KAPITTEL 2

Matematikkinnholdet i TIMSS, TIMSS Advanced og PISA

Arne Hole
Institutt for lærerutdanning og skoleforskning, UiO
Liv Sissel Grønmo
Institutt for lærerutdanning og skoleforskning, UiO
Torgeir Onstad
Institutt for lærerutdanning og skoleforskning, UiO

Den offentlige debatten om storskalaundersøkelser som PISA, TIMSS og TIMSS Advanced i Norge dreier seg i forbausende liten grad om det faglige *innholdet* i studiene, altså om *hva* de måler faglig. Et sentralt spørsmål i vurderingen av studienes relevans for norsk matematikkundervisning er i hvilken grad det de måler faktisk samsvarer med norske læreplanmål i faget. Undersøkelser om sammenheng mellom elevprestasjoner og bakgrunnsvariabler som måler elevbakgrunn, lærerkompetanse og liknende, har liten nytteverdi dersom prestasjonsdataene er målt på en måte som ikke er tilstrekkelig relevant for de læreprosessene som man ønsker skal foregå i skolen.

I dette kapitlet analyserer vi innholdet i de tre studiene TIMSS Advanced 2015, TIMSS 2011 matematikk 8. trinn og PISA 2012 matematikk ved å bruke et rammeverk for beskrivelse av matematikkinnhold. De valgte årene for studienes gjennomføring er velegnet for sammenlikning i et norsk perspektiv, siden studentkullet som ble testet i TIMSS Advanced i 2015 (13. trinn) er det samme som ble testet i PISA 2012 (10. trinn) og omtrent det samme som ble testet i TIMSS 2011 8. trinn. Vi bygger i dette kapitlet på resultater presentert på IEAs IRC-konferanser i 2015 og 2017 (Hole, Grønmo & Onstad, 2017; Hole et al., 2015). Dette gir et perspektiv på studiene som går på tvers av studienes egne underliggende rammeverk. Disse rammeverkene er svært ulike i sin natur, og vi skal starte med å se nærmere på dette. I kapittel 5 bruker vi klassifikasjonsresultatene vi får her til å undersøke ulike profiler blant land og regioner i verden når det gjelder prioritering og faglig kultur i matematikkundervisningen.

2.1 Prinsipielle ulikheter i rammeverk mellom TIMSS-studiene og PISA

Mens IEA-studiene TIMSS og TIMSS Advanced benytter seg av et rammeverk som essensielt er basert direkte på læreplanene i de deltakende landene eller utdanningssystemene (Mullis & Martin, 2014; Mullis et al., 2003), er rammeverket i PISA basert på en konkret matematikkdidaktisk teori som tar utgangspunkt i begrepet «mathematical literacy» (OECD, 2003). Røttene til PISAs rammeverk kan spores blant annet til det danske KOM-prosjektet (Kompetencer og matematiklæring) under ledelse av Mogens Niss rundt år 2000 (Niss, 1999b; Niss & Jensen, 2002), et rammeverk som igjen har fellestrekk med amerikanske *standards* fra 1980-tallet (NCTM, 1989). Her er det altså helt ulike tilnærmingsmåter som brukes i PISA og TIMSS. Mens PISA gjennom sitt teoretisk sett avklarte og spesifiserte rammeverk kan sies å innta en matematikkdidaktisk posisjon, slik at man i denne studien til og med kanskje kan sies å basere seg på et konkret «læringssyn», er det vanskelig å se noe tilsvarende i TIMSS og TIMSS Advanced. Selv om disse studiene også har et kognitivt rammeverk («cognitive domains» og liknende) som helt klart relaterer seg til matematikkdiaktiske begreper, er disse delene av studienes rammeverk så vidt nøytrale og grovkornede at de neppe kan sies å definere noen teoretisk «posisjon» på den måten PISAs rammeverk gjør. Skal man lete etter noe slikt for IEA-studiene, må man med andre ord gå til deltakerlandenes læreplaner. Det er disse som bestemmer det faglige innholdet i TIMSS og TIMSS Advanced, gjennom en form for flertallsavgjørelser. IEA-studiene har slik sett en mer pragmatisk og forsøksvis nøytral tilnærmingsmåte til matematikkdidaktikk.

Mens de faglige oppgavene i TIMSS og TIMSS Advanced altså ganske enkelt kan sees på som et uttrykk for hva som vektlegges i matematikkundervisningen blant de deltakende landene, er PISAs prestasjonsmålinger indirekte basert på et slags *anbefalt innhold* i matematikkundervisningen. Disse anbefalingene ligger implisitt i det underliggende kompetansebaserte rammeverket. I PISA modelleres begrepet «mathematical literacy» ved å splitte det opp i en liste av ulike matematiske *kompetanser*. Denne tenkningen ligger nært rammeverket til det danske KOM-prosjektet (Niss, 1999b), se også (Niss, 2015). I den norske rapporten fra PISA 2012 ble kompetansene i PISA beskrevet slik (Kjærnsli & Olsen, 2013):

1. Kommunisere med, i og om matematikk
2. Matematisere og modellere både matematiske og virkelige situasjoner
3. Representere matematiske størrelser, velge, veksle mellom og bruke matematiske representasjoner i oppgaveløsing
4. Resonnere og argumentere matematisk
5. Planlegge, velge ut og bruke problemløsingsstrategier
6. Bruke symbol- og formalspråk, regler og formelle matematiske metoder
7. Velge ut og bruke matematiske verktøy og hjelpemidler

Det kontroversielle med en slik oppsplittingsmåte ligger allerede i selve ideen om å *splitte opp* begrepet matematisk kompetanse: Noen vil kunne hevde at det ikke er mulig å arbeide meningsfylt med hver av disse syv kompetansene uten å bruke flere (alle?) på en gang, og at en oppsplitting av denne typen dermed leder oppmerksomheten bort fra det man kan hevde er et enhetlig mål som bør ligge til grunn for arbeidet med matematikk i skolen. Dette enhetlige målet kan kanskje beskrives delvis ved begreper som «matematisk forståelse», «abstraksjonsevne» og liknende – begreper som ikke passer direkte inn på denne kompetanselisten. Det er også interessant å diskutere dette i relasjon til begrepene *dybdelæring* og *progresjon*, som i skrivende stund er aktuelle i den norske læreplandebatten.

Kompetansetenkning basert på lister av enkeltstående *kompetanser* har fortsatt bred støtte i det vestlige matematikkdidaktiske miljøet, og vi har ikke til hensikt å ta stilling for eller mot en slik teoretisk vinkling her. Imidlertid er det av hensyn til våre analyser i dette kapitlet viktig at vi problematiserer det. Det er liten tvil om at tenkningen basert på listen av kompetanser gjengitt over har dratt det faglige innholdet i PISA mer i retning av såkalt hverdagsmatematikk enn hva som er tilfellet i læreplanene til mange land, og dermed også hva man finner i TIMSS og TIMSS Advanced. Dette slår tydelig ut når vi i dette kapitlet sammenlikner innholdet i PISA med IEA-studiene ved hjelp av vårt rammeverk, et rammeverk som måler avhengighet av matematisk teori.

2.2 Språk og innhold i skolematematikken: Funnet på og funnet ut

I dette delkapitlet skal vi beskrive en del av vår motivasjon for å utvikle det rammeverket for måling av matematisk teoriinnhold som vi skal bruke.

Uansett skoletrinn er det viktig at eleven klarer å skille ting matematikerne har *bestemt seg for* fra ting matematikerne har *oppdaget*. På norsk kan man si at det første området av matematisk teori består av ting som er *funnet på*, mens det andre består av ting som er *funnet ut*. Teknisk sett består området «funnet på» av definisjoner, notasjon og annen ren språkbruk. Området «funnet ut» består av matematiske *teoremer* eller *setninger*, ofte også kalt matematiske *resultater*. I området «funnet på» finner vi definisjoner av matematiske begreper som rektangel, parallellogram, kvadrattall og så videre. I området «funnet ut» ligger for eksempel Pytagoras' setning, arealformler for ulike geometriske figurer, algebraiske lover som for eksempel distributiv lov og andre elementer i matematisk teori som krever en *forklaring på hvorfor de er sanne*, formelt sett et *bevis*. I fortsettelsen vil vi bruke betegnelsene «forklaring» eller «begrunnelse» i stedet for bevis, fordi man sjelden arbeider med formelle bevis i skolen.

Røft sagt kan vi si at området «funnet på» representerer *språksiden* i matematikken, mens området «funnet ut» representerer *innholdet*. Merk at denne bruken av begrepet «innhold» fra noen synspunkter kan betraktes som misvisende; på en måte er selvsagt definisjoner også en del av «innholdet i matematikken». Likevel er vår erfaring at denne koblingen av de to områdene til «språk» og «innhold» leder tankene på en måte som i vår sammenheng er konstruktiv. Vi vil derfor referere til distinksjonen mellom «funnet på» og «funnet ut» som *LC-distinksjonen* (fra engelsk: *language and content*).

I skolen er det avgjørende at læreren hjelper eleven til å relatere seg på ulik måte til de to områdene innen matematisk teori. Hvis en lærer blir spurt «hvorfor» det er riktig at

$$a^3 = a \cdot a \cdot a$$

for alle tall a, kan ikke læreren gjøre særlig mer enn å si at dette er sånn fordi folk («matematikerne») har bestemt at det skal være sånn. Matematikerne *har blitt enige om* at skrivemåten

$$a^3$$

skal oppfattes som en forkortning for $a \cdot a \cdot a$. Så på en måte er dette noe elevene bare må akseptere, det er i L-kategorien. På den annen side bør elevene lære å forholde seg på en helt annen måte til ting som ligger i «funnet ut»-området, altså området C. Et eksempel er regelen om at

$$a^n \cdot a^m = a^{n+m}$$

for alle positive tall a og naturlige tall n og m. For denne regelen kan elevene med full rett kreve en begrunnelse for *hvorfor* det er sånn. Dette kan man forklare ut fra definisjonen av eksponentuttrykk som vi var inne på i sted. Undervisningen kan legges opp slik at elevene får arbeide med dette. Selv om en elev ikke kommer i mål med å forstå begrunnelsen, kan vissheten om at det *finnes* en begrunnelse, i seg selv være verdifull. Hvis eleven vet at dette *lar* seg begrunne, vet eleven også at det ikke er «opplagt», eller noe man forventes å forstå direkte. Dermed unngår en at eleven føler seg dum. Å bruke resultater en selv ikke har arbeidet seg gjennom begrunnelsen for, er noe også profesjonelle matematikere ofte gjør, og det er ikke nødvendigvis problematisk. Det ødeleggende problemet for elevens læring oppstår når ting i matematikken som bare «er sånn fordi de er [bestemt] sånn» blandes sammen med ting det ikke er meningen at eleven skal kunne forstå direkte.

Slik vi tenker oss inndelingen av matematisk teori i områdene L og C, er det klart at hva som havner i de to områdene, til dels avhenger av hvordan teorien bygges opp. I skolesammenheng vil grunnleggende algebraiske lover, som for eksempel distributiv lov, måtte regnes i C-området. Grunnen er at dette er regler det i skolesammenheng er naturlig å arbeide med en begrunnelse for, dette er ikke noe som i denne sammenhengen kan sies kun å være en definisjonssak. I mer avanserte kurs i reell analyse eller abstrakt algebra vil det derimot være naturlig å ta den distributive loven som et aksiom, og da ligger den formelt sett ikke i C-kategorien. Den vil da være en del av det man har bestemt som utgangspunkt for teorien, dvs. den er i kategorien «funnet på». Et annet eksempel er definisjonen av multiplikasjon. Slik tradisjonen er i norsk skole i dag, defineres $3 \cdot 4$ til å bety $4 + 4 + 4$. Andre steder defineres dette som $3 + 3 + 3 + 3$. Her har man et *valg* når det gjelder hva man bestemmer, men skal man bygge opp matematikken på en meningsfull måte for barna, må man bestemme seg for én av tingene. Man kan så *begrunne* at faktorenes orden er likegyldig; dette siste er i skolesammenheng uansett noe som bør plasseres i området «funnet ut».

Til tross for dens åpenbare viktighet i skolen er LC-distinksjonen lite diskutert i moderne matematikkdidaktikk. Man finner lite om dette i kilder som (Clements, Bishop, Keitel-Kreidt, Kilpatrick & Koon-Shing Le, 2013; English & Bussi, 2008; Niss, 2007). Distinksjonen passer ikke uten videre inn i de mest kjente rammeverkene for matematikkompetanser, for eksempel de vi var inne på i forbindelse med PISA og relaterte rammeverk. Dette har å gjøre med at verken teoriområdet L eller teoriområdet C naturlig tilsvarer noen spesiell type «kompetanse» i seg selv. Tatt i betraktning den store innflytelsen kompetansebaserte rammeverk har hatt på både vurderingsformer og læreplaner i matematikk gjennom de siste par tiårene, kan dette nettopp være noe av grunnen til at denne distinksjonen omtales så lite. Derimot finnes det mye forskning knyttet til hvert enkelt av de to områdene i LC-distinksjonen. For L-siden kan vi som eksempel ta forskning knyttet til distinksjonen mellom *concept image and concept definition* (Niss, 1999a). Merk også at generelle, mindre fagspesifikke teorier for *begrepslæring* typisk ikke vil fange opp LC-distinksjonen, fordi denne er spesiell for matematikkfaget. Blant forskning som gjelder C-siden, kan vi peke på det store forskningsfeltet knyttet til den rollen som *bevis* spiller i matematikk og matematikkundervisning på ulike nivåer (Hanna, 2000; Pedemonte, 2007; Tall, 2014) Men heller ikke disse forskningstradisjonene legger sin primære vekt på distinksjonen mellom L og C. Videre ligger LC-distinksjonen klart på et annet nivå enn prosess/objekt-dualiteten beskrevet av Anna Sfard (Sfard, 1991).

Når det gjelder betoning av LC-distinksjonen i lærebøker og andre læremidler i Norge, går det et klart skille mellom skolematematikken og matematikken på høgskole- og universitetsnivå. I lærebøker skrevet for universitetskurs er vanligvis distinksjonen mellom definisjoner og teoremer klar og eksplisitt. Dette henger sammen med at distinksjonen oppfattes som grunnleggende av profesjonelle matematikere, og at forfatterne av lærebøker på dette nivået oftest har en forskerutdanning i matematikk. I lærebøker for skolen er bildet et helt annet; der er ofte denne distinksjonen vanskelig å få øye på.

2.3 LC-rammeverket

Rammeverket vi skal bruke for å beskrive avhengighet av matematisk teori, er kalt *LC-rammeverket*. Navnet har å gjøre med at rammeverket er tenkt å fange opp både språksiden L og innholdssiden C i matematikken som beskrevet i forrige delkapittel. Utviklingen av dette rammeverket startet i 2014, og noen av resultatene vi drøfter i dette kapitlet ble presentert på IRC-konferansene i 2015 og 2017 (Hole et al., 2017; Hole et al., 2015). Rammeverket gir et mål for relevansen av matematisk teori for skriftlige tester. Vi tenker oss at testene består av en mengde enkeltoppgaver, som vi refererer til som *oppgaver* (engelsk: *items*). Oppgavene kan være flervalgsoppgaver eller åpen respons-oppgaver. Vi refererer til de siste som *åpne* oppgaver. LC-rammeverket klassifiserer oppgaver ved å dele opp mengden av oppgaver i testen ved to *dikotomier* (todelinger). Den ene dikotomien er ment å måle oppgavenes avhengighet av matematikkens L-del, og den andre av C-delen.

1. For å beskrive avhengighet av L-delen, brukes en dikotomi vi refererer til som *formel/ikke-formel*-dikotomien, eller *F/NF-dikotomien*. Siden LC-rammeverket er tenkt å måle innslaget av matematisk *teori*, ønsker vi at F/NF-dikotomien skal adressere de formelle delene av matematisk språk. I skolen operasjonaliseres dette naturlig gjennom matematiske *formler* i vid forstand, og derfor fokuserer vi på det. Kategoriene i F/NF-dikotomien er (i) mengden av oppgaver der minst én formel er involvert enten i oppgaveteksten eller i elevens forventede løsning eller løsningsmetode, og (ii) mengden av oppgaver der dette ikke er tilfelle. Vi refererer til (i) som *formelkategorien*, eller *F-kategorien*. Kategorien (ii) kaller vi *ingen formel*-kategorien, eller *NF-kategorien*.
2. For å beskrive avhengighet av C-delen, brukes en dikotomi vi refererer til som *teorem/ikke-teorem*-dikotomien, eller *T/NT-dikotomien*. Kategoriene i T/NT-dikotomien er (i) mengden av oppgaver der kjennskap til minst ett teorem («matematisk setning») er relevant for å løse oppgaven, og (ii) mengden av oppgaver der dette ikke er tilfelle. Vi refererer til (i) som *teoremkategorien, eller T-kategorien*. Kategorien (ii) kaller vi *intet teorem*-kategorien, eller *NT-kategorien*.

Til sammen gir dikotomiene F/NF og T/NT et mål for testens avhengighet av matematisk teori. Eller, om man vil, et mål for i hvilken grad kjennskap til matematisk teori hjelper eleven til å løse oppgavene i testen.

Merk at kategoriene T og F potensielt er overlappende, det går an at en oppgave kategoriseres som både F og T. For eksempel vil en oppgave der Pytagoras' setning er aktuell å bruke, typisk kunne klassifiseres som både T og F.

Begrunnelsen for å legge oppgaven i T-kategorien kan da være at Pytagoras' setning er et teorem. Oppgaven involverer altså innholdssiden (C-siden) av matematikken. Samtidig vil oppgaven kunne plasseres i kategorien F, fordi Pytagoras' setning vanligvis uttrykkes ved en formel. Det er altså ikke slik at oppgavekategoriene T og F representerer henholdsvis innholdssiden (C) og språksiden (L) i matematisk teori. Derimot er de *indikatorer på om en gitt oppgave involverer* henholdsvis C og L.

På samme måte kan en gitt oppgave også kategoriseres som både NF og NT. Dette vil da være en oppgave der ingen formler er involvert, og der ingen matematiske teoremer fra den antatte skolebakgrunnen i vesentlig grad hjelper eleven til å løse oppgaven.

Slik LC-rammeverket er konstruert, kan det også brukes på tester i andre fag enn matematikk, typisk fag der matematikk er aktuelt som et hjelpemiddel. Matematikkinnholdet vurderes på samme måte, enten vi er innenfor eller utenfor matematikkfaget selv. For eksempel bruker vi i dette kapitlet rammeverket til å analysere matematikkinnholdet i fysikkdelen av TIMSS Advanced 2015. Merk at rene «fysiske» formler som $F = ma$ og $E_k = (1/2)mv^2$ da naturlig også regnes som matematiske formler i vår forstand, fordi de involverer variabler som kan ta ulike tallverdier. Oppgaver som involverer slike formler, klassifiseres altså som F i LC-rammeverket. Tilsvarende vil det være i andre fag, som for eksempel kjemi eller økonomi.

2.4 Matematikkinnhold i TIMSS, PISA og TIMSS Advanced målt med LC-rammeverket

Klassifiseringen av oppgaver i henhold til dikotomiene F/NF og T/NT vil ha et subjektivt element i seg, og derfor må vi bruke en metodikk med grupper av kodere, måling av interkoder-reliabilitet og så videre. Dessuten trengs det presiseringer av klassifiseringskriteriene som delvis avhenger av typen tester vi ser på, i vårt tilfelle storskalaundersøkelser av typen TIMSS og PISA. For detaljene i dette henviser vi til (Hole et al., 2017; Hole et al., 2015). Vi skal gjengi noen hovedtrekk her.

Når det gjelder presisering av kriteriene for T/NT, diskuterte gruppene av kodere seg fram til følgende. For at en oppgave skal klassifiseres som T, må det finnes et teorem (minst ett) fra elevens antatte skolebakgrunn («pensum») som i vesentlig grad forenkler arbeidet med oppgaven. Hvis eleven derimot antas å måtte resonnere seg fram fra grunnen av, bare ved bruk av kjennskap til matematiske begreper (begrepsforståelse), klassifiseres oppgaven som NT. Videre framsto det som naturlig å operere med et visst «bunnfradrag». Enkle aritmetiske sammenhenger, som for eksempel $5 + 14 = 19$, ble ikke telt med som teoremer i denne sammenhengen. Begrunnelsen var at selv om disse i prinsippet er ting man har «funnet ut», slik at de altså *er* teoremer formelt sett, så behandles de ikke som dette i skolematematikken. Som eksemplifisering utarbeidet gruppene lister over teoremer som kan antas dekket i pensum for de aktuelle årstrinnene. Listene inneholdt blant annet algebraiske lover som

$$ab = ba$$

$$a(b + c) = ab + ac$$

$$a - (b + c) = a - b - c$$

$$\frac{a}{c} + \frac{b}{c} = \frac{a + b}{c}$$

$$\frac{a}{b} = \frac{ac}{bc}$$

Med på listene var også operasjoner man kan bruke for å løse likninger, som for eksempel at man kan legge til det samme tallet på begge sider av likhetstegnet. Listene inneholdt også

- former for areal, omkrets, volum og liknende for geometriske figurer
- geometriske teoremer som Pytagoras' setning, setningen om vinkelsummen i en trekant, setninger om formlikhet og så videre

Gruppene laget også en liste med eksempler på ting som mange lett kan tro representerer teoremer, men som ikke gjør det. Blant eksemplene her kan nevnes $a^0 = 1$ og $a^{-n} = 1/a^n$.

For dikotomien F/NF var utgangspunktet, som tidligere nevnt, at en *formel* er et matematisk uttrykk eller utsagn (for eksempel en likning eller en ulikhet) som inneholder minst én variabel. Variablene kan være representert syntaktisk ved bokstaver, ved andre typer symboler eller ved komplette ord. Med andre ord regner vi også

$$\text{strekning} = \text{fart} \cdot \text{tid}$$

som en formel i vår sammenheng. For at en oppgave skal klassifiseres som F, må et av disse tre kravene være oppfylt:

- Oppgaven inneholder en formel som eleven må bruke, eller
- oppgaven ber eleven om å lage en formel, eller
- en typisk elev vil kunne forventes å bruke en formel underveis i arbeidet med oppgaven.

Når det gjelder struktureringen av selve klassifiseringsarbeidet for de tre studiene, altså TIMSS matematikk 8. trinn, PISA matematikk og TIMSS Advanced, henviser vi til (Hole et al., 2017; Hole et al., 2015). Vi skal her nøye oss med å gjengi noen hovedtrekk.

For PISA 2012 og TIMSS 2011 8. trinn ble klassifiseringen gjennomført med to grupper av kodere. For hver gruppe ble interkoder-reliabiliteten målt ved Fleiss' kappa (Fleiss, Cohen & Everitt, 1969). Klassifiseringen ble for begge grupper gjennomført i to omganger, med mulighet for diskusjon av uoverensstemmelser mellom omgangene. Rollene til de to gruppene av kodere var litt ulike. Den første gruppen besto av 4 kodere, og det var hovedsakelig denne gruppen som utviklet presiseringene og tilpasningene av rammeverket beskrevet ovenfor. Den andre gruppen besto av to kodere, og denne gruppen ble brukt delvis for å sjekke overførbarheten av rammeverket via beskrivelse. Vi gjengir her kun klassifikasjonsresultatene for den kombinerte gruppen av 6 kodere. Resultatene for hver gruppe separat er gitt i (Hole et al., 2015).

Klassifikasjonen av TIMSS Advanced ble gjennomført på en litt annerledes måte. Her brukte vi i første omgang masterstudenter som arbeidet som kodere for TIMSS Advanced i Norge våren 2015, dvs. som var ansatt for å kode («sensurere») resultatene fra de norske elevene som deltok i studien. En gruppe på 4 studenter klassifiserte oppgavene fra matematikkdelen av TIMSS Advanced 2015 ved bruk av LC-rammeverket, og en annen gruppe på 4 klassifiserte oppgavene fra fysikkdelen ved det samme rammeverket. Studentene ble gitt en kort gjennomgang av rammeverket på forhånd. Deretter gjennomførte de én syklus med klassifisering, uten mulighet til diskusjon seg imellom underveis.

Dette designet ble valgt fordi det igjen var interessant å måle overførbarheten av rammeverkets kriterier. Mens klassifiseringen av fysikkoppgavene gav akseptable kappaverdier (0,70 og 0,67 for F/NF og T/NT henholdsvis), ble kappaverdiene i matematikk bare 0,53 for F/NF og 0,23 for T/NT henholdsvis.

Selv om kappaverdiene i matematikk var interessante i forbindelse med rammeverkets overførbarhet, var de for lave til å kunne gi et pålitelig bilde av matematikkoppgavene i TIMSS Advanced. Vi gjennomførte derfor en ny runde klassifisering av TIMSS Advanced matematikk våren 2017, denne gangen med en gruppe på 4 forskere som kodere. Disse gjennomførte to sykler med klassifisering, og de resulterende kappaverdiene var 0,68 for F/NF og 0,72 for T/NT. For detaljer henviser vi til (Hole et al., 2017).

Klassifikasjonsresultatene for PISA 2012 matematikk og TIMSS 2011 matematikk 8. trinn er vist i tabellene 2.1 og 2.2.

Tabell 2.1 Klassifisering av oppgaver fra PISA og TIMSS ved diktomien F/NF, i prosent av oppgaver.

	PISA matematikk 2012	TIMSS trinn 8 matematikk 2011
Klassifisert som F [bruker formler] av minst 5 av 6 kodere	18,8 %	21,7 %
Delte meninger (4–2 eller 3–3)	9,4 %	13,4 %
Klassifisert som NF [bruker ikke formler] av minst 5 av 6 kodere	71,8 %	65,0 %

Tabell 2.2 Klassifisering av oppgaver fra PISA og TIMSS ved diktomien T/NT, i prosent av oppgaver.

	PISA matematikk 2012	TIMSS trinn 8 matematikk 2011
Klassifisert som T [bruker teoremer] av minst 5 av 6 kodere	11,8 %	19,3 %
Delte meninger (4–2 eller 3–3)	2,4 %	7,4 %
Klassifisert som NT [bruker ikke teoremer] av minst 5 av 6 kodere	85,9 %	73,3 %

Tabell 2.3 viser koherensen mellom koderne for klassifikasjon av oppgaver fra PISA og TIMSS. Vi ser at klassifikasjonsresultatene framstår som ganske robuste.

Tabell 2.3 Koherens i den kombinerte gruppen av 6 kodere på TIMSS matematikk 8. trinn og PISA matematikk. Det første tallet er prosentandelen av oppgaver der alle de seks koderne var enige om klassifiseringen. Tallet i hakeparentes er prosentandel oppgaver der minst 5 av 6 kodere var enige. Tallet i parentes bakerst er koder-reliabiliteten målt ved Fleiss' kappa.

	PISA 2012	TIMSS 8. trinn 2011
T/NT klassifikasjon	88,2 % [97,6 %] (0,82)	78,8 % [92,6 %] (0,76)
F/NF klassifikasjon	77,6 % [90,6 %] (0,74)	65,4 % [86,6 %] (0,65)

Klassifikasjonsresultatene for TIMSS Advanced 2015 matematikk og fysikk er vist i tabellene 2.4 og 2.5.

Tabell 2.4 Klassifisering av oppgaver fra TIMSS Advanced ved dikotomien F/NF, i prosent av oppgaver.

	TIMSS Advanced 2015 matematikk	TIMSS Advanced 2015 fysikk
Klassifisert som F [bruker formler] av minst 3 av 4 kodere	67,0 %	31,1 %
Delte meninger (2–2)	8,7 %	7,8 %
Klassifisert som NF [bruker ikke formler] av minst 3 av 4 kodere	24,3 %	61,1 %

Tabell 2.5 Klassifisering av oppgaver fra TIMSS Advanced ved dikotomien T/NT, i prosent av oppgaver.

	TIMSS Advanced 2015 matematikk	TIMSS Advanced 2015 fysikk
Klassifisert som T [bruker teoremer] av minst 3 av 4 kodere	78,6 %	14,6 %
Delte meninger (2–2)	2,9 %	5,8 %
Klassifisert som NT [bruker ikke teoremer] av minst 3 av 4 kodere	18,4 %	79,6 %

Til sammen gir resultatene i tabellene 2.1, 2.2, 2.4 og 2.5 et grunnlag for å sammenlikne matematikkinnholdet i de fire studiene, målt gjennom LC-rammeverket. Hvis vi ordner studiene etter økende avhengighet av den matematikkfaglige innholdskomponenten, altså kategorien T, får vi følgende liste:

1. PISA 2012 matematikk: 11,8 % oppgaver T, 18,8 % F
2. TIMSS Advanced 2015 fysikk: 14,6 % oppgaver T, 31,1 % F
3. TIMSS 2011 matematikk 8. trinn: 19,3 % oppgaver T, 21,7 % F
4. TIMSS Advanced 2015 matematikk: 78,6 % oppgaver T, 67,0 % F

Merk at rangeringen ville ha blitt omtrent den samme hvis vi i stedet baserte oss på kategorien F, altså avhengighet av formelt matematisk språk. Eneste forskjell ville ha vært at TIMSS Advanced fysikk ville ha byttet plass med TIMSS matematikk 8. trinn. Dette illustrerer den naturlige avhengigheten mellom de to dikotomiene i LC-rammeverket: De måler begge innslaget av matematisk teori i studienes oppgavemateriale.

2.5 Forskjeller i studienes innhold

Fra resultatene i delkapittel 2.4 ser vi at det er store forskjeller mellom PISA og TIMSS når det gjelder hvordan studiene forholder seg til matematisk teori. Kun 11,8 % av oppgavene i PISA 2012 matematikk er slik at et eller annet matematisk resultat (teorem) som eleven kan forventes å kjenne til, i nevneverdig grad hjelper eleven når det gjelder å løse oppgaven. Kun 18,8 % av oppgavene i PISA involverer matematiske formler. Over 2/3 av PISA-oppgavene er uavhengige av både teoremer og formler. I disse oppgavene må altså elevene resonnere fra grunnen av, og det de eventuelt trenger av forkunnskaper fra matematikkfaget, er kjennskap til matematikkord («rektangel» og så videre) og andre ikke-formelle deler av matematisk språk. Utover dette er majoriteten av PISA-oppgavene rene problemløsingsoppgaver som ikke involverer matematisk teori. Også i TIMSS matematikk 8. trinn er over halvparten av oppgavene i begge kategoriene NT og NF, men relevansen av matematisk teori er langt større her. Det må også tas i betraktning at elevene som testes i TIMSS 8. trinn er 1–2 år yngre enn elevene som testes i PISA. Prosentandelen oppgaver i T-kategorien er over 19 % i TIMSS 8. trinn, mens den er under

12 % i PISA. Dette viser at disse to studiene måler kompetanse i matematikk på ganske ulike måter. Skal man vurdere betydningen av prestasjonsdata i PISA og TIMSS eller deres sammenheng med bakgrunnsvariable, er det viktig at man tar dette med i betraktning.

Ikke overraskende er matematikkinnholdet i TIMSS Advanced matematikk i særklasse det høyeste blant de fire studiene vi ser på her. Nesten 80 % av oppgavene er her i T-kategorien. Vi ser også at målt med LC-rammeverket er matematikkinnholdet i TIMSS Advanced fysikk større enn matematikkinnholdet i PISA matematikk. Vi ser dessuten at matematikkinnholdet i fysikkstudien ligger mer på språksiden (F) enn på innholdssiden (T).

I kapittel 5 undersøker vi sammenhenger mellom LC-klassifiseringene og elevprestasjoner. Vi vil da se at LC-rammeverket avdekker interessante forskjeller i undervisningstradisjoner i ulike land og ulike regioner i verden.