relation to the problem situation; or link or compare different models.

Representation

The focus of this competency is on decoding, devising, and manipulating representations of mathematical entities or linking different representations in order to pursue a solution. By 'representation of a mathematical entity' we understand a concrete expression (mapping) of a mathematical concept, object, relationship, process or action. It can be physical, verbal, symbolic, graphical, tabular, diagrammatic or figurative.

Mathematical tasks are often presented in text form, sometimes with graphic material that only helps set the context. Understanding verbal or text instructions and information, photographs and graphics does not generally belong to representation competency – that is part of the communication *competency. Similarly, working exclusively with symbolic representations lies within the* using symbols, operations and formal language *competency. On the other hand, translation between different representations is always part of the* representation *competency. For example, the act of transforming mathematical information derived from relevant text elements into a non-verbal representation is where* representation *commences to apply.*

While the representation *competency deals with representing* mathematical *entities by means of other entities (mathematical or extra-mathematical), the representation of extra-mathematical contexts by mathematical entities is dealt with under the* mathematising *competency*.

Demand for this competency increases with the amount of information to be extracted, with the need to integrate information from multiple representations, and with the need to devise representations rather than to use given representations. Demand also increases with added complexity of the representation or of its decoding, from simple and standard representations requiring minimal decoding (such as a bar chart or Cartesian graph), to complex and less standard representations comprising multiple components and requiring substantial decoding perhaps devised for specialised purposes (such as a population pyramid, or side elevations of a building).

Definition: ***Decoding***, ***translating*** between, and ***making use*** of given mathematical representations in pursuit of a solution; ***selecting*** or ***devising*** representations to capture the situation or to present one's work.

Level 0: Either no representation is involved, or read isolated values from a simple representation, for example from a coordinate system, table or bar chart; or plot such values; or read isolated numeric values directly from text.

Level 1: Use a given simple and standard representation to interpret relationships or trends, for example extract data from a table to compare values, or interpret changes over time shown in a graph; or read or plot isolated values within a complex representation; or construct a simple representation.

Level 2: Understand and use a complex representation, or construct such a representation where some of the required structure is provided; or translate between and use different simple representations of a mathematical entity, including modifying a representation.

Level 3: Understand, use, link or translate between multiple complex representations of mathematical entities; or compare or evaluate representations; or devise a representation that captures a complex mathematical entity.

Using symbols, operations and formal language

This competency reflects skill with activating and using mathematical content knowledge, such as mathematical definitions, results (facts), rules, algorithms and procedures, recalling and using symbolic expressions, understanding and manipulating formulae or functional relationships or other algebraic expressions and using the formal rules of operations (e.g. arithmetic calculations or solving equations). This competency also includes working with measurement units and derived quantities such as 'speed' and 'density'.

Developing symbolic formulations of extra-mathematical situations is part of mathematisation. *For example, setting up an equation to reflect the key elements of an extra-mathematical situation belongs to* mathematisation, *whereas solving it is part of the* using symbols, operations and formal language *competency. Manipulating symbolic expressions belongs to the* using symbols, operations and formal language *competency even though they are mathematical representations. However, translating between symbolic and other representations belongs to the* representation *competency.*

The term 'variable' is used here to refer to a symbol that stands for an unspecified number or a changing quantity, for example C and r in the formula $C = 2\pi r$.

Demand for this competency increases with the increased complexity and sophistication of the mathematical content and procedural knowledge required.

<u>Definition:</u> Understanding and ***implementing*** mathematical procedures and language (including symbolic expressions, arithmetic and algebraic operations), using the mathematical ***conventions*** and ***rules*** that govern them; ***activating*** and ***using knowledge*** of definitions, results, rules and ***formal systems***.

<u>Level 0:</u> State and use elementary mathematical facts and definitions; or carry out short arithmetic calculations involving only easily tractable numbers. For example, find the area of a rectangle given the side lengths, or write down the formula for the area of a rectangle.

<u>Level 1:</u> Make direct use of a simple mathematical relationship involving variables (for example, substitute into a linear relationship); use arithmetic calculations involving fractions and decimals; use repeated or sustained calculations from level 0; make use of a mathematical definition, fact, or convention, for example use knowledge of the angle sum of a triangle to find a missing angle.

<u>Level 2:</u> Use and manipulate expressions involving variables and having multiple components (for example, by algebraically rearranging a formula); employ multiple rules, definitions, results, conventions, procedures or formulae together; use repeated or sustained calculations from level 1.

<u>Level 3:</u> Apply multi-step formal mathematical procedures combining a variety of rules, facts, definitions and techniques; work flexibly with complex relationships involving variables, for example use insight to decide which form of algebraic expression would be better for a particular purpose.

Reasoning and argument

This competency relates to drawing valid inferences based on the internal mental processing of mathematical information needed to obtain well-founded results, and to assembling those inferences to justify or, more rigorously, prove a result.

Other forms of mental processing and reflection involved in undertaking tasks underpin each of the other competencies. For example the thinking needed to choose or devise an approach to solving a problem is dealt with under the devising strategies *competency, and the thinking involved in transforming contextual elements into a mathematical form is accounted for in the* mathematising *competency.*

The nature, number or complexity of elements that need to be brought to bear in making inferences, and the length and complexity of the chain of inferences needed would be important contributors to increased demand for this competency.

Definition: ***Drawing inferences*** by using logically rooted thought processes that explore and connect problem elements to ***form, scrutinise*** or ***justify arguments*** and conclusions.

Level 0: Draw direct inferences from the information and instructions given.

Level 1: Draw inferences from reasoning steps within one aspect of the problem that involves simple mathematical entities.

Level 2: Draw inferences by joining pieces of information from separate aspects of the problem or concerning complex entities within the problem; or make a chain of inferences to follow or create a multi-step argument.

Level 3: Use or create linked chains of inferences; or check or justify complex inferences; or synthesise and evaluate conclusions and inferences, drawing on and combining multiple elements of complex information, in a sustained and directed way.