

Hege Kaarstein og Anne-Catherine Lehre (red.)

Hva TIMSS forteller om norsk skole: Sekundæranalyser av TIMSS-undersøkelsen

Hva TIMSS forteller om norsk skole: Sekundæranalyser av TIMSS- undersøkelsen

Hege Kaarstein og Anne-Catherine Lehre (red.)

Hva TIMSS forteller om norsk skole:

Sekundæranalyser av TIMSS-undersøkelsen

CAPPELEN DAMM FORSKNING

Utgivelsesår: 2025

1. utgave

© 2025 Hege Kaarstein, Anne-Catherine Lehre, Hilde Marie Madsø Jacobsen, Fredrik Jensen, Armin Jentsch, Trude Nilsen, Rolf Vegar Olsen, Andreas Pettersen, Jelena Radišić, Manvi Aurora Rohatgi, Alexander Jonas Viktor Selling, Nani Teig, Jelena Veletić, Cecilie Weyergang, Mark White og Wangqiong Ye.

Bokens design og sats: © 2025 Cappelen Damm AS.

Dette verket omfattes av bestemmelsene i *Lov om opphavsretten til åndsverk m.v.* av 1961. Verket utgis Open Access under betingelsene i Creative Commons-lisensen CC-BY 4.0. Denne tillater tredjepart å kopiere, distribuere og spre verket i hvilket som helst medium eller format, og å remixe, endre, og bygge videre på materialet til et hvilket som helst formål, inkludert kommersielle, under betingelse av at korrekt kreditering og en lenke til lisensen er oppgitt, og at man indikerer om endringer er blitt gjort. Tredjepart kan gjøre dette på enhver rimelig måte, men uten at det kan forstås slik at lisensgiver bifaller tredjepart eller tredjeparts bruk av verket. Enhver bruk av hele eller deler av verket som input eller som treningskorpus i generative modeller som kan skape tekst, bilder, film, lyd eller annet innhold og uttrykk, er ikke tillatt uten særskilt avtale med rettighetshaverne. Lisensvilkår: <http://creativecommons.org/licenses/by/4.0/>

ISBN heftet utgave: 9788202894764
ISBN PDF: 9788202869137
ISBN EPUB: 9788202903763
ISBN HTML: 9788202903770
ISBN XML: 9788202903787
DOI: <https://doi.org/10.23865/cdf.274>



Dette er en fagfellevurdert antologi.

Cappelen Damm Forskning er redaksjonen for åpen forskningspublisering i Cappelen Damm Akademisk.

Omslagsdesign: Carine Fløystad, Cappelen Damm AS

Omslagsfoto:

Cappelen Damm Forskning
forskning@cappelendamm.no

Innhold

Forord	7
Kapittel 1 Innledning	9
<i>Hege Kaarstein og Anne-Catherine Lehre</i>	
Kapittel 2 Utvikling av ulikheter i utdanningen over tid	21
<i>Trude Nilsen, Wangqiong Ye og Nani Teig</i>	
Kapittel 3 Fra TIMSS 2015 til TIMSS 2023: et skolemiljø i endring	47
<i>Hege Kaarstein, Anne-Catherine Lehre og Mark White</i>	
Kapittel 4 Skolens og rektors betydning for elevenes prestasjoner i matematikk og naturfag	79
<i>Jelena Veletić, Jelena Radišić og Hilde Marie Madsø Jacobsen</i>	
Kapittel 5 Spredning i elevers matematikk- og naturfagprestasjoner, TIMSS 2015–2023	105
<i>Anne-Catherine Lehre og Hege Kaarstein</i>	
Kapittel 6 Kjennetegn på lavt- og høytpresterende elever i TIMSS	137
<i>Anne-Catherine Lehre og Manvi Aurora Rohatgi</i>	
Kapittel 7 Læreplan og faglig innhold i matematikk på 5. trinn i Norge	169
<i>Mark White, Alexander Jonas Viktor Selling, Andreas Pettersen og Jelena Veletić</i>	
Kapittel 8 Undervisningskvalitet og matematikkprestasjoner på 5. trinn fra TIMSS 2019 til TIMSS 2023	197
<i>Mark White, Hege Kaarstein og Armin Jentsch</i>	
Kapittel 9 Motivasjon, responstid og prestasjoner i matematikk blant norske elever på 9. trinn	217
<i>Jelena Radišić og Andreas Pettersen</i>	

Kapittel 10 Norske elevers kunnskaper og ferdigheter i algebra – den vanskelige oppgaven med å lage matematiske uttrykk	243
<i>Andreas Pettersen og Jelena Radišić</i>	
Kapittel 11 Hva viser grafen? Elevers statistiske kompetanse og kunnskap om vær og klima	267
<i>Cecilie Weyergang, Hege Kaarstein, Fredrik Jensen og Anne-Catherine Lehre</i>	
Kapittel 12 Økende ulikhet og endringer i norsk skole	293
<i>Anne-Catherine Lehre og Hege Kaarstein</i>	
Kapittel 13 Tretti år med TIMSS i Norge: Kan resultater fra 1995 sammenliknes med dagens og framtidige TIMSS-undersøkelser?	301
<i>Rolf Vegar Olsen</i>	

Forord

Våren 2023 deltok Norge i TIMSS-undersøkelsen for sjuende gang, og hovedresultatene ble offentliggjort i en kortrapport i desember 2024. TIMSS-undersøkelsen er finansiert av Utdanningsdirektoratet og gjennomføres av Institutt for lærerutdanning og skoleforskning (ILS) ved Universitetet i Oslo.

Denne vitenskapelige antologien retter seg mot forskere, lærere, skoleledere og utdanningsmyndigheter. Den skal bidra til å forstå og analysere utviklingen i norske elevers kompetanse i matematikk og naturfag, og belyse faktorer som belyse faktorer som har sammenheng med læring.

Antologien går i dybden på resultatene fra TIMSS 2023, ser på trend for resultatene i matematikk og naturfag for elevene på 5. og 9. trinn og undersøker om faktorer som er viktige for elevenes læringsutbytte har endret seg over tid så vel som hvorvidt disse endringene har sammenheng med de faglige resultatene. Noen av kapitlene konsentrerer seg om ett trinn og ett fag, mens andre ser på begge trinn og fag.

Vi vil gjerne få rette en stor takk til alle skolelederne, elevene, lærerne og foresatte som deltok i undersøkelsen. Videre vil vi også påskjønne alle som bidro til gjennomføringen av undersøkelsen og arbeidet med denne antologien. Sist, men ikke minst, vil vi gjerne få takke vår redaktør i Cappelen Damm Forskning, Maria R. Braadland, og de anonyme fagfellene som har bidratt til å heve kvaliteten på kapitlene i antologien med sine konstruktive tilbakemeldinger.

Blindern, august 2025

På vegne av forfattergruppa

Hege Kaarstein og Anne-Catherine Lehre

KAPITTEL 1

Innledning

Hege Kaarstein og Anne-Catherine Lehre

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

Våren 2023, nesten halvannet år etter at de siste pandemiestriksjonene ble opphevet, samlet Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo på vegne av Utdanningsdirektoratet inn data til den åttende runden av den internasjonale komparative storskalaundersøkelsen *Trends in Mathematics and Science Study* (TIMSS). Den 4. desember 2024 ble de første resultatene offentliggjort i en kortrapport.

Resultatene for TIMSS 2023 viste en tilbakegang for norske elever på 5. trinn i både matematikk og naturfag fra 2019 til 2023 (Kaarstein et al., 2024). Tilbakegangen skyldes en økning i andel elever på de laveste kompetansenivåene. Nedgangen i matematikk på 5. trinn kan spesielt knyttes til nedgang i jentenes prestasjoner i faget. På 9. trinn var det verken signifikante endringer i elevenes matematikk- og naturfagprestasjoner eller kjønnsforskjeller.

I et internasjonalt perspektiv rapporterte de norske TIMSS-deltakerne om et godt skolemiljø i 2023 (Kaarstein et al., 2024). Likevel har det vært en nedgang fra 2015, da det norske skolemiljøet ble oppfattet av lærerne som svært bra. Sammenlignet med lærernes rapporteringer i de andre deltakerlandene, lå de norske resultatene helt i toppen i 2015 (Mullis et al., 2016).

Formålet med denne antologien

Matematikk, naturfag og teknologi (MNT) er avgjørende for å løse samfunnsutfordringer innenfor klima, helse, energi og digitalisering. Etterspørselen etter realfagskompetanse øker, og for å sikre framtidig velferd og økonomisk vekst må befolkningen ha solid kunnskap i disse fagene (Norges forskningsråd, 2012).

Skolen har en nøkkelrolle i å vekke interesse og bygge grunnleggende forståelse hos elevene. Uten realfagskompetanse risikerer elever å møte begrensede muligheter i utdanning og arbeidsliv, og får svekket evne til å vurdere informasjon kritisk i et stadig mer komplekst samfunn (Norges forskningsråd, 2012).

TIMSS-undersøkelsen samler inn data som kan fortelle noe om matematikk- og naturfagskompetansen og læringskonteksten til norske elever på 5. og 9. trinn. Formålet med denne antologien er å analysere datamaterialet mer i dybden for å få en bedre forståelse av resultatene fra undersøkelsen. Data fra begge trinn, begge fag og fra flere sykluser av TIMSS analyseres, og resultatene presenteres i denne antologien. Den skal bidra til å forstå utviklingen i norske elevers kompetanse i matematikk og naturfag, og belyse faktorer som har sammenheng med læring.

Innholdet i antologien¹

Denne antologien setter eleven i sentrum og utforsker hvordan skolen, undervisningen og fagene har sammenheng med elevenes læring og utvikling. Med utgangspunkt i TIMSS-data og norsk skolehverdag belyses ulikheter over tid, skolemiljøets betydning og hvordan læreplaner og undervisningskvalitet former elevenes faglige reise.

Gjennom analyser av prestasjoner, motivasjon og faglig forståelse i matematikk og naturfag, får vi innsikt i hvem elevene er, hva de strever med – og hva som gir dem mulighet til å lykkes. Målet med antologien er å gi et bredere kunnskapsgrunnlag for refleksjon og videreutvikling av både læreres praksis og utdanningssystemet.

De tre første kapitlene i antologien gir et overordnet blick på skolen som system og læringsarena ved å studere utviklingen av ulikheter i utdanningen over tid (kapittel 2) og å vise hvordan endringer i skolemiljø (kapittel 3) og ledelse (kapittel 4) har sammenheng med elevenes faglige utvikling. Deretter flyttes blikket over på elevene og studier av variasjoner i elevenes prestasjoner (kapittel 5 og 6) og motivasjon (kapittel 9). To av kapitlene handler om hvordan læreplaner (kapittel 7) og undervisningspraksis (kapittel 8) legger grunnlaget for elevenes læring, og hvordan kvaliteten på undervisningen henger sammen med prestasjoner. I siste del går vi i

1 KI er benyttet til å gi en kort oppsummering av de ulike kapitlene i denne antologien.

dybden på elevenes faglige forståelse i matematikk (kapittel 10) og naturfag (kapittel 11), og ser på hvordan konkrete faglige utfordringer kan belyse elevenes kompetanse og behov. Til slutt oppsummeres funnene fra alle kapitlene (kapittel 12), og validiteten i trendanalyser som benytter data fra TIMSS-undersøkelsen drøftes (kapittel 13). En kort presentasjon av alle kapitlene finnes under.

Kort om kapitlene i antologien

Kapittel 2 viser til at utdanningsulikhet utfordrer både sosial mobilitet og demokratiske prinsipper. Forskning på langtidstrender, særlig etter covid-19-pandemien og med bredere mål på ulikhet, er begrenset. Studien undersøker utviklingen av ulikheter i matematikkprestasjoner på ungdomstrinnet i Norge ved å benytte data fra TIMSS fra 1995 til 2023. Forfatterne av kapitlet, Trude Nilsen, Wangqiong Ye og Nani Teig, analyserer sammenhengen mellom elevers sosioøkonomiske status (SØS) og prestasjoner på både elev- og klassenivå, samt endringer i andelen faglig resiliente elever, det vil si elever med lav SØS som presterer høyt. Forfatterne anvender strukturell ligningsmodellering og beregning av resiliens for å gi et mer nyansert bilde av ulikhetsutviklingen i norsk skole.

Basert på data fra TIMSS-undersøkelsene i 2015, 2019 og 2023, undersøker Hege Kaarstein, Anne-Catherine Lehre og Mark White i **kapittel 3** utviklingen i skolemiljøet i Norge med særlig vekt på trygghet, trivsel og tilfredshet blant elever, lærere og rektorer. Skolemiljø er et sammensatt begrep som inkluderer både fysiske og psykososiale aspekter, og trygghet ses som en sentral dimensjon. Kapitlet bygger på tidligere forskning som viser at et trygt skolemiljø er viktig for elevers læring og trivsel, og at lærernes opplevelse av trygghet henger sammen med undervisningskvalitet. Studien benytter spørreskjemadata fra elever, lærere og rektorer på 5. og 9. trinn, og analyserer utviklingen over tid samt sammenhenger mellom trygghet og ulike skolekjennetegn, som elevenes sosioøkonomiske bakgrunn, morsmål og innvandrerstatus. Forfatterne tar også for seg hvordan skolens beliggenhet og sammensetning har sammenheng med opplevelsen av trygghet.

I **kapittel 4** undersøker Jelena Veletić, Jelena Radišić og Hilde Marie Madsø Jacobsen hvordan skolemiljø og rektors karakteristikk henger sammen med elevenes prestasjoner i matematikk og naturfag på 9. trinn i Norge, med utgangspunkt i data fra TIMSS 2023. Studien tar utgangspunkt

i den nordiske utdanningsmodellen, som vektlegger likhet og inkludering, og ser nærmere på hvordan skolen som institusjon kan bidra til å redusere ulikheter i elevenes læringsutbytte. Et sentralt anliggende er hvordan skoleledelse og skolemiljø, som begge er kjerneområder i skolens ledelse, henger sammen med elevenes læringsutbytte. Skolemiljøet blir sett på som en viktig arena for læring, og det teoretiske rammeverket inkluderer begrepet «læringsfokus» (*school emphasis on academic success*), som måles gjennom rektorenes vurderinger av elevers, foreldres og læreres engasjement og støtte til læring. Forfatterne diskuterer også betydningen av rektors utdanningsbakgrunn, særlig mastergrad og kursbevis i utdanningsledelse, samt antall år med erfaring ved nåværende skole. Studien benytter to-nivå strukturell likningsmodellering og kontrollerer for sosioøkonomiske faktorer på både elev- og skolenivå.

Anne-Catherine Lehre og Hege Kaarstein undersøker utviklingen i spredningen av norske elevers prestasjoner i matematikk og naturfag på 5. og 9. trinn i perioden 2015 til 2023 i **kapittel 5**. Spredning i prestasjoner kan ha betydning for læreres mulighet til å differensiere undervisningen, og er relevant for diskusjoner om likeverd i skolen. Forfatterne benytter data fra TIMSS og analyserer variasjon både på tvers av faglige emneområder og mellom kjønn. Spredningen visualiseres ved hjelp av kvantildistansefunksjoner, og det benyttes statistiske tester for å vurdere forskjeller mellom grupper og over tid.

I **kapittel 6** undersøker Anne-Catherine Lehre og Manvi Aurora Rohatgi hva som kjennetegner elever på lavt og avansert mestringsnivå i matematikk og naturfag på 5. og 9. trinn, basert på data fra TIMSS-undersøkelsene i 2015, 2019 og 2023. Fokus ligger på elevsentrerte faktorer som sosioøkonomisk bakgrunn, faglig selvtillit, motivasjon, språk hjemme, læringsmiljø og fysisk velvære. Forfatterne tar utgangspunkt i internasjonale rammeverk og tidligere forskning, og benytter kvantitative analyser for å sammenligne elevgrupper på ulike mestringsnivåer over tid.

Sammenhengen mellom læreplanens innhold og undervisningen i matematikk på 5. trinn i Norge er temaet for **kapittel 7**. Med utgangspunkt i endringen fra LK06 til LK20, analyserer Mark White, Alexander Jonas Viktor Selling, Andreas Pettersen og Jelena Veletić den implementerte læreplanen, sett fra perspektivet av TIMSS-rammeverket. I kapittelet tar de for seg hvordan undervisningsinnhold har endret seg over tid og sammenligner innholdet av undervisningen i Norge med innholdet i undervisningen hos de andre nordiske landene.

Med utgangspunkt i data fra TIMSS 2019 og 2023, undersøker Mark White, Hege Kaarstein og Armin Jentsch endringer i elevers oppfatning av undervisningskvalitet i matematikk på 5. trinn i Norge i **kapittel 8**. Undervisningskvalitet regnes som en sentral og formbar faktor i elevers læringsutbytte. Forfatterne utforsker hvordan elevenes oppfatning eller opplevelse av undervisningen har utviklet seg over tid, og hvordan disse vurderingene relaterer seg til endringer i prestasjoner. Analysene bygger på elevdata og benytter regresjonsmodeller og simulering for å undersøke sammenhenger mellom undervisningskvalitet og prestasjoner.

I **kapittel 9** undersøker Jelena Radišić og Andreas Pettersen sammenhengen mellom elevers motivasjon, prestasjoner i matematikk og innsats på TIMSS-prøven, basert på data fra TIMSS 2023 for 9. trinn i Norge. Studien tar utgangspunkt i teorien om situasjonell forventning-verdi (*situated expectancy-value theory*) og benytter en personsentrert tilnærming med latent profilanalyse for å identifisere ulike motivasjonsprofiler basert på faglig selvtillit, interesse og opplevd nytteverdi. Forfatterne undersøker deretter hvordan disse profilene henger sammen med elevenes faglige prestasjoner og innsats på prøven, der innsats måles ut fra responstid på oppgavene.

Algebra utgjør en sentral del av matematisk kompetanse og spiller en viktig rolle i skolematematikken. Tidligere forskning har vist at norske elever ofte strever med algebra, særlig med å formulere algebraiske uttrykk. I **kapittel 10** undersøker Andreas Pettersen og Jelena Radišić norske 9. trinnselevers kunnskaper og ferdigheter i algebra ved å analysere deres besvarelser på tre oppgaver fra TIMSS 2023. Oppgavene krever at elevene generaliserer matematiske mønstre og uttrykker disse med algebraiske uttrykk. Gjennom en systematisk kategorisering av elevsvarene, undersøker forfatterne hva slags typer (feil)svar elevene gir og diskuterer hvilke kunnskaper og forståelse som ligger bak ulike typer svar.

Cecilie Weyergang, Hege Kaarstein, Fredrik Jensen og Anne-Catherine Lehre undersøker i **kapittel 11** hvordan elever på 9. trinn besvarer naturfagoppgaver som krever forståelse av en statistisk framstilling. I dagens samfunn formidles mye naturfaglig informasjon gjennom grafer og visuelle representasjoner, særlig innen temaer som miljø og klima. Studien bygger på teoretiske perspektiver fra feltet statistisk literacy og tar utgangspunkt i to oppgaver fra TIMSS 2023, der elevene møter en grafisk framstilling omtalt som en «klimaopplysning». For å utforske elevenes tilnærminger til oppgavene kombinerer studien innholdsanalyse av skriftlige svar med høyttenkning.

I **kapittel 12** gir redaktørene, Anne-Catherine Lehre og Hege Kaarstein, en kort oppsummering av resultatene fra kapitlene i antologien.

I **kapittel 13**, «30 år med TIMSS i Norge», drøfter Rolf Vegar Olsen hvorvidt det er mulig å sammenligne resultater fra den første TIMSS-undersøkelsen i 1995 med senere og framtidige undersøkelser. Han konkluderer med at slike sammenlikninger er metodisk utfordrende og ofte ikke gyldige, på grunn av både små og store endringer over tid. Testdesign, elevpopulasjoner, spørreskjemaer og rapporteringspraksis har endret seg betydelig, og selv små og usynlige endringer – som elevers motivasjon, teknologibruk og samfunnsforhold – kan påvirke resultatene og gjøre trendanalyser usikre. Likevel understreker Olsen at det fortsatt er svært nyttig å bruke TIMSS-data til å analysere utvikling over kortere tidsperioder. Slike analyser gir verdifull innsikt i endringer i elevers prestasjoner, undervisningskvalitet og skolemiljø, og kan bidra til informert politikktutforming og forskningsbasert utvikling av skolen. For å styrke validiteten i framtidige trendanalyser foreslår han blant annet å bevare spørreskjemaer over flere sykluser, utnytte longitudinelle design bedre og vurdere rekalkibrering av historiske data. Kapittelet avsluttes med en oppfordring til forsiktighet i tolkning av lange tidsserier, og en tro på at TIMSS vil fortsette å være relevant og nyttig i mange år framover.

Om TIMSS

Bakgrunnen for og formålet med undersøkelsen

Trends in International Mathematics and Science Study (TIMSS) er en internasjonal undersøkelse som gjennomføres hvert fjerde år. Første gang den ble gjennomført var i 1995, og Norge har deltatt hver gang siden, bortsett fra i 1999. Undersøkelsen eies av International Association for the Evaluation of Educational Achievement (IEA) og administreres i samarbeid med TIMSS & PIRLS International Study Center ved Boston College. Det overordnede målet med IEAs undersøkelser er å evaluere, forstå og forbedre utdanningen. Ved å innhente data av høy kvalitet hvert fjerde år, kan IEA følge med på den globale utviklingen i tillegg til at hvert deltakerland kan følge med på utviklingen av elevenes kompetanse i de to skolefagene matematikk og naturfag over tid innad i eget land. Videre gir undersøkelsen mulighet for å kunne sammenligne sitt eget utdanningssystem med andre utdannings-systemer det er naturlig å sammenligne med. For Norges del har dette ofte vært Danmark, Finland og Sverige.

Målgrupper og utvalg

Internasjonalt er målgruppene for TIMSS-undersøkelsen elever på 4. og 8. trinn, deres foresatte, matematikk- og naturfaglærere og rektorer. I 2015 gikk Norge fra å undersøke elever på 4. og 8. trinn til å undersøke elever på 5. og 9. trinn. Før 2015 var de norske elevene som deltok i undersøkelsen i gjennomsnitt ett år yngre enn elevene i våre nordiske naboland, og dette var grunnen til byttet av målgrupper. For mer om bakgrunnen og begrunnelsen for byttet av målgrupper, se Bergem, Kaarstein og Nilsen (2016).

TIMSS-undersøkelsen er en såkalt utvalgsundersøkelse, hvilket innebærer at elevene velges eller trekkes ut til å delta. Det trekkes et representativt utvalg elever på 5. trinn og et på 9. trinn, som for Norges del innebærer at det bare er åtte til ni prosent av elevene i målgruppene som deltar i TIMSS. Utvalgsprosessen starter med at det trekkes ut skoler og deretter trekkes en til to hele klasser ved skolene. Dataene som samles inn kan derfor organiseres hierarkisk, hvor elevene hører inn under sin klasse og klassen(e) hører inn under sin skole. For å sikre at utvalgene er representative for målgruppene, beregnes det såkalte vektorer som benyttes i statistiske analyser av de innsamlede dataene. Det er flere grunner til å benytte vektorer. En av grunnene handler om at dataene er hierarkiske. En annen er at elevene i målgruppen ikke har lik sannsynlighet for å bli trukket ut fordi de kommer fra skoler med ulike størrelser og ulik geografisk beliggenhet. Vektene kompenserer for dette, slik at resultatene er generaliserbare til alle i målgruppen. For mer om hvordan utvalget trekkes og vektet, se for eksempel Siegel og Foy (2024). I tillegg til elevene deltar også elevenes foresatte (bare på 5. trinn), elevenes matematikk- og naturfaglærere og skolens rektorer.

I TIMSS 2023 deltok 5 301 elever fra 160 skoler på 5. trinn og 6 324 elever fra 157 skoler på 9. trinn. I tillegg deltok 2 746 av elevenes foresatte på 5. trinn, 265 matematikklærere, 233 naturfaglærere og 137 rektorer. På 9. trinn deltok 266 matematikklærere, 239 naturfaglærere og 131 rektorer (Fishbein et al., 2025).

Hva undersøkelsen måler og hvilke data som samles inn

TIMSS måler elevenes faglige prestasjoner i matematikk og naturfag, basert på et internasjonalt rammeverk som ligger så tett opp til deltagerlandenes rammeverk som mulig (Mullis et al., 2021). Dette rammeverket definerer hvilke kompetanser elevene skal testes i. I naturfag måles elevenes

kompetanse i biologi, kjemi, fysikk og geofag. I matematikk måles elevenes kompetanse på områdene tall, geometri og statistikk, samt algebra på ungdomstrinnet. TIMSS måler elevenes evne til å bruke kunnskaper og ferdigheter i forskjellige situasjoner; evnen til å resonnerer, argumentere, se sammenhenger, trekke slutninger og sammenfatte.

I tillegg til faglig kunnskap, samles det inn data om elevenes læringskontekst. Den overordnede beskrivelsen av hvert deltakerlands utdanningsystem samles i en ensyklopedi (Reynolds et al., 2024). Informasjon om, for eksempel, elevenes hjemmebakgrunn, faglige motivasjon og holdninger, lærernes utdanning, undervisningskvalitet og skolemiljø hentes inn via spørreskjemaer til elevene, deres foreldre (5. trinn), faglærere og rektorer (Reynolds et al., 2021). Informasjonen fra spørreskjemaene brukes i analyser av hva som bidrar til god læring.

Når det gjelder innhenting av informasjon om elevenes hjemmebakgrunn, inkluderer spørreskjemaene spørsmål som bare kan gi en indikasjon på elevenes sosio-økonomiske status (SØS). Blant annet angir elevene hvor mange bøker de har hjemme, om de har egen pult og tilgang til internett og hvor lang utdanning de foresatte har. Disse faktorene fungerer som proxyer snarere enn at de fullt ut reflekterer de komplekse og mangfoldige aspektene ved SØS. I de kapitlene hvor SØS benyttes som en variabel, angis det alltid hvilke faktorer som inngår i variabelen.

Undersøkelsens design

Et sentralt formål med TIMSS-undersøkelsen er å følge utviklingen i elev-ers faglige prestasjoner over tid – det vi kaller *trend*. For å få til dette, benyttes en fast måleskala (med midtpunkt satt til 500 poeng og standardfeil satt til 100 poeng) som har vært uendret siden den første gjennomføringen i 1995. Det betyr at resultatene fra hver nye undersøkelse kan sammenlignes direkte med tidligere år (Mullis et al., 2021). For å sikre denne sammenlignbarheten, inneholder hver gjennomføring en stor andel såkalte trendoppgaver – oppgaver som er identiske med dem som ble brukt i forrige runde. Omtrent halvparten av oppgavene i hver undersøkelse er slike trendoppgaver, og disse holdes hemmelige for å bevare validiteten i målingen. Denne metoden gjør det mulig å analysere endringer i elevenes kunnskaper og ferdigheter over tid, og gir et solid grunnlag for å vurdere utviklingen i utdanningssystemet. For en mer teknisk beskrivelse av skaleringen og metodegrunnlaget, se von Davier et al. (2024).

For å kunne dekke bredden i det faglige rammeverket som TIMSS bygger på, må undersøkelsen inkludere et omfattende sett med oppgaver. I TIMSS 2023 for naturfag på 9. trinn, for eksempel, ble det benyttet 218 oppgaver. Ideelt sett skulle hver elev ha besvart hele oppgavesettet, men dette ville krevd uforholdsmessig mye tid og vært urealistisk i en skole-sammenheng. I stedet benyttes et *rotasjonsdesign*, der oppgavene fordeles på et antall mindre oppgavesett – typisk 14 – og hver elev får tildelt ett sett. Elevene har 45 minutter til å besvare sitt sett. Omtrent halvparten av oppgavene er flervalgsoppgaver, mens den andre halvparten krever at elevene formulerer egne svar. Siden ingen elever besvarer hele oppgavesettet, benyttes en statistisk metode for å estimere elevenes prestasjoner dersom de hadde fått og gitt en besvarelse på alle de 218 oppgavene. Estimeringen gir fem såkalte *plausible verdier*. Verdiene er statistisk genererte og representerer et sannsynlig prestasjonsnivå for hver elev, basert på de oppgavene de faktisk fikk og besvarte. I tillegg tar beregningene hensyn til elevenes bakgrunnsinformasjon (Bezirhan & von Davier, 2024; Rubin, 1987). Denne tilnærmingen gjør det mulig å analysere resultater på gruppenivå med høy presisjon, men innebærer samtidig at det ikke kan gis tilbakemelding på individnivå. For en nærmere beskrivelse av bruk og beregning av plausible verdier, se for eksempel Rutkowski, Gonzalez et al. (2010) eller Rutkowski, von Davier og Rutkowski (2010).

Om TIMSS, kort oppsummert

I tabell 1.1 er det laget en oppsummering av de sentrale aspektene ved TIMSS-studien som er omtalt i kapittelet over. Tabellen gir en enkel oversikt blant annet over hvordan studien er organisert, hvordan den gjennomføres og hvilke data som samles inn.

Tabell 1.1 Forenklet oversikt over TIMSS-studien

Målgrupper	Elever på 4. og 8. trinn internasjonalt, 5. og 9. trinn i Norge (f.o.m. 2015), deres foresatte, matematikk- og naturfaglærere og rektorer
Fag som inngår i studien	Matematikk og naturfag
Hensikt med studien	Å kartlegge faktorer som fremmer læring, følge med på utvikling i eget land (trend), kunne sammenligne med andre lands utdanningssystemer og gi informasjon til eget lands læreplanutvikling og utdanningssystem
Hva måles	Kompetanse i forhold til et omforent rammeverk basert på alle de deltakende landenes læreplaner i matematikk og naturfag. I tillegg måles kontekstuelle faktorer gjennom spørreskjema til elever (f.eks. motivasjon), lærere (f.eks. lærerkompetanse), rektorer (f.eks. skoleklima) og foreldre (f.eks. barnas tidlige læring)
Utvalgsprosedyre	Skoler trekkes først, deretter hele klasser innen skolene. Maksimalt 2 klasser per skole
Antall deltakende elever	Ca. 5000–6000 elever per trinn, utgjør ca. 8–9 % av alle elevene per trinn. Elevene som blir trukket ut, utgjør representative utvalg av målgruppene
Gjennomføres	Hvert 4. år, første gang var i 1995
Antall land med i 2023	64
Har rammeverk for	Faglige emner (inkludert kognitive nivåer) og spørreskjema
Rammeverk og oppgaver	Utformes av fageksperter på grunnlag av deltakerlandenes læreplaner og oppgaveforslag fra deltakerlandene
Tidsbruk på test for elever og på spørreskjema for elever, lærere, skoleledere og foresatte	Fagtest 4. trinn: 36 min til matematikk, 36 min til naturfag Fagtest 8. trinn: 45 min til matematikk, 45 min til naturfag Spørreskjema elever: ca. 30 min Spørreskjema lærere: ca. 35 min Spørreskjema skoleledere: ca. 30 min Spørreskjema foresatte: ca. 35 min (kun 5. trinn)
Antall fagoppgaver i 2023	Matematikk: 187 på 5. trinn og 204 på 9. trinn Naturfag: 176 på 5. trinn og 218 på 9. trinn
Oppgaveformat fagtester	Flervalgsoppgaver og åpne oppgaver, fordelt omtrent 50/50
Poengskala	Skalamidtpunkt er 500 poeng med standardavvik på 100. Skala laget ut fra gjennomsnittet for de landene som deltok i 1995

Forfatterbiografier

Hege Kaarstein er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad i matematikdidaktikk og har jobbet med TIMSS-undersøkelsen siden 2009, og vært prosjektleder fra 2018. I tillegg deltar hun i den internasjonale

ekspertgruppen for matematikk i TIMSS. Hennes interesser er blant annet lærernes undervisningskompetanse i matematikk, elevenes motivasjon for faget og svar på åpne oppgaver.

Anne-Catherine Lehre er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en bakgrunn innenfor språk og realfag, og har en doktorgrad i biostatistikk. Lehre har jobbet i TIMSS-prosjektet siden 2020.

Referanser

- Bergem, O. K., Kaarstein, H. & Nilsen, T. (2016). TIMSS 2015. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015* (s. 11–21). Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016-02>
- Bezirhan, U. & von Davier, M. (2024). TIMSS achievement scaling methodology: Item response theory and population models. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report* (s. 11.1–11.15). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs8288>
- Fishbein, B., Taneva, M. & Kowolik, K. (2025). *TIMSS 2023 user guide for the international database*. Boston College, TIMSS & PIRLS International Study Center. <https://timss2023.org/data>
- Kaarstein, H., Lehre, A.-C., Rohatgi, M. & Radišić, J. (2024). *TIMSS 2023 kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. TIMSS & PIRLS International Study Center. <http://timssandpirls.bc.edu/timss2015/international-results/>
- Mullis, I. V. S., Martin, M. O. & von Davier, M. (Red.). (2021). *TIMSS 2023 assessment framework*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Norges forskningsråd. (2012). *Realfagsrekruttering langs nye spor. Sluttrapport*. <https://kudos.dfo.no/dokument/1878/sluttrapport-realfagsrekruttering-langs-nye-spor-sluttrapport>
- Reynolds, K. A., Aldrich, C. E. A., Bookbinder, A., Gallo, A., von Davier, M. & Kennedy, A. (Red.). (2024). *TIMSS 2023 encyclopedia: Education policy and curriculum in mathematics and science*. TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs5882>
- Reynolds, K. A., Mullis, I. V. S. & Martin, M. O. (2021). TIMSS 2023 context questionnaire framework. I I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.
- Rutkowski, L., Gonzalez, E., Joncas, M. & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151. <https://doi.org/10.3102/0013189X10363170>
- Rutkowski, L., von Davier, M. & Rutkowski, D. (Red.). (2010). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Chapman & Hall/CRC.
- Siegel, P. & Foy, P. (2024). TIMSS sample design. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report* (s. 3.1–3.30). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs7952>
- von Davier, M., Fishbein, B. & Kennedy, A. (Red.). (2024). *TIMSS 2023 technical report (methods and procedures)*. Boston College, TIMSS & PIRLS International Study Center. <https://timss2023.org/methods>

KAPITTEL 2

Utvikling av ulikheter i utdanningen over tid

Trude Nilsen, Wangqiong Ye og Nani Teig

Institutt for Lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: Økende ulikhet i utdanningen utfordrer demokratiske prinsipper og sosial mobilitet, men forskningen på langtidstrender, spesielt etter covid-19-pandemien og med bredere ulikhetsmål, er begrenset. Denne studien undersøker utviklingen av ulikheter i matematikkprestasjoner på 8. og 9. trinn i norsk skole fra 1995 til 2023. Den analyserer endringer i sammenhengen mellom elev-ers sosioøkonomiske status (SØS) og prestasjoner, både på elev- og klassenivå, samt utviklingen i andelen faglig resiliente elever (elever med lav SØS, men høye prestasjoner). Studien benytter TIMSS-data fra syv sykluser (fra 1995 til 2023) og analyserer sammenhengen mellom SØS og prestasjoner i matematikk ved hjelp av ett- og tonivå strukturell ligningsmodellering og lineær regresjon, samt beregner andelen resiliente elever. Resultatene indikerer en økende sammenheng mellom SØS og matematikkprestasjoner på elevnivå og muligens økte forskjeller på klassenivå. Andelen resiliente elever sank fra ca. 18 prosent i 1995 til ca. 12 prosent i 2023. Samlet sett peker funnene mot økende ulikheter over perioden, også under pandemien, med et unntak for perioden 2003–2007. Funnene understreker et preserende behov for forsterket innsats og koordinerte politiske tiltak fra både skole og samfunn for å motvirke økende utdanningsulikhet og sikre reelt like muligheter for alle elever.¹

Nøkkelord: ulikheter i utdanningen, resiliens, sosioøkonomisk status, trend-analyser, TIMSS

1 KI er benyttet til forbedring av språket samt til å oversette sammendraget fra norsk til engelsk .

Abstract**Development of Educational Inequalities Over Time**

Rising inequality in education challenges democratic principles and social mobility, but research on long-term trends, especially post-pandemic and using broader inequality measures, is limited. This study aims to investigate the development of inequalities in mathematics achievement at grades 8/9 in Norwegian schools from 1995 to 2023. It analyzes changes in the relationship between students' socioeconomic status (SES) and achievement, at both student and class levels, as well as the development in the proportion of academically resilient students (students with low SES but high achievement). The study utilizes TIMSS data from seven cycles (from 1995 to 2023) and analyzes the relationship between SES and mathematics achievement using single- and two-level structural equation modeling and linear regression, as well as calculating the proportion of resilient students. The results indicate a strengthening association between SES and mathematics achievement at the student level and may also suggest increased disparities at the class level. The proportion of resilient students declined from approximately 18% in 1995 to approximately 12% in 2023. Overall, the findings point towards increasing inequalities throughout the period, including during the pandemic (2019–2023), with an exception for the period 2003–2007. The findings underscore an urgent need for reinforced efforts and coordinated policy measures from both schools and society to counteract rising educational inequality and ensure genuinely equal opportunities for all students.

Keywords: inequalities in education, resilience, socioeconomic status, trend analyses, TIMSS

Innledning

Ulikheter i samfunnet og i skolen kan potensielt undergrave selve grunnlaget for et velfungerende demokrati. Når store deler av befolkningen opplever at de ikke har de samme mulighetene som andre, kan det føre til økt mistillit til de demokratiske institusjonene og en følelse av at systemet ikke er rettferdig (Wilkinson & Pickett, 2009). Dette kan igjen svekke den sosiale samhörigheten og den demokratiske deltakelsen (Rege et al., 2023). Når sosiale forskjeller blir for store, kan det også føre til svekket sosial mobilitet² og redusert deltakelse i samfunnslivet (Løken et al., 2024; Rege et al., 2023).

Ulikheter i utdanningen er et komplekst begrep som kan defineres på mange forskjellige måter. Den vanligste definisjonen i de internasjonale storskalaundersøkelsene er *ulikheter i prestasjoner grunnet sosioøkonomisk status* (eller andre faktorer som språk eller kjønn) (OECD, 2023). Begrepet blir problematisert og diskutert dypere i neste delkapittel.

Tidligere forskning, blant annet fra OECD (2013), indikerer at det er en sammenheng mellom mindre ulikhet i skolen og høyere prestasjoner for alle elever. Land med god økonomi har ofte mindre sosiale ulikheter (OECD, 2024). Det er spesielt problematisk med ulikheter i realfag fordi når mange elever ikke får sjansen til å utvikle sitt potensial til å lykkes, kan dette bidra til at færre går inn i yrker i sektorer innen naturvitenskap, teknologi, ingeniørfag og matematikk (Henriksen et al., 2015). Det er et stort behov for arbeidskraft innen slike yrker, da de er viktige for bærekraftig utvikling innen miljø, økonomi og helse (Henriksen et al., 2015).

Det finnes en del forskning på hvordan ulikheter har endret seg over tid i Norge, både i samfunnet generelt og i skolen spesielt (f.eks. Nilsen et al., 2018; Sandsør et al., 2023; Aaberge et al., 2021). Bildet som tegner seg er ikke helt klart, da trenden endrer seg i forskjellige tidsepoker og avhenger både av endringer i samfunnet (som pandemien, finansielle kriser osv.), hvilken periode som undersøkes og hvilke data og metoder som benyttes.

Ulikheter i samfunnet har en tendens til å gjenspeiles i skolen. Til tross for at opplæringslovens § 1–1 har som formål å bidra til likeverdige opplæringsmuligheter, viser forskning at sosiale bakgrunnsfaktorer fortsatt har betydelig innvirkning på elevenes prestasjoner i Norge (Nilsen & Jensen, 2024; Sandsør et al., 2023). Eksisterende forskning på ulikheter i skolen har imidlertid begrensninger. Mange studier som ser på utviklingen av

2 Sosial mobilitet betyr at man har muligheten til å endre sosial status, og at sosioøkonomisk status i liten grad føres videre fra en generasjon til den neste.

ulikheter over tid, stopper før koronapandemien (f.eks. Nilsen et al., 2018; Sandsør et al., 2023), noe som gjør det vanskelig å vurdere hvordan pandemien har påvirket disse trendene. Videre har forskningen ofte et begrenset sett av indikatorer på ulikhet. Det fokuseres ofte på effekten av sosioøkonomisk status (SØS)³ på prestasjoner, uten å ta tilstrekkelig hensyn til andre viktige aspekter, som for eksempel resiliens. Resiliens, det vil si elevenes evne til å mestre skolen til tross for utfordringer (som f.eks. sosioøkonomiske forhold), er et viktig mål på ulikhet fordi det sier noe om skolens evne til å kompensere for sosiale forskjeller (Teig, 2024; Ye et al., 2024). For å kunne utvikle mer effektive tiltak for å redusere ulikhet i skolen, er det avgjørende å ha en dypere forståelse av hvordan ulikhetene har utviklet seg over tid. Denne studien tar sikte på å fylle noen av disse kunnskapshullene ved å stille følgende forskningsspørsmål:

1. Hvordan har sammenhengen mellom SØS og prestasjoner endret seg fra 1995 til 2023 på individnivå og på klassenivå i Norge?
2. Hvordan har andelen resiliente elever endret seg fra 1995 til 2023 i Norge?

Som nevnt er ulikheter i matematikk spesielt viktige med tanke på fremtidig bærekraftig utvikling, og vi fokuserer derfor på matematikk. Videre er de fleste tidligere studier på ulikheter innen matematikk på ungdomstrinn (se f.eks. Glassow et al., 2023; Nilsen & Jensen, 2024; Sandsør et al., 2023), og for å kunne sammenligne våre funn med tidligere forskning, undersøker vi derfor elever på 8. og 9. trinn i matematikk.

Som forskningsspørsmålene viser, har vi som mål både å kartlegge langsiktige trender (inkludert perioden både før og etter pandemien), samt å benytte flere mål for ulikhet (resiliens og sammenhengen mellom SØS og prestasjoner på flere nivåer). Dette vil bidra til et mer nyansert bilde. Ved å kombinere disse perspektivene, vil studien bidra til en mer helhetlig forståelse av utviklingen av ulikheter i skolen. Denne kunnskapen er avgjørende for å kunne utvikle tiltak som fremmer *like muligheter for alle elever*. Studien gir likevel ikke et komplett bilde av utviklingen av ulikheter i skolen, da mange perspektiver mangler – som for eksempel innvandringsbakgrunn og kjønn.

3 Det er viktig å understreke at faktorene som inngår i SØS-variabelen kun utgjør forenklete mål og ikke fullt ut reflekterer de komplekse og mangfoldige aspektene ved SØS. Dermed fungerer de som proxyer snarere enn fullstendige representasjoner av sosioøkonomisk bakgrunn.

Teori og tidligere forskning

Definisjoner og mål på ulikheter i utdanningen

Ulikheter i utdanningen kan defineres og måles på mange forskjellige måter (se f.eks. UNESCO, 2018). Definisjonen av ulikheter vil ofte reflektere en underliggende ideologi. Fem sentrale definisjoner på ulikheter i utdanningen er:

1. **Ulikhet i prestasjoner.** Målet er å redusere gapet mellom elever som presterer høyt og elever som presterer lavt. Dette er en egalitær ideologi som kjennetegner de nordiske landene (Frønes et al., 2020). Denne typen ulikhet måles ved for eksempel spredning i prestasjoner som standardavvik eller varians.
2. **Ulikheter i prestasjoner grunnet SØS** (eller kjønn, migrasjonsbakgrunn eller andre indikatorer). Målet er at alle skal ha like muligheter til å lykkes – at forskjeller mellom prestasjoner i minst mulig grad skal skyldes elevers SØS, kjønn, migrasjonsbakgrunn, geografisk bosted, religion og så videre. Denne typen ulikhet kan måles ved for eksempel styrken av effekten av SØS på prestasjoner, eller i hvor stor grad SØS forklarer forskjeller mellom elevers prestasjoner.
3. **Ulikheter i grunnleggende ferdigheter.** Målet er at alle elever skal oppnå grunnleggende ferdigheter, for eksempel i lesing, skriving og regning. Denne typen ulikhet kan måles ved for eksempel å dele ferdigheter eller kompetanse i nivåer og måle andel elever som oppnår det laveste kompetansenivået, slik som i PISA og TIMSS.
4. **Ulikheter i resiliens.** Målet er at skolen og samfunnet kan hjelpe elever til å lykkes på tross av utfordringer knyttet til for eksempel hjemmebakgrunn. Denne typen ulikhet kan for eksempel måles ved å finne andelen elever som presterer godt til tross for lav SØS (Agasisti et al., 2018).
5. **Ulikheter i muligheter til å lære.** Målet er at alle elever skal ha like muligheter til å lære (*opportunity to learn*, se f.eks. Rosén & Nilsen, 2024). Dette er et bredt begrep som kan omfavne hvorvidt alle elever får like god undervisningskvalitet, et godt skolemiljø uten mobbing, læringsressurser (f.eks. tilgang til PC, bøker osv.) og flere andre faktorer. Denne typen ulikhet kan for eksempel måles ved å undersøke hvorvidt ressurser eller god undervisning er jevnt fordelt mellom skoler eller klasser, eller om skoler med høy andel elever med høy SØS har bedre tilgang til for eksempel god undervisningskvalitet (se f.eks. Glassow & Jerrim, 2022; Nilsen et al., 2025).

I denne studien har vi valgt å undersøke ulikheter i prestasjoner grunnet SØS, samt ulikheter i resiliens. Av de fem typene ulikheter, så finnes det grundig dokumentasjon på de tre andre typene ulikheter, blant annet fra OECD (2023). Forskjeller i prestasjoner (i form av f.eks. spredning), og kompetansenivåer blir jevnlig publisert av de internasjonale stor-skalaundersøkelsene. Det finnes også en del forskning på ulikheter i elevenes muligheter til å lære i Norge (Blömeke, 2024; Nilsen et al., 2020, 2025), og disse dreier seg mest om hvorvidt skolene evner å minske ulikheter. I denne studien er vi mer interessert i elevenes hjemmebakrunn. I tillegg ønsker vi å undersøke resiliens, både fordi det her mangler forskning som sammenligner resiliens over tid, og fordi vi ønsker å se om disse ulikhetene følger samme mønster som ulikheter i prestasjoner grunnet SØS.

Både definisjoner og målinger av ulikheter kan påvirke resultatet når det gjelder grad av ulikhet (Mittal et al., 2020). For eksempel, ved å benytte den første av de fem definisjonene (forskjeller i prestasjoner), så viser OECD at det er forholdsvis små forskjeller i prestasjoner mellom skoler i Norge sett i et internasjonalt perspektiv (OECD, 2023b). Benytter man derimot definisjon nr. 2, altså i hvor stor grad disse forskjellene skyldes elevers hjemmebakgrunn, er ikke Norge lenger like høyt oppe på listen over land med like muligheter til å lykkes (OECD, 2023b).

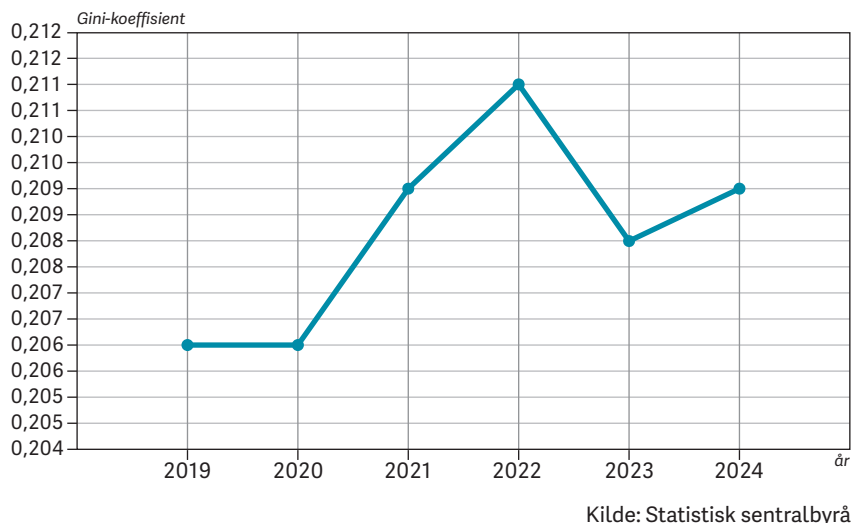
Videre vil det ha noe å si hvordan SØS måles. Oftest benyttes Bourdieus (1986) teorier for å definere SØS. Bourdieu definerte SØS ved fire aspekter: økonomiske ressurser (f.eks. inntekt), sosiale ressurser (f.eks. sosialt nettverk), kulturelle ressurser (f.eks. bøker, malerier eller piano), og symbolske ressurser (prestige, som f.eks. Abelprisen i matematikk eller en idrettsmedalje). For å måle elevers SØS benyttes som regel foreldres inntekt, foreldres utdanning, foreldres jobb, antall bøker hjemme, ressurser til bruk i utdanningen (f.eks. datamaskin, egen pult) og flere andre indikatorer (OECD, 2013). Noen ganger kan det være nyttig å studere disse indikatorene hver for seg dersom man for eksempel ønsker å undersøke hvilke faktorer som bidrar mest til ulikheter. For eksempel viser Sandsør et al. (2023) at ulikheter mellom elever har økt siden 2007 dersom kun indikatoren på SØS, foreldrenes inntekt, benyttes, mens utviklingen er tilnærmet stabil dersom man benytter indikatoren foreldrenes utdannelse. Dette gir imidlertid ikke et helhetlig bilde, fordi forskning viser at ulikheter mellom elever oppstår som en konsekvens av *sammensetningen* mellom flere faktorer, og SØS er ment for å dekke dette (Sirin, 2005).

Videre har det stor betydning om man undersøker forskjeller mellom *elevers* prestasjoner (altså analyser på elevnivå) eller mellom *skoler* (altså analyser på skolenivå) (Mittal et al., 2020). Det kan for eksempel være store forskjeller mellom elevers prestasjoner, men små forskjeller mellom skolars gjennomsnittlige prestasjoner. Forskjeller mellom skoler reflekterer i større grad forskjeller på systemnivå, og store forskjeller er tegn på det som i forskningen kalles segregering. For eksempel er det i Sverige store forskjeller mellom skoler grunnet at foreldre kan velge hvilken skole barnet deres skal gå på (Hansen & Gustafsson, 2016). Dette fører til at noen skoler har mange elever fra ressusterke hjem og hvor prestasjonene er høye, mens andre skoler har store andeler med elever fra ressursvake hjem og ofte med innvandringsbakgrunn (Hansen & Gustafsson, 2016). Dette er tegn på segregering og kan ha sammenheng med lavere sosial mobilitet (Rege et al., 2023), dårligere økonomi for landet (OECD, 2023b) og svekket demokrati (Wilkinson & Pickett, 2009). Dette er likevel kun ett eksempel, og bildet er selvsagt mer nyansert og kan variere på tvers av land.

Oppsummert vil resultater av forskning på ulikheter kunne variere avhengig av hvilken definisjon på ulikheter som benyttes, hvordan SØS måles, om man benytter enkeltfaktorer (som foreldrens inntekt), og hvilket nivå analysene blir gjort på. Det er derfor viktig at forskning er transparent og viser hvordan ulikheter blir definert, hvilke mål (f.eks på SØS) som benyttes og hvordan analysene blir gjort. For trendanalyser betyr det at det er viktig å benytte flere forskjellige mål på ulikheter for å undersøke om man får samme mønster over tid; om analysene peker mot økte likheter eller ulikheter.

Ulikheter i utdanningen i Norge

Når man skal undersøke hvordan ulikheter i *utdanningen* har endret seg over tid, kan det være nyttig å også undersøke hvordan ulikheter i *samfunnet* generelt har endret seg. Dette fordi ulikheter i samfunnet smittes over til skolen (Løken et al., 2024). Statistisk sentralbyrå (SSB) har flere artikler som viser at ulikheter i Norge har økt (se f.eks. Grini et al., 2023; Aaberge et al., 2021). Aaberge et al. (2021) viste at utviklingen i ulikhet i fordelingen av inntekt etter skatt økte fra 2001 til 2018, og det samme gjorde ulikheter i formue. Med andre ord har ulikheter i samfunnet økt i denne perioden. Statistikkbanken deres viser videre at ulikhetene i fordelingen av inntekt etter skatt økte fra 2020 til 2024, men med et fall mellom 2022 og 2023 (se figur 2.1). Økningen mellom 2020 og 2022 har høyst sannsynlig sammenheng med pandemien.



Figur 2.1. Lønnsfordeling ved Gini⁴-koeffisient.

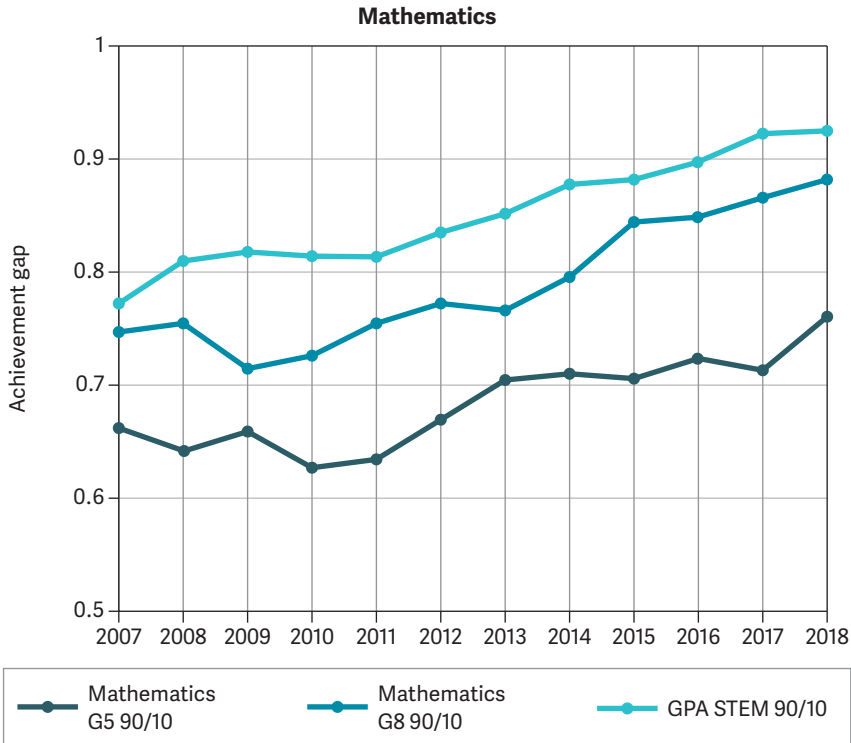
Videre har sosial mobilitet minket i løpet av de siste tiårene (Markussen & Røed, 2023). Sosial mobilitet er muligheten til å bevege seg mellom forskjellige sosiale klasser eller grupper av samfunnet, slik at barn født av foreldre med lav SØS har muligheten til selv å oppnå høy SØS. Lav sosial mobilitet er forbundet med høy ulikhet (Markussen & Røed, 2023).

Disse resultatene om økt ulikhet i inntekt, formue og sosial mobilitet de siste tiårene samsvarer med resultatene til Sandsør et al. (2023), som viser at gapet mellom prestasjonen til elever med foreldre med høy inntekt og de med lav inntekt, økte fra 2007 til 2018 (se figur 2.2). Elevers prestasjoner ble målt ved å benytte resultater fra de nasjonale prøvene i matematikk på 5. og 8. trinn, og karaktersnittet av standpunktarakter og eksamen i matematikk og naturfag på 10. trinn.

Funnene samsvarer også med en studie som har analysert data fra PISA og TIMSS (Nilsen et al., 2018). Denne viser at SØS kan forklare en økende andel av forskjellene i elevers prestasjoner fra 1995 til 2015, selv om resultatene fra TIMSS var klarere enn de fra PISA. Til sammen indikerer disse studiene at ulikheter i samfunnet ser ut til å smitte over til skolen.

Nylig ble funn fra PISA 2022 publisert. Bildet som tegner seg er noe uklart. I PISA var matematikk et hovedområde i 2022, 2012 og 2003.

4 Gini-koeffisienten er et mål for spredning. Koeffisienten har en tallverdi fra 0 til 1, hvor 0 betyr at alle innbyggerne har akkurat like stor inntekt eller formue, mens verdien 1 indikerer at én person eier all inntekt eller formue.



Figur 2.2. Fra Sandsør et al. (2023). Forskjeller i prestasjoner på nasjonale prøver på 5. og 8. trinn, og karakterer i 10. trinn på matematikk og naturfag for elever med foreldre med høy og lav inntekt.

For å få mest reliable resultater er det disse syklusene som burde sammenlignes dersom matematikk er i fokus. Det er fordi det er langt høyere antall matematikkoppgaver i disse syklusene, enn i de syklusene hvor lesing eller naturfag er i fokus. OECD (2023a) viste at endringen i forklart varians⁵ i matematikkprestasjoner av SØS på elevnivå fra 2012 til 2022 ikke var signifikant. I 2022 kunne 9,6 prosent (standardfeil, SE = 0,9) av variansen på elevnivå forklares av SØS, i 2012 var det 7,2 prosent (SE = 1,0), hvor endringen ikke var signifikant (2,3, SE = 1,3). Derimot var endringer i i gjennomsnittsprestasjoner mellom de 25 prosent av elevene med høyest og lavest SØS fra 2012 til 2015 signifikant. Endringer i gjennomsnittresultater mellom de 25 prosent av elevene med høyest og lavest SØS fra 2018 til 2022 var derimot ikke signifikant i matematikk.

5 Et mål på forskjeller i prestasjoner mellom elever.

Nilsen og Jensen (2024) fant heller ingen signifikant forskjell i forklart varians mellom 2018 og 2022 på *elevnivå*, men derimot økte gapet mellom *skoler* i denne perioden. Alle disse resultatene kan tyde på at endringer i ulikheter for det meste er på skolenivået, altså at forskjellene mellom skolene har økt.

Når det gjelder resiliente elever, er det svært få som har undersøkt endringer i et langtidsperspektiv. Agasti et al. (2018) brukte data fra PISA 2006, 2009, 2012 og 2015 for å undersøke endringer i andel resiliente elever. De fant at andelen resiliente elever i matematikk, naturfag og lesing *sammenlagt*, økte i Norge i denne perioden. Det er usikkert hva dette resultatet hadde blitt dersom dette ble undersøkt i kun matematikk, fordi grad av resiliens (motstandsdyktighet) kan variere på tvers av fag.

Oppsummert viser de forskjellige studiene ulike perspektiver og ulike mål på ulikheter, og fra perioder som ikke er helt sammenfallende. Selv om disse forskningsstudiene til sammen muligens tegner et bilde av økte ulikheter i samfunnet og skolen, er det ikke et entydig bilde. Ingen av studiene undersøker en langtidstrend for skolen, inkludert perioden med pandemi. De studiene som undersøkte ulikheter i skolen, ble enten gjort på skolenivå fra 2018 til 2022 (Nilsen & Jensen, 2024), eller på elevnivå fra 2007 til 2018 (Sandsør et al., 2023). I tillegg er det ingen studier på klassenivå. Det er viktig å adressere dette kunnskapshullet, for det er indikasjoner på at forskjeller mellom klasser innen skoler kan være store (Nilsen & Bergem, 2016; Nilsen et al., 2025). Forskjeller mellom klasser kan blant annet skyldes at elever med behov for tilrettelagt undervisning eller ekstra norskundervisning ofte går i samme klasse (Haugen, 2020). Det er med andre ord behov for kunnskap om endringer i ulikheter i et langtidsperspektiv hvor perioden med pandemi er innkludert, og det er behov for å undersøke dette på flere nivåer (inkludert klassenivå), og med flere perspektiver. Dette vil gjøre det mulig å sammenligne ulikhet målt på forskjellige måter, og gi et nansert bilde av ulikheter i skolen. Denne studien adresserer disse behovene.

Metode

Data og utvalg

Vi benytter TIMSS-data fra 1995, 2003, 2007, 2011, 2015, 2019 og 2023. Utvalget som inkluderes i analysene er elever på 9. trinn (i 2023, 2019 og 2015) og 8. trinn (i 1995, 2003, 2007 og 2011). Vi fokuserer på matematikkprestasjoner.

Mål på SØS og prestasjoner

Vi benytter elevspørreskjema og bruker følgende mål på SØS:

- Antall bøker i hjemmet. Elevene får fem svaralternativer, fra «Ingen eller veldig få (0–10 bøker)» til «Nok til å fylle tre eller flere reoler (flere enn 200)».
- Foreldrenes utdanning: Elevene får syv valg, fra «Gått litt på barneskole eller ungdomsskole eller har ikke gått på skolen» til «Universitet eller høyskole, minst 5 år (mastergrad eller doktorgrad)».
- Ressurser i hjemmet for utdanning: hvorvidt elevene hadde «Din egen pult eller et bord som du gjør lekser ved» og hvorvidt de hadde «PC eller nettbrett».

Antall bøker hjemme har ikke endret seg siden 1995. Foreldrenes utdanning ble rekodet slik at variablene fikk verdiene 1 dersom høyeste utdanning var videregående skole eller lavere, 2 dersom foreldrene hadde utdanning utover videregående, men ikke universitetsutdanning, og 3 dersom foreldrene hadde graden bachelor eller høyere.

Disse målene på SØS har ikke endret seg nevneverdig, og blir derfor benyttet til analysene. Indikatorene blir satt sammen til en samlevariabel for SØS.

Matematikkprestasjoner blir angitt ved fem plausible variabler (se kapittel 1), og alle fem blir benyttet i analysene ved imputasjon (altså ved at analysene blir gjort for hver plausible variabel, og deretter blir gjennomsnittet av estimatene fra de fem analysene kalkulert).

Analysen

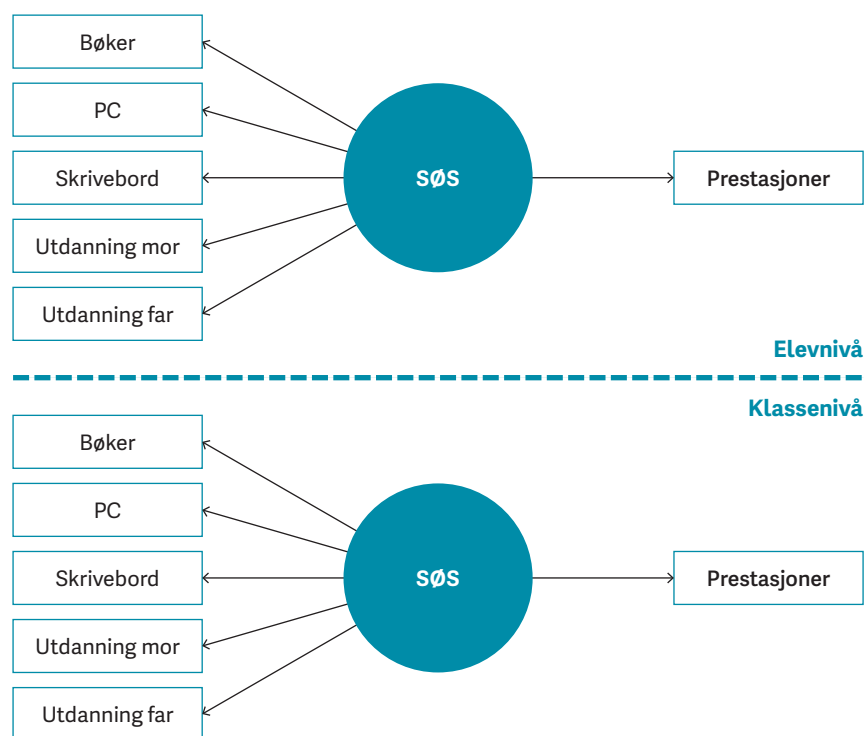
For å besvare forskningsspørsmål 1, benyttes både vanlig lineær regresjon og strukturell ligningsmodellering (SEM). Vi benytter begge disse typene regresjon for validering, for å sjekke hvor robuste resultatene er. SEM er en type regresjonsanalyse, hvor man også får mål på reliabilitet og validitet, og hvor man kan gjøre analysene på to nivåer samtidig (f.eks elev- og klassenivå). Vi undersøker også målingsinvarians på tvers av tid, for å se om målet på SØS er sammenlignbart over tid.

Først undersøker vi sammenhengen mellom SØS og prestasjoner på *ett nivå* (elevnivå). Deretter benyttes *to-nivå* lineær regresjon og SEM hvor sammenhengen mellom SØS og prestasjoner estimeres på elev- og klassenivå samtidig. Vi estimerer også forklart varians (R^2), altså hvor mange prosent av variansen («forskjellen» i prestasjoner) mellom *elev*er som kan

forklares av SØS, og hvor mange prosent av variansen mellom *klasser* som kan forklares av SØS.

I de lineære regresjonsmodellene benytter vi faktorskår for å måle SØS. Det betyr at vi legger sammen de fem indikatorene for SØS beskrevet i forrige delkapittel. Disse blir vektet i henhold til resultater av konfirmatorisk faktoranalyse (som blant annet gir oss faktor-ladinger). Resultatene av disse analysene er enklere å tolke enn de fra SEM-modellene, og blir derfor inkludert i teksten.

I SEM-modellene måles SØS som en latent variabel bestående av de fem indikatorene beskrevet over. Den konseptuelle modellen for tonivå-analysene er vist i figur 2.3.



Figur 2.3. SEM-modell for tonivåanalyser

Resultatene fra SEM-modellene, inkludert faktorladinger, modell tilpasning og målingsinvarianse, er inkludert i henholdsvis vedlegg 2.1, 2.2, 2.3 og 2.4.

For å besvare forskningsspørsmål 2, altså andelen resiliente elever over tid, så identifiserer vi elever med lav SØS og høy matematikkprestasjon. Elever som er blant de med 25 prosentene med lavest SØS i Norge, men

som samtidig tilhører de 25 prosentene av elevene som skårer høyest på matematikk i Norge, defineres her som resiliente elever.

Resultater

Forskningsspørsmål 1: SØS og prestasjoner

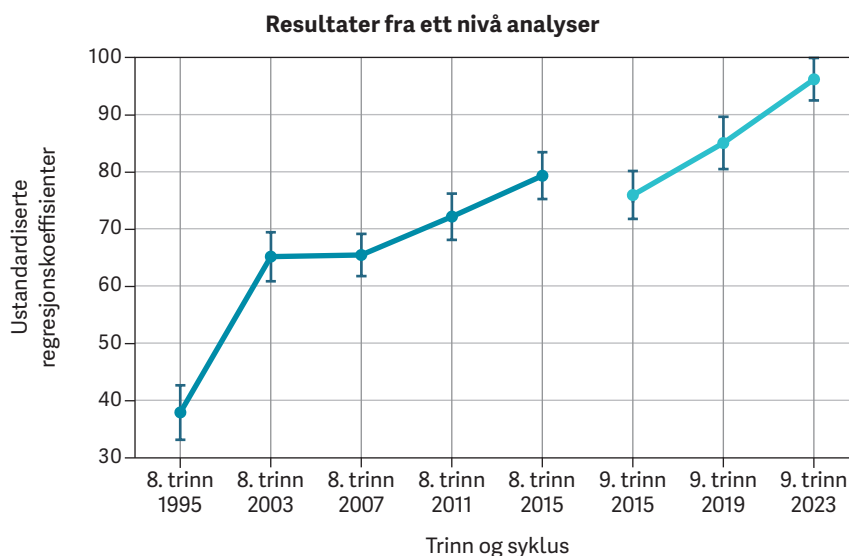
Tabell 2.1 viser resultatene av de lineære regresjonsanalysene som er gjort på ett nivå. I 2015 gikk Norge fra å måle elever på 8. trinn til 9. trinn, og dette året har vi data fra begge trinnene. Tabell 2.1 viser både de ustandardiserte og standardiserte regresjonskoeffisientene og standardfeil (SE), samt andel forklart varians. Standardiserte og ustandardiserte regresjonskoeffisienter er begge mål på styrken og retningen av forholdet mellom en uavhengig variabel (f.eks. SØS) og en avhengig variabel (f.eks. prestasjoner) i en regresjonsanalyse, men de uttrykkes i ulike skalaer. Ustandardiserte koeffisienter måles i de opprinnelige enhetene til variablene, mens standardiserte koeffisienter er skala-uavhengige og måles i standardavvik. Standardiserte regresjonskoeffisienter ligger mellom 0 og 1, og jo høyere de er, jo sterkere er sammenhengen. For eksempel, i 1995 var den ustandardiserte regresjonskoeffisienten omtrent 38, som betyr at en økning med 1 på SØS-skalaen tilsvarer en økning med 38 poeng i matematikkprestasjoner. I 2023 har denne økt til omtrent 96. Den er altså nesten tre ganger så stor i 2023 som i 1995. Siste kolonne til høyre viser forklart varians (R^2). I 1995 kunne 2,4 prosent av variansen (forskjellene i prestasjoner mellom elever) forklares med SØS, mens i 2023 kunne SØS forklare omtrent 12 prosent av variansen.

Tabell 2.1. Analyser på ett nivå (elev-nivå)

Trinn og år	Ustandardiserte regresjonskoeffisienter	S.E.	Standardiserte regresjonskoeffisienter	S.E.	R^2
8. trinn 1995	37,86	4,75	0,15	0,02	0,02
8. trinn 2003	65,12	4,30	0,27	0,02	0,08
8. trinn 2007	65,43	3,71	0,29	0,02	0,08
8. trinn 2011	72,13	4,05	0,32	0,02	0,10
8. trinn 2015	79,30	4,10	0,33	0,02	0,11
9. trinn 2015	75,92	4,19	0,31	0,02	0,10
9. trinn 2019	85,03	4,58	0,31	0,02	0,09
9. trinn 2023	96,18	3,73	0,35	0,01	0,12

SE = standardfeil.

Figur 2.4 viser en graf av de ustandardiserte regresjonskoeffisientene fra tabell 2.1.



Figur 2.4. Regresjonskoeffisienter for sammenhengen mellom SØS og prestasjoner (y-aksen) over tid (x-aksen). Standardfeil er vist for hvert punkt. Mørkeblå linje viser 8. trinn, lyseblå linje viser 9. trinn.

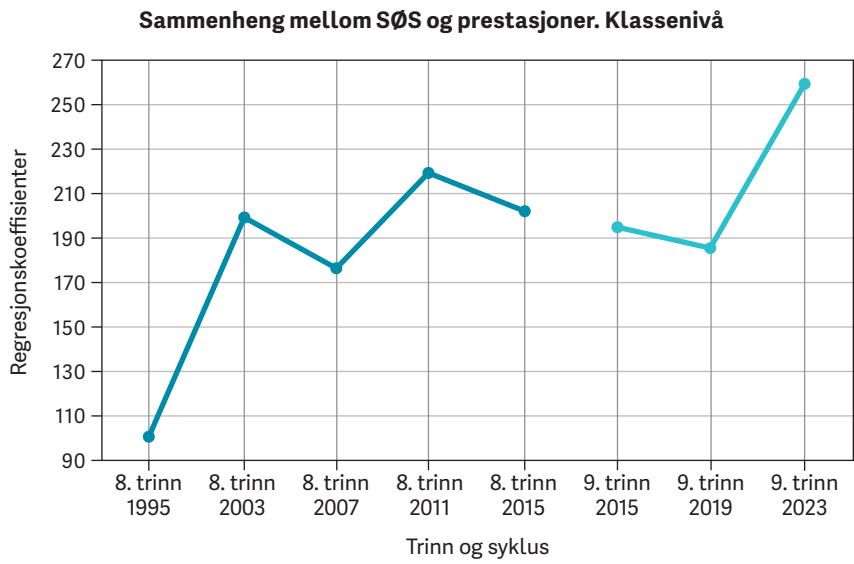
De mørkeblå linjene viser 8. trinn og de lyseblå viser 9. trinn. Selv om ikke alle endringer er statistisk signifikante, er det en statistisk signifikant økning fra 1995 til 2015 for 8. trinn, og fra 2015 til 2023 for 9. trinn. Resultatene viser en trend hvor ulikheter i prestasjoner relatert til SØS ser ut til øke over tid. Det ble også gjort to-nivå lineære regresjonsanalyser (elev- og klassenivå), og tabell 2.2 viser resultatene. Figur 2.5 viser resultatene fra klassenivå (ustandardiserte regresjonskoeffisienter).

Resultatene på klassenivå (figur 2.5) fra to nivå analysene viser ikke like klare resultater som analysene på ett nivå. Det er heller ikke like enkelt å beregne hvorvidt endringer over tid er statistisk signifikante, i og med at estimatene på klassenivå er kontrollert for individnivå, og i og med at det er langt færre klasser enn elever. Som for resultatene fra analysene på ett nivå, er endringene størst fra 1995 til 2003 og fra 2019 til 2023. Dersom man også tar forklart varians med i betraktning, kan trenden muligens peke på økte ulikheter.

Tabell 2.2. To-nivå analyser (elever og klasser)

To-nivå-analyser		Ustandardiserte regresjons-koeffisienter	S.E.	Standardiserte regresjons-koeffisienter	S.E.	R ²
8. trinn 1995	Elevnivå	33,34	4,59	0,14	0,019	0,02
	Klassenivå	101,34	41,41	0,42	0,152	0,17
8. trinn 2003	Elevnivå	53,96	4,43	0,23	0,018	0,05
	Klassenivå	199,43	25,77	0,70	0,073	0,49
8. trinn 2007	Elevnivå	57,64	3,41	0,26	0,015	0,07
	Klassenivå	176,57	26,78	0,62	0,082	0,38
8. trinn 2011	Elevnivå	60,18	3,47	0,27	0,074	0,07
	Klassenivå	219,27	26,71	0,71	1,486	0,51
8. trinn 2015	Elevnivå	63,67	3,73	0,27	0,015	0,07
	Klassenivå	202,27	19,43	0,76	0,054	0,58
9. trinn 2015	Elevnivå	62,30	3,73	0,26	0,015	0,07
	Klassenivå	194,61	20,93	0,76	0,063	0,58
9. trinn 2019	Elevnivå	73,64	4,14	0,27	0,015	0,07
	Klassenivå	185,16	24,86	0,62	0,065	0,39
9. trinn 2023	Elevnivå	82,78	3,72	0,31	0,013	0,09
	Klassenivå	258,81	21,10	0,78	0,045	0,61

Forklart varians i analysene på ett nivå er allerede veldig liten, og det kan forklare hvorfor den minker lite når variansen på klassenivået blir inkludert i tonivåanalysene. Forklart varians på klassenivå er, som forventet, større enn på elevnivå og på høyde med det som er vanlig i tilsvarende studier (se f.eks. Nilsen & Jensen, 2024). I flernivåanalyser av elevprestasjoner ligger brorparten av den totale variansen vanligvis på *elevnivå*, heller enn mellom klasser. Ettersom denne store variansen innad i klassene påvirkes av mange unike, individuelle faktorer, kan en enkelt prediktor på elevnivå bare forklare en beskjeden andel av den. I motsetning til dette gjenspeiler den mindre *variansen mellom klasser* mer systematiske forskjeller mellom læringsmiljøer. Følgelig har aggregerte faktorer på systemnivå, slik som gjennomsnittlig SØS i klassen, større forklaringskraft, ettersom de jevner ut individuell støy og dermed kan forklare en mye høyere andel av disse systematiske forskjellene mellom klasser.



Figur 2.5. Regresjonskoeffisienter av SØS på prestasjoner på klassenivå. Mørkeblå linje viser 8. trinn, lyseblå viser 9. trinn.

Analysene fra SEM-modellene på ett og to nivåer, samt målingsinvarians, er vist i vedlegg 2.1 til 2.4. Mønsteret i endringer av regresjonskoeffisienter er tilsvarende som for de lineære regresjonsmodellene. Resultatene fra målingsinvariansen viste metrisk (på engelsk: *metric*) invarians, som betyr at regresjonskoeffisienter er sammenlignbare over tid, mens gjennomsnittet av SØS ikke er sammenlignbart over tid.

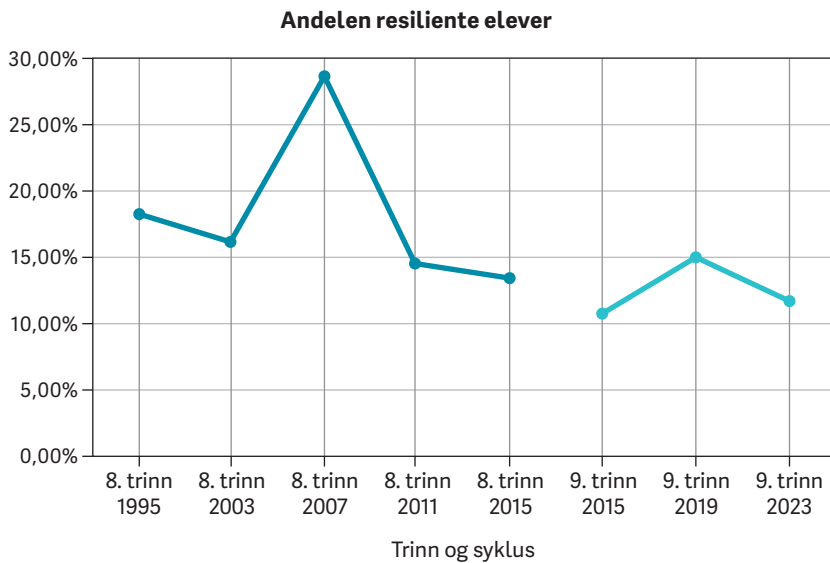
Forskningsspørsmål 2. Ulikheter med resiliens som mål

I forskningsspørsmål 2 undersøker vi prosentandelen resiliente elever fra 1995 til 2023. Dersom andelen resiliente elever har økt, viser det at det er flere elever som lykkes (presterer bra i matematikk) til tross for lav SØS.

Tabell 2.3. Prosentandel resiliente elever

Syklus og trinn	8. trinn 1995	8. trinn 2003	8. trinn 2007	8. trinn 2011	8. trinn 2015	9. trinn 2015	9. trinn 2019	9. trinn 2023
Andel resiliente elever	18,12 %	16,16 %	28,53 %	14,60 %	13,35 %	10,82 %	14,95 %	11,64 %

Tabell 2.3 viser andelen resiliente elever over tid, og det gjør også figur 2.6. Andelen minker over tid, men det er ingen jevn nedgang. Fra 2003 til 2007 er det en oppgang i prosentandel resiliente elever. Dersom man ser på hele perioden, var det i 1995 cirka 18 prosent resiliente elever, mens det i 2023 var cirka 12 prosent. Det betyr altså at det er færre elever over tid som presterer bra i matematikk til tross for lav SØS. Dette reflekterer økende grad av ulikhet i utdanningen.



Figur 2.6. Andel resiliente elever (y-aksen) over tid (x-aksen). Mørkeblå linje viser 8. trinn, lyseblå linje viser 9. trinn.

Diskusjon

Tolkning og diskusjon i lys av tidligere forskning

I denne studien har vi undersøkt ulikheter ved å se på endringer i sammenhengen mellom SØS og prestasjoner (med ett - og to nivåanalyser) og andelen resiliente elever fra 1995 til 2023. Resultatene er ikke helt entydige, men sammenlagt kan det se ut som de peker i samme retning: Ulikhetene øker i skolen i hele perioden, og de øker også mellom 2019 og 2023 (altså under pandemi-perioden). Disse funnene samsvarer med funn fra SSB (Aaberge et al., 2021), og med resultatene fra nasjonale prøver mellom 2007 og 2018 (Sandsør et al., 2023). Det samsvarer også

med Markussen og Røed (2023), som viser at sosial mobilitet har minket i denne perioden.

Resultatene på klassenivå var minst tydelige, men indikerer muligheten at forskjellene mellom klasser har økt over tid. Dette kan tyde på økt segresjon. I media har det vært skrevet en del om at flere skoler plasserer elever med innvandrerbakgrunn eller lavtpresterende elever i klasser for seg, og høytpresterende elever i andre klasser (se f.eks. Aftenposten, 2025; Ertesvåg, 2012). Ofte kan dette skje uten intensjon, fordi elever som var venner i barnehagen, og venner på barneskolen, eller bor i samme nabolag, ofte får komme i samme klasse (Haugen, 2020). Imidlertid kan ett nabolag ha annerledes SØS og demografi enn et annet nabolag i nærheten. Nyere studier indikerer også at det er store gap i prestasjoner mellom klasser som følge av SØS og innvandrerbakgrunn, spesielt i de store byene (Nilsen et al., 2025), og at dette kan skyldes skjult segregering. Våre resultater på klassenivå ligner også resultater fra PISA, hvor endringer i sammenhengen mellom SØS og prestasjoner i matematikk på skolenivå på 10. trinn økte fra 2018 til 2022 (Nilsen & Jensen, 2024).

Det er videre foruroligende at andelen resiliente elever minker. Dette indikerer at skolene ikke klarer å hjelpe de svakeste. Det kan skyldes økt heterogenitet i skolene, blant annet på grunn av innvandring. Men det kan også skyldes økt segregering (Nilsen & Jensen, 2024); at noen skoler og klasser får flere og flere lavtpresterende elever, mens andre får elever med stadig høyere SØS og prestasjoner.

Et unntak fra trenden med økende ulikheter, er perioden fra 2003 til 2007. I denne perioden var det flere resiliente elever, samtidig som sammenhengen mellom SØS og prestasjoner var svakere på klassenivå. Begge disse resultatene peker på synkende ulikheter i denne perioden. Våre data er ikke egnet til å gi en forklaring på dette, men en mulig forklaring kan være innføringen av Kunnskapsløftet i 2006. Denne reformen kan ha hatt en samlende effekt på skolen: den akademiske optimismen i skolen steg; lærere, skoleledere, og foreldre sammen hadde et økt fokus på læring og det ble satt inn mye ressurser til etterutdanning av lærere (Nilsen & Gustafsson, 2014; Sivesind, 2013). Det kan også skyldes økonomiske svingninger i samfunnet, eller andre faktorer.

TIMSS ble gjennomført rett før pandemien (i 2019) og rett etter (i 2023). Våre resultater viser at sammenhengen mellom SØS og

matematikkprestasjoner økte på elevnivå og muligens også på klassenivå. Samtidig minket andelen resiliente elever i denne perioden. Med andre ord er resultatene sammenfallende og peker på økte ulikheter under pandemien. Våre resultater for denne perioden er sterke og klare. De er også i overensstemmelse med SSB, som viser at ulikhetene økte under pandemien i samfunnet (Aaberge et al., 2021), og med funn fra Progress in International Reading Literacy Study (PIRLS) (Kennedy & Strietholt, 2023) og PISA (Nilsen & Jensen, 2024), som viser at ulikhetene også økte i skolen.

Svakheter ved studien

De to største svakhetene ved denne studien er at TIMSS ikke har data om foreldrenes inntekt, og at TIMSS er en tverrsnittsstudie. En longitudinell studie, derimot, følger de samme elevene over tid, og basert på slike data kan man trekke mer robuste slutninger. Det er ikke mulig å trekke kausale slutninger basert på tverrsnittsdata. Vi kan for eksempel ikke si at flere bøker i hjemmet *fører til* bedre prestasjoner. Antagelig er det én eller flere medierende variabler. Kanskje vil foreldrenes kulturelle kapital og hjemmebakgrunn føre til høyere utdanning som igjen fører til flere bøker hjemme. Det er usikkert hva variabelen «antall bøker hjemme» egentlig måler; det kan være et mål på kulturell kapital, men muligens også økonomiske ressurser.

En annen svakhet er indikatoren «PC», som måler hvorvidt elever har en PC hjemme. Prosentandelen som ikke har PC hjemme er forholdsvis stabil mellom 2003 og 2023 og ligger på mellom 1 og 3 prosent. I 1995 var det imidlertid 34 prosent som ikke hadde PC. I 1995 var den også sterkere korrelert med andre SØS-indikatorer. Dette kan tyde på at PC i større grad målte SØS, og muligens hadde sterkere sammenheng med foreldrenes inntekt, i 1995 enn i senere sykluser av TIMSS. Man skal derfor være mer forsiktig med tolkningen av trender fra 1995 til 2023, enn fra 2003 til 2023.

Denne studien klarer ikke å fange opp alle typer ulikheter over tid. For å få en fullstendig oversikt, hadde det vært ønskelig med en bok som analyserer ulikheter når det gjelder innvandringsbakgrunn, kjønn, mental og fysisk helse, bosted, skolens og lærerernes bidrag til likhet, sosial mobilitet og andre typer ulikheter.

Implikasjoner og veien videre

Til tross for at diskusjoner i media og forskning om økte forskjeller i samfunnet og i skolen startet for flere år siden, har ulikhetene fortsatt å øke i begge arenaer. Og begge arenaer er ansvarlige for å forsøke å redusere ulikhetene. Forskning har vist at skolene kan redusere gapet mellom elever, og at lærerne spiller en viktig rolle i utjevning av ulikheter (Gustafsson et al., 2016; Nilsen & Bergem, 2020; Nilsen & Pettersen, 2024). Imidlertid kan skolene også bidra til å øke gapet mellom elevene, blant annet ved segregering (Atlay et al., 2019; Blömeke, 2024; Cardichon et al., 2020; Glassow et al., 2023). Likevel, selv om skolen skal være med på å utjevne ulikheter i følge opplæringsloven, så klarer ikke skolen dette alene. Det må politiske tiltak til. Å utjevne ulikheter i samfunn og utdanning er svært komplisert, men et samarbeid mellom ulike aktører kan lette byrden. Ekspertgruppen om barn i fattige familier (Rege et al., 2023) har foreslått en rekke tiltak, og det har også ekspertgruppen om betydningen av barnehage, skole og SFO for sosial utjevning og sosial mobilitet (Løken et al., 2024).

Konklusjon

Selv om ikke alle resultater var like klare, indikerer våre funn økte ulikheter mellom elever fra 1995 til 2023. Vi har undersøkt hvorvidt dette mønsteret er likt uavhengig av definisjon på ulikheter (sammenhengen mellom SØS og prestasjoner versus andel resiliente elever), og uavhengig av analysemetode (analyser på ett nivå versus analyser på elev- og klasse-nivå). Videre samsvarer vår forskning med forskning på feltet. Det er nå viktig at mønsteret med økte ulikheter under covid-19-pandemien brytes, og at samfunn og skole jobber for at alle elever skal ha like muligheter til å lykkes.

Forfatterbiografier

Trude Nilsen er seniorforsker ved Institutt for lærerutdanning og skoleforskning og ved Senter for forskning på likhet i utdanning (CREATE) ved Universitetet i Oslo. Hun har deltatt og deltar i internasjonale ekspertgrupper for TIMSS og TALIS. Hun har mange anerkjente internasjonale og nasjonale publikasjoner innen ulikheter i utdanningen, lærernes undervisning og kompetanse, skolemiljø og anvendt metode.

Wangqiong Ye er postdoktor ved Senter for forskning på likhet i utdanning (CREATE) ved Universitetet i Oslo. Hun ser spesielt på hvordan vi kan få til å redusere prestasjonforskjeller mellom ulike grupper elever for å oppnå mer likeverd i skolen. I sine analyser bruker hun data fra storskalaundersøkelser som TIMSS, PIRLS og PISA, og norske registerdata.

Nani Teig er professor ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hennes forskning tar utgangspunkt i naturfagdidaktikk og storskalaundersøkelser, med fokus på utforskende undervisning, vitenskapelig resonnering, undervisningskvalitet og akademisk resiliens.

Referanser

- Agasisti, T., Avvisati, F., Borgonovi, F. & Longobardi, S. (2018). *Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA*. OECD. [https://one.oecd.org/document/EDU/WKP\(2018\)3/En/pdf](https://one.oecd.org/document/EDU/WKP(2018)3/En/pdf)
- Aftenposten. (2025, 23. januar). Ser utlendinger på høylys dag [Leder]. *Aftenposten*. <https://www.aftenposten.no/meninger/leder/i/JbGXnR/ser-utlendinger-paa-hoeylys-dag>
- Atlay, C., Tieben, N., Hillmert, S. & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction*, 63, Artikkel 101211. <https://doi.org/10.1016/j.learninstruc.2019.05.008>
- Blömeke, S. (2024). Equality in content coverage in the Nordic countries? I N. Teig, T. Nilsen & K. Y. Hansen (Red.), *Effective and equitable teacher practice in mathematics and science education: A Nordic perspective across time and groups of students* (s. 231–250). Springer. https://doi.org/10.1007/978-3-031-49580-9_8
- Bourdieu, P. (1986). The forms of capital I J. Richardson (Red.), *Handbook of theory and research for the sociology of education* (s. 241–258). Greenwood.
- Cardichon, J., Darling-Hammond, L., Yang, M., Scott, C., Shields, P. M. & Burns, D. (2020). *Inequitable opportunity to learn: Student access to certified and experienced teachers* Learning Policy Institute. <https://learningpolicyinstitute.org/product/crdc-teacher-access>
- Ertesvåg, F. (2012, 23. januar). Går i klasse med bare innvandrere. Og bare gutter. *Verdens Gang*. <https://www.vg.no/nyheter/i/wa8MM/gaar-i-klasse-med-bare-innvandrere-og-bare-gutter>
- Frønes, T. S., Pettersen, A., Radišić, J. & Buchholtz, N. (2020). *Equity, equality and diversity in the Nordic model of education*. Springer Nature.
- Glassow, L. N., Hansen, K. Y. & Gustafsson, J.-E. (2023). Does socioeconomic sorting of teacher qualifications exacerbate mathematics achievement inequity? Panel data estimates from 20 years of TIMSS. *Studies in Educational Evaluation*, 77, Artikkel 101255. <https://doi.org/10.1016/j.stueduc.2023.101255>
- Glassow, L. N. & Jerrim, J. (2022). Is inequitable teacher sorting on the rise? Cross-national evidence from 20 years of TIMSS. *Large-scale Assessments in Education*, 10, Artikkel 6. <https://doi.org/10.1186/s40536-022-00125-9>
- Grini, K. H., Røv, V. & Konci, E. (2023). *Månedslønn og ulikhet gjennom 25 år. Avstanden har blitt større de siste årene* (SSB-rapport 2023/33). https://www.ssb.no/arbeid-og-lonn/lonn-og-arbeidskraftkostnader/artikler/manedslonn-og-ulikhet-gjennom-25-ar/_attachment/inline/2dacfbcf-767b-4578-93c5-f7d19f35d03d:aa1d3cda9b05c8b90a9dd6f6dacedfd2e300bd0d/RAPP2023-33.pdf

- Gustafsson, J. E., Nilsen, T. & Hansen, K. Y. (2016). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*, 57, 16–30. <https://doi.org/10.1016/j.stueduc.2016.09.004>
- Hansen, K. Y. & Gustafsson, J.-E. (2016). Causes of educational segregation in Sweden – school choice or residential segregation. *Educational Research and Evaluation*, 22(1–2), 23–44. <https://doi.org/10.1080/13803611.2016.1178589>
- Haugen, C. R. (2020). Teachers experiences of school choice from marginalised and privileged public schools in Oslo. *Journal of Education Policy*, 35(1), 68–94. <https://doi.org/10.1080/02680939.2019.1587519>
- Henriksen, E. K., Dillon, J. & Ryder, J. (2015). *Understanding student participation and choice in science and technology education*. Springer.
- Kennedy, A. I. & Strietholt, R. (2023). School closure policies and student reading achievement: Evidence across countries. *Educational Assessment, Evaluation and Accountability*, 35(4), 475–501.
- Løken, K. V., Sandsør, A. M. J., Drugli, M. B., Leiulfstrud, H., Bhuller, M., Størksen, I. & Torvik, F. A. (2024). *Et jevnere utdanningsløp. Barnehage og skole/SFO som innsats mot ulikhet blant barn*. <https://www.regjeringen.no/contentassets/52d541c8ff94f3b84740a0e260c4597/no/pdfs/et-jevnere-utdanningslop.pdf>
- Markussen, S. & Røed, K. (2023). The rising influence of family background on early school performance. *Economics of Education Review*, 97, Artikkel 102491. <https://doi.org/10.1016/j.econedurev.2023.102491>
- Mittal, O., Nilsen, T. & Björnsson, J. K. (2020). Measuring equity across the Nordic education systems– conceptual and methodological choices as implications for educational policies. I T. S. Frones, A. Pettersen, J. Radišić & N. Buchholtz (Red.), *Equity, equality and diversity in the Nordic model of education* (s. 43–71). Springer. https://doi.org/10.1007/978-3-030-61648-9_3.
- Nilsen, T. & Bergem, O. K. (2016). Hjemmebakgrunn. En likeverdig skole. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015*. Universitetsforlaget.
- Nilsen, T. & Bergem, O. K. (2020). Teacher competence and equity in the Nordic countries. Mediation and moderation of the relation between SES and achievement. *Acta Didactica Norden*, 14(1). <https://doi.org/10.5617/adno.7946>
- Nilsen, T., Björnsson, J. K. & Olsen, R. V. (2018). Hvordan har likeverd i norsk skole endret seg de siste 20årene? I J. K. Björnsson & R. V. Olsen (Red.), *Tjue år med TIMSS og PISA i Norge* (s. 150–172). Universitetsforlaget. <https://doi.org/10.18261/9788215030067-2018-08>
- Nilsen, T. & Gustafsson, J.-E. (2014). School emphasis on academic success: Exploring changes in science performance in Norway between 2007 and 2011 employing two-level SEM. *Educational Research and Evaluation*, 20(4), 308–327. <https://doi.org/10.1080/13803611.2014.941371>
- Nilsen, T. & Jensen, F. (2024). Like muligheter til matematisk kompetanse? I A. Pettersen & F. Jensen (Red.), *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (s. 167–185). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.222>
- Nilsen, T. & Pettersen, A. (2024). Matematikkundervisning i norske og nordiske klasserom – klasseledelse, støttende undervisning og faglige utfordringer. I A. Pettersen & F. Jensen (Red.), *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (s. 139–165). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.222>
- Nilsen, T., Scherer, R., Gustafsson, J.-E., Teig, N. & Kaarstein, H. (2020). Teachers’ role in enhancing equity: A multilevel structural equation modeling with mediated moderation. I T. S. Frones, A. Pettersen, J. Radišić & N. Buchholtz (Red.), *Equity, equality and diversity in the Nordic model of education* (s. 173–196). Springer. https://doi.org/10.1007/978-3-030-61648-9_7

- Nilsen, T., Senden, B., Ye, W., Jentsch, A., Teig, N. & König, J. (2025). Mathematics teaching quality in classrooms of different compositions in Norway. *ZDM–Mathematics Education*, 1–15. <https://doi.org/10.1007/s11858-025-01722-y>
- OECD. (2013). *PISA 2012 results: Vol 2. Excellence through equity: Giving every student the chance to succeed*. PISA, OECD Publishing.
- OECD. (2023a). *PISA 2022 results: Vol. 1. The state of learning and equity in education*. PISA, OECD Publishing.
- OECD. (2023b). *PISA 2022 results: Vol. 1. The state of learning and equity in education*. PISA, OECD Publishing. <https://doi.org/https://doi.org/10.1787/53f23881-en>
- OECD. (2024). *Society at a glance 2024: OECD social indicators*. OECD.
- Rege, M., Drange, N. E., Salvanes, K. G., Bøe, T., Zachrisson, H. D., Glemmestad, H. S., Hugvik, K. & Fløtten, T. (2023). *En barndom for livet: Økt tilhørighet, mestring og læring for barn i fattige familier*. Barne- og familiedepartementet. <https://www.regjeringen.no/no/dokumenter/en-ny-barndom-for-livet/id3000835/>
- Rosén, M. & Nilsen, T. (2024). Are changes in content coverage related to changes in achievement over time? I N. Teig, T. Nilsen & K. Y. Hansen (Red.), *Effective and equitable teacher practice in mathematics and science education: A Nordic perspective across time and groups of students* (s. 187–205). Springer. https://doi.org/10.1007/978-3-031-49580-9_6
- Sandsør, A. M. J., Zachrisson, H. D., Karoly, L. A. & Dearing, E. (2023). The widening achievement gap between rich and poor in a Nordic country. *Educational Researcher*, 52(4), 195–205. <https://doi.org/10.3102/0013189X221142596>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Sivesind, K. (2013). Læreplanene i Kunnskapsløftet: Et internasjonalt sammenliknende perspektiv. *Norsk pedagogisk tidsskrift*, 97(6), 370–387. <https://doi.org/10.18261/ISSN1504-2987-2007-05-05>
- Teig, N. (2024, 5.–7. juni). What role do teachers play in fostering academic resilience? Global insight across 58 countries. I S. Pétursdóttir (Red.), *Revisioning STEAM education in times of climate change. Proceedings of the 14th Nordic Research Symposium on Science Education* (s. 272–276). NFSUN.
- UNESCO. (2018). *Handbook on measuring equity in education*. UNESCO Institute for Statistics.
- Wilkinson, R. G. & Pickett, K. (2009). *The spirit level: Why more equal societies almost always do better*. Allen Lane London.
- Ye, W., Teig, N. & Blömeke, S. (2024). Systematic review of protective factors related to academic resilience in children and adolescents: unpacking the interplay of operationalization, data, and research method. *Frontiers in Psychology*, 15, Artikkel 1405786. <https://doi.org/10.3389/fpsyg.2024.1405786>
- Aaberge, R., Mogstad, M., Vestad, O. L. & Vestre, A. (2021). *Økonomisk ulikhet i Norge i det 21. århundre* (SSB-rapport 2021/33). <https://www.ssb.no/inntekt-og-forbruk/inntekt-og-formue/artikler/okonomisk-ulikhet-i-norge-i-det-21.arhundre>

Vedlegg

Vedlegg 2.1. Faktorladninger for SØS

Standardiserte faktor-ladninger					
Årstall	Trinn	Indikator	Faktor-ladning	standardfeil	P-verdi
1995	8.trinn	Bøker	0,622	0,032	0
		Utdanning mor	0,426	0,036	0
		Utdanning far	0,403	0,036	0
		PC	0,339	0,028	0
		Eget skrivebord	0,225	0,032	0
2003	8.trinn	Bøker	0,576	0,027	0
		Utdanning mor	0,52	0,031	0
		Utdanning far	0,456	0,032	0
		PC	0,231	0,031	0
		Eget skrivebord	0,214	0,031	0
2007	8.trinn	Bøker	0,602	0,03	0
		Utdanning mor	0,497	0,031	0
		Utdanning far	0,434	0,031	0
		PC	0,114	0,022	0
		Eget skrivebord	0,161	0,023	0
2011	8.trinn	Bøker	0,725	0,029	0
		Utdanning mor	0,493	0,028	0
		Utdanning far	0,477	0,031	0
		PC	0,064	0,026	0,014
		Eget skrivebord	0,143	0,023	0
2015	8.trinn	Bøker	0,66	0,025	0
		Utdanning mor	0,544	0,024	0
		Utdanning far	0,515	0,026	0
		PC	0,038	0,022	0,082
		Eget skrivebord	0,202	0,022	0
2015	9.trinn	Bøker	0,702	0,025	0
		Utdanning mor	0,51	0,024	0
		Utdanning far	0,461	0,027	0
		PC	0,057	0,023	0,013
		Eget skrivebord	0,206	0,019	0
2019	9.trinn	Bøker	0,712	0,029	0
		Utdanning mor	0,511	0,025	0
		Utdanning far	0,433	0,028	0
		PC	0,072	0,022	0,001
		Eget skrivebord	0,106	0,021	0

Standardiserte faktor-ladninger					
Årstall	Trinn	Indikator	Faktor-ladning	standardfeil	P-verdi
2023	9.trinn	Bøker	0,577	0,023	0
		Utdanning mor	0,493	0,023	0
		Utdanning far	0,499	0,025	0
		PC	0,13	0,019	0
		Eget skrivebord	0,166	0,019	0

Vedlegg 2.2. Modeltilpasning. Analyser på ett nivå

Modeltilpasning. SES om latent variabel. SEM på ett nivå					
Årstall	Trinn	CFI	TLI	RMSEA	SRMR
1995	8.trinn	0,988	0,975	0,027	0,019
2003	8.trinn	0,998	0,997	0,007	0,01
2007	8.trinn	0,99	0,978	0,018	0,013
2011	8.trinn	0,993	0,985	0,016	0,013
2015	8.trinn	0,989	0,976	0,025	0,017
2015	9.trinn	0,997	0,993	0,014	0,01
2019	9.trinn	1	1,003	0,001	0,008
2023	9.trinn	0,98	0,957	0,033	0,024

Vedlegg 2.3. Modeltilpasning. Analyser på to nivåer

Modeltilpasning. SES om latent variabel. SEM på to nivåer						
Årstall	Trinn	CFI	TLI	RMSEA	SRMR_W	SRMR_B
1995	8.trinn	0,983	0,963	0,024	0,02	0,09
2003	8.trinn	0,988	0,973	0,018	0,012	0,085
2007	8.trinn	0,987	0,971	0,017	0,016	0,081
2011	8.trinn	0,991	0,98	0,015	0,015	0,062
2015	8.trinn	0,98	0,957	0,023	0,021	0,102
2015	9.trinn	0,995	0,989	0,012	0,012	0,082
2019	9.trinn	0,999	0,999	0,003	0,009	0,073
2023	9.trinn	0,97	0,936	0,031	0,028	0,092

Vedlegg 2.4. Målingsinvarians

Resultater fra tradisjonell målingsinvarians gir metrisk invarians. Her var $\Delta cfi = 0.016$ og $\Delta rmsea = 0.014$. Med tradisjonell målingsinvarians menes: 1) konfigurasjonsinvarians med samme struktur, 2) metrisk invarians med samme faktorladninger, og 3) skalarinvarians med samme skjæringspunkt («intercept»).

Imidlertid fungerer ikke denne metoden like bra for kategoriske data (Svetina et al., 2020) som vi har. Derfor gjorde vi også målingsinvarians med metoden for kategoriske data og fulgte den anerkjente metoden til Svetina og kollegaer (2020). Denne metoden følger følgende kriterier: 1) konfigurasjonsinvarians med samme struktur, 2) metrisk invarians hvor man kontrollerer for skjæringspunkt før faktorladninger, og 3) skalarinvarians med samme faktorladninger. Med denne metoden fant vi skalarinvarians med $\Delta CFI < 0.01$ og $\Delta RMSEA < 0.015$ som vist i tabellen:

	Kji-kvadrat skalert	skalert frihetsgrader	P-verdi skalert	RMSEA skalert	CFI skalert	TLI skalert
Konfigurasjons- invarians	417,344	40	0	0,067	0,979	0,958
Metrisk invarians	509,115	54	0	0,064	0,975	0,962
Skalarinvarians	646,246	82	0	0,058	0,969	0,969

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using M plus and the lavaan/semtools packages. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(1), 111–130.

KAPITTEL 3

Fra TIMSS 2015 til TIMSS 2023: et skolemiljø i endring

Hege Kaarstein, Anne-Catherine Lehre og Mark White

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: Skolemiljøet i Norge er i endring. Kortrapporten for TIMSS 2023 viste at andel elever og lærere som rapporterer om et veldig trygt skolemiljø har blitt redusert fra 2015 til 2023. I dette kapittelet undersøkes den delen av begrepet skolemiljø som er knyttet til trygghet, trivsel og tilfredshet nærmere. Det overordnede bildet av resultatene av undersøkelsen viser at skolemiljøet fremdeles er relativt trygt. Elevene rapporterer imidlertid om en økende tendens til mobbing, og lærerne rapporterer om dårligere oppførsel blant elevene i 2023 enn i 2015. Selv om rektorene rapporterer om økende grad av problemer knyttet til hærverk og trusler både blant elevene og overfor skolens ansatte, blir problemene ikke karakterisert som mer enn moderate i 2023.¹

Nøkkelord: skolemiljø, trygghet, elever, lærere, rektorer, TIMSS

Abstract

From TIMSS 2015 to TIMSS 2023: A Changing School Climate

The school environment in Norway is changing. A new TIMSS report indicates that the proportion of students and teachers reporting a very safe school environment has decreased from 2015 to 2023. This chapter explores the aspect of the school environment related to safety, well-being, and satisfaction. The overall findings of the survey suggest that the school environment remains relatively safe; however,

1 KI er benyttet i oversettelse fra norsk til engelsk og til forbedring av språket.

students report an increasing tendency towards bullying, and teachers report a decline in student behavior in 2023 compared to 2015. Although school principals report an increasing prevalence of issues related to vandalism and threats both among students and towards school staff, these problems are categorized as no more than moderate in severity in 2023.

Keywords: school climate, safety, students, teachers, principals, TIMSS

Innledning

I opplæringslova heter det at «alle elever har rett til eit trygt og godt skolemiljø som fremjar helse, inkludering, trivsel og læring» (Opplæringslova, 2024, § 12–2), og et av formålene til arbeidsmiljøloven er «å sikre et arbeidsmiljø som gir grunnlag for en helsefremmende og meningsfylt arbeidssituasjon, som gir full trygghet mot fysiske og psykiske skadevirkninger» (Arbeidsmiljøloven, 2006, § 1–1). Skolene er med andre ord pålagt å sørge for et godt skolemiljø for alle elever og ansatte.

Skolemiljø er et begrep som omfatter relativt mange aspekter. I forskningslitteraturen defineres skolemiljø noe ulikt i ulike settinger, men ofte inkluderer definisjonen komplekse faktorer som læringsmiljø, skoleledelse, orden, trygghet, fravær av mobbing og elevers følelse av skoletilhørighet (se f.eks. Grazia & Molinari, 2021; Nilsen & Teig, 2022; Wang & Degol, 2016). Ifølge Kutsyuruba et al. blir skolemiljø «ofte forbundet med og sett i perspektiv av trygghet på skolen» (2015, s. 107, vår oversettelse).

Tidligere forskning har vist at et trygt skolemiljø har sammenheng med høye prestasjoner både internasjonalt (se f.eks. Erdem & Kaya, 2024; Mullis et al., 2020) og i de nordiske landene (se f.eks. Scherer & Nilsen, 2016). Videre kan et skolemiljø med lite mobbing, god orden og disiplin og en følelse av trygghet blant både elever og lærere fremme elevers læring og motivasjon (Nilsen & Kaarstein, 2021; Rutkowski & Rutkowski, 2016; Wang & Degol, 2016). I et godt og trygt skolemiljø vil både lærere og elever trives, og ikke overraskende er det en sammenheng mellom lærernes og elevenes trivsel også (Harding et al., 2019).

I undersøkelsen *Trends in International Mathematics and Science Study* (TIMSS) 2015, rapporterte de norske ungdomsskolelærerne om det tryggeste skolemiljøet blant alle deltagerlandene (Martin et al., 2016). Norske barne- og skoleelever hadde også lærere som rapporterte om et veldig godt skolemiljø, helt i toppen av målingene over deltagerlandenes skolemiljø. Videre, med bruk av data fra både TIMSS og PISA², fant Hatlevik et al. (2018) at norske rektorer i tidsrommet fra 2003 til 2015 rapporterte om en positiv utvikling i skoleklimaet. Elevundersøkelsen i 2015 viste den samme tendensen, med lav frekvens av mobbing i norsk skole (Wendelborg et al., 2016), og fra TALIS³-undersøkelsene kom det rapporter om at «norske ungdomsskoler

2 PISA: Programme for International Student Assessment, se pisa.no eller <https://www.oecd.org/en/about/programmes/pisa.html>

3 TALIS: Teaching and Learning International Survey, se <https://www.iea.nl/studies/additional/TALIS>

kjennetegnes av gode personlige relasjoner» (Throndsen et al., 2019, s. 25). Noe av dette kan kanskje tilskrives store satsinger på å fremme et godt skolemiljø i Norge (Meld. St. 21 (2016–2017)). Ett resultat av satsingene er endringen i kapittel 9A i opplæringsloven som ble innført i 2017 (Endringslov til opplæringslova og friskolelova, 2017). Den nå omdiskuterte endringen førte til strengere krav til skolenes arbeid mot mobbing og sterkere makt til Fylkesmannen (nå Statsforvalteren) til å gripe inn dersom skolene ikke oppfyller sine forpliktelser. Diskusjonen handler i stor grad om hvor godt lovendringen fungerer og hvor komplisert det er å følge opp skolemiljøsaker (se f.eks. Skutlaberg, 2023; Utdanningsdirektoratet, 2023b; Walker, 2023). I tillegg ga Utdanningsdirektoratet støtte til skoler som ønsket å implementere skolemiljøprogrammer, som for eksempel Olweus-programmet (Eriksen et al., 2014; Olweus & Limber, 2010). Olweus-programmet er forskningsbasert og tar for seg alle nivåer i skolen (elever, foresatte, lærere og øvrig personale på skolen) med mål om å redusere og forebygge problemer ved å øke bevisstheten og kunnskapen om mobbing (Olweus & Limber, 2019).

Siden 2015, har det trygge skolemiljøet som ble beskrevet i Norge blitt gradvis redusert (se f.eks. Kaarstein et al., 2024; Nilsen & Kaarstein, 2021). Andelen elever som rapporterer om mobbing har økt (Wendelborg et al., 2020). *Ungdata 2024*, en undersøkelse blant et representativt utvalg av 150 600 norske ungdommer om hvordan det er å være ung, viser at mobbing er et økende problem, at andelen som er utsatt for vold fra jevnaldrende er økende og at elevene trives dårligere og dårligere på skolen (Bakken, 2024). Videre peker PISA 2022 på en tydelig sammenheng mellom elevenes faglige prestasjoner og deres rapportering av mobbing (Jensen et al., 2023). Tidligere analyser av TIMSS-data har vist at reduksjonen i andelen lærere og elever som oppfatter skolemiljøet sitt som veldig trygt er med på å forklare noe av nedgangen i elevenes prestasjoner i matematikk og naturfag fra 2015 til 2019 på 9. trinn (Nilsen et al., 2022; Nilsen & Kaarstein, 2021).

Denne reduksjonen av det gode og trygge skolemiljøet i Norge de senere årene, har medført flere medieoppslag (se f.eks. Frengstad, 2023; Olafsen, 2022), politisk engasjement og offentlige diskusjoner (f.eks. Arendalsuka, 2017) om hva som bør gjøres for å hindre ytterligere nedgang. Målet for dette kapittelet er derfor å se nærmere på skolemiljø knyttet til trygghet blant elever, lærere og rektorer ved å stille følgende forskningsspørsmål:

- 1 Hvordan har elevenes, lærernes og rektorenes oppfatning av trygghet utviklet seg over tid?

- 2 Hva kjennetegner trygge skoler?
- 3 Hva kjennetegner skoler som fortsatt oppleves som trygge til tross for de observerte endringene i skolemiljøet?

Forskningsspørsmålene belyses ved å bruke data fra TIMSS-undersøkelsen på 5. og 9. trinn. I TIMSS-undersøkelsen får et representativt utvalg elever, deres lærere og rektorer spørsmål knyttet til blant annet skolemiljøet. TIMSS gjennomføres hvert fjerde år, noe som gir mulighet til å se på utvikling over tid. For mer informasjon om TIMSS-undersøkelsen, se introduksjonkapittelet i denne boken.

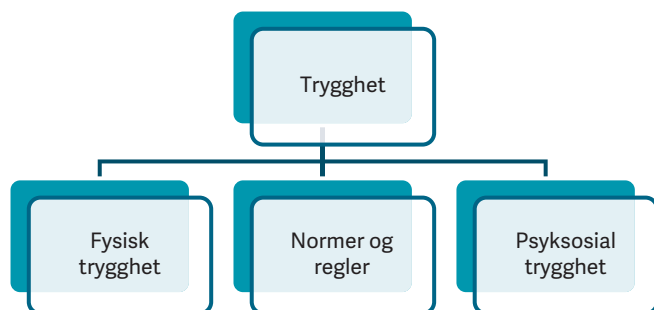
Bakgrunn

I poetiske ordelag forteller Marraccini et al. (2020) at *skolemiljø* har blitt beskrevet som skolens «hjerter og sjel» fordi det handler om hvordan skolelivet oppleves både i kvalitet og karakter. Begrepet sies videre å inneholde flere dimensjoner (se f.eks. Grazia & Molinari, 2021; Nilsen & Teig, 2022; Wang & Degol, 2016). I tillegg til at det handler om mellommenneskelige interaksjoner som påvirker elevenes kognitive, sosiale og psykososiale utvikling, handler det også om organisatoriske strukturer og ressurser (Bradshaw et al., 2021). I en ofte benyttet modell, presentert av Wang og Degol (2016), deles skolemiljø inn i fire dimensjoner som handler om (1) skolens fokus på læring (i skoleledelse, undervisning og etter- og videreutdanningsaktivitet), (2) god samhandling mellom alle «medlemmene» av skolesamfunnet, hvor samhandling inkluderer god kommunikasjon, følelse av tilhørighet, respekt for mangfold og samarbeid med nærmiljøet og hjemmene, (3) skolens fysiske rammer og ressurser, og (4) både fysisk og psykisk trygghet – ofte i form av fravær av vold og mobbing/trusler (Bradshaw et al., 2021) – og trygghet i form av «ro og orden».

Mens modellens første dimensjon inkluderes i denne antologiens kapittel om skolens og rektors betydning for elevenes prestasjoner (se Veletić, Radišić & Jacobsen, 2025), ser vi i dette kapittelet spesielt på trygghet i skolehverdagen som inngår i modellens fjerde dimensjon. Tidligere forskning har funnet positive sammenhenger mellom elevenes opplevelse av trygghet og trivsel og læringsutbytte (Grazia & Molinari, 2021; Nilsen & Kaarstein, 2021; Nilsen & Teig, 2022; Wang & Degol, 2016). I tillegg rapporteres det om en sammenheng mellom lærernes trygghet og tilfredshet med jobben (se f.eks. Jentsch et al., 2023; Toropova et al., 2021), som igjen

henger sammen med deres undervisningskvalitet (Harrison et al., 2023). Et trygt skolemiljø er med andre ord, og ikke overraskende, essensielt for elevenes læringsutbytte (Kutsyuruba et al., 2015; Wang & Degol, 2016).

Forskningsoppsummeringen til Lewno-Dumdie et al. (2020) benyttes her for å gi et raskt og enkelt innblikk i faktorer som kan inngå i dimensjonen som handler om trygghet. Se figur 3.1 for en forenklet versjon av deres oversikt.



Figur 3.1. Forenklet skisse av oversikt over innholdet i dimensjonen trygghet presentert i Lewno-Dumdie et al. (2020).

Oversikten bygger på tidligere studier som har inkludert ulike faktorer knyttet til det Lewno-Dumdie et al. overordnet har valgt å kalle trygghet. I deres oversikt (forenklet i figur 3.1) er trygghet videre inndelt i tre underkategorier. Den første kategorien handler om fysisk trygghet. På en skole hvor det fysiske aspektet av trygghet er bra, er det for eksempel lite mobbing og aggresjon. I den andre kategorien, som handler om trygghet på det psykososiale planet, inngår blant annet en tro på at klare regler for oppførsel fungerer, og at skolen derfor oppfattes som trygg. Den siste kategorien handler om at de normene og (de klare) reglene som er satt håndheves konsistent slik at skolehverdagen oppleves som trygg. Overordnet sett handler trygghet om samhandling, relasjoner og samhold mellom alle som, på en eller annen måte, er involvert i skolehverdagen (elever, foreldre, lærere og skolens ledelse) og at det er ro og orden på skolen. Elever og lærere ved en skole hvor samhandlingen er og oppfattes som god, både trives bedre og føler oftere sterkere tilhørighet til skolen enn ved skoler hvor samhandlingen ikke er så god.

Ifølge tidligere forskning kan for eksempel kommunikasjonsproblemer vanskeliggjøre god samhandling og føre til utrygghet (Kutsyuruba et al.,

2015). Problemene kan blant annet oppstå fordi elevene har ulik kulturell og/eller språklig bakgrunn (Peterson et al., 2024; Walsh et al., 2016). Da Walsh et al. (2016) undersøkte sammenhengen mellom andel innvandrere ved barne- og ungdomsskoler i 11 land (deriblant Danmark, Nederland og Storbritannia) og fysisk vold blant elevene, fant de at en høyere andel innvandrere var relatert til oftere fysisk vold og mobbing blant både innvandrere og ikke-innvandrere. I en norsk studie knyttet til den språklige bakgrunnen til elever på barnetrinnet, fant Peterson et al. (2024) ikke bare at de språklige utfordringene kunne bidra til mindre positive interaksjoner, men også at de språklige utfordringene spilte en større rolle for det sosiale klimaet i klassene enn elevenes sosioøkonomiske status.

Kommunikasjonsproblemer og følelse av utrygghet kan videre føre til et lavere engasjement for skole og utdanning og dermed også øke muligheten for å droppe ut av skolen, eller ikke fullføre høyere studier (Archambault et al., 2022; Ripamonti, 2018). Fordi høyt frafall fra skole og høyere studier har vært og fremdeles er et problem i Norge, er det et politisk mål at ni av ti elever gjennomfører og består videregående opplæring i 2030 (Meld. St. 21 (2020–2021)). Uten videregående har det blitt vanskeligere og vanskeligere å komme inn i arbeidsmarkedet, fordi videregående opplæring gir ungdommer og unge voksne flere muligheter ved valg av høyere studier og det å lære seg å lære gir mulighet til å tilpasse seg «raskt skiftende behov i samfunnet» (Meld. St. 21 (2020–2021), s. 7).

Metode

For å svare på forskningsspørsmålene i dette kapittelet, benyttes norske data hentet fra spørreskjemaene i TIMSS. Målgruppene for spørreskjemaene er elever på 5. og 9. trinn, samt deres matematikk- og naturfaglærere og rektorer. På 5. trinn innhentes det også informasjon fra elevenes foresatte i TIMSS-undersøkelsen, men disse dataene benyttes ikke her. I dette kapittelet hentes data for elever, lærere og rektorer fra undersøkelsene som ble gjennomført i 2015, 2019 og 2023.⁴ Se tabell 3.1 for antall elever, lærere og rektorer som svarte i hvert år.

4 Selv om norske elever har deltatt i TIMSS-undersøkelsen helt siden 1995, ble målgruppene endret i Norge fra 4. til 5. trinn og fra 8. til 9. trinn i 2015 (Bergem et al., 2016). Trenddata for elevene på 5. og 9. trinn kan derfor bare hentes fra gjennomføringene i 2015, 2019 og 2023.

Tabell 3.1. Antall respondenter i målgruppene per trinn per TIMSS-undersøkelse

	2015	2019	2023
5. trinn			
Elever	4 329	3 951	5 301
Lærere	280	326	395
Rektorer	140	150	160
9. trinn			
Elever	4 697	4 575	6 324
Lærere	300	346	430
Rektorer	143	157	157

For mer informasjon om TIMSS-undersøkelsen, se bokens introduksjonskapittel (Kaarstein & Lehre, 2025). Data fra Norge og alle andre deltakerland i TIMSS er offentlig tilgjengelige og kan lastes ned fra <https://timssandpirls.bc.edu/>.

I rammeverket som ligger til grunn for spørreskjemaene, beskrives de konseptene som inkluderes i TIMSS-undersøkelsen (Reynolds et al., 2021). Konseptene er valgt på grunnlag av deres kjente positive sammenheng med elevenes prestasjoner i matematikk og naturfag (Reynolds et al., 2021). De konseptene i TIMSS-spørreskjemaene som kan knyttes til trygghet handler om tilhørighet, trivsel, mobbing, regler og orden. For elevene handler trygghet i TIMSS i stor grad om fravær av mobbing, men de blir også spurt om de trives på skolen og har en følelse av å høre til. For lærerne inkluderes trygghet og, på samme måte som for rektorene, deres oppfatning av skolens evne til å opprettholde ro og orden.

Spørreskjemaene i TIMSS-undersøkelsen er i utgangspunktet designet for å være så konsistente som mulig fra en undersøkelse til den neste, men de må likevel oppdateres for å reflektere samfunnsendringer (Reynolds et al., 2021). Dette resulterer noen ganger i at enkelte variabler eller utsagn fjernes og nye legges til. Når man ønsker å studere endringer over tid, er det avgjørende at utsagnene forblir de samme fra gang til gang. Derfor inkluderes bare de utsagnene fra TIMSS 2023 som er mulig å sammenligne med tilsvarende utsagn inkludert TIMSS 2019 og/eller TIMSS 2015 i analysene til dette kapittelet (se tabell 3.2). Tabellen angir svaralternativer og utsagn knyttet til hvert spørsmål, målgruppe for spørsmålet samt hvilket trinn spørsmålet er inkludert for.

Tabell 3.2. Spørsmål og utsagn knyttet til trygghet slik det er målt i TIMSS

Målgruppe ¹	Spørsmål	Svaralternativer ²	Utsagn ²	Trinn
Elever	I løpet av dette skoleåret, hvor ofte har andre elever fra skolen din gjort noe av dette mot deg personlig, gjennom meldinger eller på sosiale medier?	Minst én gang i uka En eller to ganger i md. Noen få ganger i året Aldri	Gjort narr av meg eller kalt meg stygge ting	5
			Ikke latt meg få være med på leker eller aktiviteter – « Ekskludert .»	5
			Slått meg eller skadet meg (f.eks. skubbet veldig hardt eller sparket)	5
			Spredd løgner om meg	5 og 9
			Stjålet noe fra meg	5 og 9
			Med vilje ødelagt noe som var mitt	5 og 9
			Presset meg til å gjøre ting jeg ikke ønsket	5 og 9
			Sendt meg stygge eller sårende meldinger	5 og 9
			Delt stygge eller sårende ting om meg på internett	5 og 9
			Delt bilder av meg på internett som har gjort meg flau	5 og 9
			Truet meg	5 og 9
			Holdt meg utenfor (f.eks. fra fester, gruppesamtaler) – « Ekskludert .»	9
			Skadet meg fysisk	9
			Sagt stygge ting om utseendet mitt (f.eks. håret mitt, størrelsen min)	9
			Latt være å snakke med meg	9
TRYGGHET som fravær av mobbing	Hva synes du om skolen din? Fortell hvor enig du er i disse utsagnene. TRYGGHET, trivsel og skoletilhørighet	Svært enig Litt enig Litt uenig Svært uenig	Sagt sårende ting til meg eller om meg på grunn av min kulturelle bakgrunn (f.eks. hudfarge, etnisitet og religion)	9
			Delt mine hemmeligheter med andre	9
			Jeg føler meg trygg på skolen	5 og 9
			Jeg liker å være på skolen	5 og 9
			Jeg føler at jeg hører til på denne skolen	5 og 9
			(Forts.)	

Tabell 3.2. (Forts.)

Målgruppe ¹	Spørsmål	Svaralternativer ²	Utsagn ²	Trinn
Lærere	Hvor enig er du i de følgende utsagnene om den skolen du jobber på?	Svært enig Litt enig Litt uenig Svært uenig	Jeg føler meg trygg på denne skolen Skolens plan og praksis for sikkerhet er god nok Elevene oppfører seg ordentlig Elevene har respekt for lærerne Elevene respekterer skolens eiendom Skolen har klare regler for elevenes oppførsel Skolens regler opprettholdes på en rettferdig og konsistent måte	5 og 9 5 og 9 5 og 9 5 og 9 5 og 9 5 og 9 5 og 9
	TRYGGHET, ro og orden			
			Hærverk	5 og 9
			Tyveri	5 og 9
Rektorer	I hvilken grad er det følgende et problem blant elevene på 5./9. trinn?	Ikke et problem Et mindre problem Et moderat problem Et alvorlig problem	Trusler eller skjellsord blant elevene (inkludert SMS, e-post, sosiale medier osv.) Fysiske konflikter blant elevene Trusler eller skjellsord overfor lærere eller andre ansatte (inkludert SMS, e-post, sosiale medier osv.) Fysiske skader påført lærere eller andre ansatte	5 og 9 5 og 9 5 og 9 5 og 9
	TRYGGHET som ro og orden			
				9

Note. ¹ Kolonnen viser hvem som har besvart spørsmålene. ² De uthevede ordene benyttes som forkortelser når resultatene fra analysene vises i figurene 3.2 til 3.5.

Analyser

For å svare på det første forskningsspørsmålet om hvordan elevenes, lærernes og rektorenes oppfatning av trygghet har utviklet seg over tid, ble svaralternativene til utsagnene i tabell 3.2 tilordnet stigende verdier. Eksempelvis ble svaralternativene til utsagnet «Jeg føler meg trygg på denne skolen» gitt verdier som følger: *Svært uenig* = 0, *Litt uenig* = 1, *Litt enig* = 2 og *Svært enig* = 3. For utsagnene som rektorene svarte på, for eksempel «Fysiske konflikter blant elevene», fikk svaralternativene følgende verdier: *Ikke et problem* = 0, *Et mindre problem* = 1, *Et moderat problem* = 2, *Et alvorlig problem* = 3. For utsagnene elevene fikk om trygghet (som fravær av mobbing), ble svaralternativene gitt følgende verdier: *Aldri* = 0, *Noen få ganger i året* = 1, *En eller to ganger i måneden* = 2, *Minst én gang i uka* = 3. For hvert utsagn er det beregnet et gjennomsnitt av alle respondentenes valg med tilhørende 95 prosent-konfidensintervall.

Alle analyser er utført i analyseprogrammet R (R Core Team, 2024). Som anbefalt av van Buuren og Groothuis-Oudshoorn (2011), ble manglende data håndtert ved å bruke R-pakken *MICE* til «multiple imputation» før analysene ble gjennomført. For å ta hensyn til det komplekse designet til TIMSS, der elevene blir vektet fordi de er en del av en klasse som igjen er en del av en skole (se f.eks. Siegel & Foy, 2024), ble *survey*-pakken i R benyttet for å ta hånd om elevenes vektor (Kim et al., 2006). Det er det vektete gjennomsnittet for hvert utsagn i tabell 3.2 som presenteres i resultatkapittelet som svar på forskningsspørsmål 1.

For å svare på forskningsspørsmålet om hva som kjennetegner trygge skoler, ble det gjennomført regresjonsanalyser (Lumley, 2011). For å redusere antall regresjonsanalyser, ble et vektet gjennomsnitt for alle utsagn tilhørende spørsmålene i tabell 3.2 beregnet og benyttet som avhengig variabel: Til spørsmålet om hvor ofte elevene har opplevd mobbing hører det med elleve utsagn på 5. trinn og fjorten på 9. trinn (variabelnavn: *Elevrapportert fravær av mobbing*). Til spørsmålet om hva elevene synes om skolen sin hører det med tre utsagn på begge trinn (variabelnavn: *Elevrapportert tilhørighet*). Til lærernes spørsmål om trygghet, ro og orden på skolen hører det med syv utsagn på begge trinn (variabelnavn: *Lærer trygghet, ro og orden*), og når rektorene blir bedt om å angi grad av problemer ved skolen, blir de bedt om å ta stilling til fem utsagn på 5. trinn og seks på 9. trinn (variabelnavn: *Rektor trygghet som ro og orden*). Dette førte til åtte regresjonsanalyser totalt for de avhengige variablene, fire på hvert trinn.

De uavhengige variablene i regresjonsanalysene hentes fra spørreskjemaene til rektorene og elevene. Fra spørreskjemaet til rektorene hentes det ut informasjon om elevenes sosioøkonomiske status (SØS)⁵ ved å spørre etter hvor stor andel av skolens elever som kommer fra økonomisk vanskeligstilte og økonomisk velstående hjem. I tillegg benyttes andel elever som har norsk som morsmål og rektorenes beskrivelse av skolens beliggenhet. Fra spørreskjemaet til elevene hentes det ut informasjon om elevenes innvandrersstatus. Elever som ikke er født i Norge og har to foreldre som er ikke født i Norge kalles her førstegenerasjonsinnvandrere, mens elever som er født i Norge av utenlandske foreldre kalles her andregenerasjonsinnvandrere. I analysene settes TIMSS 2015 som referanseår, mens TIMSS 2019 og TIMSS 2023 dummy-kodes.

For å svare på det siste forskningsspørsmålet om hva som kjennetegner skoler som fortsatt oppleves som trygge til tross for de observerte endringene i skolemiljøet, ble regresjonsanalysene beskrevet over gjennomført på nytt med interaksjoner mellom gjennomføringsår for TIMSS og kjennetegn ved skoler og elever (f.eks. skolens beliggenhet og andel elever med lav SØS) i tillegg. Dersom interaksjonene er statistisk signifikante, tyder det på at skoler med visse kjennetegn (som f.eks. beliggenhet og stor andel elever med lav SØS) har større eller mindre endringer i trygghet enn de gruppene som settes som referanse. Alle interaksjonene ble testet én om gangen for å unngå multikollinearitet (Gelman & Hill, 2006).

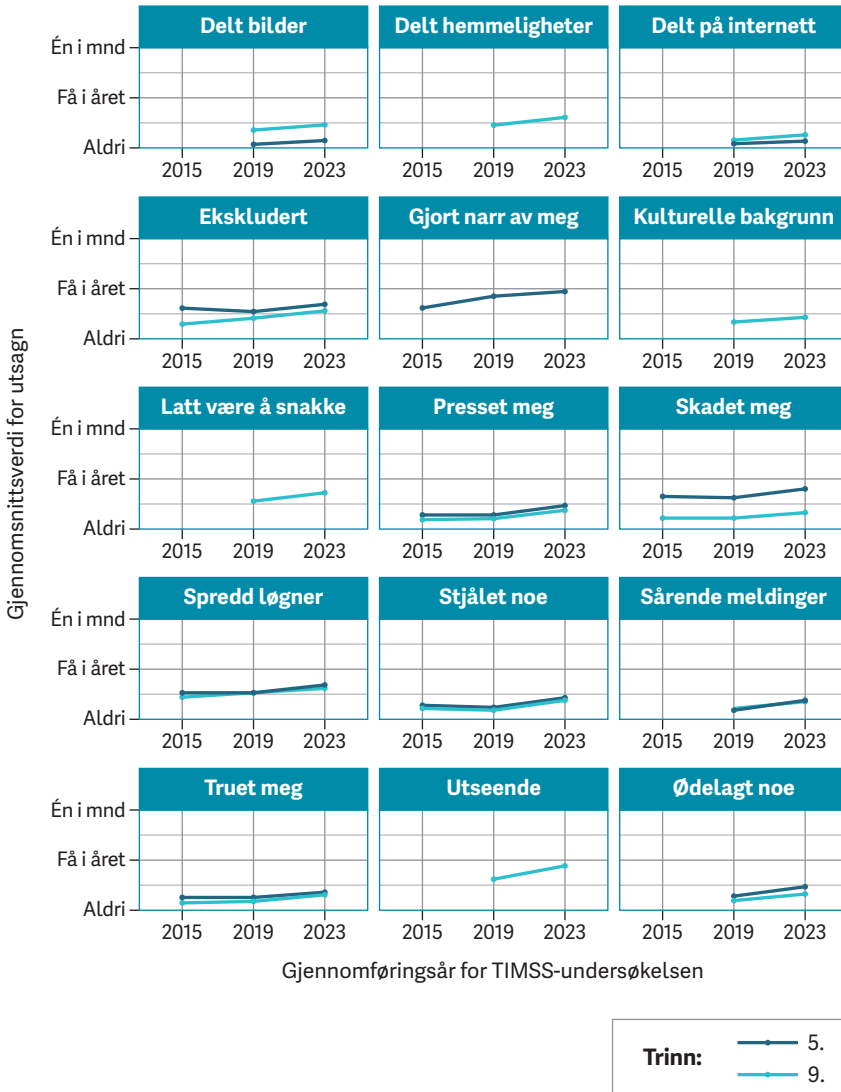
Resultater

Utviklingen av elevenes, lærernes og rektorenes oppfatning av trygghet fra TIMSS 2015 til TIMSS 2023

Resultatene for det første forskningsspørsmålet, om hvordan elevenes, lærernes og rektorenes oppfatning av trygghet har utviklet seg over tid, vises som gjennomsnitt av respondentenes svar på hvert utsagn (vist i tabell 3.2) for hver gjennomføring av TIMSS fra 2015 til 2023, som separate del-illustrasjoner i figurene 3.2, 3.3, 3.4 og 3.5.

I figur 3.2 vises resultatene for utsagnene knyttet til trygghet rapportert av elever som fravær av mobbing. Elevene har svart på hvor ofte andre elever ved skolen har gjort det som utsagnet beskriver, mot dem. «Overskriftene»

5 Det er viktig å understreke at faktorene som inngår i SØS-variabelen kun utgjør forenklede mål og ikke fullt ut reflekterer de komplekse og mangfoldige aspektene ved SØS. Dermed fungerer de som proxyer snarere enn fullstendige representasjoner av sosio-økonomisk bakgrunn.



Figur 3.2. Gjennomsnittlige verdier for utsagn knyttet trygghet rapportert av elever som fravær av mobbing.

Note. Fullstendig tekst for svarkategoriene i figuren er «Aldri», «Noen få ganger i året», «En eller to ganger i måneden» og «Minst én gang i uka», se tabell 3.2. Svarkategorien «Minst én gang i uka» er ikke inkludert for å forenkle figuren.

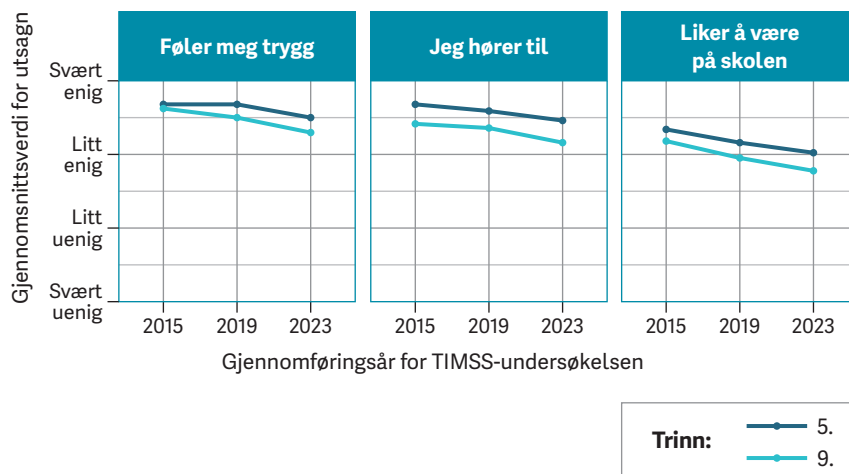
på hver del-illustrasjon i figur 3.2 er hentet fra kolonnen «Utsagn» i tabell 3.2, hvor de uthevede ordene i utsagnene utgjør overskriftene til del-illustrasjonene. Eksempelvis kommer overskriften «Delt bilder» øverst til venstre i figur 3.2 fra utsagnet «**Delt bilder** av meg på internett» som har

gjort meg flau». Dette utsagnet er inkludert i elevenes spørreskjema på begge trinn og det var inkludert i TIMSS 2019 og 2023, men ikke i 2015.

Det overordnede bildet som kommer frem fra figur 3.2 er at elevene i høy grad føler seg trygge på skolen, men at dette er i endring. For så godt som alle utsagnene knyttet til mobbing er elevenes gjennomsnittlige verdi i TIMSS 2023 på vei bort fra *Aldri* og mot oftere enn i de foregående TIMSS-undersøkelsene.

Elevene på 5. og 9. trinn ser i all hovedsak ut til å være enige, men det er to utsagn som skiller seg ut: «Delt bilder av meg (på internett som har gjort meg flau)» foregår oftere på 9. trinn enn på 5. trinn, og «Skadet meg (fysisk)» foregår oftere på 5. trinn enn på 9. trinn.

I figur 3.3 vises resultatene for utsagnene knyttet hva elevene synes om sin egen skole. For hvert utsagn har elevene angitt grad av enighet.

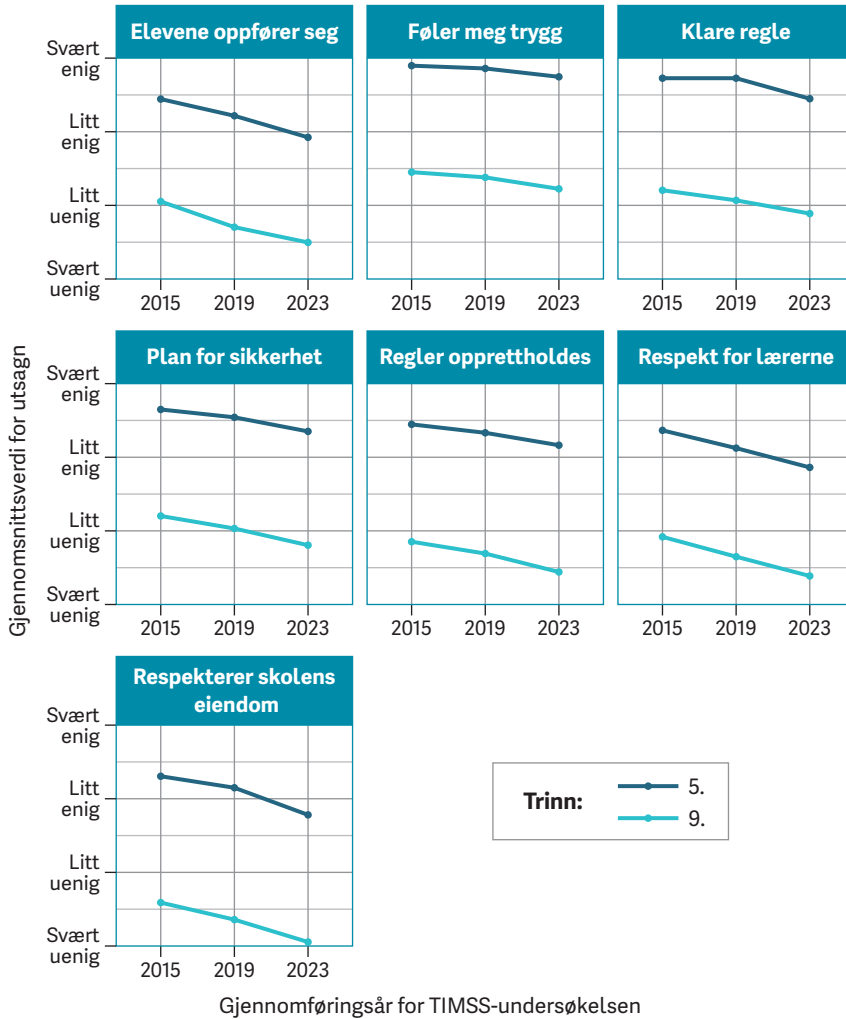


Figur 3.3. Gjennomsnittlige verdier for utsagn knyttet til hva elever på 5. og 9. trinn føler om skolen sin.

Ut fra resultatene vist i figur 3 kan det sies at elevene rapporterer om høy grad av enighet med hvert utsagn på begge trinn. Elevene på 9. trinn rapporterer om noe mindre enighet enn elevene på 5. trinn og gjennomsnittet for alle tre utsagnene synker etter hvert som tiden går.

I figur 3.4 vises resultatene knyttet til hvor enige lærerne er om trygghet, ro og orden ved den skolen de jobber på.

Ut fra resultatene vist i figur 3.4 går lærernes enighet ned på begge trinn. På 5. trinn er lærerne enige i at skolen er trygg og at det er ro og orden i alle tre målepunktene, men de har gått fra å være svært enige i TIMSS

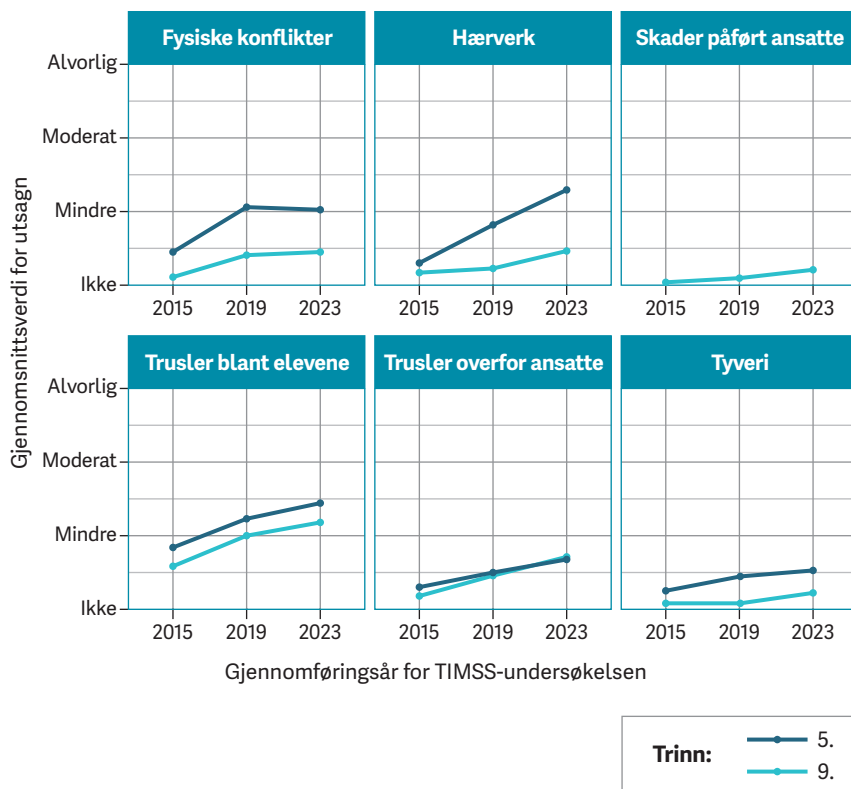


Figur 3.4. Gjennomsnittlige verdier for utsagn knyttet til hvor enige lærerne på 5. og 9. trinn er om trygghet, ro og orden ved egen skole.

2015 til å være litt enige i TIMSS 2023. På 9. trinn, derimot, har lærerne i alle målepunktene vært uenige. Her har deres oppfatninger, for de fleste av utsagnene, gått fra litt uenig til svært uenig.

I figur 3.5 vises resultatene for utsagnene rektorene svarer på knyttet til trygghet, målt som ro og orden. Rektorene angir hvor høy grad av problemer innholdet i de ulike utsagnene skaper for skolen.

Ut fra resultatene vist i figur 3.5, rapporterer rektorene på begge trinn om økende grad av problemer. Bortsett fra fysiske konflikter (som ser ut til å opptre oftere på 5. trinn enn på 9. trinn) og trusler overfor ansatte (hvor



Figur 3.5. Gjennomsnittlige verdier for utsagn knyttet til grad av problemer rektorer på 5. og 9. trinn rapporterer om ved egen skole.

Note. Fullstendig tekst for svarkategoriene i figuren er «Ikke et problem», «Et mindre problem», «Et moderat problem» og «Et alvorlig problem», se tabell 3.2.

gjennomsnittet for 5. og 9. trinn er så godt som likt), rapporterer rektorene på 9. trinn om høyere grad av problemer enn rektorene på 5. trinn for de resterende utsagnene. Det er viktig å notere her at til tross for at det er økende grad av problemer over tid, viser resultatene at rektorene i gjennomsnitt kun rapporterer om mindre problemer, ikke moderate eller alvorlige.

Kjennetegn på trygge skoler

Forenklede oversikter over resultatene for det andre forskningsspørsmålet, om hva som kjennetegner trygge skoler, vises i tabell 3.3 og 3.4. I tabell 3.3 vises resultatene fra regresjonsanalysene som er utført for variablene *Elevrapportert fravær av mobbing* og *Elevrapportert tilhørighet* på 5. og 9. trinn. I tabell 3.4 vises resultatene for variablene *Lærer trygghet*, *ro og orden*

Tabell 3.3. Endring av elevrapportert trygghet og tilhørighet over tid og sammenheng mellom elevrapportert trygghet og tilhørighet og kjennetegn ved skolen

	Variabel			
	Elevrapportert fravær av mobbing		Elevrapportert følelse av tilhørighet	
	5. trinn	9. trinn	5. trinn	9. trinn
TIMSS 2015	Referanseår	Referanseår	Referanseår	Referanseår
TIMSS 2019	—	—	Mindre tilhørighet	Mindre tilhørighet
TIMSS 2023	Økt mobbing	Økt mobbing	Mindre tilhørighet	Mindre tilhørighet
Lav SØS (< 10 %)	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Lav SØS (11–25 %)	—	—	—	—
Lav SØS (> 25 %)	—	—	—	—
Referansegruppe for skolens beliggenhet¹	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Beliggenhet er tettsted	—	—	Mindre tilhørighet	Mindre tilhørighet
Beliggenhet er landlig	Mer mobbing	—	Mindre tilhørighet	Mindre tilhørighet
Norsk morsmål (> 90 %)	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Norsk morsmål (76–90 %)	—	—	—	—
Norsk morsmål (51–75 %)	—	—	—	—
Norsk morsmål (< 50 %)	—	—	—	—
Variabel knyttet til andel elever med innvandrerbakgrunn				
Førstegenerasjonsinnvandrere	—	Større andel henger sammen med økt mobbing	—	Større andel henger sammen med mindre tilhørighet
Andre generasjonsinnvandrere	—	—	—	—

Note. Lav SØS: prosentandel elever med lav SØS ved skolen. Norsk morsmål: prosentandel elever med norsk som morsmål. 1 Svaralternativene tettbefolket byområde, forstad eller utkantsområde og mellomstor eller stor by er slått sammen som referansegruppe for skolens beliggenhet. Tegnet «—» indikerer at gruppens variabelverdi ikke er signifikant forskjellig fra referansegruppens.

Tabell 3.4. Endring av lærer- og rektor-rapportert trygghet, ro og orden over tid og sammenheng mellom lærer- og rektor-rapportert trygghet, ro og orden og kjennetegn ved skolen

	Variabel			
	Lærer trygghet, ro og orden		Rektor trygghet som ro og orden	
	5. trinn	9. trinn	5. trinn	9. trinn
TIMSS 2015	Referanseår	Referanseår	Referanseår	Referanseår
TIMSS 2019	Mindre trygghet	Mindre trygghet	Mindre trygghet	Mindre trygghet
TIMSS 2023	Mindre trygghet	Mindre trygghet	Mindre trygghet	Mindre trygghet
Lav SØS (< 10 %)	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Lav SØS (11–25 %)	—	—	—	—
Lav SØS (> 25 %)	Mindre trygghet	—	Mindre trygghet	—
Referansegruppe for skolens beliggenhet¹	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Beliggenhet er tettsted	—	—	—	—
Beliggenhet er landlig	—	—	—	—
Norsk morsmål (> 90 %)	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Norsk morsmål (76–90 %)	Mindre trygghet	—	Mindre trygghet	Mindre trygghet
Norsk morsmål (51–75 %)	—	Mindre trygghet	—	Mindre trygghet
Norsk morsmål (< 50 %)	—	—	—	Mindre trygghet
Variabel knyttet til andel elever med innvandrerbakgrunn				
Førstegenerasjonsinnvandrere	—	Større andel henger sammen med mindre trygghet	—	—
Andre generasjonsinnvandrere	—	—	—	—

Note. Lav SØS: prosentandel elever med lav SØS ved skolen. Norsk morsmål: prosentandel elever med norsk som morsmål. 1.Svaralternativene tettbefolket byområde, forstad eller utkantsområde og mellomstor eller stor by er slått sammen som referansegruppe for skolens beliggenhet. Tegnet «—» indikerer at gruppens variabelverdi ikke er signifikant forskjellig fra referansegruppens.

og *Rektor trygghet som ro og orden*. I tabell V1 i vedlegget vises resultatene for regresjonsanalysene som danner grunnlaget for oversikten gitt i tabell 3.3. Tabell V2 viser grunnlaget for tabell 3.4.

I tabell 3.3 og 3.4 angis et referanseår eller en referansegruppe som benyttes til sammenligning. Eksempelvis settes TIMSS 2015 som referanseår for målingen av variabelen *Elevrapportert tilhørighet* i tabell 3.3. Verdien for variabelen *Elevrapportert tilhørighet* fra TIMSS 2019 sammenlignes med verdien for TIMSS 2015, og resultatet av analysen angir om verdiene fra de to gjennomføringsårene er signifikant forskjellige og i hvilken retning forskjellen eventuelt peker. Tilsvarende gjøres for sammenligningen mellom TIMSS 2023 og TIMSS 2015. Videre viser begge tabellene resultatet av regresjonsanalysene når variablene beregnes for ulike grupperinger av skoler. Grupperingene er basert på kjennetegn ved skolene som for eksempel hvor stor andel elever med norsk som morsmål som går på skolen (*Norsk morsmål*), eller hvor stor andelen elever med lav SØS er (*Lav SØS*). For disse kjennetegnene er referansegruppene henholdsvis skoler hvor andelen elever med norsk som morsmål er over 90 prosent og skoler hvor andelen elever med lav SØS er under 10 prosent.

Resultatene i tabell 3.3 og 3.4 viser at elever over tid rapporterer om mer mobbing og mindre tilhørighet på både 5. og 9. trinn. Samtidig rapporterer lærere og rektorer om en nedgang i følelse av trygghet, ro og orden.

Når skolene deles inn etter kjennetegn, viser det seg at elevenes rapportering om mobbing og tilhørighet ikke har sammenheng med andelen elever med lav SØS eller andelen elever med norsk som morsmål (tabell 3.3). For lærere og rektorer på 5. trinn er det derimot en sammenheng mellom deres opplevelse av trygghet, ro og orden og andelen elever med lav SØS og norsk som morsmål. De rapporterer om mindre trygghet på skoler med disse kjennetegnene (tabell 3.4).

For lærere og rektorer på 9. trinn viser tabell 3.4 at det kun er andelen elever med norsk som morsmål som har sammenheng med deres følelse av trygghet. På skoler hvor færre enn 90 prosent av elevene har norsk som morsmål, rapporterer både lærere og rektorer om mindre trygghet sammenlignet med referanseskolene hvor over 90 prosent av elevene har norsk som morsmål.

Når skolene deles inn etter beliggenhet, rapporterer kun elever på 5. trinn ved landlige skoler om mer mobbing enn referansegruppen. Elever på både 5. og 9. trinn føler mindre tilhørighet ved skoler i tettsteder og landlige områder sammenlignet med referansegruppen. For lærere og rektorer har

skolens beliggenhet ingen sammenheng med deres opplevelse av trygghet, ro og orden.

På 5. trinn har ikke elevenes følelse av mobbing og tilhørighet sammenheng med andelen første- eller andregenerasjonsinnvandrere. På 9. trinn er det en sammenheng mellom økt andel førstegenerasjonsinnvandrere og økt rapportering av mobbing samt redusert følelse av tilhørighet blant elevene. Blant de voksne er det kun lærerne på 9. trinn som rapporterer om mindre trygghet ved skoler med høyere andel førstegenerasjonsinnvandrere.

Kjennetegn på skoler som oppleves som trygge til tross for observerte endringer i skolemiljøet

Som beskrevet i analyseseksjon, for å svare på forskningsspørsmål 3 om hva som kjennetegner skoler som fortsatt oppleves som trygge til tross for observerte endringer i skolemiljøet, ble det kjørt tilsvarende regresjonsanalyser som for forskningsspørsmål 2 med en interaksjon mellom gjennomføringsår for TIMSS-undersøkelsen og de ulike kjennetegnene ved skolene i tillegg for hver regresjonsmodell.

Med fem interaksjonsledd for hver av de åtte regresjonsmodellene, beskrives bare hovedtrekkene fra resultatene her da det er en stor risiko for at noen av regresjonskoeffisientene blir signifikante ved ren tilfeldighet bare fordi analysen inneholder så mange modeller (Simmons et al., 2011). Å fokusere på hovedtrekk, eller mønster på tvers av modeller, reduserer denne risikoen.

Ingen av analysene resulterte i signifikante interaksjoner – ikke for noen av trinnene og ikke for noen av respondentgruppene. Det er to mulige forklaringer på dette. Den første er knyttet til den relativt jevne / like store nedgangen i rapportert trygghet for alle skoler (uansett kjennetegn). Den andre forklaringen kan knyttes til manglende statistisk kraft til å oppdage at skoler med ett eller flere kjennetegn har større eller mindre nedgang enn de andre. Begge forklaringene kan opptre samtidig. Derfor, sagt med litt andre ord: Med utgangspunkt i dette datamaterialet og disse analysene er det ikke mulig å identifisere skoler med gitte kjennetegn som fremdeles oppleves trygge til tross for de observerte endringene i skolemiljøet.

Diskusjon

Ifølge tidligere forskning har et trygt skolemiljø en positiv sammenheng med elevenes læringsutbytte (se f.eks. Grazia & Molinari, 2021; Nilsen &

Teig, 2022). Videre er det i Norge lovfestet at elever, lærere og rektorer har rett til en trygg og god arbeidsplass. Derfor har målet for dette kapittelet vært å se nærmere på skolemiljøet, med fokus på elevers, læreres og rektors opplevelse av trygghet i skolehverdagen.

Med norske data fra TIMSS-undersøkelsene gjennomført i 2015, 2019 og 2023 som grunnlag, viser resultatene fra denne studien i all hovedsak at elever, lærere og rektorer føler seg trygge på skolen. I tillegg viser resultatene at det er tendenser til endringer som peker mot et mindre trygt skolemiljø i 2023 enn det som ble rapportert i 2015. Resultatene fra denne studien viser at elevene rapporterer om hyppigere episoder av mobbing nå enn tidligere (se figur 3.2). De er heller ikke like enige i utsagnene om å høre til og føle seg trygge på skolen og at de liker mindre og mindre det å være på skolen (se figur 3.3). Dette er i samsvar med funnene fra blant annet *Ungdata 2024* (Bakken, 2024).

Som elevene, rapporterer også lærerne om mer og mer uenighet med utsagn som handler om at skolen oppleves som trygg og at det er ro og orden der. For disse utsagnene er det dog signifikant forskjell i oppfatningene til lærerne på 5. og 9. trinn. Mens lærerne på 5. trinn har oppfattet og fremdeles oppfatter miljøet som relativt trygt og ordentlig, har lærerne på 9. trinn stort sett gått fra å være litt uenige i utsagnene til å være svært uenige (se figur 3.4). Rektorene melder stort sett om liten grad av problemer knyttet til ro og orden på skolene (se figur 3.5), men hærverk ser ut til å være et voksende problem. På 9. trinn har gjennomsnittsverdien for utsagnet gått fra at hærverk ikke utgjør et problem i 2015, til at det utgjør et mindre til moderat problem i 2023. Dette samsvarer med det lærerne på det samme trinnet mener om elevenes respekt for skolens eiendom (se figur 3.4). Lærerne på 5. trinn oppfatter elevenes respekt for skolens eiendom som bedre enn på 9. trinn, men også på 5. trinn blir lærerne mer og mer uenige i at elevene respekterer skolens eiendom. I tillegg opplever lærerne på begge trinn at elevenes oppførsel er dårligere og at elevene har mindre respekt for lærerne i 2023 enn i 2015. Dersom denne trenden fortsetter er det, basert på tidligere forskning, en reell sannsynlighet for at elevenes engasjement og læringsutbytte blir dårligere (Kutsyruba et al., 2015; Ripamonti, 2018; Wang & Degol, 2016). Utrygge elever kan oppleve mindre trivsel på skolen, noe som kan hindre dem i å yte sitt beste. Samtidig kan utrygghet blant lærerne resultere i redusert undervisningskvalitet (Holzberger & Schiepe-Tiska, 2021). Alt dette kan i sin tur gjøre det vanskeligere å oppnå regjeringens mål om høyere gjennomføringsrate

for videregående opplæring (Meld. St. 21 (2020–2021)), og det kan også hindre ungdommenes muligheter til å lykkes i høyere utdanning og arbeidsliv.

Resultatene fra analysene for sammenheng mellom kjennetegn på trygge skoler og trygghet, spesielt på 9. trinn, sammenfaller med tidligere forskning når det gjelder skoler med høyere andel førstegenerasjonsinnvandrere (Peterson et al., 2024; Walsh et al., 2016). Ved de norske skolene har både elevenes, lærernes og rektorenes rapportering om opplevelse av trygghet en moderat, negativ sammenheng med høyere andel førstegenerasjonsinnvandrere ved skolen (se tabell V1 og V2 i vedlegget).

Siden regresjonsanalysene for dette kapittelet kun har inkludert ett kjennetegn ved skolene per analyse, kan ikke denne studien si noe om skoler med høyere andel førstegenerasjonsinnvandrere også har høyere andeler elever med lav SØS og et annet morsmål enn norsk. Disse to faktorene er også kjent fra tidligere forskning for å ha en negativ påvirkning på skolemiljøet og elevenes prestasjoner (se f.eks. Peterson et al., 2024; Wang & Degol, 2016). Men den svake til moderate sammenhengen mellom trygghet og skoler med eksempelvis høyere andel elever med lav SØS, ikke har norsk som morsmål eller høyere andel førstegenerasjonsinnvandrere, kan også skyldes helt andre faktorer som ikke er inkludert i denne studien. Det kan for eksempel handle om faktorer utenfor skolen, eller det kan handle om hvilke læringsressurser skolen har til rådighet og hvordan disse benyttes, eller det kan være en blanding av dette. I tillegg ligger det også en begrensning for studien knyttet til målgruppenes subjektive tolkninger og vurderinger av spørsmålene i spørreskjemaet. Det er også knyttet usikkerhet til om elever, lærere og rektorer har forstått og tolket spørsmålene og utsagnene (i tabell 3.2) på samme måte i 2015, 2019 og 2023. For å oppnå mer nøytrale/objektive datasett kunne feltobservasjoner eller videoopptak vært tatt i bruk, men det hadde vært for ressurskrevende å gjennomføre for representative utvalg ved tre målepunkter. For å komme nærmere et mer nyansert bilde av dette, trengs det mer forskning.

Resultatene tilhørende det tredje og siste forskningsspørsmålet antyder at det, ut fra det valgte datagrunnlaget for denne studien, ikke er mulig å finne fellestrekk ved skoler som oppleves som trygge til tross for de observerte endringene i skolemiljøet. Her kan det også være flere forklaringer, og en av disse kan henge sammen med at elever, lærere og rektorer ved *alle* skolene rapporterer om tilnærmet lik tilbakegang for skolemiljøet. En annen forklaring kan henge sammen med at representativiteten i dataene er

på elevnivå, ikke på skolenivå (hvilket også inkluderer representativiteten til lærerne og rektoren).

Avsluttende kommentarer

Til tross for at resultatene av denne studien ikke kan beskrives som urovekkende, viser de tendenser til forverring som det er viktig å ta tak i før det blir et stort problem. Derfor er det viktig å fortsette arbeidet med å opprettholde, forbedre eller skape gode skolemiljø. Det finnes allerede ulike ressurser som er tilgjengelige for skolene. For eksempel har Utdanningsdirektoratet laget en egen nettside med «Råd og kunnskap om skolemiljø» (Utdanningsdirektoratet, u.å.) og utviklet kompetansepakker som skolene kan benytte seg av i arbeidet med skolemiljøet (Utdanningsdirektoratet, 2023a). Regjeringen har opprettet en tilskuddsordning for styrking av skolemiljøarbeidet som kommunene kan søke på (Meld. St. 34 (2023–2024); Regjeringen, 2025), og Helsedirektoratet har utgitt en forskningsbasert rapport som handler om ulike faktorer som fremmer elevers trivsel på skolen (Helsedirektoratet, 2015).

Forfatterbiografier

Hege Kaarstein er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad i matematikkdidaktikk og har jobbet med TIMSS-undersøkelsen siden 2009, og vært prosjektleder fra 2018. I tillegg deltar hun i den internasjonale ekspertgruppen for matematikk i TIMSS. Hennes interesser er blant annet lærernes undervisningskompetanse i matematikk, elevenes motivasjon for faget og svar på åpne oppgaver.

Anne-Catherine Lehre er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en bakgrunn innenfor språk og realfag, og har en doktorgrad i biostatistikk. Lehre har jobbet i TIMSS-prosjektet siden 2020.

Mark White er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han leder prosjektet *Teaching and Learning International Survey (TALIS) 2024* i Norge. Hans forskningsinteresse er blant annet undervisningskvalitet, lærerprofesjonalitet og kvantitative metoder.

Referanser

- Arbeidsmiljøloven. (2006). *Lov om arbeidsmiljø, arbeidstid og stillingsvern mv.* (LOV-2005-06-17-62). Lovdata. [https://lovdata.no/lov/2005-06-17-62/\\$1-1](https://lovdata.no/lov/2005-06-17-62/$1-1)
- Archambault, I., Janosz, M., Olivier, E. & Dupéré, V. (2022). Student engagement and school dropout: Theories, evidence, and future directions. I A. L. Reschly & S. L. Christenson (Red.), *Handbook of research on student engagement* (s. 331–355). Springer International Publishing. https://doi.org/10.1007/978-3-031-07853-8_16
- Arendalsuka. (2017, 16. august). *Hadde vi ikke verdens beste skole – eller? En debatt om skolemiljø og kunnskap* [Video]. YouTube. <https://youtu.be/ZRSJ8qotCaA?feature=shared>
- Bakken, A. (2024). *Ungdata 2024. Nasjonale resultater* (NOVA-rapport 6/24). <https://hdl.handle.net/11250/3145138>
- Bergem, O. K., Kaarstein, H. & Nilsen, T. (2016). TIMSS 2015. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015* (s. 11–21). Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016-02>
- Bradshaw, C. P., Cohen, J., Espelage, D. L. & Nation, M. (2021). Addressing school safety through comprehensive school climate approaches. *School Psychology Review*, 50(2–3), 221–236. <https://doi.org/10.1080/2372966X.2021.1926321>
- Endringslov til opplæringslova og friskolelova. (2017). *Lov om endringer i opplæringslova og friskolelova* (LOV-2017-06-09-38). Lovdata. <https://lovdata.no/LTI/lov/2017-06-09-38>
- Erdem, C. & Kaya, M. (2024). The relationship between school and classroom climate, and academic achievement: A meta-analysis. *School Psychology International*, 45(4), 380–408. <https://doi.org/10.1177/01430343231202923>
- Eriksen, I. M., Hegna, K., Bakken, A. & Lyng, S. T. (2014). *Felles fokus. En studie av skolemiljøprogrammer i norsk skole* (NOVA-rapport 15/2014). Velferdsforskningsinstituttet NOVA.
- Frengstad, J. M. (2023, 29. mars). Mobbing i Innlandet øker. *Oppland Arbeiderblad*, 6.
- Gelman, A. & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Grazia, V. & Molinari, L. (2021). School climate multidimensionality and measurement: A systematic literature review. *Research Papers in Education*, 36(5), 561–587. <https://doi.org/10.1080/02671522.2019.1697735>
- Harding, S., Morris, R., Gunnell, D., Ford, T., Hollingworth, W., Tilling, K., Evans, R., Bell, S., Grey, J., Brockman, R., Campbell, R., Araya, R., Murphy, S. & Kidger, J. (2019). Is teachers' mental health and wellbeing associated with students' mental health and wellbeing? *Journal of Affective Disorders*, 242, 180–187. <https://doi.org/10.1016/j.jad.2018.08.080>
- Harrison, M. G., King, R. B. & Wang, H. (2023). Satisfied teachers are good teachers: The association between teacher job satisfaction and instructional quality. *British Educational Research Journal*, 49(3), 476–498. <https://doi.org/10.1002/berj.3851>
- Hatlevik, O. E., Rohatgi, A. & Björnsson, J. K. (2018). Skoleledernes syn på skoleklima: Noen erfaringer fra PISA og TIMSS fra 2003 og 2015. I J. K. Björnsson & R. V. Olsen (Red.), *Tjue år med TIMSS og PISA i Norge* (s. 173–198). Universitetsforlaget. <https://doi.org/10.18261/9788215030067-2018-09>
- Helsedirektoratet. (2015). *Trivsel i skolen* (Rapport IS-2345). https://www.helsedirektoratet.no/rapporter/trivsel-i-skolen/Trivsel%20i%20skolen.pdf/_/attachment/inline/9e76ad2b-14eb-4c5f-bf8b-e157f5efcc06:176ea3b766b3d06407dbd9395c3b1776682c5ab7/Trivsel%20i%20skolen.pdf
- Holzberger, D. & Schiepe-Tiska, A. (2021). Is the school context associated with instructional quality? The effects of social composition, leadership, teacher collaboration, and school climate. *School Effectiveness and School Improvement*, 32(3), 465–485. <https://doi.org/10.1080/09243453.2021.1913190>
- Jensen, F., Pettersen, A., Frønes, T. S., Eriksen, A., Løvgren, M. & Narvhus, E. K. (2023). *PISA 2022. Norske elevers kompetanse i matematikk, naturfag og lesing*. Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.205>

- Jentsch, A., Hoferichter, F., Blömeke, S., König, J. & Kaiser, G. (2023). Investigating teachers' job satisfaction, stress and working environment: The roles of self-efficacy and school leadership. *Psychology in the Schools*, 60(3), 679–690. <https://doi.org/10.1002/pits.22788>
- Kim, J. K., Brick, J. M., Fuller, W. A. & Kalton, G. (2006). On the bias of the multiple-imputation variance estimator in survey sampling. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(3), 509–521. <https://doi.org/10.1111/j.1467-9868.2006.00546.x>
- Kutsyruba, B., Klinger, D. A. & Hussain, A. (2015). Relationships among school climate, school safety, and student achievement and well-being: A review of the literature. *Review of Education*, 3(2), 103–135. <https://doi.org/10.1002/rev3.3043>
- Kaarstein, H. & Lehre, A.-C. (2025) Innledning. I H. Kaarstein & A.-C. Lehre (Red.), *Hva TIMSS forteller om norsk skole. Sekundæranalyser av TIMSS-undersøkelsen* (s. 9–19). Cappelen Damm. <https://doi.org/10.23865/cdf.274>
- Kaarstein, H., Lehre, A.-C., Rohatgi, M. & Radišić, J. (2024). *TIMSS 2023. Kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Lewno-Dumdie, B. M., Mason, B. A., Hajovsky, D. B. & Villeneuve, E. F. (2020). Student-report measures of school climate: A dimensional review. *School Mental Health*, 12(1), 1–21. <https://doi.org/10.1007/s12310-019-09340-2>
- Lumley, T. (2011). *Complex surveys: A guide to analysis using R*. John Wiley & Sons.
- Marraccini, M. E., Fang, Y., Levine, S. P., Chin, A. J. & Pittleman, C. (2020). Measuring student perceptions of school climate: A systematic review and ecological content analysis. *School Mental Health*, 12(2), 195–221. <https://doi.org/10.1007/s12310-019-09348-8>
- Martin, M. O., Mullis, I. V. S., Foy, P. & Hooper, M. (2016). *TIMSS 2015 international results in mathematics and science*. TIMSS & PIRLS International Study Center. <http://timssandpirls.bc.edu/timss2015/international-results/>
- Meld. St. 21 (2016–2017). *Lærelyst – tidlig innsats og kvalitet i skolen*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-21-20162017/id2544344/>
- Meld. St. 21 (2020–2021). (2021). *Fullføringsreformen – med åpne dører til verden og fremtiden*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-21-20202021/id2840771/>
- Meld. St. 34 (2023–2024). *En mer praktisk skole Bedre læring, motivasjon og trivsel på 5.–10. Trinn*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-34-20232024/id3052898/>
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L. & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/international-results/>
- Nilsen, T. & Kaarstein, H. (2021). Skolemiljø, motivasjon og naturfagprestasjoner fra TIMSS 2015 til TIMSS 2019. I T. Nilsen & H. Kaarstein (Red.), *Med blikket mot naturfag* (s. 143–164). Universitetsforlaget. <https://doi.org/10.18261/9788215045108-2021-06>
- Nilsen, T., Kaarstein, H. & Lehre, A.-C. (2022). Trend analyses of TIMSS 2015 and 2019: School factors related to declining performance in mathematics. *Large-scale Assessments in Education*, 10(1), Artikkel 15. <https://doi.org/10.1186/s40536-022-00134-8>
- Nilsen, T. & Teig, N. (2022). A systematic review of studies investigating the relationships between school climate and student outcomes in TIMSS, PISA, and PIRLS. I T. Nilsen, A. Stancel-Piatak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (s. 1053–1086). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_34-1
- Olafsen, M. J. (2022, 27. juni). – Ungene er redde for å dra på skolen. *Avisa Nordland*.
- Olweus, D. & Limber, S. P. (2010). The Olweus Bullying Prevention Program: Implementation and evaluation over two decades. I S. R. Jimerson, S. M. Swearer & D. L. Esplage (Red.), *Handbook of bullying in schools. An international perspective* (s. 377–401). Routledge.
- Olweus, D. & Limber, S. P. (2019). The Olweus Bullying Prevention Program (OBPP). New evaluations and current status. I P. K. Smith (Red.), *Making an impact on school bullying. Interventions and recommendations* (s. 22). Routledge.

- Opplæringslova. (2024). *Lov om grunnskoleopplæringa og den vidaregåande opplæringa* (LOV-2023-06-09-30). Lovdata. [https://lovdata.no/lov/2023-06-09-30/\\$12-2](https://lovdata.no/lov/2023-06-09-30/$12-2)
- Peterson, J. M., Foldnes, N. & Jensen, M. T. (2024). Relations between linguistic diversity, classroom social climate, and reading comprehension in Norwegian elementary classrooms. *Scandinavian Journal of Educational Research*, 69(4), 682–697. <https://doi.org/10.1080/00313831.2024.2330937>
- R Core Team. (2024). *R: A language and environment for statistical computing* [Programvare]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Regjeringen. (2025, 21. mars). *Kraftig satsing på bedre skolemiljø* [Pressemelding]. Kunnskapsdepartementet. <https://www.regjeringen.no/no/aktuelt/kraftig-satsing-pa-bedre-skolemiljo/id3093012/>
- Reynolds, K. A., Mullis, I. V. S. & Martin, M. O. (2021). TIMSS 2023 context questionnaire framework. I I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Ripamonti, E. (2018). Risk factors for dropping out of high school: A review of contemporary, international empirical research. *Adolescent Research Review*, 3(3), 321–338. <https://doi.org/10.1007/s40894-017-0075-y>
- Rutkowski, L. & Rutkowski, D. (2016). The relation between students' perceptions of instructional quality and bullying victimization. I T. Nilsen & J.-E. Gustafsson (Red.), *Teacher quality, instructional quality and student outcomes*. Springer. https://doi.org/10.1007/978-3-319-41252-8_6
- Scherer, R. & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. I T. Nilsen & J.-E. Gustafsson (Red.), *Teacher quality, instructional quality and student outcomes*. Springer. https://doi.org/10.1007/978-3-319-41252-8_3
- Siegel, P. & Foy, P. (2024). TIMSS sample design. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report* (s. 3.1–3.30). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs7952>
- Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2011). False-positive psychology. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Skutlaberg, K. B. (2023). *Å være rektor i arbeid med langvarige saker knyttet til elevers psykososiale skolemiljø* [Monografi, Universitetet i Oslo]. <https://www.duo.uio.no/handle/10852/109965>
- Thronsdén, I., Carlsten, T. C. & Björnsson, J. K. (2019). *Talis 2018. Første hovedfunn fra ungdomstrinnet*. NIFU.
- Toropova, A., Myrberg, E. & Johansson, S. (2021). Teacher job satisfaction: The importance of school working conditions and teacher characteristics. *Educational Review*, 73(1), 71–97. <https://doi.org/10.1080/00131911.2019.1705247>
- Utdanningsdirektoratet. (2023a). *Kompetansepakker om inkludering, læring, trivsel og lek*. <https://www.udir.no/kvalitet-og-kompetanse/kompetansepakker-om-inkludering/>
- Utdanningsdirektoratet. (2023b). *Undersøkelse av brudd på opplæringsloven kapittel 9A*. <https://www.udir.no/tall-og-forskning/finn-forskning/tema/saker-meldt-til-fylkesmennene-om-skolemiljo/undersokelse-av-brudd-pa-opplaringsloven-kapittel-9a/>
- Utdanningsdirektoratet. (u.å.). *Råd og kunnskap om skolemiljø*. <https://www.udir.no/laring-og-trivsel/skolemiljo/>
- van Buuren, S. & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Veletić, J., Radišić, J. & Madsø Jacobsen, H. M. (2025). Skolens og rektors betydning for elevenes prestasjoner i matematikk og naturfag. I H. Kaarstein & A.-C. Lehre (Red.), *Hva TIMSS forteller om norsk skole. Sekundæranalyser av TIMSS-undersøkelsen* (s. 79–104). Cappelen Damm. <https://doi.org/10.23865/cdf.274>

- Walker, M. J. (2023, 21. september). *Har innføringen av kapittel 9A virket motsatt?* Utdanningsnytt. <https://www.utdanningsnytt.no/mette-johnsen-walker-mobbing-paragraf-9a/har-innforingen-av-kapittel-9a-virket-motsatt/373658>
- Walsh, S. D., De Clercq, B., Molcho, M., Harel-Fisch, Y., Davison, C. M., Rich Madsen, K. & Stevens, G. W. J. M. (2016). The relationship between immigrant school composition, classmate support and involvement in physical fighting and bullying among adolescent immigrants and non-immigrants in 11 countries. *Journal of Youth and Adolescence*, 45(1), 1–16. <https://doi.org/10.1007/s10964-015-0367-0>
- Wang, M.-T. & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, 28(2), 315–352. <https://doi.org/10.1007/s10648-015-9319-1>
- Wendelborg, C., Røe, M. & Caspersen, J. (2016). *Elevundersøkelsen 2015. Analyse av Elevundersøkelsen 2015* (Rapport 2016). NTNU Samfunnsforskning. <https://samforsk.no/uploads/files/Publikasjoner/Elevundersokelsen-2015-WEB.pdf>
- Wendelborg, C., Dahl, T., Røe, M. & Buland, T. H. (2020). *Elevundersøkelsen 2019: Analyse av Utdanningsdirektoratets brukerundersøkelser* (Rapport 2020). NTNU Samfunnsforskning. <https://hdl.handle.net/11250/2689870>

Elevrapportert fravær av mobbing				Elevrapportert tilhørighet			
5. trinn	9. trinn	5. trinn	9. trinn				
Referansegruppe beliggenhet ¹	Referanse- gruppe	Referanse- gruppe	Referanse- gruppe				
Beliggenhet er tettsted	0 (0,02)	15 % -0,01 (0,01)	8 % -0,06* (0,03)	15 % -0,06* (0,02)	16 % -0,06* (0,03)		
Beliggenhet er landlig	-0,07** (0,02)	5 % -0,02 (0,02)	8 % -0,11*** (0,02)	6 % -0,08* (0,03)	6 % -0,08* (0,03)		
Andel elever med norsk som morsmål:							
Prosentandel med norsk som morsmål er > 90 %	Referanse- gruppe	Referanse- gruppe	Referanse- gruppe	Referanse- gruppe	Referanse- gruppe		
Prosentandel med norsk som morsmål er 76-90 %	0 (0,02)	4 % 0,01 (0,02)	6 % -0,02 (0,02)	6 % -0,02 (0,02)	3 % -0,02 (0,02)		
Prosentandel med norsk som morsmål er 51-75 %	-0,01 (0,03)	4 % -0,01 (0,02)	6 % -0,02 (0,04)	3 % -0,02 (0,04)	13 % 0,01 (0,04)		
Prosentandel med norsk som morsmål < 50 %	-0,03 (0,04)	7 % 0,02 (0,04)	5 % 0 (0,05)	20 % 0,06 (0,06)	6 % 0,06 (0,06)		
Andel elever med innvandrerbakgrunn:							
Prosentandel første- generasjonsinnvandrere	0,04 (0,15)	1 % -0,45** (0,16)	0 % -0,18 (0,21)	0 % -0,6** (0,22)	1 % 0,17 (0,14)		
Prosentandel andre- generasjonsinnvandrere	-0,11 (0,08)	2 % 0,06 (0,09)	2 % -0,02 (0,11)	3 % 0,17 (0,14)	4 % 0,17 (0,14)		

Noter. ***Signifikansnivå p<0,001, **Signifikansnivå p<0,01, *Signifikansnivå p<0,05. 1Svaralternativene tettbefolket byområde, forstad eller utkantsområde og mellomstor eller stor by er slått sammen som referansegruppe for skolens beliggenhet. Beta viser verdien til regresjonskoeffisientene, standardfeil i parentes (SF).

Tabell V2. Endring av lærer- og rektor-rapportert trygghet, ro og orden over tid og sammenheng mellom lærer- og rektor-rapportert trygghet, ro og orden og kjennetegn ved skolen, samt andel manglende informasjon

Variabel	Lærer trygghet, ro og orden			Rektor trygghet som ro og orden		
	5. trinn	9. trinn	5. trinn	9. trinn	Andel manglende informasjon	Andel manglende informasjon
(Skjæringspunkt)	Beta (SF)	Andel manglende informasjon	Beta (SF)	Andel manglende informasjon	Beta (SF)	Andel manglende informasjon
	2,6*** (0,05)	24 %	1,15*** (0,04)	1 %	2,77*** (0,05)	5 %
					2,75*** (0,04)	4 %
Endring over tid:						
TIMSS 2015	Referanseår	Referanseår	Referanseår	Referanseår		
TIMSS 2019	-0,12** (0,05)	11 %	-0,16*** (0,04)	7 %	-0,24*** (0,06)	2 %
						11 %
TIMSS 2023	-0,29*** (0,05)	13 %	-0,35*** (0,04)	5 %	-0,33*** (0,05)	6 %
						4 %
Andel lav-SØS elever:						
Prosentandel lav-SØS elever < 10 %	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe
Prosentandel lav-SØS elever 11-25 %	-0,03 (0,05)	18 %	-0,03 (0,04)	11 %	-0,1 (0,05)	9 %
						8 %
Prosentandel lav-SØS elever > 25 %	-0,16* (0,07)	17 %	-0,08 (0,07)	1 %	-0,22* (0,09)	3 %
						13 %
Skolens beliggenhet:						
Referansegruppe beliggenhet ¹	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe	Referansegruppe

Lærer trygghet, ro og orden			Rektor trygghet som ro og orden		
	5. trinn	9. trinn	5. trinn	9. trinn	
Beliggenhet er tettsted	-0,02 (0,05)	15 %	-0,11 (0,06)	14 %	-0,02 (0,04) 3 %
Beliggenhet er landlig	0,02 (0,05)	20 %	0,04 (0,06)	4 %	0,01 (0,06) 1 %
Andel elever med norsk som morsmål:					
Prosentandel med norsk som morsmål er >90 %	Referanse- gruppe	Referanse- gruppe	Referanse- gruppe	Referanse- gruppe	
Prosentandel med norsk som morsmål er 76-90 %	-0,12** (0,05)	5 %	-0,02 (0,04)	18 %	-0,1* (0,05) 1 %
Prosentandel med norsk som morsmål er 51-75 %	-0,15 (0,08)	20 %	-0,17* (0,08)	18 %	-0,31*** (0,09) 1 %
Prosentandel med norsk som morsmål < 50 %	-0,17 (0,09)	21 %	-0,05 (0,12)	9 %	-0,04 (0,1) 15 %
Andel elever med innvandrerbakgrunn:					
Prosentandel første- generasjonsinnvandrere	0,17 (0,38)	9 %	-0,7* (0,31)	5 %	-0,02 (0,38) 3 %
Prosentandel andre- generasjonsinnvandrere	-0,05 (0,19)	10 %	-0,12 (0,22)	25 %	0 (0,28) 9 %
					0,31 (0,21) 8 %

Noter. ***Signifikansnivå p<0,001, **Signifikansnivå p<0,01, *Signifikansnivå p<0,05. 1.Svaralternativene tettbefolket byområde, forstad eller utkantsområde og mellomstor eller stor by er slått sammen som referansegruppe for skolens beliggenhet. Beta viser verdien til regresjonskoeffisientene, standardfeil i parentes (SF).

KAPITTEL 4

Skolens og rektors betydning for elevenes prestasjoner i matematikk og naturfag

Jelena Veletić, Jelena Radišić og Hilde Marie Madsø Jacobsen

Institutt for Lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: I dette kapittelet undersøker vi sammenhengen mellom kjennetegn ved rektorer, skolemiljø og elevenes prestasjoner i matematikk og naturfag på 9. trinn, med særlig vekt på faktorer på skolenivå. Ved bruk av to-nivåanalyse viser resultatene at omtrent 8–9 prosent av variasjonen i elevenes prestasjoner kan forklares ved forskjeller mellom skolene, noe som understreker betydningen av skolekonteksten på den enkelte skole. På skolenivå ble det funnet en positiv sammenheng mellom rektorenes erfaring og utdanningsnivå (det å ha en mastergrad i utdanningsledelse) og høyere elevprestasjoner. Studien viste imidlertid begrenset støtte for at læringsfokus (*school emphasis on academic success*) hadde sterk sammenheng med elevenes prestasjoner. Hvor ofte man snakker norsk hjemme og det gjennomsnittlige nivået på læringsressurser hjemmefra var i vår studie ikke signifikant assosiert med forskjeller i prestasjoner mellom skolene når vi kontrollerte for de samme faktorene på elevnivå. Funnene tyder på at selv om hjemme-konteksten er avgjørende for den enkelte elev, har likevel faktorer på skolenivå – spesielt kvalifiserte og erfarne skoleledere – stor betydning for skolens prestasjoner. Resultatene understreker behovet for kvalifiserte og erfarne skoleledere som kan skape et skolemiljø som støtter elevenes læring og utvikling.¹

1 KI ble benyttet brukt som verktøy for språkhjelp.

Nøkkelord: læringsfokus, skolemiljø, utdanningsledelse, elevers prestasjoner, fler-nivåanalyse

Abstract

The Schools' and Principals' role in Students' Achievement in Mathematics and Science

This study explores the relationship between principal characteristics, school climate, and student achievement in mathematics and science among ninth-grade students in Norway, emphasising the role of school-level factors. Using a multi-level modelling approach, the results reveal that approximately 8–9 per cent of the variance in student achievement can be attributed to differences between schools, highlighting the importance of school context. At the school level, principal characteristics such as years of experience and holding a master's degree in educational leadership were positively associated with higher student outcomes at the school level. However, the study found limited evidence linking schools' emphasis on academic success to higher student performance. How frequently students speak Norwegian at home and the average home educational resources of students at the school level, did not significantly account for differences in academic achievement between schools after controlling for the same factors at the student level. These findings suggest that while home context plays a crucial role for individual students, school-level factors—particularly the presence of experienced and educated school principals—are essential for supporting the academic success of schools. The results further emphasise the need for well-qualified and experienced school leaders to cultivate environments that encourage student learning and development.

Keywords: school emphasis on academic success, school climate, school leadership, students' achievement, multilevel hierarchical modelling

Introduksjon

Norge har lenge vært kjent for sitt likeverdige skolesystem (Frønes et al., 2020). Ifølge Boarini (2009) viste PISA-undersøkelsen² i 2009 at forskjellene i elevenes prestasjoner i norske skoler var blant de minste i OECD-landene. I Norge var forskjellen i gjennomsnittlig prestasjon mellom skolene litt over 10 prosent, mens OECD-gjennomsnittet var 33 prosent. I Tyskland var forskjellen over 60 prosent, og i Japan over 50 prosent (Boarini, 2009). Den høye graden av likhet mellom skolene i Norge er tett knyttet til den nordiske utdanningsmodellen (Frønes et al., 2020), som vektlegger inkludering og viktige samfunnsverdier som likhet og samarbeid. Den nordiske utdanningsmodellen legger grunnlaget for at alle elever, uavhengig av bakgrunn, kan få like muligheter gjennom skolen, som har en unik posisjon til å bidra til å redusere ulikheter ved å tilby støtte, læringsressurser og god undervisning (Blossing et al., 2014; Good & Weinstein, 1986).

Et positivt skolemiljø er avgjørende for å støtte elevenes prestasjoner og generelle trivsel i skolen. Ifølge de siste resultatene fra *Trends in International Mathematics and Science Study* (TIMSS) 2023 i Norge, rapporterer norske elever om høy grad av trivsel og tilhørighet på skolen sammenlignet med andre land. Likevel er dette lavere enn i forrige TIMSS-syklus i 2019 (Kaarstein et al., 2024). I tillegg gir nye funn fra *Lærerundersøkelsen*, gjennomført av Utdanningsdirektoratet, viktige innsikter i dagens tilstand i norske skolemiljø (Smedsvik & Wendelborg, 2024). Mens lærere generelt rapporterer om høy relevans i undervisningen, elevengasjement og lærerstøtte, uttrykkes det bekymringer om synkende elevmotivasjon og problemer med disiplin i klasserommet, særlig i ungdomsskolen (Smedsvik & Wendelborg, 2024). Spesifikt viser rapporten at skoler med flere lærere som er tilfredse med ledelsen er assosiert med sterkere skolemiljø, tydeligere regler og bedre elevmotivasjon (Smedsvik & Wendelborg, 2024). Slike funn kan sees i sammenheng med påstanden om at ledelse regnes som en av de viktigste faktorene for elevprestasjoner etter undervisningen i klasserommet (Leithwood et al., 2020). Skoleledelsens viktigste rolle er å skape et støttende skolemiljø som hjelper lærerne med å yte sitt beste. Det er innenfor dette området (bidrag til skolemiljø) vi finner de tydeligste resultatene som viser betydningen av skoleledelse (Hallinger, 2005; Robinson et al., 2008).

2 PISA = Programme for International Student Assessment

Skolemiljøet og kjennetegn ved rektorene er grunnleggende faktorer for studier av skoleledelse. Skoler og rektorer som klarer å fremme et positivt miljø, med tydelige mål, støttende relasjoner og oppmerksomhet rettet mot akademisk prestasjon, oppnår ofte høyere elevprestasjoner (Hallinger, 2005; Schneider & Bryk, 2002). Samtidig er det viktig å understreke at det ikke bare er rektorens ansvar å skape og opprettholde et slikt miljø (Schneider & Bryk, 2002; Veletić et al., 2023). Modeller for delt ledelse, som involverer ulike skoleaktører som lærere, administrasjon og foreldre, får økt anerkjennelse særlig i land som Norge, hvor samarbeid og delt ansvar er grunnleggende samfunnsverdier (Abrahamsen & Aas, 2016; Taneja-Johansson & Powell, 2024; Aas et al., 2024).

Videre peker UNESCOs nye *Global Education Monitoring Report* (2024) på betydningen av kjennetegn ved rektorers lederpraksis som igjen kan påvirke elevenes prestasjoner. I skoler der rektorene har mer erfaring, presterer elevene bedre i matematikk (Gümüş et al., 2024). Profesjonelle utviklingsprogrammer for skoleledere har vist seg å styrke positive kjennetegn/kompetanse hos lederne og å forbedre skolemiljø (Grissom et al., 2021). Norge har ligget bak for eksempel USA når det gjelder krav om utdanning av skoleledere, og i en sammenlignende studie (2006–2008) av skoleledelse i regi av OECD skilte Norge seg ut som et av få land som ikke hadde definert tydelige krav til formell utdanning av skoleledere (Pont, Nusche & Hopkins, 2008; Pont, Nusche & Moorman, 2008). Resultatet av studien bidro til en politisk endring gjennom utforming av sentrale retningslinjer for skolelederutdanning og en uttrykt forventning om at skoleledere tar skolelederutdanning (Møller, 2016). På begynnelsen av 2000-tallet ble det etablert masterprogram for skoleledelse ved flere universiteter og høyskoler i Norge, og i 2009 ble det opprettet nasjonal skolelederutdanning på 30 studiepoeng som norske skoleledere ble oppmuntret til å ta (Møller, 2016). Sistnevnte tilbud er blitt kalt rektorutdanningen. Den nasjonale rektorutdanningen har siden vist seg å styrke skolelederes praksis ved å forberede dem på å håndtere skoleutvikling og fremme en positiv skolekultur (Aas et al., 2024). Videre har den nasjonale skolelederutdanningen, sammen med andre kontekstuelle faktorer, styrket skolelederes evne til å anvende forskningsbasert kunnskap i beslutningstaking, fremme et delt profesjonelt språk i skolene sine og hjelpe dem i å lede skoleutviklingsprosjekter og håndtere motstand mot endring (Aas et al., 2024).

Ved å belyse rektorenes og skolemiljøets betydning for elevprestasjoner, har denne studien som mål å bidra til diskusjonen om utdanningsledelse,

skolemiljø og elevprestasjoner i den norske konteksten, samt å peke på viktige implikasjoner for praksis.

På bakgrunn av dette søker vi å besvare to forskningsspørsmål:

- 1 Hvor mye av variasjonen i prestasjoner i matematikk og naturfag på 9. trinn er innenfor og mellom skolene i Norge?

Det første forskningsspørsmålet gir oss innsikt i hvor stor del av variasjonen i elevenes prestasjoner i matematikk og naturfag som skyldes forskjeller mellom individuelle elever, og hvor mye som kan forklares med forskjeller mellom skolene. Enkelt sagt ønsker vi å finne ut om det finnes systematiske forskjeller mellom norske skoler når det gjelder elevenes prestasjoner i disse fagene.

- 2 I hvilken grad er det en sammenheng mellom kjennetegn ved rektorene (erfaring, utdanning), de ulike aspektene ved læringsfokus (elevens, for-elderens, lærerens) og elevenes prestasjoner i matematikk og naturfag på 9. trinn?

Det andre forskningsspørsmålet søker å identifisere hvilke skolerelaterte faktorer som kan bidra til å forklare forskjellene i elevenes prestasjoner mellom skolene. Vi ser spesielt på rektorenes erfaring og utdanning, samt aspekter ved skolemiljøet, og kontrollerer samtidig for elevenes sosioøkonomiske status (SØS) på både elev- og skolenivå. Når det gjelder skolemiljøet, analyserer vi mer konkret hvordan rektorens egne vurderinger av skolens læringsfokus (*school emphasis on academic success*) henger sammen med prestasjonsforskjeller i matematikk og naturfag mellom skolene.

Teoretisk bakgrunn

Faktorer som påvirker elevenes prestasjoner

Elevenes prestasjoner regnes som en viktig indikator på hvor gode og effektive skoler er, og har derfor motivert omfattende forskning på faktorer som påvirker prestasjoner på ulike nivåer i utdanningssystemet (Jarl et al., 2021; Kyriakides & Tsangaridou, 2008). Ifølge Creemers og Kyriakides (2007) er målet med forskning på utdanningens effektivitet å identifisere faktorer i undervisning, læreplaner og læringsmiljøer på forskjellige

nivåer – klasse, skole og overordnet nivå – som kan forklare forskjeller i elevenes prestasjoner.

Disse faktorene inkluderer både de som kan ha direkte og indirekte påvirkning, og tar også hensyn til kjennetegn ved elevenes bakgrunn, som individuelle trekk, SØS og tidligere prestasjoner i fag. Dette er fordi SØS er blant de viktigste prediktorene for elevprestasjoner på individnivå (Liu et al., 2022; Sirin, 2005). Samtidig viser forskning blandede resultater når det gjelder betydningen av skolens SØS. Effektene er generelt mindre når de måles på skole- eller kohortnivå enn på klassenivå, og forskere som ikke kontrollerer for tidligere prestasjoner eller mulige utelatte variabler overvurderer betydningen av slike komposisjonseffekter (van Ewijk & Sleegers, 2010).

I en tid preget av økende sosiale og demografiske ulikheter er det avgjørende å ta slike endringer på alvor. For eksempel viser den norske rapporten *En barndom for livet* (Rege, 2023) at elever med innvandrerbakgrunn, som også ofte har lavere SØS, har en tendens til å prestere dårligere enn sine norske jevnaldrende. For å møte disse utfordringene i skolene kan én tilnærming være å prioritere forbedring av skolemiljøet, da dette kan bidra til mer støttende og læringsfremmende omgivelser for både undervisning og læring (Maxwell et al., 2017).

Nyere studier har pekt på den viktige rollen faktorer som ledelse, forvaltning av ressurser og det generelle skolemiljøet spiller (se f.eks. Corral-Granados et al., 2023; Jarl et al., 2021; Özdemir et al., 2024). For eksempel har forskning fra norske skoler vist at skoler med flere ressurser, som høyere antall datamaskiner per elev, eller skoler som ligger i et spesifikt område der skolens ressurser generelt er bedre, har en tendens til å rapportere om høyere elevprestasjoner (Opheim et al., 2010). I tillegg er det mer sannsynlig at skoler med et positivt og støttende læringsmiljø fremmer bedre resultater for elevene på tvers av alle trinn (Opheim et al., 2010). Samtidig viser enkelte studier at skoler har utfordringer med å tilrettelegge for et best mulig skolemiljø til alle elever, uavhengig av sosial bakgrunn eller innvandrerbakgrunn (Bendixsen & Danielsen, 2020).

Skolemiljø – en vei til betydningsfull skoleledelse

Rektorer samarbeider stadig mer med lærere og andre involverte på skolen for å forbedre undervisning og læring, primært ved å skape de riktige forholdene innenfor skolen (Schneider & Bryk, 2002; Veletić et al., 2023).

Slike prosesser innebærer ifølge forskning blant annet å sette en klar visjon for skolen, forbedre undervisningsoppleggene og utvikle et positivt og støttende miljø som fremmer undervisning og læring (Hallinger, 2005; Printy & Marks, 2006; Zhan et al., 2023). I sin modell for pedagogisk ledelse viste Hallinger (2005) at rektorer som legger vekt på undervisning av høy kvalitet og mål knyttet til faglige prestasjoner og læring, som koordinerer læringsarbeidet, som vurderer og veileder undervisningen, samt støtter bruken av vurdering i klasserommet, anses som sterke pedagogiske ledere.

God ledelse og et godt skolemiljø er to sammenkoblede faktorer som begge har betydning. Skolemiljøet er en integrert del av ledelsen, og sammen påvirker de både skolens aktivitet og elevenes prestasjoner (se f.eks Pepper & Thomas, 2002). I en stor studie av skolemiljø (trygghet, fysisk miljø, akademisk miljø og felleskap), viser Wang og Degol (2016) at den viktigste faktoren for elevenes læring har å gjøre med akademisk miljø. Dette omfatter ledelse som er støttende overfor lærerne, og undervisningspraksiser som fokuserer på læring gjennom vurderinger, lærernes forventninger, elevenes motivasjon, muligheter for profesjonell utvikling samt evaluering av undervisningspraksiser (Wang & Degol, 2016).

I TIMSS måles akademisk miljø med begrepet *school emphasis on academic success* (SEAS), som vi i denne studien oversetter til læringsfokus.³ Dette innebærer at alle medlemmer av skolesamfunnet (elever, ansatte, skoleledere og foreldre) støtter læring (Yin & Reynolds, 2024). Dette kan gjøres ved å satse på lærerens forståelse av målene i læreplanen, foreldrenes engasjement i skoleaktiviteter, elevenes ønske om å gjøre det bra på skolen og så videre. Dette målet for læringsfokus i TIMSS er tilgjengelig i spørreskjemaene til både rektorer og lærere. I disse spørreskjemaene rapporterer rektorer og lærere om sine perspektiv på læringsfokus, inkludert deres vurderinger av elevers, foreldres og læreres læringsfokus (Yin & Reynolds, 2024). I denne studien har vi valgt å gå nærmere inn på rektorenes svar angående tre dimensjoner av læringsfokus: lærerens, elevens og foreldrenes læringsfokus.

Tidligere forskning har vist at læringsfokus samlet sett har vært positivt assosiert med elevenes prestasjon (se f.eks. Nilsen & Gustafsson, 2014). Når læringsfokus ble undersøkt på tvers av tre dimensjoner – lærere, foreldre

3 Vi bruker i denne studien begrepet «læringsfokus» for *school emphasis on academic success* i mangel på et bedre begrep på norsk. «Læringsfokus» er også i tråd med formålet med den norske utdanningen, og legger ikke så stor vekt på akademiske prestasjoner, men heller på læring og utvikling av kompetanse.

og elever – var det kun foreldredimensjonen som hadde en sterk sammenheng med elevenes prestasjoner (Gustafsson & Nilsen, 2016), til tross for noe variasjon der både foreldredimensjonen og elevdimensjonen hadde betydning (Badri, 2019). Lærerdimensjonen ved læringsfokus var imidlertid ikke signifikant. På skoleklassenivå fant Ye et al. (2024) til og med en negativ sammenheng mellom lærerens læringsfokus og elevenes presentasjoner.

Kjennetegn ved skoleledere

Den økende ansvarliggjøringen i en internasjonal utdanningskontekst, der rektorer stilles til ansvar for elevers prestasjoner, har bidratt til flere formelle krav til utdanning og profesjonell utvikling for skoleledere (Bush, 2018; Earley, 2016). I USA har utdanning av skoleledere i regi av egne utdanninger for skoleledere vært et satsingsområde siden 1950-tallet (Bush 2018; Gümüş et al., 2024; Jensen, 2023). Samtidig har erfaring som lærer tradisjonelt vært det viktigste kompetansekravet i rekruttering til rektorrollen (Bush, 2018).

Forskning på sammenhenger mellom skolelederes erfaring som lærer og elevers resultater finner ingen statistisk sammenheng som tilsier at rektors erfaring som lærer har betydning for elevers resultater (Gümüş et al., 2024). Videre har en rekke studier (se Gümüş et al., 2024) undersøkt sammenhenger mellom elevers prestasjoner og ulike faktorer som rektors høyeste formelle utdanningsnivå, rektors erfaring i andre mellomlederroller i skolen og antall års erfaring som rektor. Funnene fra slike studier er imidlertid preget av lav signifikans når det gjelder sammenheng, og dessuten av mangel på konsistente funn på tvers av studier (Gümüş et al., 2024). Det studier har funnet, er en tilsynelatende sammenheng mellom at rektor deltar i opplegg for kontinuerlig profesjonell utvikling (*continuous professional development*) rettet mot rektorrollen og spesifikke situasjoner som rektor står i (Gümüş et al., 2024). Slike opplegg kan gi muligheter for refleksjon over egen rolle, deltagelse i nettverk med andre eller utvikling av rektors praksis, som for eksempel observasjon og coaching. Resultatene er imidlertid også her varierende.

Tilsvarende som i en internasjonal sammenheng har det også i Norge vært tradisjon for at rektorer har vært erfarne, respekterte og dyktige lærere som påtar seg rollen som skoleleder (Møller, 2016). Som en følge av utvikling av masterprogram og rektorutdanning i Norge, har innstillingen til skolelederes erfaring og utdanning endret seg (Gunnulfsen & Skedsmo,

2024; Møller, 2016). I en ny evaluering av den nasjonale rektorutdanningen i Norge finner Aas et al. (2024) at skolelederutdanning «gir avtrykk i praksis» særlig når det gjelder ledelse av skoleutvikling og personalet. Rektorer spiller altså en viktig rolle ved å tilrettelegge for viktige prosesser som kan bidra til å skape et godt skolemiljø, og skolelederutdanning kan se ut til å ha betydning for dette.

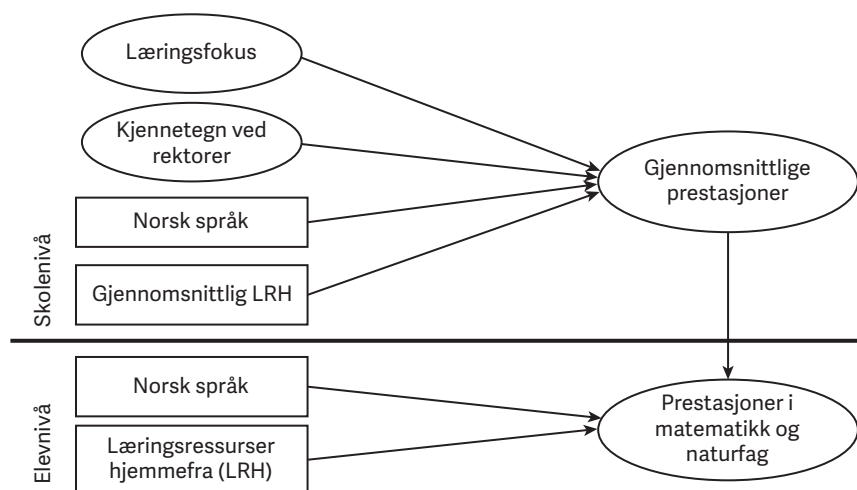
Denne studien

I dette kapittelet undersøker vi om det finnes forskjeller mellom skolene når det gjelder gjennomsnittlig prestasjon i matematikk og naturfag blant elever på 9. trinn i Norge. Ved hjelp av data fra TIMSS 2024, ser vi på om disse forskjellene kan forklares med skolemiljø og rektorenes erfaring og utdanning mens vi kontrollerer for elevenes sosioøkonomiske status både på individ- og skolenivå.

Mer spesifikt fokuserer vi på elevers, foreldres og læreres læringsfokus slik det rapporteres av rektorer, og hvordan dette henger sammen med elevenes prestasjoner på skolenivå. Videre undersøker vi kjennetegn ved ledere (antall år med erfaring, det å ha mastergrad i utdanningsledelse eller kursbevis⁴ i utdanningsledelse) og hvordan disse er knyttet til elevprestasjoner på skolenivå. Siden både skolemiljø og skoleledelse er skolefenomener, har vi i våre analyser brukt to-nivå strukturell likningsmodellering (SEM, *structural equation modelling*). Denne flernivåmetoden gjør det mulig å ta hensyn til det hierarkiske designet i TIMSS som reflekterer at elever tilhører klasser som hører inn under skoler, som igjen hører inn under land (Rutkowski et al., 2013). Modellen gjør det også mulig å analysere variabler på flere nivåer slik at relasjonen mellom ulike faktorer på forskjellige nivåer kan undersøkes samtidig og sees i forhold til prestasjon (se figur 4.1).

Siden vårt anliggende er forskjellene mellom skoler, retter vi søkelyset i analysene på rektorens rapporter om læringsfokus, fremfor lærernes rapportering som er også tilgjengelig i TIMSS. Vi kontrollerer også for elevenes bakgrunn på både elev- og skolenivå ved å inkludere elevenes hjemmespråk og læringsressurser hjemmefra i modellene våre.

4 Kursbevis kan være både studiepoenggivende kurs som rektorutdanning, eller det kan være andre typer utdanningsledelseskurs ved universiteter og høyskoler, eller utdanningslederkurs i regi av skoleeiere. Dette er ikke spesifisert i spørreskjemaet.



Figur 4.1. Den konseptuelle modellen for elev- og skolefaktorer som predikerer/indikerer prestasjoner i matematikk og naturfag

Metode

Data og utvalg

TIMSS er en internasjonal undersøkelse av elevers kompetanse i naturfag og matematikk på 5. og 9. trinn. Undersøkelsen samler inn elevenes prestasjoner og bakgrunnsdata, samt kontekstuell informasjon gjennom spørreskjemaer rettet mot foreldre, lærere og rektorer. For en mer detaljert beskrivelse av studien, vennligst se TIMSS sin internasjonale nettside og bakgrunnsinformasjon (<https://www.iea.nl/studies/iea/timss>) og kapittel 1 i denne boken.

Den foreliggende studien inkluderer et representativt utvalg av norske 9. trinnselever ($N = 6\,324$) og rektorer ved skolene deres ($N = 157$) som deltok i TIMSS 2023. Studien inkluderer både matematikk- og naturfagsresultater og rektorenes besvarelser på spørreskjema. 46,9 prosent av elevutvalget består av jenter. 73,5 prosent av elevene i utvalget snakker alltid norsk hjemme, 14,3 prosent snakker nesten alltid norsk hjemme, 8,3 prosent snakker noen ganger norsk hjemme, og 3,8 prosent snakker aldri norsk hjemme.

På grunn av TIMSS sitt hierarkiske design har vi i tillegg til flernivåmodellering også brukt vektorer for å sikre at resultatene er representative for alle elever i Norge.

Variabler

Alle variablene som er inkludert i studien, er beskrevet i tabell 4.1 og 4.2. Elevnivået representerer nivå 1, mens skolenivået representerer nivå 2.

Tabell 4.1. Variabler og konstrukter inkludert i studien på elevnivå (nivå 1)

Variabel	Beskrivelse	Eksempel-spørsmål	Skala	Verdier eller type skala
Prestasjoner i matematikk (avhengig variabel)	Fem plausible verdier	/	/	Numerisk, jo høyere verdi, desto høyere prestasjon
Prestasjoner i naturfag (avhengig variabel)	Fem plausible verdier	/	/	Numerisk, jo høyere verdi, desto høyere prestasjon
SØS⁵				
Norsk språk	Frekvens på bruk av norsk språk hjemme	Hvor ofte snakker du norsk hjemme?	1 = Alltid 2 = Nesten alltid 3 = Av og til 4 = Aldri	Jo høyere verdi, desto oftere snakker elevene norsk*
Læringsressurser hjemmefra	Skala basert på elevenes rapporter om antall bøker og annen studiestøtte i hjemmet, samt foreldrenes høyeste utdanningsnivå	Har du noe av dette hjemme hos deg? Ditt eget rom	1 = Ja 2 = Nei	Numerisk, jo høyere verdi, desto mer læringsressurser hjemmefra

*reversert kodet

Tabell 4.2. Variabler og konstrukter inkludert i studien på skolenivå (nivå 2)

Variabel	Beskrivelse	Eksempel-spørsmål	Skala	Verdier eller type skala
SØS				
Læringsressurser hjemmefra	Elevenes gjennomsnitt på skolenivå	Har du noe av dette hjemme hos deg? Ditt eget rom	1 = Ja 2 = Nei	Numerisk, jo høyere verdi, desto mer læringsressurser hjemmefra

(Forts.)

5 I denne studien benyttes norsk språk og tilgjengelighet på læringsressurser hjemmefra som indikatorer på sosio-økonomisk status (SØS). Det er imidlertid viktig å understreke at disse faktorene kun utgjør forenklede mål og ikke fullt ut reflekterer de komplekse og mangfoldige aspektene ved SØS. Dermed fungerer de som proxyer snarere enn fullstendige representasjoner av sosio-økonomisk bakgrunn.

Tabell 4.2. (Forts.)

Variabel	Beskrivelse	Eksempel-spørsmål	Skala	Verdier eller type skala
Norsk språk	Elevens gjennomsnitt på skolenivå	Hvor ofte snakker du norsk hjemme	1 = Alltid 2 = Nesten alltid 3 = Av og til 4 = Aldri	Numerisk, jo høyere verdi, desto mer snakkes norsk hjemme*
Kjennetegn ved rektorer				
Erfaring på skolen	Antall år rektorene har vært ved en bestemt skole	Hvor mange år har du vært rektor på denne skolen ved slutten av dette skoleåret		Numerisk
Kursbevis i utdanningsledelse	Hvis rektor har kursbevis i utdanningsledelse	Har du følgende kvalifikasjoner eller attester i utdanningsledelse? Kursbevis	1 = Ja 2 = Nei	Nominal
Master i utdanningsledelse	Hvis rektor har mastergrad i utdanningsledelse	Har du følgende kvalifikasjoner eller attester i utdanningsledelse? Mastergrad	1 = Ja 2 = Nei	Nominal
Læringsfokus				
Lærerens læringsfokus	Skala (latent variabel) basert på rektorens rapportering av lærerens læringsfokus	Hvordan vil du karakterisere det følgende på skolen? Lærernes forståelse av målene i læreplanen.	1 = Svært høy 2 = Høy 3 = Middels 4 = Lav 5 = Svært lav	Numerisk, jo høyere verdi, desto sterkere fokus*
Foreldrenes læringsfokus	Skala (latent variabel) basert på rektorens rapportering av foreldrenes læringsfokus	Hvordan vil du karakterisere det følgende på skolen? Foreldrenes engasjement i skoleaktiviteter	1 = Svært høy 2 = Høy 3 = Middels 4 = Lav 5 = Svært lav	Numerisk, jo høyere verdi, desto sterkere fokus*
Elevenes læringsfokus	Skala (latent variabel) basert på rektorens rapportering av elevenes læringsfokus	Hvordan vil du karakterisere det følgende på skolen? Elevenes ønske om å lykkes på skolen		Numerisk, jo høyere verdi, desto sterkere fokus*

*reversert kodet

Den primære avhengige variabelen i denne studien er TIMSS matematikk- og naturfagsprestasjoner, som er standardisert til et gjennomsnitt på 500 poeng og et standardavvik på 100 poeng. TIMSS bruker et rotasjonsdesign

for oppgavene, noe som gjør at undersøkelsene dekker et mye bredere spekter av fagenes emneområder enn det elevene faktisk svarte på. Videre bruker TIMSS *item response theory* (IRT) for å håndtere at ikke alle elever svarer på hver prestasjonsoppgave. Det innebærer å generere fem plausible verdier for å estimere et spekter av elevers karakteristikk basert på både elevenes skår på oppgavene i TIMSS-testen og på bakgrunnsinformasjon som for eksempel SØS (Wu, 2005).

Statistiske analyser

IEAs IDB Analyzer er en programvare som først ble brukt for å slå sammen student- og skoledata og alle analyser ble gjennomført ved hjelp av Mplus versjon 8.4 (L. K. Muthén & Muthén, 2017) med *robust maximum likelihood (MLR) estimator*. De manglende dataene ble håndtert ved bruk av *full information maximum likelihood procedure* (FIML), som er implementert i Mplus. For å ta hensyn til TIMSS' hierarkiske design og uforholdsmessige utvalgssannsynligheter, har vi brukt to-nivåmodellering og videre utvalgsvekter både på skole- og studentnivå (Rutkowski et al., 2013).

Preliminære analyser inkluderte *konfirmerende faktoranalyse* (CFA). CFA brukes for å vurdere strukturen på teoretiske konstrukter (faktorer eller underliggende dimensjoner) som ikke kan observeres eller måles direkte (Brown, 2015). For eksempel bruker vi mange indikatorer (observerte svar på spørsmål) som til sammen kan dekke det vi mener med læringsfokus. For å evaluere hvorvidt modellen passer til observert data refererte vi til kjente retningslinjer (Hu & Bentler, 1999).⁶

Videre analyserte vi relasjoner mellom faktorer ved bruk av fler-nivå strukturell ligningsmodellering (MSEM) (B. Muthén & Asparouhov, 2011). For å svare på det første forskningsspørsmålet startet vi med å kjøre en ukondisjonert (tom) modell, også kjent som nullmodellen, for å dekomponere den totale variansen i elevprestasjon (matematikk og naturfag i to forskjellige modeller) til variansen på elevnivået og variansen på skolenivået. Nullmodellen gir oss en første pekepinn på om skolene i Norge skiller seg fra hverandre når det gjelder elevenes prestasjoner, eller om forskjellene stort sett finnes mellom elevene på samme skole.

6 Det vil si *chi-square* (χ^2), *comparative fit index* (CFI) $\geq 0,95$; *Tucker-Lewis index* (TLI) $\geq 0,95$; *root mean square error of approximation* (RMSEA) $\leq 0,08$; *standardized root mean square residual* (SRMR) $\leq 0,08$ (Hu & Bentler, 1999).

For å svare på det andre forskningsspørsmålet ble nullmodellen brukt som utgangspunkt, og deretter utvidet ved å inkludere prediktorer (uavhengige variabler) på både elevnivå og skolenivå. Deretter inkluderte vi først SØS-proxy-variabel på elevnivå, og utviklet modellen ved å inkludere SØS-proxy-variabel på skolenivå. Deretter la vi til rektorers erfaring og utdanning, og til slutt tre aspekter ved læringsfokus, i tre separate modeller. Modellen ble bygget gradvis, noe som gav oss mulighet til å vurdere effekten av de ulike prediktorene på varianskomponentene samt på modellens tilpasning (Dedrick et al., 2009).

Resultater

Resultatene er organisert i tre seksjoner. Først presenterer vi deskriptiv statistikk og målemodeller for våre tre latente variable (se tabell 4.2), deretter resultatene for hver av forskningsspørsmålene. Siden dataene våre inkluderer sammenslåtte data for elever og rektorer, skal vi når vi presenterer og diskuterer disse, for eksempel når det gjelder hvor mange rektorer som har mastergrad i utdanningsledelse, merke oss at disse ikke representerer antall rektorer med godkjent mastergrad i utdanningsledelse, men heller antall elever som er i skoler der rektor har en slik kvalifisering. I de andre avsnittene beskriver vi resultatene på en lignende måte.

Deskriptiv statistikk og målemodeller

Undersøkelsen av deskriptive statistikkene for læringsfokus-indikatorer viste ingen sterke tak- eller bunneffekter, men den viste at rektorer generelt svarte mer positivt på utsagnene (for detaljer, se vedlegg 4.1). For eksempel rapporterte rektorene at lærerne har god forståelse av målene i læreplanen, og at de er flinke til å sette disse målene ut i praksis. Det som også er interessant, er at rektorene pekte på at lærerne har høye forventninger til elevenes prestasjoner, og at de også har en sterk evne til å inspirere elevene. Når det gjelder foreldrene, er bildet litt annerledes. Rektorene nevnte at foreldrene har middels høye forventninger til elevenes prestasjoner, og de gir også en moderat mengde støtte til elevenes faglige utvikling. Samtidig er de relativt engasjerte i skoleaktiviteter. Men når det gjelder viljen til å ta ansvar for at elevene møter godt forberedt på skolen, er ikke den så høy, selv om det fortsatt er et visst engasjement. Elevene selv, derimot, ifølge rektorenes rapportering, har et moderat til høyt ønske om å gjøre det bra på skolen og en god evne til å nå skolens faglige mål.

Når det gjelder utdanning og erfaring for norske rektorer, viser dataene at rundt 76 prosent av elevene går på skoler hvor rektor har kursbevis i utdanningsledelse, og 43,6 prosent går på skoler der rektor har mastergrad i utdanningsledelse. Det kan også være mulig at en rektor har *både* kursbevis og mastergrad i utdanningsledelse. Rektorene har i gjennomsnitt 5,55 års erfaring i den nåværende skolen, med et standardavvik på 4,25.

Resultatene fra faktoranalysene samsvarer med det vi tenkte å måle i utgangspunktet, og det er tre dimensjoner av læringsfokus: læreres, foreldres og elevers læringsfokus. Ifølge de vanlige passformindeksene viste alle modellene seg å passe bra til dataene, inkludert for lærere ($\chi^2 = 2,414$; $df = 2$; CFI = 0,997; TLI = 0,992; RMSEA = 0,040; SRMR = 0,020), elever ($\chi^2 = 0,000$; $df = 0$; CFI = 1,000; TLI = 1,000; RMSEA = 0,000; SRMR = 0,020) og foreldre ($\chi^2 = 0,769$; $df = 2$; CFI = 1,000; TLI = 1,000; RMSEA = 0,040; SRMR = 0,012).

Hvor mye av variansen i prestasjoner i matematikk og naturfag på 9. trinn er innenfor og mellom skolene i Norge?

For å svare på forskningsspørsmål 1 kjørte vi en nullmodell for å undersøke hvor mye av variansen i matematikk- og naturfagprestasjoner som ligger mellom skolene, sammenlignet med den innen skolene. Nullmodellen er et enkelt utgangspunkt som ikke inkluderer noen forklaringsvariabler, men som hjelper oss å se hvordan variasjonen i matematikk- og naturfagprestasjoner fordeler seg før vi begynner å lete etter årsaker. Analysene viste at det er en tydelig variasjon i matematikk- og naturfagprestasjoner, med en betydelig andel av variasjonen som ligger mellom skolene.

Spesifikt viste analysene at *intraclass correlation* (ICC), altså andelen av den totale variansen som kan tilskrives forskjeller mellom skoler, var 0,081 for matematikkprestasjoner (modell 1, tabell 4.3) og 0,087 for naturfagprestasjoner (modell 1, tabell 4.4). Dette betyr at litt over 8 prosent (matematikk) og 9 prosent (naturfag) av variasjonen i elevenes prestasjoner kan relateres til systematiske forskjeller mellom skolene.

Selv om 8–9 prosent kanskje høres lite ut, er det faktisk en betydelig andel når vi ser på store elevgrupper. Det betyr at hvilken skole en elev går på, faktisk kan ha en viss betydning for hvordan de presterer, selv om det meste av variasjonen (90 %) finnes mellom elevene på samme skole. Med andre ord: Elever på en og samme skole kan prestere veldig ulikt, men det finnes også forskjeller mellom skolene som er verdt å undersøke nærmere.

Dette funnet stemmer godt overens med de tidligere PISA-resultatene, som viser at de i Norge, sammenlignet med andre land, er relativt små forskjeller mellom skolene (Boarini, 2009). Likevel er forskjellene som faktisk finnes, betydningsfulle nok til at det er verdt å undersøke hva som kan forklare dem. Dette ser vi nærmere på i de videre analysene.

I hvilken grad er det en sammenheng mellom kjennetegn ved rektorene (erfaring, utdanning), de tre dimensjonene ved læringsfokus (elevens, forelderens, lærers), og elevenes prestasjoner i matematikk og naturfag på 9. trinn?

I alle våre modeller kontrollerer vi for variablene læringsressurser hjemmefra og hvor ofte det ble snakket norsk hjemme som vi brukte som proxy for SØS. Læringsressurser hjemmefra hadde en sterk positiv effekt på matematikk- og naturfagsprestasjoner og indikerer at elever med flere utdanningsressurser hjemme presterte bedre i begge fagene. Oftere bruk av norsk språk hjemme var en signifikant prediksjon på bedre matematikk- og naturfagsprestasjoner, men i mindre grad enn læringsressurser hjemmefra. På elevnivå kunne rundt 14 prosent av variansen i matematikkprestasjoner og 15 prosent i naturfagprestasjoner forklares av disse to prediktorene (se tabell 4.3 og 4.4).

På skolenivå kontrollerer vi for gjennomsnittet av læringsressurser hjemmefra og gjennomsnittet for hvor ofte norsk ble snakket hjemme, men ingen av disse var signifikante (se modell 2 i tabell 4.3; modell 2 i tabell 4.4). Med andre ord kunne ikke gjennomsnittet av læringsressurser hjemmefra og gjennomsnittet for hvor ofte norsk ble snakket hjemme på skolenivå signifikant forklare forskjeller mellom skolene i matematikk og naturfag mens vi kontrollerer for SØS på individnivå.

Blant kjennetegn ved rektorer var år med erfaring i nåværende skolen og mastergrad i utdanningsledelse positivt og signifikant relatert til matematikk- og naturfagsprestasjoner også når man kontrollerte for individuell- og skole-SØS (modell 3 i tabell 4.3; modell 3 i tabell 4.4.) Imidlertid predikerte rektorenes kursbevis i utdanningsledelse ikke signifikant når det gjelder verken matematikk- eller naturfagsprestasjoner (se modell 2 og 3 i tabell 4.3; modell 2 og 3 i tabell 4.4).

Lærernes læringsfokus som rektorer rapporterte om var ikke signifikant assosiert med matematikkprestasjoner, eller naturfagprestasjoner når

Tabell 4.3. Estimerte standardiserte regresjonskoeffisienter for sammenhengen mellom matematikkprestasjoner og elev- og skolenivå-prediktorer (utvalgte modeller)

Matematikk	MODELL 1	MODELL 2	MODELL 3	MODELL 4	MODELL 5	MODELL 6
	Nullmodell	N1: SØS N2: SØS	Kun rektorens variabel på N2	N1: SØS N2: SØS, kjennetegn ved rektorer og lærerens læringsfokus	N1: SØS N2: SØS, kjennetegn ved rektorer og forelderens læringsfokus	N1: SØS N2: SØS, kjennetegn ved rektorer og elevenes læringsfokus
	β	β	β	β	β	β
	SE	SE	SE	SE	SE	SE
Intercept (5PV matematikk)	501,35***					
Elevnivå						
Læringsressurser hjemmefra	0,360**	0,020		0,360**	0,020	0,360**
Norsk språk	0,053**	0,017		0,001**	0,018	0,045**
Skolenivå						
Læringsressurser hjemmefra	0,237	0,100		0,171	0,174	0,164
Norsk språk	0,103	0,174		0,006	0,149	-0,017
Rektorens erfaring			0,360**	0,365**	0,133	0,392**
Kursbevis i utdanningsledelse			0,044	0,101	0,115	0,083
Mastergrad i utdanningsledelse			0,257**	0,113	0,127	0,226
Lærerens læringsfokus				0,225	0,157	
Forelderens læringsfokus					0,292*	0,128
Elevenes læringsfokus						0,291
Forklart andel av varians (R2)						
Mellom elevene	0,137***	0,013		0,137***	0,014	0,137***
Mellom skolene	0,083	0,082	0,214*	0,308**	0,103	0,325**
ICC		0,081			0,117	

Note. N1 – elevnivå, N2 – skolenivå; B = standardisert regresjonskoeffisient, SE = standardfeil av et gjennomsnitt ***p < .001; **p < .01; *p < .05.

Tabell 4.4. Estimerte standardiserte regresjonskoeffisienter for sammenhengen mellom naturfagprestasjoner og elev- og skolenivå prediktorer (utvalgt modeller)

Naturfag	MODELL 1	MODELL 2	MODELL 3	MODELL 4	MODELL 5	MODELL 6
	Null- modell	N1: SØS N2: SØS	Kun rektorens variabel på N2	N1: SØS N2: SØS, kjennetegn ved rektorer og lærerens læringsfokus	N1: SØS N2: SØS, kjennetegn ved rektorer og forelderens læringsfokus	N1: SØS N2: SØS, kjennetegn ved rektorer og eleveneslæringsfokus
Intercept (5PV naturfag)						
	β	β	SE	β	SE	β
	SE	SE		β	SE	SE
Elevnivå						
Læringsressurser hjemmefra	0,37***	0,016		0,37***	0,019	0,37***
Norsk språk	0,061***	0,020		0,061**	0,024	0,061**
Skolenivå						
Læringsressurser hjemmefra	0,323	0,173		0,283	0,17	0,283
Norsk språk	0,202	0,161		0,088	0,132	0,099
Rektorens erfaring			0,339**	0,321**	0,13	0,329**
Kursbevis i utdanningsledelse			-0,004	0,028	0,031	0,04
Mastergrad i utdanningsledelse			0,345**	0,321**	0,115	0,295*
Lærerens læringsfokus			0,105	0,008	0,125	
Forelderens læringsfokus					0,129	
Elevenes læringsfokus					0,114	
					0,16	0,136
Forklart andel av varians (R2)						
Mellom elevene	0,151***	0,013		0,149***	0,015	0,149***
Mellom skolene	0,183	0,099		0,364***	0,104	0,381**
ICC	0,087		0,113	0,254***	0,104	0,111

Note. N1 – elevnivå, N2 – skolenivå; B = standardisert regresjonskoeffisient, SE = standardfeil av et gjennomsnitt, ***p < .001, **p < .01, *p < .05.

SØS og rektorens variabler ble inkludert i modellen samtidig (modell 4 i tabell 4.3; tabell 4.4). Disse faktorene sammen forklarte rundt 14 prosent av variansen i matematikkprestasjoner på elevnivå og rundt 30 prosent av variansen i matematikkprestasjoner mellom skolene. For naturfagprestasjoner ble omtrent 15 prosent av variansen forklart på elevnivå, mens 36 prosent av variansen ble forklart mellom skolene av de samme faktorene. Videre inkluderte vi (1) foreldres læringsfokus (modell 5 i tabell 4.3; tabell 4.4). Denne faktoren var signifikant assosiert med matematikkprestasjoner, men ikke med naturfagprestasjoner. Vi inkluderte også (2) elevers læringsfokus (modell 6 i tabell 4.3; tabell 4.4), som ikke var signifikant relatert til verken matematikk eller naturfag. Det er også viktig å påpeke at mastergrad i utdanningsledelse ikke lenger er signifikant når foreldrenes læringsfokus er inkludert i modellen.

Diskusjon

Denne studien har hatt som mål å utforske om det er forskjeller mellom norske skoler når det gjelder gjennomsnittlig prestasjon i matematikk og naturfag, samt å videre utforske forholdet mellom kjennetegn ved rektorer, skolemiljø og sammenheng mellom skolene når det gjelder 9. trinnselevers matematikk- og naturfagprestasjoner. Funnene våre understreker viktigheten av både individuelle elevfaktorer og skolefaktorer når det gjelder å forstå og legge til rette for elevers prestasjon, og fremhever spesielt betydningen av rektorenes profesjonelle bakgrunn. Selv om vi ikke direkte fokuserte på å undersøke påvirkningen av SØS på elevenes prestasjoner, kontrollerte vi for elevers SØS i alle modellene våre, slik at vi inkluderte læringsressurser hjemmefra og frekvensen av norsk som snakkes hjemme som kontrollvariabler.

Analysen viste at en betydelig andel av variansen i matematikkprestasjoner (8 %) og naturfagprestasjoner (9 %) kan relateres til systematiske forskjeller mellom norske skoler. Selv om andelen av variansen som ligger mellom skolene ikke er stor sammenlignet med andre europeiske land, er den stor nok til å kunne konkludere med at skolens kontekst har en reell betydning for elevenes matematikk- og naturfagsprestasjoner. Det reflekterer også forskjeller mellom skolene, som kan skyldes ulikheter i ledelse, miljø, ressurser og så videre (Blossing et al., 2014; Good & Weinstein, 1986; Opheim et al., 2010).

I neste steg av analysen undersøkte vi forskjellene mellom skolene ved å se på kjennetegn ved rektorer, som erfaring og kvalifikasjoner innen

utdanningsledelse, og vi fant en positiv sammenheng mellom kjennetegn ved rektorene og elevenes prestasjoner i både matematikk og naturfag. Særlig var rektors erfaring ved skolen og en mastergrad i utdanningsledelse signifikante prediktorer for bedre elevresultater, noe som understreker betydningen av erfarne og godt kvalifiserte ledere for å fremme læring. Dette resultatet er ulikt tidligere studier, som viser varierende sammenhenger mellom rektors erfaring og elevresultater (Gümüş et al., 2024; Teig et al., 20204).

Funnene våre viser at rektorer med mer erfaring, spesielt de som har vært lenge på samme skole, ser ut til å bidra til forskjeller i skoleprestasjoner i Norge. Dette understreker viktigheten av tid når det gjelder skoleledelse – det tar tid å etablere gode rutiner og å utøve god ledelse. Tilsvarende fant Branch et al. (2013) at skoler med gode elevprestasjoner med høy sannsynlighet hadde hatt samme rektor over flere år. Basert på våre funn, er en implikasjon at det kan være hensiktsmessig å beholde rektorer på samme skole i lengre perioder.

Videre viser resultatene at en mastergrad i utdanningsledelse er en viktig prediktor på prestasjonsforskjellene mellom skolene. Vi fant imidlertid at kun 43,6 prosent av norske elever går på skoler der rektor har en mastergrad i utdanningsledelse. Derimot har mange rektorer kursbevis i utdanningsledelse (76 %), som man kan anta at de har fra den nasjonale rektorutdanningen (30 stp.), eller fra lederkurs iverksatt av skoleeiere, men vi fant at dette ikke var en relevant faktor for å forklare forskjellene mellom skolene. En mulig forklaring på dette kan være at rollen som skoleleder har blitt mer kompleks over tid (Earley, 2016), og at kortere kurs eller studier ikke er tilstrekkelige for å håndtere de utfordringene som følger med lederansvaret. Lengre utdanning, som en mastergrad, har vist seg å være viktig for å kunne møte disse utfordringene effektivt. En annen mulig forklaring er at rektorer som har mastergrad, kanskje kan være mer motiverte eller dedikerte, ettersom dette ikke er et obligatorisk utdanningsprogram i Norge. Dette kan føre til en selvseleksjonseffekt, hvor de som velger å ta mastergrad, skiller seg fra de som ikke gjør det – på måter som også kan ha betydning for deres kjennetegn som rektorer og elevers skoleprestasjoner. Dermed kan den observerte effekten av masterutdanning delvis skyldes underliggende forskjeller i motivasjon eller kompetanse, mer enn selve graden (Gümüş et al., 2024).

Til tross for betydningen av kjennetegn ved rektorene, som for eksempel mastergrad i utdanningsledelse, fant studien begrenset bevis for direkte sammenhenger mellom læringsfokus og elevers prestasjoner. Vi fant at

blant de tre dimensjonene ved læringsfokus, var det kun foreldrenes fokus på prestasjon som viste seg å være signifikant. Dette er i samsvar med funn fra tidligere forskning i den norske konteksten (Gustafsson & Nilsen, 2016; Nilsen & Gustafsson, 2014) og interessant nok kun for matematikk. En mulig forklaring kan være at foreldrene legger mer press på elevene for å gjøre det bra i matematikk enn i naturfag, ettersom naturfag ikke er en del av eksamen (Utdanningsdirektoratet, 2017). Inkluderer vi både rektors mastergrad i utdanningsledelse og foreldrenes læringsfokus, forsvinner likevel mastergradens betydning. Dette betyr at mastergrad i utdanningsledelse og foreldrenes læringsfokus henger sammen på en eller annen måte. Det kan være at foreldrenes læringsfokus fungerer medierende for effekten av mastergrad i utdanningsledelse på matematikk og naturfag. Enkelt sagt kan det være at rektorer med en mastergrad i utdanningsledelse skaper et mangfold av kommunikasjonsmåter og finner effektive måter å bidra til at foresatte legger mer vekt på prestasjoner, noe som resulterer i bedre elevprestasjoner (Bendixsen & Danielsen, 2020). Alternativt kan et slikt resultat også reflektere at noen rektorer jobber i fylker som både støtter masterutdanning og fremmer foreldrenes bidrag til skolen, noe som kan gi en pekepinn på at at rektorene må finne kommunikasjonsmetoder som når ut til alle foresatte. Dette er noen mulige forklaringer, men slike slutninger ligger utenfor rammene av dette kapittelet. Det er likevel noe som kan undersøkes nærmere.

Skolefaktorer på skolenivå, som læringsressurser hjemmefra og språk snakket hjemme, forklarte ikke signifikant forskjeller i faglige prestasjoner mellom skolen når vi kontrollerer for det samme på elevnivå. Imidlertid ble det også funnet at individuelle elevfaktorer, særlig læringsressurser hjemmefra og språk snakket hjemme, spiller en viktig rolle. Læringsressurser hjemmefra ble konsekvent identifisert som sterke prediktorer for prestasjoner i begge fagene, noe som fremhever betydningen av elevenes hjemmeforhold. Disse funnene samsvarer med litteraturen om viktigheten av sosial kontekst i utdanningsresultater (se for eksempel Kyriakides & Tsangaridou, 2008; Rege, 2023).

Vår forskning viser på samme måte som tidligere veletablert forskning at sosioøkonomisk status og læringsressurser hjemmefra er blant de viktigste prediktorene for elevprestasjoner på individnivå (Liu et al., 2022; Sirin, 2005). I dette kapittelet har vi i tillegg vist den viktige rollen lederne spiller. Elevenes bakgrunn er en betydelig del av deres forutsetninger, men skolen har likevel et stort handlingsrom for å fremme elevenes læring og trivsel.

Funnene i denne studien understreker verdien av å investere i erfarne og godt kvalifiserte skoleledere, ettersom deres bakgrunn kan ha stor betydning på skolemiljøet og for elevenes resultater. Videre kan det være nyttig å oppmuntre og støtte rektorer til å ta videreutdanning, for eksempel gjennom mastergradsprogrammer, som kan bidra til å styrkederes kjennetegn som ledere, deres kommunikasjon og pedagogiske forståelse. Et støttende skolemiljø og gode ledelsespraksiser kan ha en betydelig positiv innvirkning på elevenes utvikling og læring.

Forfatterbiografier

Jelena Veletić er postdoktor ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har doktorgrad i utdanningsmåling og bakgrunn fra psykologi. Hun arbeider særlig med analyser av storskala utdanningsdata, og forskningsinteressene hennes omfatter skoleledelse, skolemiljø, læreres jobbtilfredshet og andre forhold som påvirker lærere.

Jelena Radišić er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Der har hun vært siden 2016. Hennes arbeid er knyttet til forskning på elev-, lærer- og skolekarakteristika som påvirker elevenes resultater. Hun er spesielt opptatt av hvordan elevenes læring påvirkes av motivasjon, selvtillit i matematikk, akademiske følelser, læreres holdninger og undervisningspraksis. I sin forskning benytter Radišić en rekke ulike tilnærminger som inngår i kvantitativ, kvalitativ og blandet metodeforskning.

Hilde Marie Madsø Jacobsen er universitetslektor ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har skrevet en doktorgrad om veiledning av nyutdannede og ledelse i skolen. Hun underviser i utdanningsledelse og veiledning, og har forskningsinteresser knyttet til veiledning av nyutdannede lærere og skoleledelse. Hilde har gitt ut flere artikler om temaet veiledning av nyutdannede lærere.

Referanser

- Abrahamsen, H. & Aas, M. (2016). School leadership for the future: Heroic or distributed? Translating international discourses in Norwegian policy documents. *Journal of Educational Administration and History*, 48(1), 68–88. <https://doi.org/10.1080/00220620.2016.1092426>

- Badri, M. (2019). School emphasis on academic success and TIMSS science/math achievements. *International Journal of Research in Education and Science*, 5(1), 176–189.
- Bendixsen, S. & Danielsen, H. (2020). Great expectations: Migrant parents and parent–school cooperation in Norway. *Comparative Education*, 56(3), 349–364. <https://doi.org/10.1080/03050068.2020.1724486>
- Blossing, U., Imsen, G. & Moos, L. (2014). Nordic Schools in a time of change. I U. Blossing, G. Imsen & L. Moos (Red.), *The Nordic education model: "A school for all" encounters neo-liberal policy* (s. 1–14). Springer Netherlands. https://doi.org/10.1007/978-94-007-7125-3_1
- Boarini, R. (2009). *Making the most of Norwegian schools*. OECD Publishing. <https://doi.org/10.1787/227333554837>
- Branch, G. R., Hanushek, E. A. & Rivkin, S. G. (2013). School leaders matter: Measuring the impact of effective principals. *Education Next*, 13(1), 62–69. <https://hanushek.stanford.edu/publications/school-leaders-matter-measuring-impact-effective-principals>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2. utg.). The Guilford Press.
- Bush, T. (2018). Preparation and induction for school principals: Global perspectives. *Management in Education*, 32(2), 66–71. <https://doi.org/10.1177/0892020618761805>
- Corral-Granados, A., Rapp, A. C. & Smeplass, E. (2025). Nordic challenges related to exclusion and local responses in Swedish, Finnish, and Norwegian urban compulsory education. *Education Inquiry*, 16(1), 37–68. <https://doi.org/10.1080/20004508.2022.2163002>
- Creemers, B. & Kyriakides, L. (2007). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. <https://doi.org/10.4324/9780203939185>
- Dedrick, R. F., Ferron, J. M., Hess, M. R., Hogarty, K. Y., Kromrey, J. D., Lang, T. R., Niles, J. D. & Lee, R. S. (2009). Multilevel modeling: A review of methodological issues and applications. *Review of Educational Research*, 79(1), 69–102. <https://doi.org/10.3102/0034654308325581>
- Earley, P. (2016). Global trends and challenges for school leaders: Keeping the focus on learning. *Journal of Educational, Cultural and Psychological Studies*, (14), Artikkel 14. <https://doi.org/10.7358/ecps-2016-014-earl>
- Frønes, T. S., Pettersen, A., Radišić, J. & Buchholtz, N. (Red.). (2020). *Equity, equality and diversity in the Nordic model of education*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-61648-9>
- Good, T. L. & Weinstein, R. S. (1986). Schools make a difference: Evidence, criticism, and new directions. *American Psychologist*, 41(10), 1090–1097. <https://doi.org/10.1037/0003-066X.41.10.1090>
- Grissom, J., Egalite, A. & Lindsay, C. (2021). *How principals affect students and schools* (Rapport). Wallace Foundation.
- Gunnulfsen, A. E. & Skedsmo, G. (2023). School leadership in Norway: Key characteristics and current challenges. I A. E. Gunnulfsen, H. Årlestig & M. Storgaard (Red.), *Education and democracy in the Nordic countries: Making sense of school leadership, policy, and practice* (s. 61–73). Springer International Publishing. https://doi.org/10.1007/978-3-031-33195-4_5
- Gustafsson, J. E. & Nilsen, T. (2016). The impact of school climate and teacher quality on mathematics achievement: A difference-in-differences approach. I T. Nilsen & J.-E. Gustafsson (Red.), *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (s. 81–95). Springer International Publishing. https://doi.org/10.1007/978-3-319-41252-8_4
- Gümüş, S., Bellibaş, M. S., Şen, S. & Hallinger, P. (2024). Finding the missing link: Do principal qualifications make a difference in student achievement? *Educational Management Administration & Leadership*, 52(1), 28–51. <https://doi.org/10.1177/17411432211051909>
- Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools*, 4(3), 221–239. <https://doi.org/10.1080/15700760500244793>

- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jarl, M., Andersson, K. & Blossing, U. (2021). Organizational characteristics of successful and failing schools: A theoretical framework for explaining variation in student achievement. *School Effectiveness and School Improvement*, 32(3), 448–464. <https://doi.org/10.1080/09243453.2021.1903941>
- Jensen, F., Pettersen, A., Frønes, T. S., Eriksen, A., Løvgren, M. & Narvhus, E. K. (2023). *PISA 2022. Norske elevers kompetanse i matematikk, naturfag og lesing*. Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.205>
- Kaarstein, H., Lehre, A.-C., Radisic, J. & Rohatgi, M. A. (2024). *TIMSS 2023. Kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Kyriakides, L. & Tsangaridou, N. (2008). Towards the development of generic and differentiated models of educational effectiveness: A study on school and teacher effectiveness in physical education. *British Educational Research Journal*, 34(6), 807–838. <https://doi.org/10.1080/01411920802041467>
- Leithwood, K., Harris, A. & Hopkins, D. (2020). Seven strong claims about successful school leadership revisited. *School Leadership & Management*, 40(1), 5–22. <https://doi.org/10.1080/13632434.2019.1596077>
- Liu, J., Peng, P., Zhao, B. & Luo, L. (2022). Socioeconomic status and academic achievement in primary and secondary education: A meta-analytic review. *Educational Psychology Review*, 34(4), 2867–2896. <https://doi.org/10.1007/s10648-022-09689-y>
- Maxwell, S., Reynolds, K. J., Lee, E., Subasic, E. & Bromhead, D. (2017). The impact of school climate and school identification on academic achievement: Multilevel modeling with student and teacher data. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.02069>
- Muthén, B. & Asparouhov, T. (2011). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. I J. Hox & J. K. Roberts (Red.), *Handbook for advanced multilevel analysis* (s. 15–40). Routledge.
- Muthén, L. K. & Muthén, B. O. (2017). *Mplus user's guide* (8. utg.). Muthén & Muthén. https://www.statmodel.com/html_ug.shtml
- Møller, J. (2016). Kvalifisering som skoleleder i en norsk kontekst: Et historisk tilbakeblikk og perspektiver på utdanning av skoleledere. *Acta Didactica Norge*, 10(4), Artikkel 4. <https://doi.org/10.5617/adno.3871>
- Nilsen, T. & Gustafsson, J.-E. (2014). School emphasis on academic success: Exploring changes in science performance in Norway between 2007 and 2011 employing two-level SEM. *Educational Research and Evaluation*, 20(4), 308–327. <https://doi.org/10.1080/13803611.2014.941371>
- Opheim, V., Grøgaard, B. J. & Næss, T. (2010). *De gamle er eldst? Betydning av skoleressurser, undervisningsformer og læringsmiljø for elevenes prestasjoner på 5., 8. og 10. trinn i grunnopplæringen* (NIFU STEP-rapport 34/2010). <http://hdl.handle.net/11250/279405>
- Pepper, K. & Thomas, L. H. (2002). Making a change: The effects of the leadership role on school climate. *Learning Environments Research*, 5(2), 155–166. <https://doi.org/10.1023/A:1020326829745>
- Pont, B., Nusche, D. & Hopkins, D. (Red.). (2008). *Improving school leadership: Vol. 2. Case studies on system leadership*. OECD.
- Pont, B., Nusche, D. & Moorman, H. (Red.). (2008). *Improving school leadership: Vol. 1. Policy and practice*. OECD.
- Printy, S. M. & Marks, H. M. (2006). Shared leadership for teacher and student learning. *Theory Into Practice*, 45(2), 125–132. https://doi.org/10.1207/s15430421tip4502_4
- Rege, M. (2023). *En barndom for livet* (Rapport fra Ekspertgruppen om barn i fattige familier). Barne- og familiedepartementet. <https://www.regjeringen.no/no/dokumenter/en-ny-barndom-for-livet/id3000835/>

- Robinson, V. M. J., Lloyd, C. A. & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, 44(5), 635–674. <https://doi.org/10.1177/0013161X08321509>
- Rutkowski, L., Gonzalez, E., von Davier, M. & Zhou, Y. (2013). Assessment design for international large-scale assessments. I L. Rutkowski, M. von Davier & D. Rutkowski (Red.), *Handbook of international large-scale assessment*. Chapman and Hall/CRC.
- Schneider, B. & Bryk, A. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation. <https://muse.jhu.edu/pub/207/monograph/book/15670>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Smedsvik, B. & Wendelborg, C. (2024). *Lærerundersøkelsen og Foreldreundersøkelsen. Analyse av to av Udirs brukerundersøkelser skoleåret 2023/24*. NTNU Samfunnsforskning.
- Taneja-Johansson, S. & Powell, J. J. W. (2024). Confronting the idealised 'Nordic model' in education with contemporary realities of special education in Sweden. *European Journal of Special Needs Education*, 39(6), 866–881. <https://doi.org/10.1080/08856257.2024.2425501>
- Teig, N., Veletic, J., Price, H., Moore, R. & Ye, W. (2024). *Navigating resilient schools: Examining principals' characteristics and leadership practices in eleven countries*. UNESCO. <https://doi.org/10.54676/JZEG2765>
- UNESCO. (2024). *UNESCO global education monitoring report 2024. Leadership in education: Lead for learning*. <https://unesdoc.unesco.org/ark:/48223/pf0000391758>
- Utdanningsdirektoratet. (2017). *Rammeverk for eksamen LK06*. Utdanningsdirektoratet.
- van Ewijk, R. & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134–150. <https://doi.org/10.1016/j.edurev.2010.02.001>
- Veletić, J., Price, H. E. & Olsen, R. V. (2023). Teachers' and principals' perceptions of school climate: The role of principals' leadership style in organizational quality. *Educational Assessment, Evaluation and Accountability*, 35(4), 525–555. <https://doi.org/10.1007/s11092-023-09413-6>
- Wang, M.-T. & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, 28(2), 315–352. <https://doi.org/10.1007/s10648-015-9319-1>
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2), 114–128. <https://doi.org/10.1016/j.stueduc.2005.05.005>
- Ye, W., Scherer, R. & Blömeke, S. (2024). Teachers' and principals' perceptions of school emphasis on academic success: Measurement invariance, agreement, and relations to student achievement. *Large-Scale Assessments in Education*, 12(1), Artikkel 19. <https://doi.org/10.1186/s40536-024-00207-w>
- Yin, L. & Reynolds, K. A. (2024). Creating and interpreting the TIMSS context questionnaire scales. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report. Methods and procedures* (s. 15.1–15.14). IEA. <https://timss2023.org/methods/chapter-15/>
- Zhan, X., Anthony, A. B., Goddard, R. & Beard, K. S. (2023). Development, factor structure, and reliability of the Shared Instructional Leadership Scale in public secondary schools. *Educational Management Administration & Leadership*, 51(1), 75–94. <https://doi.org/10.1177/1741143220963103>
- Özdemir, N., Gümüş, S., Kılınç, A. Ç. & Bellibaş, M. Ş. (2024). A systematic review of research on the relationship between school leadership and student achievement: An updated framework and future direction. *Educational Management Administration & Leadership*, 52(5), 1020–1046. <https://doi.org/10.1177/1741143222118662>
- Aas, M., Ballangrud, B. B., Vennebo, K. & Andersen, F. C. (2024). *Forskning på den nasjonale skolelederutdanningen En utdanning som gir avtrykk i praksis: Delrapport 3*. OsloMet – storbyuniversitetet.

Vedlegg 4.1. Postene i variablene om læringsfokus

Variabel	Spørsmål og svaralternativ til rektorer	Gjennomsnitt	Standardavvik
Skolens læringsfokus	Hvordan vil du karakterisere hvert av de følgende på skolen din? (1 – Svært lav, 5 – Svært høy)		
Lærernes aspekt	Lærernes forståelse av målene læreplanen	4,10	0,61
	Hvor godt lærerne lykkes med å iverksette læreplanen	3,75	0,65
	Lærernes forventninger til elevenes prestasjoner	3,94	0,60
	Lærernes evne til å inspirere elevene	3,70	0,57
Foreldrenes aspekt	Foreldrenes engasjement i skoleaktiviteter	3,18	0,70
	Foreldrenes vilje til å ta ansvar for at elevene stiller godt forberedt på skolen	3,15	0,73
	Foreldrenes forventninger til elevenes prestasjoner	3,66	0,67
	Foreldrenes støtte til elevenes prestasjoner	3,32	0,65
Elevenes aspekt	Elevenes ønske om å gjøre det bra på skolen	3,49	0,60
	Elevenes evne til å nå skolens faglige mål	3,40	0,58
	Elevenes respekt for medelever som utmerker seg på skolen	3,45	0,64
Språk		3,57	0,80
HER		11,18	1,42
		Ja	Nei
Kursbevis i utdanningsledelse		76,0 %	23,0 %
Mastergrad i utdanningsledelse		43,6 %	56,4 %

KAPITTEL 5

Spredning i elevers matematikk- og naturfagprestasjoner, TIMSS 2015–2023

Anne-Cathrine Lehre og Hege Kaarstein

Institutt for Lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: Differensiering kan være en utfordring for lærere om det er stor spredning i elevenes prestasjoner innenfor en og samme klasse. I dette kapittelet undersøkes det om spredningen i norske elevers matematikk- og naturfagprestasjoner på 5. og 9. trinn har endret seg i perioden fra 2015 til 2023. Resultatene viser en økning i spredningen i matematikkprestasjonene fra 2015 til 2019 på 5. trinn, og at denne økningen kan stamme fra en større spredning i resultatene innen emneområdet tall og en større spredning blant jenter. I naturfag på 5. trinn var det en signifikant økning i spredningen i prestasjoner for begge kjønn og innenfor alle emneområdene (biologi, fysikk/kjemi og geofag) for hver påfølgende syklus fra 2015 til 2023. For 9. trinn er spredningen i matematikkprestasjonene signifikant størst i 2023, med signifikant økning av spredningen innenfor emnene tall og geometri. For naturfag viste analysene at spredningen var signifikant større i 2019 enn i 2015 og 2023, og at spredningen innenfor biologi og fysikk har økt signifikant fra en TIMSS-syklus til den neste. Generelt viser resultatene at gutter på 9. trinn har signifikant større spredning i sine naturfags- og matematikkprestasjoner enn jenter.

Nøkkelord: trend, spredning, matematikk, naturfag, kjønn, TIMSS

Abstract¹**Variation in Students' Mathematics and Science Achievement, TIMSS 2015–2023**

Differentiation can pose a challenge for educators when there is considerable variation in student performance within a single class. This chapter examines whether the variability in the performance of Norwegian students in mathematics and natural sciences at grades 5 and 9 has changed between 2015 and 2023. The findings indicate an increase in variability in mathematics performance from 2015 to 2019 at grade 5, which may be attributed to a greater dispersion in results within the content domain number and a more pronounced spread among female students. In natural sciences at grade 5, there was a significant increase in performance variability for both genders across all content domains (biology, physics/chemistry, and earth sciences) for each successive cycle from 2015 to 2023. For grade 9, the variability in mathematics performance was significantly greater in 2023, with a notable increase observed in the content domains numbers and geometry. In natural sciences, analyses revealed that variability was significantly higher in 2019 compared to both 2015 and 2023, with significant increases in variability noted in biology and physics from one TIMSS cycle to the next. Overall, the results demonstrate that male students in grade 9 exhibit significantly greater variability in their performance in natural sciences and mathematics compared to their female counterparts.

Keywords: trend, variability, mathematics, natural science, gender, TIMSS

1 KI er benyttet til oversettelse fra norsk til engelsk.

Introduksjon

Differensiering – det å tilpasse undervisningen til elevenes potensial, motivasjon, kunnskapsnivå og interesse (Kunnskapsdepartementet, 2017a) – kan være en større utfordring for lærere om det er store forskjeller mellom de høyt- og lavtpresterende elevene i en klasse. I dette kapittelet undersøkes det om spredningen i elevers matematikk- og naturfagprestasjoner endrer seg i perioden fra 2015 til 2023. Hvor stor spredningen i elevers prestasjoner er, er også en måte å studere likeverd på. Likeverd er en grunntanke i det norske skolesystemet (Frønes & Jensen, 2020; Nilsen et al., 2018; Opplæringslova, 2023), og refererer til prinsippet om at alle mennesker har samme verdi. Alle elever har samme rett til utdanning, uavhengig av «hjemmebakgrunn, kjønn, bosted og etnisitet» (St.meld. nr. 16 (2006–2007)). Det er forventet at skolen bidrar til utjevning av sosiale forskjeller, og at alle elever har like muligheter til å lykkes (St.meld. nr. 16 (2006–2007)). Selv om all opplæring skal tilpasses evnene og forutsetningene til den enkelte elev (NOU 2009: 18, s. 15), er det også forventet at det kun er små forskjeller mellom de høyt- og lavtpresterende elevenes resultater i en likeverdig skole (Strietholt, 2014). Å tilpasse opplæringen til hver enkelt elev, og dermed leve opp til forventningen om en likeverdig opplæring, krever forskjellsbehandling/differensiering av elevene (NOU 2009: 18, s. 15). Tidligere forskning har vist at spredningen i matematikkprestasjoner først sank og deretter økte i perioden 1995 til 2015 for norske barne- og ungdomsskoleelever (Nilsen et al., 2018). I dette kapittelet undersøkes norske elevers spredning i prestasjoner i naturfag og matematikk på 5. og 9. trinn fra 2015 til 2023.

Skolesystemet er komplekst. For å kunne belyse slike komplekse situasjoner kan det være nyttig å kunne oppsummere dataene i enkle effektmål, som gjennomsnitt eller median. For eksempel benytter TIMSS-undersøkelsen blant annet gjennomsnittlige poengsummer til å oppsummere elevers prestasjoner i matematikk og naturfag. Dersom en er interessert i forskjeller mellom to grupper, vil forskjeller i gjennomsnitt eller medianer gi en enkel beskrivelse av forskjellene. Ulempen med kun å studere forskjeller i gjennomsnitt eller medianer er at dette gir et veldig forenklet bilde av situasjonen. Ved å se på hele distribusjonen og variasjonen i fordelingen, kan eventuelle forskjeller mellom to grupper lettere oppdages (for eksempel kan like gjennomsnitt maskere ekstrempunkter). I dette kapittelet studeres spredningen i norske elevers prestasjoner ved å visualisere forskjeller

i hele distribusjonen ved hjelp av kvantildistansefunksjoner (*Q-distance*) (Laake et al., 1985).

I tillegg til å studere spredningen i matematikk- og naturfagprestasjoner i perioden fra 2015 til 2023, undersøkes også spredningen for gutter og jenter. Følgende forskningsspørsmål undersøkes:

- 1) Hvordan har spredningen i matematikk- og naturfagprestasjoner, inkludert innenfor ulike emneområder, endret seg fra TIMSS 2015 til TIMSS 2019 og til TIMSS 2023?
- 2) Hvordan har spredningen i matematikk- og naturfagprestasjoner endret seg for gutter og jenter i perioden 2015–2023?

Bakgrunn og teori

Emneområder i TIMSS-rammeverkene

TIMSS-undersøkelsen har rammeverk i naturfag og matematikk som definerer hvilke kunnskaper og ferdigheter barne- og ungdomsskole-elevne testes i (Kaarstein et al., 2024; Mullis et al., 2021). Rammeverkene tar utgangspunkt i og bygger på deltakerlandenes læreplanverk. De overordnede rammeverkene deles inn i faglige emneområder. Hvert faglig emneområde er definert og beskrevet for hver syklus med små endringer mellom hver syklus (Mullis et al., 2021), og de faglige oppgavene tilhører gitte emner.

Matematikkrammeverket for 5. trinn er delt inn i tre emneområder: tall, geometri og statistikk, hvor den prosentvise fordelingen innenfor hver av disse emnene er henholdsvis 50, 30 og 20 prosent (Kaarstein et al., 2024; Mullis et al., 2021). I 2023 besto emnet tall av oppgaver som omhandler heltall, enkle ligninger og sammenhenger, brøk og desimaltall. Geometri handler om geometri og målinger, mens statistikk går ut på å lese, tolke og fremstille data, og å bruke data for å løse oppgaver. Innenfor hvert av disse emnene var det en nedgang i prestasjonene fra 2015 til 2023, hvor den største nedgangen var innenfor geometri (Kaarstein et al., 2024).

Naturfagrammeverket på 5. trinn består av emnene biologi (45 %), fysikk og kjemi (35 %) og geofag (20 %) (Mullis et al., 2021). Det var en nedgang i prestasjonene for norske elever innenfor hver av de tre emneområdene fra 2015 til 2023 (Kaarstein et al., 2024). I 2023 inkluderte emnet biologi oppgaver om organismers kjennetegn og livsprosesser, livssyklus,

reproduksjon og arv, økosystemer og helse. Fysikk og kjemi testet elevene i oppgaver i klassifisering, egenskaper og endringer i stoff, energiformer og -overganger samt krefter og bevegelse. Oppgavene i geofysikk handlet om jordas fysiske egenskaper, ressurser og historie, vær og klima og jorda i solsystemet.

På 9. trinn er rammeverket i matematikk delt inn i fire emner: tall, algebra, geometri og statistikk og sannsynlighet, hvor den prosentvise fordelingen innenfor hver av disse emnene er henholdsvis 30, 30, 20 og 20 prosent (Kaarstein et al., 2024; Mullis et al., 2021). På ungdomstrinnet handlet oppgavene i emnet tall om heltall, brøk og desimaltall, forhåndstall, proporsjonalitet og prosent. Emnet algebra inneholdt oppgaver om algebraiske uttrykk, sammenhenger og funksjoner, og i geometri var fokuset på geometriske former og målinger. Emnet statistikk og sannsynlighet handler nettopp om statistikk og sannsynlighet. Norske ungdomsskoleelever har hatt en nedgang i prestasjonene i emnet tall og oppgang i både algebra og geometri fra 2015 til 2019. For statistikk og sannsynlighet har de norske elevene hatt en oppgang i prestasjonene fra 2019 til 2023, men de er ikke på samme høye nivå som i 2015 (Kaarstein et al., 2024).

Naturfagrammeverket på 9. trinn består av emnene biologi (35 %), fysikk (25 %), kjemi (20 %) og geofag (20 %). Norske ungdomsskoleelever har hatt en nedgang innenfor alle disse emnene fra 2015 til 2023 (Kaarstein et al., 2024; Mullis et al., 2021). I 2023 handler emnet biologi om organismers kjennetegn og livsprosesser, celler og celledfunksjoner, livssyklus, reproduksjon og arv, mangfold, tilpasning og naturlig utvalg, økosystemer og helse. I fysikk ble elevene prøvd i oppgaver innenfor materiens fysiske tilstander og endringer, energiformer og -overganger, lys og lyd, elektrisitet og magnetisme samt krefter og bevegelse. Kjemien inneholdt oppgaver innenfor stoffers oppbygning, stoffers egenskaper og kjemiske endringer, mens emnet geofag handlet om jordas struktur og fysiske egenskaper, jordas prosesser, sykluser og historie, jordas ressurser, bruk og gjenbruk, og jorda i solsystemet og i universet.

Mer detaljerte rammeverk for hver syklus finnes på nettsidene timss.no og timssandpirls.bc.edu/.

Likeverd og differensiering

Ifølge Utdanningsdirektoratet (2023) handler likeverd om at «alle mennesker er like mye verdt og skal ha like muligheter til å utvikle seg og

delta i fellesskapet». Dette inkluderer både juridisk og sosial likestilling, og er et sentralt prinsipp i norsk utdanning og velferdspolitik (Kunnskapsdepartementet, 2017a, 2017b). Grunnskolen skal reflektere verdier knyttet til likeverd da skolen har ansvar for alle elever (Frønes & Jensen, 2020; Nilsen et al., 2018; Opplæringslova, 2023). Likeverd i skole-sammenheng handler om å sikre at alle elever får de samme mulighetene til å lykkes (St.meld. nr. 16 (2006–2007)). Det innebærer å skape et inkluderende læringsmiljø hvor hver elev blir verdsatt og respektert, og hvor forskjeller blir sett på som en ressurs (NOU 2009: 18). Høy likeverd fremmer læring (Creemers & Kyriakides, 2008, 2013), og store forskjeller i prestasjoner kan indikere at ikke alle elever har lik tilgang til læringsressurser og støtte (NOU 2009: 18). Ulikheter i skoleprestasjoner kan ha langsiktige konsekvenser for elevenes fremtidige utdannings- og yrkesmuligheter. Dette kan bidra til å opprettholde sosiale og økonomiske ulikheter i samfunnet.

Elever har ulike styrker, svakheter, interesser og læringsstiler (Thomsen, 2024; Tomlinson & Imbeau, 2010). Dette fører til variasjon i hvordan de presterer i ulike fag og aktiviteter. Variasjon i prestasjoner blant elever er en naturlig del av skolehverdagen, og differensiering er en nøkkelstrategi for å håndtere dette på en måte som fremmer likeverd. Differensiering handler om forskjellsbehandling for å gi elevene individuell opplæring ut fra elevens evner og forutsetninger (NOU 2009: 18), og skal bidra til at elever utvikler seg innenfor og på tvers av fag (Brevik & Gunnulfsen, 2016). Ved å tilpasse undervisningen til elevenes individuelle behov, kan lærere hjelpe hver elev med å nå sitt fulle potensial. Dette kan innebære å gi ekstra støtte til de som trenger det, og utfordringer til de som er klare for det. Differensiering kan praktiseres på mange måter, for eksempel gjennom varierte undervisningsmetoder, tilpassede oppgaver og individuelle læringsmål. Dette skaper et inkluderende læringsmiljø hvor alle elever føler seg sett og verdsatt. Tomlinson (2014) understreker lærerens betydning ved differensiering, men differensiering kan være utfordrende for lærere å ivareta (Brevik & Gunnulfsen, 2016).

I Norge er det mindre forskjeller i elevenes prestasjoner enn i mange andre vesteuropeiske land (Mullis et al., 2016; OECD, 2016). Sett i et internasjonalt perspektiv er forskjellene mellom skoler i Norge små, men økende (OECD, 2016). I en OECD-utredning som omhandler kjønnsforskjeller i utdanningsløpet trekkes det frem at det er relativt små kjønnsforskjeller i Norge tidlig i skoleløpet sammenlignet med andre land, men at forskjellene øker i løpet av ungdomsskolen (Borgonovi et al., 2018). Studier viser

at gutter har lavere leseferdigheter enn jenter ved 15-årsalderen (Björnsson & Olsen, 2018; Borgonovi et al., 2018; Stoet & Geary, 2013), mens jenter ofte har lavere motivasjon og selvoppfatning i realfagene, til tross for at de presterer like godt eller bedre enn guttene (Jensen et al., 2019). Forskjellene forsterkes ytterligere av sosiale bakgrunnsvariabler, som foreldrenes utdanningsnivå og skolens pedagogiske praksis (Jensen et al., 2019). En rapport fra NOVA viser at karakterforskjeller mellom gutter og jenter er stabile gjennom ungdomsskolen, men at jenter generelt har større faglig fremgang enn gutter (Bakken, 2021). Et inkluderende læringsmiljø og varierte undervisningsformer fremmer motivasjon og mestring for alle elever, uavhengig av kjønn (Bakken, 2010; Wollscheid et al., 2018), hvilket sikrer likeverd og like muligheter for alle.

Spredning i prestasjoner og kjønn

I standardiserte intelligenstagstester er det ingen signifikante kjønnsforskjeller i gjennomsnittsprestasjoner (Feingold, 1994; Geary, 1998), men innenfor noen ulike kognitive områder er det noe variasjon mellom kjønnene (Hunt, 2010). For eksempel har det blitt funnet at jenter i gjennomsnitt presterer bedre enn gutter på leseforståelse, essayskriving og ordforråd, mens gutter presterer i gjennomsnitt bedre enn jenter på tester relatert til romlig visualisering og mental rotasjon av tredimensjonale objekter (Hunt, 2010). Som gruppe har gutter/menn større spredning enn jenter/kvinner innenfor mange akademiske og kognitive områder (se for eksempel Deary et al., 2003; Feingold, 1994; Halpern & LaMay, 2000; Hedges & Nowell, 1995). Større variasjon blant menn enn kvinner er ikke en ny observasjon, men har eksistert siden slutten av 1800-tallet, og har blitt stadfestet gjennom tiden (se f.eks. Baye & Monseur, 2016; Ellis, 1894; Feingold, 1994; Lehre et al., 2009).

Analyser av data fra internasjonale storskalaundersøkelser viser også at gutter som regel har større spredning i sine prestasjoner enn jenter (Baye & Monseur, 2016; Meinck & Brese, 2019). Baye og Monseur (2016) studerte kjønnsforskjeller i spredningen av prestasjoner i matematikk, naturfag og lesing med data hentet fra TIMSS 1995–2007, PISA 2000–2012 og *Progress in International Reading Literacy Study* (PIRLS) 2001–2011. Av alle deltakerlandene i disse undersøkelsene, var det kun på ungdomstrinnet for TIMSS 2019 at det fantes deltakerland hvor spredningen i prestasjonene var størst blant jenter. Dette var i matematikk og kun blant 3 prosent av landene. Ingen andre land, fag, sykluser eller utvalg av elever i disse tre internasjonale

undersøkelsene hadde signifikant større spredning i prestasjonene blant jenter enn gutter. Forfatterne fant derimot en økende andel jenter blant de lavtpresterende elevene i matematikk for begge trinn fra TIMSS 1995 til TIMSS 2007. I samme periode var det en høyere andel gutter blant de høytpresterende i matematikk. For PISA-undersøkelsen var det også flest gutter blant de høytpresterende i matematikk, men en ganske lik fordeling av gutter og jenter blant de lavtpresterende i perioden 2000 til 2012.

Analyser av barneskoleelever som deltok i TIMSS 2011 og PIRLS 2011, viste at gutter var overrepresentert blant de lavtpresterende og de høytpresterende i lesing, matematikk og naturfag (Bergold et al., 2017). Prestasjonene til 5. trinnselever i den dinariske regionen (Albania, Bosnia-Herzegovina, Kroatia, Kosovo, Montenegro, Nord-Makedonia og Serbia) som deltok i TIMSS 2019 viste at andelen gutter blant de høytpresterende i matematikk er høyere enn andelen jenter, og at gutter er overrepresentert blant de lavtpresterende i naturfag (Vrapi et al., 2022). Norske resultater fra PISA 2022 viser at i matematikk er 15-åringer gutter overrepresentert blant de som presterte på et lavt nivå (Jensen & Nilsen, 2024) og blant de som presterer på et høyt nivå (Jensen et al., 2023). Videre har forskjellen i spredning i skoleprestasjoner for norske gutter og jenter blitt fremhevet av Stoltenberg-utvalgets rapport (NOU 2019: 3), som påpeker at variasjonen i gutters skoleprestasjoner er betydelig større enn blant jenter.

Metode

Data

I dette kapittelet benyttes norske data fra TIMSS 2015, TIMSS 2019 og TIMSS 2023. Utvalget består av elever på både 5. og 9. trinn og begge fag. Tabell 5.1 viser en oversikt over totalt antall deltakende elever og fordelingen av gutter og jenter innenfor hver TIMSS-syklus.

Siden det er fire år mellom syklusene i TIMSS-undersøkelsen, og utvalgene er representative for populasjonen, tilhører utvalget av 5. trinnselever i en syklus og utvalget av 9. trinnselever i påfølgende syklus samme populasjon.

Variabler

Variablene som brukes i dette kapittelet er kjønn (tabell 5.1) og TIMSS' faglige rammeverk (se bakgrunn og teoridelen for detaljer om emneområdene). TIMSS' faglige rammeverk definerer hvilke emneområder elevene får oppgaver i innenfor matematikk og naturfag (Kaarstein et al., 2024;

Mullis et al., 2021). For å få dekket innholdet i de faglige rammeverkene med tilhørende emneområder, har hvert fag for hvert trinn omtrent to hundre oppgaver. Disse oppgavene fordeles i 14 oppgavesamlinger, og en elev løser en oppgavesamling for naturfag og en oppgavesamling for matematikk (kalt *matrix item sampling design*). Omtrent halvparten av oppgavene er flervalgsoppgaver, og de resterende oppgavene er åpne oppgaver hvor eleven formulerer svaret selv. At elevene ikke svarer på alle oppgavene i det faglige rammeverket kompenseres ved at det beregnes plausible verdier (Jewsbury et al., 2024). For nærmere beskrivelse av plausible verdier, se for eksempel Rutkowski et al. (2010) eller Jewsbury et al. (2024).

Tabell 5.1 Antall deltakende elever per TIMSS-syklus, totalantall og delt på kjønn

Syklus	5. trinn			9. trinn		
	Totalt	Gutter	Jenter	Totalt	Gutter	Jenter
2015	4 329	2 183	2 146	4 697	2 345	2 352
2019	3 951 ¹	1 993	1 957	4 575 ²	2 315	2 252
2023	5 301 ¹	2 650	2 650	6 324 ³	3 340	2 982

Note. ¹ fjernet 1 elev; ² fjernet 8 elever; ³ fjernet 2 elever som ikke hadde data for kjønn. Derfor blir ikke summen av antall jenter og gutter nøyaktig lik totalsummen.

Statistisk metode

Dersom en ønsker å sammenlikne spredning i resultater for to utvalg er det flere ulike spredningsmål som kan benyttes, og veldig ofte benyttes utvalgenes standardavvik. Standardavviket forteller hvor nært eller langt fra en typisk verdi i utvalget ligger i forhold til utvalgets gjennomsnitt. Lave verdier for standardavviket betyr liten spredning rundt gjennomsnittet for utvalget, og høye verdier angir stor/større spredning. Standardavvik er et gjennomsnittsmål som gir liten informasjon om hvordan spredningen varierer på tvers av hele utvalget. I dette kapittelet studeres derfor spredningen i norske elevers prestasjoner ved å visualisere forskjeller i hele distribusjonen ved hjelp av kvantildistansefunksjoner (*Q-distance*) (Laake et al., 1985).

Programmet *Q-distance* (versjon 1.6.0) ble brukt til å regne ut kvantildistansefunksjonen, 95 prosent konfidensintervall, Levenes test for likhet i variansen (Levene, 1960) og *Students t-test* (Student, 1908). P-verdien ble satt til 0,05 som grense mellom signifikante og ikke-signifikante forskjeller mellom utvalgene som ble sammenlignet. Grafene ble glattet ved hjelp av en ikke-vektet LOWESS algoritme i BioArray Software Environment,

versjon 1.2.16. For elevens prestasjoner ble én plausibel verdi (PV1) benyttet i visualiseringene og beregningene.

Kvantildistansefunksjonen er basert på en ikke-parametrisk tilnærming og konfidensintervallene er basert på asymtotiske resultater (Laake et al., 1985). Metoden er velegnet for studier av spredningen i datasett. For å vise hvordan spredningen i de to utvalgene er i forhold til hverandre, benyttes differansen mellom de to utvalgenes resultater. Differansen plottes i et koordinatsystem hvor det så er mulig å se om spredningen i utvalgene er like, om det ene utvalget har større eller mindre spredning enn det andre, og om det er en jevn eller ujevn spredning på tvers av utvalgene. Størrelsen på utvalgene trenger ikke å være like da analysene benytter seg av prosentil-verdiene i utvalgene for å sammenligne dem.

For å illustrere prosessen, kan vi tenke oss at vi har tre utvalg. Selv om utvalgene kan bestå av et ulikt antall personer, vil hele utvalget alltid utgjøre 100 prosent (eller 1,0). Først sorteres hvert av utvalgenes resultater i stigende rekkefølge, deretter identifiseres utvalgenes resultater for hver 0,002 prosentil. Tabell 5.2 vises resultatene for de tre tenkte utvalgene for hver tierprosentil og differansene mellom utvalgene for/i disse prosentilene. Det er differansen for hver prosentil som plottes som en graf (kvantildistansefunksjonen; figur 5.1). For grafen beregner og plotter også Q-distance-programmet 95 prosent konfidensintervallgrafer (disse vises ikke i figur 5.1).

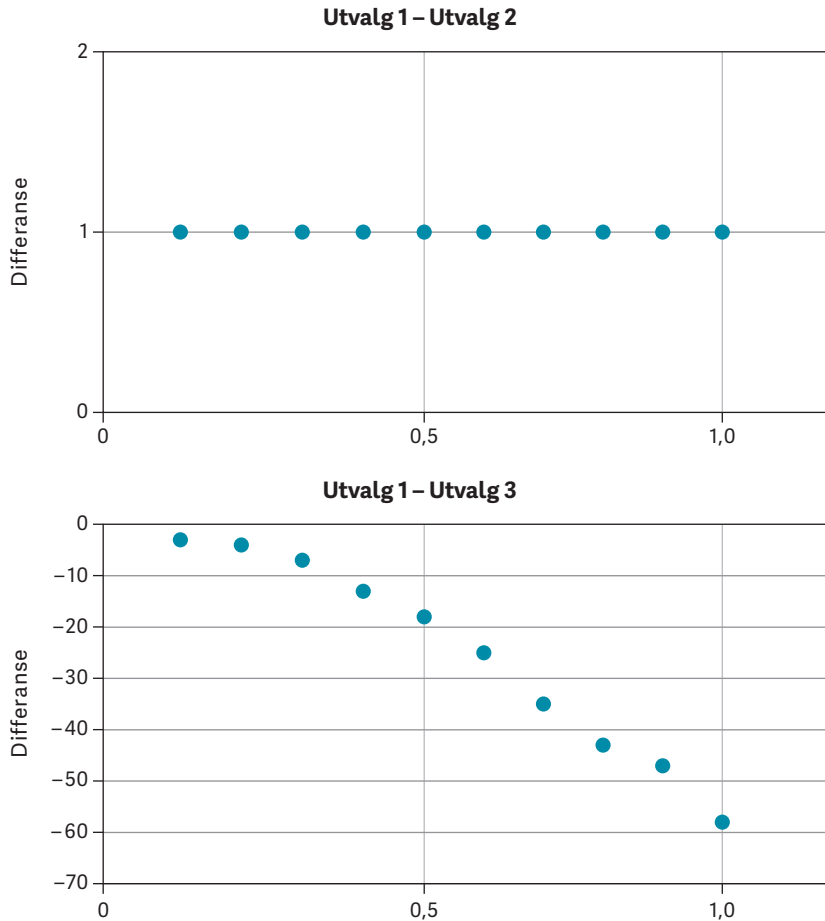
Tabell 5.2. Resultat for hver tiende prosentil i utvalg 1, 2 og 3 samt differansene mellom resultatene per prosentil for utvalg 1 og de to andre utvalgene

Prosentil	Resultat i prosentil									
	10	20	30	40	50	60	70	80	90	100
Utvalg 1	52	54	55	59	63	66	67	68	71	78
Utvalg 2	51	53	54	58	62	65	66	67	70	77
Utvalg 3	55	58	62	72	81	91	102	111	118	136
Differanse (utvalg 1–utvalg 2)	1	1	1	1	1	1	1	1	1	1
Differanse (utvalg 1–utvalg 3)	-3	-4	-7	-13	-18	-25	-35	-43	-47	-58

Differansen mellom utvalg 1 og 2, og 1 og 3, kan illustreres som vist i figur 5.1.

Når grafen for differansene mellom utvalgene blir tilnærmet en horisontal linje, som den til venstre i figur 5.1, betyr dette at det er like stor spredning i de to utvalgene. Dersom utvalgene har ulik spredning, vil kvantildistansefunksjonen stige eller synke. Den høyre grafen i figur 5.1 viser

et eksempel på en slik graf hvor utvalg 3 har større spredning enn utvalg 1. Dersom utvalg 1 hadde hatt større spredning enn utvalg 3, ville grafen ha steget mot høyre. Videre viser figuren at det er liten variasjon i spredningen blant de lavtpresterende (differansen er liten mellom de to utvalgene, ligger nært 0) og stor variasjon blant de høytpresterende (differansen mellom utvalgene er omtrent 60 poeng).



Figur 5.1. Differanse mellom prosentilene i utvalg 1 og 2 (øverst) og i utvalg 1 og 3 (nederst)

Når alle differansene for de to utvalgene er positive (over x-aksen) eller negative (under x-aksen, slik som i figur 5.1), kan det tyde på at resultatene til det ene utvalget gjennomgående er høyere enn det andre utvalgets resultater. Dette kan videre tyde på at utvalgenes gjennomsnittresultat er signifikant forskjellig, men dette må alltid sjekkes med *Student's t-test*.

Resultater

Først presenteres en sammenligning av prestasjonsspredningen i matematikk, de tilhørende emneområdene, for hvert av kjønnene mellom de tre syklusene TIMSS 2015, TIMSS 2019 og TIMSS 2023. Deretter sammenlignes prestasjonsspredningen for gutter og jenter i matematikk innenfor hver syklus. Så presenteres de tilsvarende resultatene for naturfag. Resultatene for 5. trinn presenteres før resultatene for 9. trinn.

Tabell 5.3 viser resultatene for matematikkprestasjonene på 5. trinn, både totalskår, skår per emneområde og totalskår for hvert kjønn. Den viser i hvilken av de to sammenlignede syklusene spredningen er størst. For eksempel, når totalskåren for TIMSS 2019 sammenlignes med totalskåren for TIMSS 2015, er det en signifikant større spredning i 2019 enn i 2015 ($p < 0,001$). Dette er markert med 19 (for TIMSS 2019) i tabellen. Ved sammenligning av spredningen i totalskår for 2023 og 2015, er det størst spredning i 2023, som er markert med 23 i tabellen. Der det ikke er signifikant forskjell i spredningen mellom to sykluser, er dette markert med ordet «lik» (for eksempel er det ikke forskjell i spredning for totalskår fra 2019 til 2023). Den samme trenden som vises for totalskåren på tvers av syklusene, vises også for emnet tall. For emnene geometri og statistikk har spredningen i prestasjoner blitt redusert fra 2015 til 2023. For deskriptiv statistikk for matematikkprestasjoner, se tabell 5.A1 (vedlegg).

Tabell 5.3. Syklus med størst spredning i matematikkresultatene, når to og to sykluser sammenlignes på 5. trinn

	TIMSS-sykluser som sammenlignes		
	T23 og T19	T19 og T15	T23 og T15
Matematikk, totalskår	lik ($p = 0,128$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Emneområde			
Tall	lik ($p = 0,920$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Geometri	T19 ($p < 0,001$)	T15 ($p = 0,014$)	T15 ($p < 0,001$)
Statistikk	T19 ($p < 0,001$)	lik ($p = 0,321$)	T15 ($p < 0,001$)
Delt på kjønn			
Gutter	lik ($p = 0,099$)	T19 ($p < 0,001$)	lik ($p = 0,243$)
Jenter	lik ($p = 0,357$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)

Note. Den syklusen med størst spredning i resultatene, er angitt i tabellen. «Lik» betyr at det ikke er en signifikant forskjell i spredningen for de to syklusene som er sammenlignet. I parentes, p -verdien for Levene's test for equality of variances mellom de to syklusene som er sammenliknet.

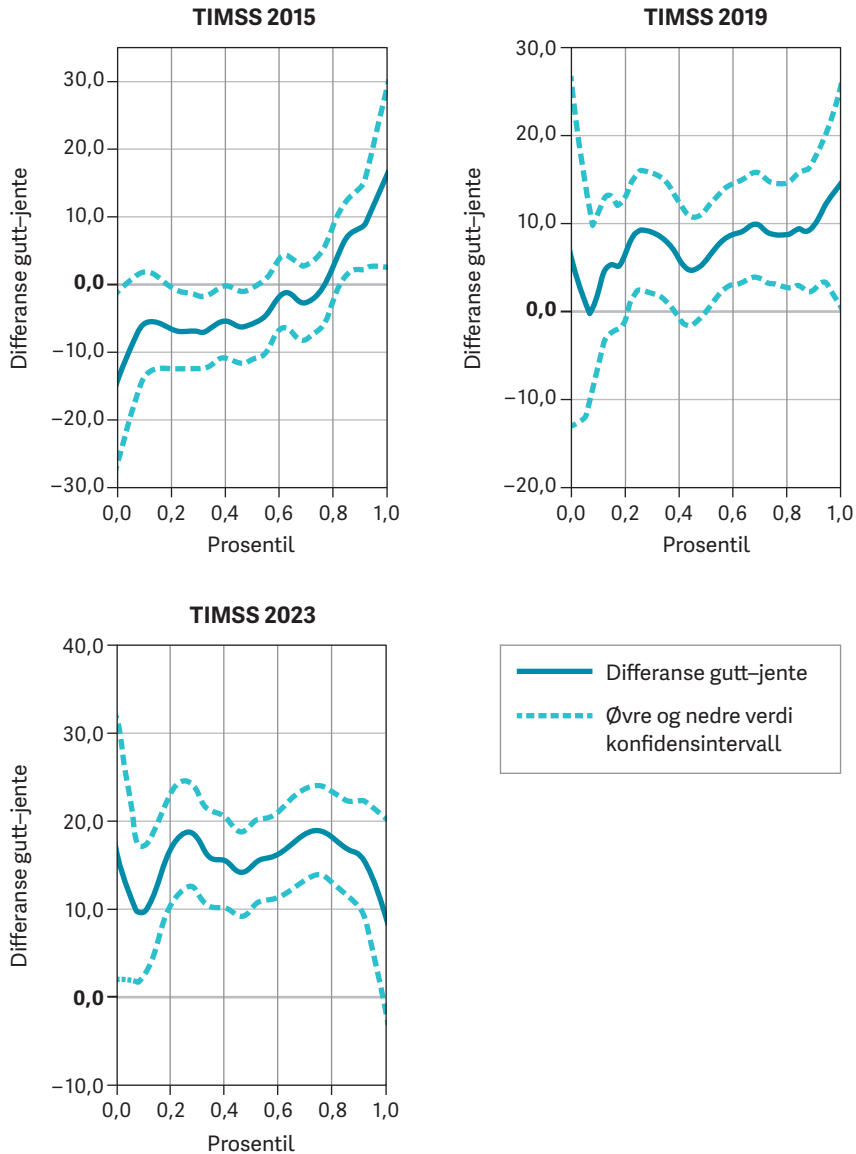
Sammenligninger av spredningen i matematikkprestasjoner for jenter, viser at spredningen økte signifikant fra 2015 til 2019/2023 (tabell 5.3). Guttenes resultater viser en signifikant økning i spredningen av prestasjoner fra 2015 til 2019, mens forskjellene i spredningen når man sammenligner 2023-syklusen med enten 2015 eller 2019 viser ingen signifikante forskjeller. Dette vises også ved at standardavviket for gutters matematikkprestasjoner var høyere i 2019 (76) enn i 2015 (71,8) og 2023 (73,7), se tabell 5.A1 (vedlegg).

For å kunne sammenligne spredningen i resultater for gutter med spredningen i resultater for jenter, må man se på prestasjonene innenfor hver av syklusene. Figur 5.2 viser differansen i spredningen av gutter og jenters matematikkprestasjoner for 5. trinnselever. Kurven for TIMSS 2015 stiger mye brattere mot høyre enn de tilsvarende kurvene for TIMSS 2019 og TIMSS 2023, og kurven for 2015 viser en signifikant større spredning i prestasjonene til gutter sammenlignet med prestasjonene til jenter (*Levene's test* $p < 0,001$). I 2019 og 2023 er det ikke signifikante forskjeller i spredningen mellom kjønnene (*Levene's test* for 2019 $p = 0,193$ og for 2023 $p = 0,475$).

Kvantildistansefunksjonen gir innsikt i hvordan kjønnsforskjeller varierer langs hele prestasjonsspekteret, og hvordan disse forskjellene har utviklet seg over tid. Kurven for 2015 viser en signifikant forskjell i spredningen mellom kjønnene. I den nedre delen av prestasjonsspekteret presterer de svakeste guttene svakere enn de svakeste jentene. Denne forskjellen avtar gradvis, og ved omtrent 80-prosentilen (0,8 i figuren) skifter mønsteret; her presterer de sterkeste guttene bedre enn de sterkeste jentene. I midtre prosentiler (ca. 0,4–0,6 i figuren) ligger forskjellen på rundt 5 poeng i jentenes favør.

I 2019 observeres det en endring i mønsteret sammenlignet med mønsteret fra 2015. I 2019 presterer de svakeste guttene bedre enn de svakeste jentene. I midtre prosentiler har forskjellen mellom kjønnene blitt mindre, og guttene presterer nå noe bedre enn jentene, med en (grovt sett) forskjell på rundt 5 poeng i guttenes favør.

I resultatene fra 2023 observeres et nytt skifte i toppen av prestasjonsspekteret. De sterkeste jentene presterer nå bedre enn de sterkeste guttene, i kontrast til mønstrene fra 2015 og 2019. I midtre prosentiler har guttene fortsatt en fordel, og forskjellen har økt sammenlignet med 2019. Hele kurven for 2023 ligger også på oversiden av x-aksen, hvilket viser seg å utgjøre en signifikant forskjell i gjennomsnittsprestasjonene til gutter og jenter (*Student's t-test* $p < 0,001$), hvor gutter presterer bedre enn



Figur 5.2. Differanse i spredningen av gutter og jenters matematikkprestasjoner, 5. trinn (gutt – jente)

Note. Den mørkeblå linjen er kvantildistansefunksjonen, og de stiplede lyseblå linjene viser grensene for 95 % konfidensintervallet for kvantildistansefunksjonen. Kurven for TIMSS 2015 viser en signifikant forskjell i spredningen mellom kjønnene, med størst spredning for gutter (kurven stiger mot høyre). Kurvene for TIMSS 2019 og TIMSS 2023 viser ingen signifikante forskjeller mellom kjønnene.

jenter (resultatene for gutter minus resultatene for jenter er positive i hele distribusjonen). Tabell 5.A5 i vedlegget viser resultatene av t-testene for alle sammenlikninger av gjennomsnittsprestasjoner for gutter og jenter på 5. og 9. trinn.

For 5. trinnselever har spredningen i naturfagsprestasjonene økt signifikant fra 2015 til 2023 (tabell 5.4). Den samme trenden vises for alle emnene (biologi, fysikk/kjemi og geofag). For deskriptiv statistikk for naturfagsprestasjoner, se tabell 5.A2 (vedlegg).

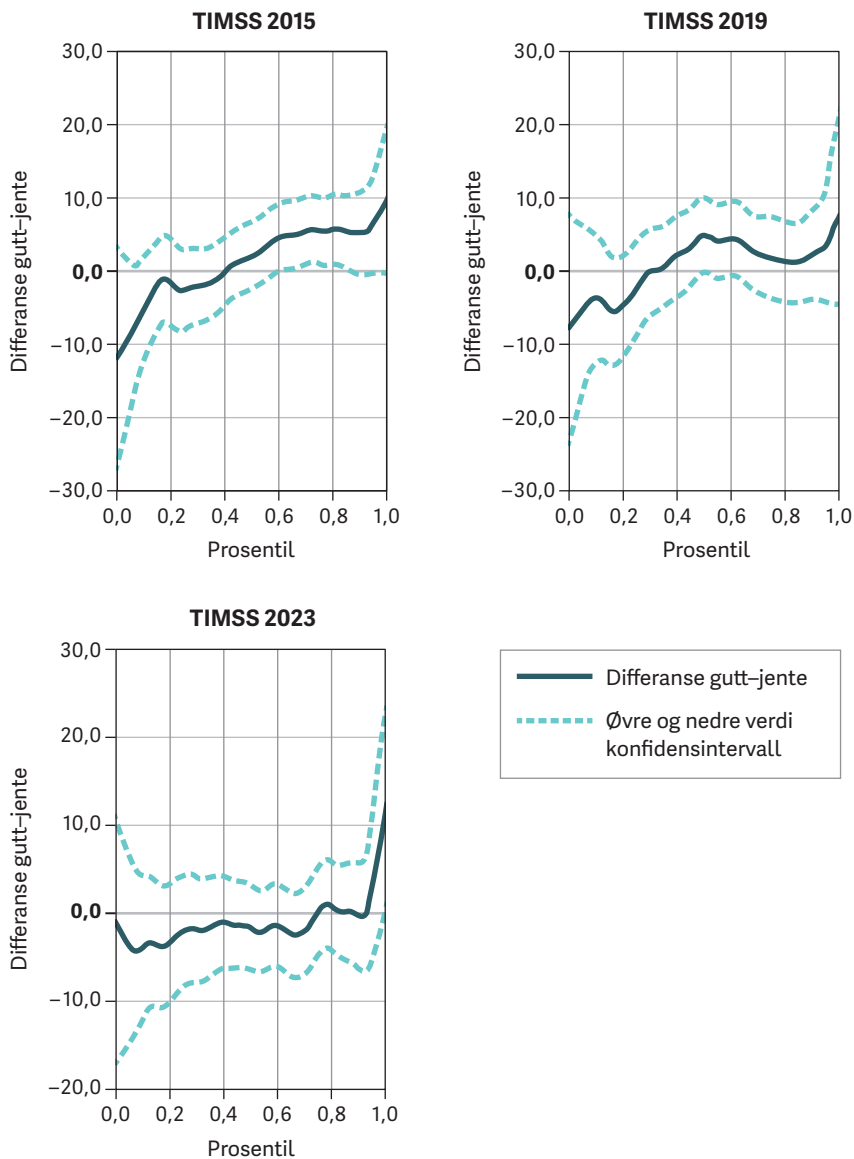
Tabell 5.4. Syklus med størst spredning i naturfagresultatene, når to og to sykluser sammenlignes på 5. trinn

	TIMSS-sykluser som sammenlignes		
	T23 og T19	T19 og T15	T23 og T15
Naturfag, totalskår	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Emneområde			
Biologi	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Fys/kjemi	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Geofag	T23 ($p < 0,001$)	lik ($p = 0,634$)	T23 ($p < 0,001$)
Delt på kjønn			
Gutter	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Jenter	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)

Note. Den syklusen med størst spredning i resultatene, er angitt i tabellen. «Lik» betyr at det ikke er en signifikant forskjell i spredningen for de to syklusene som er sammenlignet. I parentes, p -verdien for *Levene's test for equality of variances* mellom de to syklusene som er sammenliknet.

Sammenlikninger av spredningen i naturfagsprestasjoner for jenter mellom de ulike syklene, viser at spredningen økte signifikant fra 2015 til 2019, og fra 2019 til 2023 (tabell 5.4). Spredningen i naturfagsprestasjoner for gutter økte også signifikant fra TIMSS 2015 til TIMSS 2019 og til TIMSS 2023. Standardavvikene øker for hver syklus for begge kjønn (tabell 5.A2). Innenfor hver syklus (figur 5.3) er differansen mellom jenter og gutter kun signifikant forskjellige i 2015 (*Levene's test* $p < 0,001$). For 2019 og 2023 er det ingen signifikante forskjeller i spredningen (*Levene's test* for 2019 $p = 0,056$ og for 2023 $p = 0,321$).

Resultatene fra 2015 viser at de svakeste jentene presterer bedre enn de svakeste guttene (kurven ligger under x-aksen i den nedre delen av prestasjonsspekteret). I den øvre delen av spekteret er det derimot guttene som presterer best, med en tydelig forskjell til deres fordel. I 2019 er det



Figur 5.3. Differanse i spredningen av gutter og jenters naturfagprestasjoner, 5. trinn (gutt – jente)

Note. Den mørkegrønne linjen er kvantildistansefunksjonen, og de stiplede lysegrønne linjene viser grensene for 95 % konfidensintervallet for kvantildistansefunksjonen. Kurven for TIMSS 2015 viser en signifikant forskjell i spredningen mellom kjønnene, med størst spredning for gutter (kurven stiger mot høyre). Kurvene for TIMSS 2019 og TIMSS 2023 viser ingen signifikante forskjeller mellom kjønnene.

ikke lenger signifikante forskjeller i spredning mellom kjønnene. Likevel vedvarer mønsteret fra 2015, der de sterkeste guttene presterer bedre enn de sterkeste jentene. De svakeste jentene presterer fortsatt bedre enn de svakeste guttene, men forskjellen er noe redusert sammenlignet med 2015. I 2023 er forskjellene i spredning ytterligere redusert, og kvantildistansefunksjonen ligger nær null gjennom nesten hele prestasjonsspekteret med unntak av de høyest presterende elevene. De svakeste guttene og jentene presterer mer likt i 2023 enn i 2015 og 2019. Samtidig vedvarer mønsteret i den øvre delen av spekteret: De sterkeste guttene presterer fortsatt bedre enn de sterkeste jentene, med omtrent samme differanse som i 2015 og 2019.

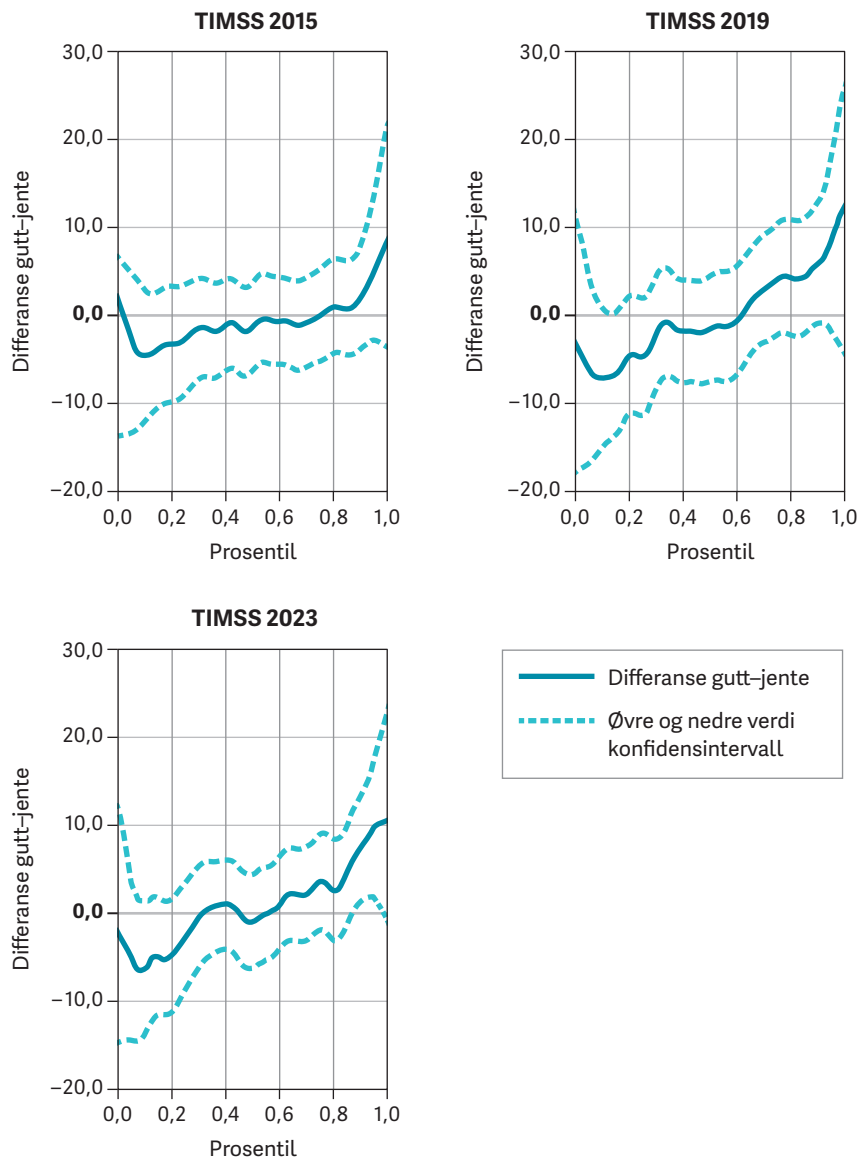
For 9. trinnselever har spredningen i matematikkprestasjonene økt signifikant fra 2015 til 2019 til 2023 (tabell 5.5). For emnene tall og geometri er det signifikant større spredning i prestasjonene i 2023 enn i de tidligere TIMSS-syklusene. Prestasjonene i algebra har størst spredning i 2019, mens det for statistikk ikke er noen signifikante forskjeller i prestasjonsspredningen mellom de tre syklene. For deskriptiv statistikk for matematikkprestasjoner, se tabell 5.A3 (vedlegg).

Tabell 5.5. Syklus med størst spredning i matematikkresultatene, når to og to sykluser sammenlignes på 9. trinn

	TIMSS-sykluser som sammenlignes		
	T23 og T19	T19 og T15	T23 og T15
Matematikk, tot.	T23 ($p = 0,016$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Emneområde			
Tall	T23 ($p < 0,001$)	lik ($p = 0,213$)	T23 ($p < 0,001$)
Geometri	lik ($p = 0,213$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Statistikk og sannsynlighet	lik ($p = 0,072$)	lik ($p = 0,195$)	lik ($p = 0,677$)
Algebra	T19 ($p < 0,001$)	T19 ($p < 0,001$)	lik ($p = 0,906$)
Delt på kjønn			
Gutter	lik ($p = 0,150$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Jenter	lik ($p = 0,061$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)

Note. Den syklusen med størst spredning i resultatene, er angitt i tabellen. «Lik» betyr at det ikke er en signifikant forskjell i spredningen for de to syklusene som er sammenlignet. I parentes, p -verdien for *Levene's test for equality of variances* mellom de to syklusene som er sammenliknet.

Sammenligninger av spredningen i matematikkprestasjoner mellom syklene, viser at spredningen økte signifikant fra 2015 til 2019/2023 for både gutter og jenter (tabell 5.5). Innenfor hver syklus (figur 5.4; tabell A3)



Figur 5.4. Differanse i variasjon for gutter og jenters matematikkprestasjoner, 9. trinn (gutt – jente)

Note. Den mørkeblå linjen er kvantildistansefunksjonen, og de stiplede lyseblå linjene viser grensene for 95 % konfidensintervallet for kvantildistansefunksjonen. Kurven for TIMSS 2015 viser en ingen signifikant forskjell i spredningen mellom kjønnene. Kurvene for TIMSS 2019 og TIMSS 2023 viser signifikante forskjeller mellom kjønnene, og større spredning blant gutter (kurvene stiger mot høyre).

er differansen i spredningen mellom jenter og gutters prestasjoner signifikant forskjellige i 2019 (*Levene's test* $p = 0,007$) og 2023 (*Levene's test* $p = 0,009$), men ikke i 2015 (*Levene's test* $p = 0,199$). Figur 5.4 viser en stigning mot høyre, hvilket viser at spredningen er større for gutters prestasjoner i matematikk enn de tilsvarende for jenter.

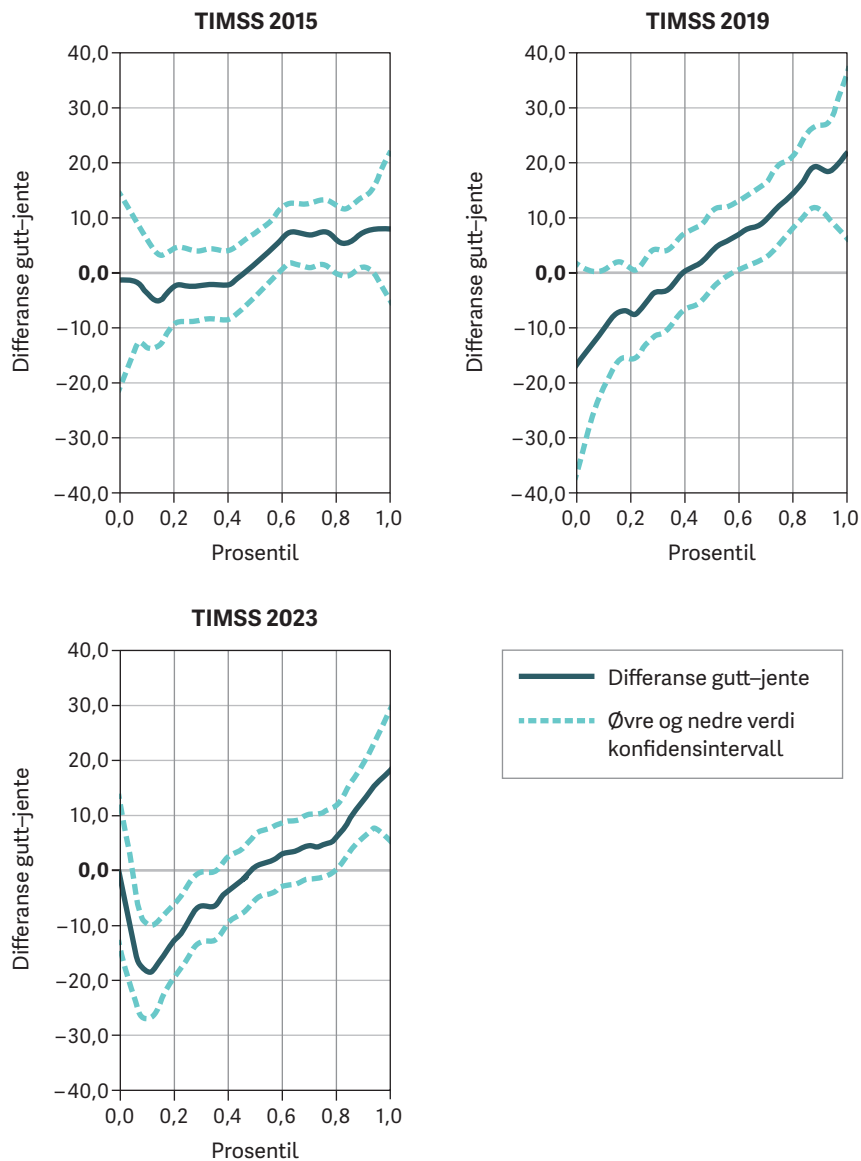
Til tross for at alle kurvene er relativt like i de tre syklusene, viser utviklingen fra 2015 til 2023 en økende tendens til større spredning blant gutter. Det er små forskjeller mellom de svakeste guttene og jentene, men denne forskjellen er noe mindre i 2015 enn i 2019 og 2023. I de midtre prosentilene er det ingen kjønnsforskjeller i prestasjonene (kurvene ligger nær 0) for noen av syklusene. Kjønnsforskjellene i variasjon er mest fremtredende blant de høyest presterende elevene, hvor guttene presterer bedre enn jentene i hele perioden.

For 9. trinnselever har spredningen i naturfagsprestasjonene økt signifikant fra 2015 til 2019, mens det ikke er en signifikant forskjell i spredningen i prestasjoner i 2019 og 2023 (tabell 5.6). I emnene biologi og fysikk her det vært signifikante økninger i spredningen av prestasjonene fra 2015 til 2019 til 2023. I kjemi og geofag var prestasjonsspredningen størst i 2019, og spredningen var signifikant større enn i 2015 og i 2023. For deskriptiv statistikk for matematikkprestasjoner, se tabell 5.A4 (vedlegg).

Tabell 5.6. Syklus med størst spredning i naturfagresultatene, når to og to sykluser sammenlignes på 9. trinn

	TIMSS-sykluser som sammenlignes		
	T23 og T19	T19 og T15	T23 og T15
Naturfag, tot.	lik ($p = 0,562$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Emneområde			
Biologi	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Fysikk	T23 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Kjemi	T19 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Geofag	T19 ($p < 0,001$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Delt på kjønn			
Gutter	lik ($p = 1,000$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)
Jenter	lik ($p = 0,553$)	T19 ($p < 0,001$)	T23 ($p < 0,001$)

Note. Den syklusen med størst spredning i resultatene, er angitt i tabellen. «Lik» betyr at det ikke er en signifikant forskjell i spredningen for de to syklusene som er sammenlignet. I parentes, p -verdien for *Levene's test for equality of variances* mellom de to syklusene som er sammenliknet.



Figur 5.5. Differanse i variasjon for gutter og jenters na turfagsprestasjoner, 9. trinn (gutt – jente)

Note. Den mørkegrønne linjen er kvantildistansefunksjonen og de stiplede lysegrønne linjene viser grensene for 95 % konfidensintervallet for kvantildistansefunksjonen. Kurvene for TIMSS 2015, TIMSS 2019 og TIMSS 2023 viser signifikante forskjeller mellom kjønnene, og større spredning blant gutter (kurvene stiger mot høyre).

Sammenligninger av spredningen i naturfagsprestasjoner mellom syklene, viser at spredningen økte signifikant fra 2015 til 2019/2023 for både gutter og jenter (tabell 5.6). Innenfor hver syklus (figur 5.5; tabell A4), er differansen i spredningen mellom jenter og gutters prestasjoner, signifikant forskjellige i 2015 (*Levene's test* $p = 0,003$), 2019 (*Levene's test* $p < 0,001$) og 2023 (*Levene's test* $p < 0,001$). Figur 5.5 viser en stigning mot høyre hvilket viser at spredningen er større for gutters prestasjoner i matematikk enn de tilsvarende for jenter.

Kurven for 2015 viser at er resultatene til gutter og jenter tilnærmet like i den nedre halvdel av prestasjonsspekteret. I den øvre halvdel presterer guttene noe bedre enn jentene, men forskjellen er moderat. Sammenlignet med senere år er kurven mindre bratt, noe som tyder på mindre variasjon i prestasjoner mellom kjønnene i 2015 enn i 2019 og 2023. I 2019 observeres en tydeligere forskjell i spredning mellom kjønnene. Kvantildistansefunksjonen stiger brattere mot høyre enn i 2015, og det er større kjønnsforskjeller både blant de elevene som presterer svakest og de elevene som presterer sterkest. Gutter presterer svakere enn de svakeste jentene, og de sterkeste guttene presterer bedre enn de sterkeste jentene. Resultatene fra 2023 viser et mønster som ligner på 2019. De sterkeste guttene presterer bedre enn de sterkeste jentene, med omtrent samme differanse som i 2019. Blant de elevene som presterer svakest, er det frem til 20-prosentilen (0,2 i grafen) en endring i mønsteret ved at det er en noe mindre forskjell mellom prestasjonene til de svakeste jentene og de svakeste guttene.

En sammenligning av prestasjonene til 5. trinnselever i 2019 og 9. trinnselever i 2023, viser en signifikant større spredning i prestasjoner på ungdomstrinnet enn på barnetrinnet.

Diskusjon

I dette kapittelet ble det undersøkt om spredningen i norske elevers matematikk- og naturfagprestasjoner på 5. og 9. trinn endrer seg i perioden fra 2015 til 2023. Resultatene antyder at den økte spredningen i matematikkprestasjoner på 5. trinn fra 2015 til 2019 kan skyldes en større økning i spredningen av jentenes prestasjoner enn hos guttene og/eller økt spredning for prestasjoner innenfor emneområdet tall. I naturfag økte spredningen i prestasjoner for begge kjønn og for alle emner (biologi, fysikk/kjemi og geofag) signifikant fra 2015 til 2019 til 2023. For 9. trinn er spredningen i

matematikkprestasjonene signifikant størst i 2023, med signifikant økning i spredningen innenfor tall og geometri. For naturfag viste analysene at spredningen var signifikant større i 2019 enn i 2015 og 2023, og at spredningen innenfor biologi og fysikk har økt signifikant fra en TIMSS-syklus til den neste. Videre har gutter på 9. trinn signifikant større spredning i sine naturfags- og matematikkprestasjoner enn jenter, mens det på 5. trinn ikke er signifikante kjønnsforskjeller. Det er generelt større spredning i resultatene på 9. trinn enn på 5. trinn.

Det er en økning i spredningen av elevenes resultater i flere emner innenfor matematikk og naturfag. I matematikk gjelder dette særlig for emnet tall. At spredningen er større innenfor dette emnet enn de andre emnene i matematikk, kan skyldes at omtrent halvparten av oppgavene i TIMSS faller inn under dette emnet, og at det er en økning i andelen lavt-presterende elever i matematikk (Kaarstein et al., 2024; Mullis et al., 2021). Norske elever har de siste årene vist en nedgang i matematikkprestasjoner, hvilket også er dokumentert i PISA 2022. Her presterte norske 15-åringer lavere enn noen gang tidligere, og andelen elever på laveste mestringsnivå økte fra 19 prosent i 2018 til 31 i 2022 (Jensen et al., 2023). Et eksempel på en type oppgaver som hører til emnet tall i TIMSS, er brøkoppgaver. Brøk er et tema som mange elever strever med (Ni & Zhou, 2005). For å kunne tilrettelegge må læreren kjenne elevene sine godt og vite noe om hvilke oppgaver enkeltelever finner lette og vanskelige. Det en elev opplever som en abstrakt oppgave, kan en annen elev oppleve som en lett oppgave. Læreren vil da spille en nøkkelrolle i å tilpasse undervisningen til elevenes ulike nivåer (Tomlinson, 2014). Differensiering kan være utfordrende (Brevik & Gunnulfsen, 2016), men en oppgavetype som kan egne seg til differensiering, er såkalte listeoppgaver (NRICH, 2013). Disse oppgavene gir elevene en åpen problemstilling der de skal lage flere løsninger eller eksempler. Ved å åpne for flere riktige svar og ulike tilnærminger, bryter listeoppgaver med den tradisjonelle ideen om én fasit og oppfordrer til utforskning, kreativitet og refleksjon. Dette gir elevene mulighet til å utvikle sin matematiske tenkning og styrker deres evne til å argumentere og kommunisere faglig (Akselsen, 2020). Samtidig gir det læreren innsikt i elevenes strategier og forståelse, noe som er avgjørende for å kunne differensiere undervisningen på en meningsfull måte.

I naturfag har norske lærere fått en mer utfordrende jobb med differensiering siden det har vært en signifikant økende spredning for både gutter og jenter på begge trinn fra 2015 til 2023. Dette gjelder innenfor

alle emner på 5. trinn og for biologi og fysikk på 9. trinn. Norske elever har også i den samme perioden hatt en nedgang i gjennomsnittsprestasjonene i alle naturfaglige emner (Kaarstein et al., 2024; Mullis et al., 2021). Dette kan indikere at det generelle kunnskapsnivået i naturfag har sunket i perioden, og at det i tillegg har blitt større variasjon i elevmassen. Det er vanskelig å si noe om hva dette skyldes, da det er mange faktorer som kan bidra til den økte spredning – for eksempel endringer i elevmassens hjemmebakgrunn, holdninger og interesser, lærerens kompetanse og/eller undervisningsstrategier, mengden praktiske aktiviteter, tilgang på ressurser og laboratorieutstyr eller faktorer relatert til læringsmiljøet/disiplin. Det kreves en helhetlig tilnærming som tar hensyn til skolens kompleksitet, og forskning på hvilke faktorer som driver økningen i prestasjonsforskjellene mellom elever, for å kunne ta tak i de økte ulikhetene i faget.

I 2015 hadde gutter på 5. trinn en signifikant større spredning i sine matematikk- og naturfagsprestasjoner enn jenter, men dette var ikke tilfelle i de to neste undersøkelsene. Denne endringen kan ha sammenheng med en signifikant økning i spredningen i jenters prestasjoner fra 2015 til 2019, og ikke en tilsvarende endring for spredningen i guttenes prestasjoner i samme periode. For matematikk er det mulig at denne økningen i variasjon blant jenter kan skyldes en trend med økende andel jenter blant de lavtpresterende elevene i matematikk for begge trinn i TIMSS (Baye & Monseur, 2016). Analysene viser en signifikant større spredning i prestasjoner på ungdomstrinnet enn den på barnetrinnet. Dette samsvarer med tidligere undersøkelser om kjønnsforskjeller i utdanningsløpet hvor det vises relativt små kjønnsforskjeller i Norge tidlig i skoleløpet, men at forskjellene øker i løpet av ungdomsskolen (Borgonovi et al., 2018). På ungdomstrinnet er guttenes prestasjoner i matematikk og i naturfag mer variable enn prestasjonene til jentene. Dette stemmer overens med tidligere studier som viser «greater male variability» (for eksempel Ellis, 1894; Feingold, 1994; Lehre et al., 2009), og da særlig innenfor mange akademiske og kognitive områder (for eksempel Deary et al., 2003; Feingold, 1994; Halpern & LaMay, 2000; Hedges & Nowell, 1995) og også i studier av internasjonale storskaladata (Baye & Monseur, 2016; Meinck & Brese, 2019). Det er verdt å merke seg at variabiliteten i guttenes prestasjoner ikke er ensartet på tvers av land (Baye & Monseur, 2016), noe som kan indikere at kulturelle og utdanningsmessige kontekster spiller en rolle. Dette antyder at internasjonalt samarbeid og sammenligning kan gi verdifulle innsikter i beste praksis, slik som er formålet med mange internasjonale storskalaprojekter, som for eksempel TIMSS.

Funnene fra TIMSS 2015, 2019 og 2023 gir et noe mer nyansert bilde av kjønnsforskjeller i skoleprestasjoner enn det som ofte kommer frem i den offentlige debatten. Selv om gutter som gruppe har lavere gjennomsnittlige skoleprestasjoner (NOU 2019: 3), viser analysene her at de også er overrepresentert blant de høytpresterende, særlig i matematikk og naturfag. Dette samsvarer med tidligere forskning som peker på intern variasjon innenfor kjønn (for eksempel Baye & Monseur, 2016; Borg, 2014). I norsk utdanningspolitisk kontekst har kjønnsforskjeller vært et sentralt tema, særlig etter Stoltenberg-utvalgets rapport (NOU 2019: 3), som fremhever at gutter er overrepresentert blant dem som presterer svakt og faller ut av videre utdanning. Samtidig viser TIMSS-data at det finnes en betydelig gruppe gutter som presterer svært godt, noe som utfordrer mediebildet om at «gutter sliter i skolen». Disse funnene understreker behovet for en differensiert pedagogisk tilnærming i norsk skole, der både lavt- og høytpresterende elever får tilpasset opplæring. Rapporter har fremhevet betydningen av et inkluderende læringsmiljø og varierte undervisningsformer som fremmer motivasjon og mestring for alle elever, uavhengig av kjønn (Bakken, 2010; Wollscheid et al., 2018). For å møte utfordringene som både lavt- og høytpresterende elever står overfor, bør det utvikles strategier som fanger opp hele prestasjonsspekteret. Dette innebærer å skape rom for faglig utfordring og utvikling, samtidig som man ivaretar elevenes trivsel og engasjement i læringsprosessen. En slik tilnærming kan bidra til å redusere kjønnsforskjeller over tid og sikre at både gutter og jenter får mulighet til å realisere sitt faglige potensial.

Begrensninger

I denne studien ble det benyttet én plausibel verdi for hver elev selv om TIMSS beregner fem plausible verdier per elev. Det er anbefalt å bruke alle plausible verdier i utregningene, men simuleringstudier viser at det veldig ofte er godt nok å kun benytte en slik plausibel verdi (Wu, 2005). Metoden brukt i dette kapittelet er en visualisering, og det er greit å velge en plausibel verdi i grafer så lenge man benytter samme plausible verdi i alle grafer (Jewsbury et al., 2024). Videre brukes plausible verdier for å estimere karakteristika for en gruppe, ikke for enkeltindivider, hvilket også er tilfelle her. Det er derfor antatt at grafene vist i dette kapittelet viser et godt nok estimat av differensene i spredningen i datasettene på et gruppenivå.

Avsluttende kommentarer

Resultatene i dette kapittelet viser en økning i forskjellene mellom elever, og da spesielt i naturfag på barnetrinnet og i matematikk på ungdomstrinnet. Blant de lavtpresterende er det en endring fra 2015 til 2023 fra at gutter presterer lavest til at jenter presterer lavest i matematikk på 5. trinn. For matematikk på 9. trinn og naturfag på begge trinn er det ingen kjønnsforskjeller i prestasjoner blant de lavtpresterende. Blant de høytpresterende, presterer gutter bedre enn jenter på begge trinn og i begge fag, en trend som har holdt seg fra 2015 til 2023. Selv om variasjon i prestasjoner blant elever er en naturlig del av skolehverdagen, er det fremdeles, og kanskje enda mer, viktig for naturfaglærere på barnetrinnet og matematikklærere på ungdomstrinnet å jobbe med differensiering i klasserommet for å oppnå målet om at hver elev skal få mulighet til å nå sitt fulle potensial.

Forfatterbiografier

Anne-Catherine Lehre er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en bakgrunn innenfor språk og realfag, og har en doktorgrad i biostatistikk. Lehre har jobbet i TIMSS-prosjektet siden 2020.

Hege Kaarstein er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad i matematikkdidaktikk og har jobbet med TIMSS-undersøkelsen siden 2009, og vært prosjektleder fra 2018. I tillegg deltar hun i den internasjonale ekspertgruppen for matematikk i TIMSS. Hennes interesser er blant annet lærernes undervisningskompetanse i matematikk, elevenes motivasjon for faget og svar på åpne oppgaver.

Referanser

- Akselsen, T. (2020). «Du får tenke og utforsk selv, det gir mestringsfølelse.» En fenomenologisk studie om hvordan tre elever i den videregående opplæring opplever og erfarer LIST som didaktisk metode i matematikkundervisningen (Masteroppgave, Nord universitetet). Nord Open. <https://hdl.handle.net/11250/2682761>
- Bakken, A. (2010). Prestasjonsforskjeller i Kunnskapsløftets første år – kjønn, minoritetsstatus og foreldres utdanning (NOVA-rapport 9/10). <https://hdl.handle.net/20.500.12199/3334>
- Baye, A. & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-Scale Assessments in Education*, 4(1), Artikkel 1. <https://doi.org/10.1186/s40536-015-0015-x>

- Bergold, S., Wendt, H., Kasper, D. & Steinmayr, R. (2017). Academic competencies: Their interrelatedness and gender differences at their high end. *Journal of Educational Psychology*, 109(3), 439–449. <https://doi.org/10.1037/edu0000140>
- Björnsson, J. K. & Olsen, R. V. (Red.). (2018). *Tjue år med TIMSS og PISA i Norge*. Universitetsforlaget. <https://doi.org/10.18261/9788215030067-2018>
- Borg, E. (2014). *Beyond a dual understanding of gender differences in school achievement. A study of the gender gap among youth in Oslo secondary schools* [Doktorgradsavhandling]. Universitetet i Oslo.
- Borgonovi, F., Ferrara, A. & Maghnouj, S. (2018). *The gender gap in educational outcomes in Norway*. OECD. <https://doi.org/10.1787/f8efl489-en>
- Brevik, L. M. & Gunnulfsen, A. E. (2016). Differensiert undervisning for høytpresterende elever med stort læringspotensial. *Acta Didactica Norge*, 10(2), 212–234. <https://doi.org/10.5617/adno.2554>
- Creemers, B. & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. <https://doi.org/10.4324/9780203939185>
- Creemers, B. P. M. & Kyriakides, L. (2013). *Improving quality in education: Dynamic approaches to school improvement*. Routledge. <https://doi.org/10.4324/9780203817537>
- Deary, I. J., Thorpe, G., Wilson, V., Starr, J. M. & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish mental survey 1932. *Intelligence*, 31(6), 533–542. [https://doi.org/10.1016/S0160-2896\(03\)00053-9](https://doi.org/10.1016/S0160-2896(03)00053-9)
- Ellis, H. (1894). *Man and woman: A study of human secondary sexual characters*. London.
- Feingold, A. (1994). Gender differences in variability in intellectual abilities: A cross-cultural perspective. *Sex Roles*, 30(1), 81–92. <https://doi.org/10.1007/BF01420741>
- Frønes, T. S. & Jensen, F. (2020). Introduksjon: Like muligheter til god leseforståelse? 20 år med lesing i PISA. I T. S. Frønes & F. Jensen (Red.), *Like muligheter til god leseforståelse?* (s. 10–20). Universitetsforlaget. <https://doi.org/10.18261/9788215040066-2020-01>
- Geary, D. C. (1998). *Male, female: The evolution of human sex differences*. American Psychological Association. <https://doi.org/10.1037/10370-000>
- Halpern, D. F. & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, 12(2), 229–246. <https://doi.org/10.1023/A:1009027516424>
- Hedges, L. V. & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269(5220), 41–45. <https://doi.org/10.1126/science.7604277>
- Hunt, E. (2010). *Human intelligence*. Cambridge University Press.
- Jensen, F. & Nilsen, T. (2024). Lavtpresterende elever i matematikk. I A. Pettersen & F. Jensen (Red.), *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (s. 187–211). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.222.ch8>
- Jensen, F., Pettersen, A., Frønes, T. S., Eriksen, A., Løvgren, M. & Narvhus, E. K. (2023). *PISA 2022. Norske elevers kompetanse i matematikk, naturfag og lesing*. Cappelen Damm Forskning. <https://doi.org/10.23865/noasp.205>
- Jensen, F., Pettersen, A., Frønes, T. S., Kjærnsli, M. & Rohatgi, A. (2019). *PISA 2018. Norske elevers kompetanse i lesing, matematikk og naturfag*. Universitetsforlaget.
- Jewsbury, P. A., Jia, Y. & Gonzalez, E. J. (2024). Considerations for the use of plausible values in large-scale assessments. *Large-Scale Assessments in Education*, 12(1), Artikkel 24. <https://doi.org/10.1186/s40536-024-00213-y>
- Kunnskapsdepartementet. (2017a). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Fastsatt som forskrift ved kongelig resolusjon. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/overordnet-del/>
- Kunnskapsdepartementet. (2017b). *Rammeplan for barnehagen: Forskrift om rammeplan for barnehagens innhold og oppgaver*. <https://www.udir.no/laring-og-trivsel/rammeplan-for-barnehagen/>

- Kaarstein, H., Lehre, A.-C., Radišić, J. & Rohatgi, M. A. (2024). *TIMSS 2023. Kortrapport*. Universitetet i Oslo.
- Lehre, A., Lehre, K. P., Laake, P. & Danbolt, N. C. (2009). Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Developmental Psychobiology*, 51(2), 198–206. <https://doi.org/10.1002/dev.20358>
- Levene, H. (1960). Robust tests for equality of variances. I I. Olkin (Red.), *Contributions to probability and statistics* (s. 278–292). Stanford University Press.
- Laake, P., Laake, K. & Aaberge, R. (1985). On the problem of measuring the distance between distribution functions: Analysis of hospitalization versus mortality. *Biometrics*, 41(2), 515–523.
- Meinck, S. & Brese, F. (2019). Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large-Scale Assessments in Education*, 7(1), Artikkel 8. <https://doi.org/10.1186/s40536-019-0076-3>
- Mullis, I. V. S., Martin, M. O. & Loveless, T. (2016). *20 years of TIMSS: International trends in mathematics and science achievement, curriculum, and instruction*. IEA TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O. & von Davier, M. (Red.). (2021). *TIMSS 2023 assessment frameworks*. TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Ni, Y. & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, 40(1), 27–52. https://doi.org/10.1207/s15326985ep4001_3
- Nilsen, T., Björnsson, J. K. & Olsen, R. V. (2018). Hvordan har likeverd i norsk skole endret seg de siste 20 årene? I J. K. Björnsson & R. V. Olsen (Red.), *Tjue år med TIMSS og PISA i Norge: Trender og nye analyser* (s. 150–172). Universitetsforlaget. <https://doi.org/10.18261/9788215030067-2018>
- NOU 2009: 18. (2009). *Rett til læring*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2009-18/id570566/>
- NOU 2019: 3. (2019). *Nye sjanser – bedre læring. Kjønnforskjeller i skoleprestasjoner og utdanningsløp*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2019-3/id2627718/>
- NRICH. (2013, 5. september). *Low threshold high ceiling – an introduction*. <https://nrich.maths.org/articles/low-threshold-high-ceiling-introduction>
- OECD. (2016). *PISA 2015 results: Vol. I. Excellence and equity in education*. <https://doi.org/10.1787/9789264266490-en>
- Opplæringslova. (2023). *Lov om grunnskoleopplæringa og den vidaregåande opplæringa* (LOV-2023-06-09-30). Lovdata. <https://lovdata.no/lov/2023-06-09-30>
- Rutkowski, L., Gonzalez, E., Joncas, M. & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151. <https://doi.org/10.3102/0013189X10363170>
- St.meld. nr. 16 (2006–2007). ... *Og ingen sto igjen. Tidlig innsats for livslang læring*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/stmeld-nr-16-2006-2007-/id441395/>
- Stoet, G. & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PLOS One*, 8(3), Artikkel e57988. <https://doi.org/10.1371/journal.pone.0057988>
- Strietholt, R. (2014). Studying educational inequality: Reintroducing normative notions. I R. Strietholt, W. Bos, J. E. Gustafsson & M. Rosén (Red.), *Educational policy evaluation* (s. 51–58). Waxman.
- Student. (1908). The probable error of the mean. *Biometrika*, 6(1), 1–25. <https://doi.org/10.2307/2331554>
- Thomsen, F. S. (2024). Spesialpedagogisk didaktikk som nøkkelen til en mer inkluderende opplæring. *Spesialpedagogikk*, 2, 22–32.

- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2. utg.). ASCD.
- Tomlinson, C. A. & Imbeau, M. B. (2010). *Leading and managing a differentiated classroom*. ASCD. <http://archive.org/details/leadingmanagingd0000toml>
- Utdanningsdirektoratet. (2023). *Likestilling og likeverd. Rammeplan for barnehagen*. <https://www.udir.no/laring-og-trivsel/rammeplan-for-barnehagen/verdigrunnlag/likestilling-og-likeverd/>
- Vrapi, R., Alia, A. & Brese, F. (2022). Characteristics of high- and low-performing students. I B. J. Pavešić, P. Koršňáková & S. Meinck (Red.), *Dinaric perspectives on TIMSS 2019: Teaching and learning mathematics and science in South-Eastern Europe* (s. 191–212). Springer International Publishing. https://doi.org/10.1007/978-3-030-85802-5_9
- Wollscheid, S., Hjetland, H. N., Rogde, K. & Skjelbred, S.-E. (2018). *Årsaker til og tiltak mot kjønns- forskjeller i skoleprestasjoner. En kunnskapsoversikt* (NIFU-rapport 2018:25). NIFU.
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2), 114–128. <https://doi.org/10.1016/j.stueduc.2005.05.005>

Vedlegg

Tabell 5.A1. Deskriptiv statistikk, 5. trinn, matematikk

	Syklus	Gjennomsnittsskår	SD	Min.	Median	Maks.
Matematikk	2015	550,3	68,9	288,6	551,6	781,7
	2019	546,1	74,9	280,5	550,6	791,1
	2023	531,8	73,6	267,4	535,1	777,6
Tall	2015	544,1	70,8	270,7	545,9	781,5
	2019	543,8	76,0	233,1	546,1	813,0
	2023	531,5	76,0	189,4	534,6	880,5
Geometri	2015	561,9	81,9	188,8	563,1	820,5
	2019	549,4	78,7	257,4	552,5	846,3
	2023	524,4	75,2	238,6	527,8	832,0
Statistikk	2015	567,1	84,7	90,3	567,3	995,0
	2019	551,2	86,2	214,3	553,9	892,5
	2023	537,9	79,2	153,3	541,1	872,8
Gutter	2015	549,2	71,8	288,6	548,8	781,7
	2019	549,9	76,0	280,5	553,6	791,1
	2023	539,6	73,7	267,4	542,8	777,6
Jenter	2015	551,5	65,9	307,3	554,8	770,0
	2019	542,4	73,6	285,0	547,6	744,4
	2023	524,0	72,8	270,3	528,3	754,1

Note. SD = standardavvik, Min. = minimumsverdi, Maks. = maksimumsverdi

Tabell 5.A2. Deskriptiv statistikk, 5. trinn, naturfag

	Syklus	Gjennomsnittsskår	SD	Min.	Median	Maks.
Naturfag	2015	539,5	61,0	297,3	542,2	800,5
	2019	543,3	67,5	295,8	547,9	751,3
	2023	532,8	74,5	189,3	538,2	802,2
Biologi	2015	547,2	63,7	196,3	550,0	752,6
	2019	553,3	68,8	275,1	556,0	759,3
	2023	536,7	74,1	164,2	540,7	791,9
Fys/kjemi	2015	524,6	62,7	260,6	528,9	742,2
	2019	527,1	70,1	250,6	530,8	742,1
	2023	524,1	80,3	190,6	529,6	761,6
Geofag	2015	548,2	76,5	169,7	550,8	916,8
	2019	551,9	77,4	128,8	556,2	813,6
	2023	545,4	84,3	180,9	550,0	854,2
Gutter	2015	540,1	63,2	313,2	543,1	730,5
	2019	543,6	68,8	308,0	550,3	751,3
	2023	532,2	75,5	198,4	537,7	802,2
Jenter	2015	538,9	58,8	297,3	540,7	800,5
	2019	543,0	66,0	295,8	545,3	738,6
	2023	533,4	73,5	189,3	538,9	730,0

Note. SD = standardavvik, Min. = minimumsverdi, Maks. = maksimumsverdi

Tabell 5.A3. Deskriptiv statistikk, 9. trinn, matematikk

	Syklus	Gjennomsnittsskår	SD	Min.	Median	Maks.
Matematikk	2015	512,2	70,0	246,0	515,5	739,4
	2019	506,1	78,7	238,5	506,1	770,6
	2023	500,9	81,5	227,1	502,6	741,0
Tall	2015	528,9	76,2	264,1	529,7	789,6
	2019	509,8	77,5	240,2	511,3	735,7
	2023	495,6	81,1	218,9	497,3	756,6
Geometri	2015	499,1	82,3	209,7	499,7	758,5
	2019	528,9	76,2	264,1	529,7	789,6
	2023	509,8	77,5	240,2	511,3	735,7
Statistikk	2015	542,3	92,5	217,0	547,3	814,2
	2019	518,6	94,5	161,0	522,4	808,6
	2023	521,2	91,9	181,3	526,2	853,4
Algebra	2015	473,9	85,0	199,1	477,0	743,9
	2019	479,7	89,1	159,5	481,1	752,4
	2023	485,4	84,5	190,6	487,2	777,2
Gutter	2015	511,8	71,0	246,0	514,6	739,4
	2019	506,0	80,9	247,6	505,5	770,6
	2023	501,2	83,4	227,1	502,0	735,2
Jenter	2015	512,5	69,1	248,7	515,9	732,7
	2019	506,3	76,5	238,5	507,7	746,3
	2023	500,7	79,3	264,1	503,5	741,0

Note. SD = standardavvik, Min. = minimumsverdi, Maks. = maksimumsverdi

Tabell 5.A4. Deskriptiv statistikk, 9. trinn, naturfag

	Syklus	Gjennomsnittsskår	SD	Min.	Median	Maks.
Naturfag	2015	509,6	79,8	142,1	514,9	759,9
	2019	496,0	91,2	167,0	502,4	754,8
	2023	490,0	91,2	196,3	493,1	783,6
Biologi	2015	502,3	79,4	107,7	506,2	818,5
	2019	487,9	89,5	88,5	493,1	737,1
	2023	483,2	93,1	121,8	485,6	787,5
Fysikk	2015	511,5	87,1	154,1	514,8	841,5
	2019	491,1	99,6	5,0	498,4	785,2
	2023	494,6	103,3	127,6	497,4	891,7
Kjemi	2015	503,2	84,7	194,5	505,9	805,8
	2019	495,6	105,9	78,0	499,9	857,1
	2023	481,1	98,8	98,7	484,5	833,1
Geofag	2015	522,9	84,5	144,2	526,2	817,7
	2019	519,4	100,3	94,4	526,3	818,1
	2023	505,4	89,9	206,9	509,5	809,7
Gutter	2015	510,6	81,5	229,6	515,5	755,6
	2019	497,9	95,9	167,0	504,8	754,8
	2023	489,5	95,2	196,3	493,6	783,6
Jenter	2015	508,6	78,0	142,1	514,5	759,9
	2019	494,3	85,9	177,0	500,6	732,5
	2023	490,7	86,3	204,0	492,7	748,4

Note. SD = standardavvik, Min. = minimumsverdi, Maks. = maksimumsverdi

Tabell 5.A5. Resultater fra Students t-test for sammenligning av gutters og jenters prestasjoner i matematikk og naturfag på 5. og 9. trinn for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

	Matematikk			Naturfag		
	2015	2019	2023	2015	2019	2023
5. trinn	p = 0,257	p = 0,002	p < 0,001	p = 0,510	p = 0,756	p = 0,543
9. trinn	p = 0,738	p = 0,889	p = 0,803	p = 0,388	p = 0,186	p = 0,623

KAPITTEL 6

Kjennetegn på lavt- og høytpresterende elever i TIMSS

Anne-Catherine Lehre og Manvi Aurora Rohatgi

Institutt for lærerutdanning og skoleutvikling, Universitetet i Oslo

Sammendrag: Andelen norske elever på et lavt mestringsnivå i TIMSS-undersøkelsen har økt fra 2015 til 2019 til 2023, mens andelen elever på et avansert mestringsnivå har vært relativt stabilt. Dette kapittelet undersøker hva som kjennetegner elever på et lavt mestringsnivå (definert som lavt og under lavt nivå i TIMSS) sammenlignet med elever som presterer på et avansert nivå i matematikk og naturfag på 5. og 9. trinn. Gruppen med elever på et lavt mestringsnivå kjennetegnes ved lavere sosioøkonomisk status, lavere faglig selvtilit, lavere indre motivasjon (9. trinn), er oftere sultne (5. trinn) og snakker mindre norsk hjemme (naturfag 5. trinn). Med unntak av matematikk på 5. trinn, hvor det er en større andel jenter blant de lavtpresterende (56 %) og en lik kjønnsfordeling på avansert nivå, er andelen gutter og jenter lik blant de lavtpresterende og andelen gutter noe høyere enn andelen jenter på avansert nivå (55–61 %). Det er noen små generelle endringer i elevers gjennomsnittlige karakteristika fra 2015 til 2023 (oftere sultne, lavere fravær, noe redusert motivasjon), men disse endringene er ganske like på tvers av mestringsnivå, fag og trinn.

Nøkkelord: likeverd, mestringsnivå, kontekstuelle faktorer, TIMSS

Abstract

Characteristics of Low- and High-Performing Students in TIMSS

The share of Norwegian students performing at the low benchmarks in TIMSS has increased from 2015 to 2019 to 2023 while the share of students performing at the advanced benchmarks have been relatively stable. This chapter investigates

what characterizes low performing students (at or below low benchmark in TIMSS) compared to students performing at an advanced benchmark in mathematics and natural sciences in grade 5 and grade 9. The characteristics of low performing students are lower socioeconomic status, lower academic self-esteem, lower intrinsic motivation (grade 9), higher levels of hunger (grade 5), and speaking less Norwegian at home (science, grade 5). Except for mathematics in grade 5, where there is a higher share of girls among the low-performing students (56 %) and an equal share of boys and girls at the advanced level, the share of boys and girls is equal among the low-performing students, and the share of boys is somewhat higher than that of girls at the advanced level (55–61 %). There are some general changes in students' average characteristics from 2015 to 2023 (more often hungry, less absenteeism, slightly reduced motivation), but these changes are quite similar across proficiency levels, subjects, and grades.

Keywords: equity, benchmarks, contextual factors, TIMSS

Kompetanse i MNT-fagene (matematikk, naturfag og teknologi) er nøkkel til å løse mange store samfunnsutfordringer innenfor blant annet energi, helse, digitalisering, klima og bærekraft. Det er økende etterspørsel etter MNT-kompetanse i arbeidslivet, både nasjonalt og internasjonalt (Norges forskningsråd, 2012). For å kunne opprettholde en vekst i økonomien og velferden, kreves det at befolkningen har solid kunnskap. For storsamfunnet vil det derfor være viktig at befolkningen har grunnleggende kunnskaper i MNT-fagene. Skolen er en viktig arena for å både vekke interessen for realfag og gi grunnleggende realfagskunnskap til elever. Elever som ikke mestrer grunnleggende matematikk, eller som ikke forstår elementære naturfaglige begreper, vil på lengre sikt for eksempel kunne oppleve færre utdannings- og karrieremuligheter og streve med kritisk vurdering av informasjon innenfor helse, teknologi og økonomi (Norges forskningsråd, 2012).

Å redusere andelen elever som presterer på et lavt nivå, og samtidig øke andelen elever som presterer på et høyt nivå, er et viktig utdanningsmål (Meld. St. 34 (2023–2024); Vrapı et al., 2022). Resultatene fra *Trends in International Mathematics and Science Study* (TIMSS) 2023, viser at andelen norske elever på et lavt mestringsnivå har økt siden forrige gjennomføring av TIMSS i 2019 (Kaarstein et al., 2024). Denne trenden vises for både 5. og 9. trinn elever i matematikk og naturfag. Samtidig viser resultatene at andelen norske elever på et avansert mestringsnivå er noe redusert i perioden, spesielt i matematikk på 5. trinn (Kaarstein et al., 2024). Det kan derfor være interessant å få mer informasjon om hva som kjennetegner elevgrupper på ulike mestringsnivå.

Tidligere forskning har vist at gutter ofte er overrepresentert både blant de som strever og de som gjør det svært godt (Baye & Monseur, 2016; Lehre et al., 2009; Vrapı et al., 2022). I de fleste deltagerlandene i TIMSS-undersøkelsen, har gutter vært overrepresentert blant de elevene som presterer best i matematikk helt siden den første syklusen i 1995 (Vrapı et al., 2022). I norsk skole har særlig gutter med minoritetsbakgrunn og fra lav-SØS-familier utfordringer (NOU 2019: 3). Høy sosioøkonomisk status (SØS) er assosiert med høyere prestasjoner (von Davier et al., 2024).

I dette kapitlet er hovedmålet å karakterisere elevgruppene som presterer på et lavt eller avansert mestringsnivå i TIMSS-undersøkelsen. Å undersøke prestasjonene til elevene i ytterpunktene av prestasjonsskalaen har et potensial til å gi større innsikt i faktorer som bidrar til elevers ulike mestringsnivå. Noen av disse faktorene er det vanskelig å gjøre noe

med, mens andre faktorer vil kunne manipuleres gjennom skolebaserte tiltak, og dermed kan kunnskapen brukes til å skreddersy løsninger for å støtte både de akademisk sterke elevene og de elevene som sliter med de grunnleggende begrepene innen matematikk og naturfag (Meinck & Brese, 2019). Det er ikke alltid lett å bestemme hvilke faktorer som utgjør en forskjell. I dette kapittelet undersøkes hovedsakelig elevsentrerte faktorer som kjønn, SØS, språk hjemme, indre motivasjon og faglig selvtillit i matematikk og naturfag, læringsmiljø (trygghet, tilhørighet), og om eleven er klar for undervisning (tretthet, sult, fravær). Følgende forskningsspørsmål undersøkes:

- 1 Hva kjennetegner elevgruppen med et *lavt* mestringsnivå sammenlignet med elever som presterer på et *avansert* nivå i TIMSS?
- 2 Hvordan endrer karakteriska av elevgruppene med et *lavt* mestringsnivå seg fra TIMSS 2015 til TIMSS 2023?
- 3 Hvordan endrer karakteriska av elevgruppene med et *avansert* mestringsnivå seg fra TIMSS 2015 til TIMSS 2023?

Teori

Kontekstuelle faktorer

Tidligere forskning har vist at ulike faktorer er relatert til elevers prestasjoner, som for eksempel elevenes hjemmebakgrunn, ressurser på skolen, skoleklima, undervisningsmetoder og elevers motivasjon (Mullis et al., 2020; Nilsen & Pettersen, 2024). Disse faktorene ligger på ulike nivå i utdanningssystemet (Teddle & Reynolds, 2000), noe som også gjenspeiles i Creemers og Kyriakides' (2008) helhetlige rammeverk for utdanningseffektivitet, «Den dynamiske modellen». Den dynamiske modellen refererer til faktorer på ulike nivå i skolesystemet (elev, lærer, klasserom, skole og kontekst/system) som påvirker undervisnings- og læringssituasjonen (Creemers & Kyriakides, 2010). Noen av disse faktorene blir berørt i dette kapittelet med hovedfokus på elevsentrerte faktorer som har vist seg viktige for elevers læringsutbytte, og inkluderer SØS, kjønn, hvor ofte eleven snakker norsk hjemme, motivasjon for matematikk og naturfag, læringsmiljø (trygghet, tilhørighet), fravær og elevens fysiske velvære (tretthet, sult).

Forskning viser at SØS, til tross for at SØS blir definert noe ulikt fra studie til studie, har en sammenheng med elevers akademiske prestasjoner

(Mullis et al., 2020; OECD, 2023; Sirin, 2005; White et al., 1993). I TIMSS blir SØS målt noe ulikt for de to elevgruppene som deltar, som foreldrerapportert sosioøkonomisk status for barneskoleelever og som elevenes egenrapporterte læringsressurser hjemme for ungdomsskoleelever, men begge viser at høyere SØS er assosiert med høyere prestasjoner (von Davier et al., 2024). I tillegg til SØS, har andre demografiske faktorer en betydning for elevers prestasjoner, som for eksempel minoritetsbakgrunn (OECD, 2023; von Davier et al., 2024), hvilket språk man snakker hjemme (Kaarstein et al., 2024; Lehre & Nilsen, 2021) og kjønn (Lehre et al., 2009; Meinck & Brese, 2019; Rosén et al., 2020).

Elevers motivasjon og selvtillit er viktige for læring og prestasjoner, og ofte presterer motiverte elever bedre akademisk (Bandura, 1997; Eccles & Wigfield, 2002; Eklöf, 2020). Hva som motiverer en elev til å jobbe med skolefag, kan for eksempel være en indre interesse for å lære mer, et ytre press fra foresatte, et ønske om å få tilhøre en gruppe eller et ønske om å unngå fiasko (Ryan & Deci, 2000). I sin selvbestemmelsesteori beskriver Ryan og Deci (2000) indre og ytre motivasjon, og de hevder at mennesker er mest motiverte når de har kontroll over egen læring, hvilket inkluderer autonomi, kompetanse og tilhørighet. I TIMSS svarer alle elever på spørsmål knyttet til indre motivasjon og faglig selvtillit i naturfag og matematikk, og elever på ungdomstrinnet svarer i tillegg på spørsmål om ytre motivasjon for fagene (Mullis & Martin, 2017). Selv om barneskoleelever rapporterer høyere faglig motivasjon enn ungdomsskoleelever (Eccles et al., 1993; Hascher & Hagenauer, 2010), viser undersøkelser av norske elever som deltok i TIMSS 2015 at den positive sammenhengen mellom prestasjoner og motivasjon og selvtillit var sterkere blant norske ungdomsskoleelever enn blant barneskoleelever (Kaarstein, 2017). Internasjonalt viser faglig selvtillit sterkere positiv sammenhengen med elevers prestasjoner i matematikk og naturfag enn hva indre og ytre motivasjon viser (Ivanova & Michaelides, 2022; Pavešić et al., 2022).

I tillegg til demografiske faktorer og motivasjon, har læringsmiljøet til elevene også en stor betydning for elevers prestasjoner (Nilsen & Teig, 2022). Et godt læringsmiljø inkluderer et godt akademisk miljø, trygghet og lite mobbing, tilhørighet og gode relasjoner samt nok skoleressurser (Thapa et al., 2013; Wang & Degol, 2016). I dette kapittelet fokuseres det på elevens følelse av trygghet og om eleven føler tilhørighet til skolen. Disse faktorene har stor betydning for elevers prestasjoner (Nilsen & Teig, 2022). Videre viser forskningen at elever som ikke føler tilhørighet på skolen eller

som blir mobbet, har høyere sannsynlighet for å ikke møte opp på skolen (Korpershoek et al., 2020; Lee, 2012). Stort fravær kan føre til at elevene svekkes faglig da de går glipp av undervisning og læringsmuligheter, men det holder ikke å bare være fysisk til stede på skolen. For å lære noe og prestere må en elev også være «klar for undervisningen». Fysisk velvære i form av å ha dekket grunnleggende fysiske behov som søvn og mat kan påvirke læringssituasjoner (Alfonsi et al., 2020; Vik et al., 2022b; Vrapı et al., 2022). For eksempel viste en studie av Lin et al. (2020) en direkte sammenheng mellom tenåringsjenters rapporterte antall timers søvn og deres prestasjoner i matematikk. Undersøkelser av norske 9. trinnselever som deltok i TIMSS 2015 og i TIMSS 2019, har vist en signifikant positiv sammenheng mellom det å spise frokost og prestasjoner i matematikk og naturfag (Vik et al., 2022a), og en negativ sammenheng mellom søvnunderskudd og prestasjoner (Vik et al., 2022b).

Tidligere forskning på mestringsnivå

Flertallet av studier som bruker data fra storskalaundersøkelser fokuserer på gjennomsnittsforskjeller i prestasjoner, og det er færre studier som har fokusert på kjønnsforskjeller innenfor ulike mestringsnivå (Meinck & Brese, 2019). Forskning på elever innenfor ulike mestringsnivå har ofte sett på absolutt andel gutter og jenter blant de lavt- eller høytpresterende (f.eks. Hedges & Nowell, 1995; Strand et al., 2006). Felles for disse studiene er at de ofte finner at gutter er overrepresentert i ytterpunktene, hvilket betyr at det er flere gutter enn jenter blant både de som presterer lavest og de som presterer høyest, med unntak av verbale ferdigheter, hvor andelen jenter er høyere enn andelen gutter blant de høytpresterende (Strand et al., 2006).

Baye og Monseur (2016) har sett på andelen jenter og andelen gutter blant lavt- og høytpresterende elever i internasjonale storskalaundersøkelser i matematikk, naturfag og lesing fra 1995 til 2012. For TIMSS-undersøkelsen viser trenden fra 1995 til 2007 en økende andel jenter blant de lavtpresterende i matematikk for både barne- og ungdomsskoleelevene, mens det er flest gutter blant de høytpresterende i hele perioden. Analyser av barneskoleelever som deltok i TIMSS 2011 og PIRLS 2011, viste at gutter var overrepresentert blant de lavtpresterende og de høytpresterende i lesing, matematikk og naturfag (Bergold et al., 2017). For Programme for International Student Assessment (PISA), som undersøker 15-åringers

kompetanse i lesing, matematikk og naturfag, fant Baye og Monseur (2016) at det i perioden 2000–2012 var flest gutter blant høytpresterende og ganske lik kjønnsfordeling blant lavtpresterende. I Norge viste analyser av PISA 2022 at elevgruppen av lavtpresterende i matematikk kjennetegnes ved at det er en noe større andel gutter enn jenter, en større andel elever med innvandrerbakgrunn, lavere sosioøkonomisk status, lavere utholdenhet og mestringsforventning i faget og mer matematikkangst (Jensen & Nilsen, 2024).

Metode

Data

I dette kapittelet benyttes norske data fra TIMSS 2015, TIMSS 2019 og TIMSS 2023. Alle dataene for alle deltagerlandene i TIMSS er offentlig tilgjengelig fra nettsiden <https://timssandpirls.bc.edu/>. I Norge deltok det henholdsvis 4 329, 3 951 og 5 301 elever på 5. trinn og 4 697, 4 575 og 6 324 elever på 9. trinn i TIMSS 2015, TIMSS 2019 og TIMSS 2023 (Bergem et al., 2016; Kaarstein et al., 2020, 2024). Fokuset har vært på elever som presterte på et lavt eller et avansert mestringsnivå i TIMSS-undersøkelsen.

Mestringsnivå

For å lettere kunne tolke resultatene til elever som deltar i TIMSS-undersøkelsen, har TIMSS beskrevet fire ulike internasjonale mestringsnivå (lavt, middels høyt, høyt og avansert) for hvert trinn og hvert fag (matematikk og naturfag). På en skala fra 0 til 1 000, er mestringsnivåene delt inn med følgende poenggrenser 400–474 (lavt), 475–549 (middels høyt), 550–624 (høyt) og 625–1 000 (avansert) (Mullis et al., 2020). I tillegg beregnes andelen elever under lavt nivå (de som ikke oppnår 400 poeng). Hvert av disse nivåene er beskrevet i detalj i de internasjonale rapportene (se f.eks. Mullis et al., 2020; von Davier et al., 2024), men felles for disse er at elevene viser begrenset fagkunnskap på lavt nivå, har økende fagkunnskap og kan anvende denne kunnskapen på de neste to nivåene, og viser forståelse for faget på det avanserte nivået (Mullis et al., 2020; von Davier et al., 2024). Andelen elever som oppnår de ulike mestringsnivåene varierer fra land til land. Tabell 6.1 viser en oversikt over andelen norske 5. trinnselever som presterer på de ulike mestringsnivåene i matematikk og naturfag. Tabell 6.2 viser det tilsvarende for norske elever på 9. trinn.

Tabell 6.1. Andel 5. trinnselever som presterer på de ulike mestringsnivåene i matematikk og naturfag i TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Mestringsnivå	Matematikk: prosentandel elever			Naturfag: prosentandel elever		
	2015 ¹	2019 ²	2023 ³	2015 ⁴	2019 ²	2023 ³
Avansert	14 %	13 %	10 %	7 %	9 %	9 %
Høyt	36 %	35 %	31 %	37 %	37 %	34 %
Middels høyt	36 %	34 %	36 %	41 %	37 %	35 %
Lavt	12 %	15 %	18 %	13 %	14 %	17 %
Under lavt	2 %	3 %	5 %	2 %	3 %	5 %

Note. ¹ Bergem, 2016a ² Kaarstein et al., 2020 ³ Kaarstein et al., 2024 ⁴ Bergem, 2016b

Tabell 6.2. Andel 9. trinnselever som presterer på de ulike mestringsnivåene i matematikk og naturfag i TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Mestringsnivå	Matematikk: prosentandel elever			Naturfag: prosentandel elever		
	2015 ¹	2019 ²	2023 ³	2015 ⁴	2019 ²	2023 ³
Avansert	5 %	5 %	6 %	6 %	6 %	6 %
Høyt	25 %	24 %	23 %	25 %	22 %	26 %
Middels høyt	40 %	36 %	34 %	37 %	33 %	31 %
Lavt	24 %	25 %	25 %	23 %	25 %	26 %
Under lavt	6 %	10 %	12 %	9 %	14 %	17 %

Note. ¹ Bergem, 2016a ² Kaarstein et al., 2020 ³ Kaarstein et al., 2024 ⁴ Bergem, 2016b

Alle oppgaver gitt i TIMSS-undersøkelsen er relatert til et mestringsnivå, og eksempler på oppgaver på de ulike nivåene finnes blant annet i TIMSS 2019 kortrapport (Kaarstein et al., 2020) og TIMSS 2023 kortrapport (Kaarstein et al., 2024).

Variabler

I TIMSS-undersøkelsen blir alle variablene som er inkludert i spørreskjemaene beskrevet i et rammeverk (Reynolds et al., 2021). TIMSS-undersøkelsen ser på trend, og derfor er det kun små endringer i variablene fra syklus til syklus. For å reflektere endringer i samfunnet kan noen variabler fjernes eller nye legges til (Reynolds et al., 2021), og dette kan påvirke utvalget når variabler studeres på tvers av sykluser. Variablene benyttet i dette kapittelet er beskrevet i tabell 6.3.

Tabell 6.3. Oversikt over TIMSS-variablene benyttet i denne studien

Variabel	Utsagn	Svaralternativer
Kjønn ¹		Jente Gutt
SØS	Omtrent hvor mange bøker er det hjemme hos deg?	Ingen eller veldig få (0–10 bøker) Nok til å fylle en bokhyllerad (11–25 bøker) Nok til å fylle en bokhylle (26–100 bøker) Nok til å fylle to bokhyller (101–200 bøker) Nok til å fylle tre eller flere bokhyller (mer enn 200 bøker)
Språk hjemme	Hvor ofte snakker du norsk hjemme?	Alltid Nesten alltid Av og til Aldri
Indre motivasjon	Jeg liker å lære matematikk/naturfag	Svært enig Litt enig Litt uenig Svært uenig
Faglig selvtillit	Jeg gjør det vanligvis bra i matematikk/naturfag	Svært enig Litt enig Litt uenig Svært uenig
Læringsmiljø	Jeg føler meg trygg på skolen	Svært enig Litt enig Litt uenig Svært uenig
	Jeg føler at jeg hører til på denne skolen	Svært enig Litt enig Litt uenig Svært uenig
Klar for undervisning	Når du kommer på skolen, hvor ofte føler du deg trøtt ²	Hver dag Nesten hver dag Av og til Aldri
	Når du kommer på skolen, hvor ofte føler du deg sulten ³	Hver dag Nesten hver dag Av og til Aldri
	Omtrent hvor ofte er du borte fra skolen?	En gang i uka eller mer En gang annenhver uke En gang i måneden En gang annenhver måned ⁴ Aldri eller nesten aldri

Note. ¹ Variabelen kjønn har tre svaralternativer i TIMSS-2023-undersøkelsen (gutt / jente / ingen av delene), men har blitt kodet dikotomt basert på elevdata som ble samlet inn fra skolene før skolene deltok i undersøkelsen. I TIMSS 2015 og TIMSS 2019 var variabelen dikotom. ² Variabelen finnes ikke i spørreskjemaet i TIMSS 2015. ³ I 2015 var spørsmålet: «Hvor ofte spiser du frokost de dagene du er på skolen?» ⁴ Dette svaralternativet fantes kun i 2023.

For de inkluderte variablene i denne studien, er det en økning i prosentandelen manglende elevsvar i perioden. Det er også generelt en lavere prosentandel elevsvar blant de lavtpresterende. For prosentandeler manglende elevsvar, se tabellene 6.A1–6.A4 i vedlegget.

Analyser

For å sammenligne karakteristika for elever på et lavt mestringsnivå med elever på et avansert mestringsnivå, ble det beregnet gjennomsnitt og statistisk signifikans for ulike variabler for elevgruppene på forskjellige mestringsnivåene (Lorah, 2022) ved å benytte IEAs IDB Analyser (versjon 5.0.39) og IBM SPSS (versjon 30.0.0.0(172)). Før beregningen av gjennomsnitt, ble alle variablenes svaralternativer rekodet slik at de mest positive svaralternativene fikk høyeste verdi og de minst positive svaralternativene fikk verdien 0. For eksempel ble svaralternativene til «Når du kommer på skolen, hvor ofte føler du deg trøtt» rekodet til følgende verdier: *Hver dag* = 0, *Nesten hver dag* = 1, *Av og til* = 2 og *Aldri* = 3. Gjennomsnittsverdiene for variablene og sammenligningen av elevgruppene ble visualisert med radarkart laget i Microsoft® Excel® for Microsoft 365 MSO (Versjon 2408, 16.0.17928.20336).

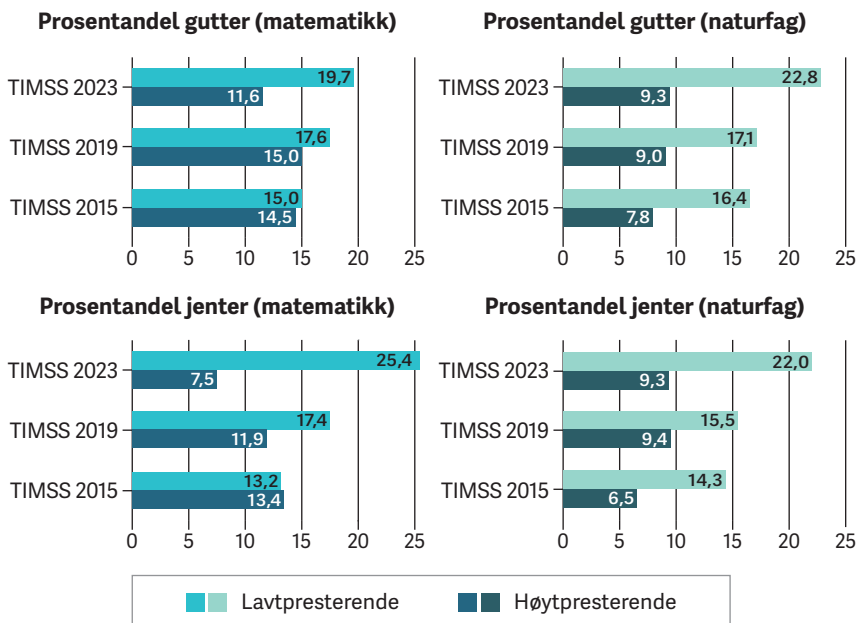
Resultater

Først presenteres gjennomsnittlige karakteristika av 5. trinnselever med et lavt mestringsnivå i matematikk og i naturfag, sammenlignet med elever på et avansert mestringsnivå, for TIMSS 2015, TIMSS 2019 og TIMSS 2023. Så presenteres gjennomsnittlige karakteristika av 9. trinnselever med et lavt mestringsnivå i matematikk og i naturfag, sammenlignet med elever på et avansert mestringsnivå, for TIMSS 2015, TIMSS 2019 og TIMSS 2023. Deretter presenteres endringer i karakteristika for 5. og 9. trinnselever på et lavt mestringsnivå, og til slutt presenteres endringer i karakteristika for elever på et avansert mestringsnivå.

Forskningsspørsmål 1: Kjennetegn for elevgruppene på lavt og avansert mestringsnivå

Figur 6.1 gir en oversikt over hvor stor andel av guttene og jentene som presterte på et lavt eller et avansert mestringsnivå i matematikk og naturfag

på 5. trinn. Fra 2015 til 2023 var det en økning i andelen elever som presterte på et lavt mestringsnivå, fra rett under 15 prosent av elevmassen i 2015 til ca. 22 prosent av elevmassen i 2023. I 2015 var det noen flere gutter enn jenter som presterte på et lavt mestringsnivå i begge fagene (53 % gutter). I 2023 var det en jevn fordeling av gutter og jenter som presterte på et lavt mestringsnivå i naturfag, mens det var en overvekt av jenter som presterte på et lavt mestringsnivå i matematikk (56 % jenter).



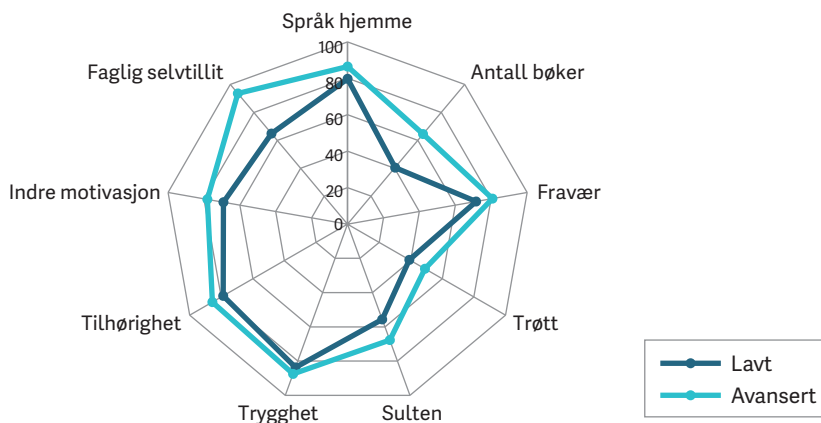
Figur 6.1. Oversikt over hvor stor andel av guttene og jentene som presterte på et lavt eller et avansert mestringsnivå i matematikk og naturfag på 5. trinn i TIMSS 2015, TIMSS 2019 og TIMSS 2023.

Andelen elever som presterte på et avansert nivå i matematikk har sunket i perioden 2015–2023, og da særlig blant jenter i 2023 (figur 6.1). Det er relativt sett flere gutter enn jenter blant de høytpresterende i matematikk, og i 2023 utgjorde gutter 61 prosent av de som presterte på et avansert nivå i faget. I naturfag har andelen elever på et avansert nivå steget noe fra 2015 til 2023: fra 6,5 prosent (jenter) og 7,8 prosent (gutter) til 9,3 prosent (begge). Videre har spredningen i både gutters og jenters naturfagsprestasjoner økt fra 2015 til 2023 ved at det er blitt flere elever av begge kjønn på et lavt og et avansert mestringsnivå i 2023.

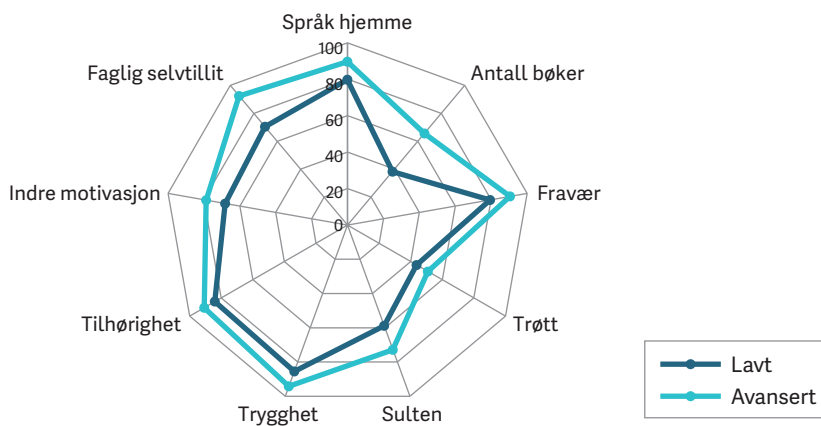
Figur 6.2 viser gjennomsnittlige karakteristika av 5. trinnselever på et lavt mestringsnivå sammenlignet med de på et avansert mestringsnivå i matematikk. Høye verdier peker på et mer positivt gjennomsnitt enn lave verdier. De største forskjellene i karakteristika av elevgruppene er «antall bøker» (proxy for SØS) og «faglig selvtilitt», hvor elevgruppen på avansert nivå rapporterer mer positive verdier. For de andre variablene er det små forskjeller mellom elevgruppene.

Dersom differansen i gjennomsnittene i 2023 rangeres for de fire gjennomsnittene med størst forskjell mellom elevgruppene, fra størst til minst, viser «faglig selvtilitt» størst differanse med SØS rett bak, og deretter kommer «sulten» og «trøtt». Det er ingen store endringer i hvilke variabler

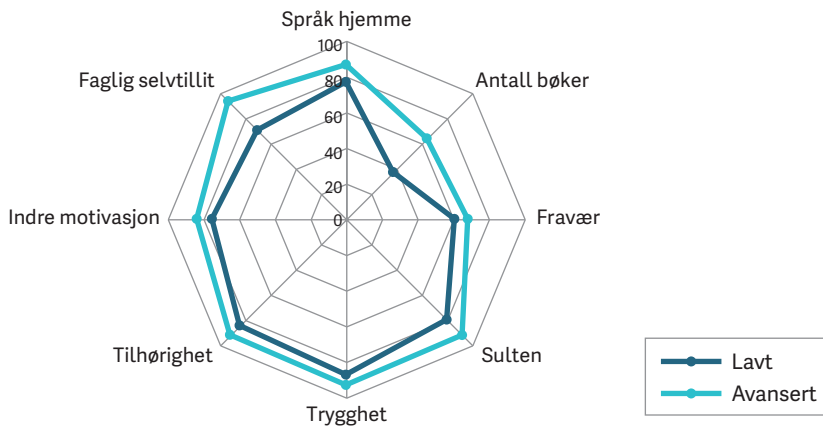
Matematikk, T23 (5. trinn)



Matematikk, T19 (5. trinn)



(Forts.)

Matematikk, T15 (5. trinn)

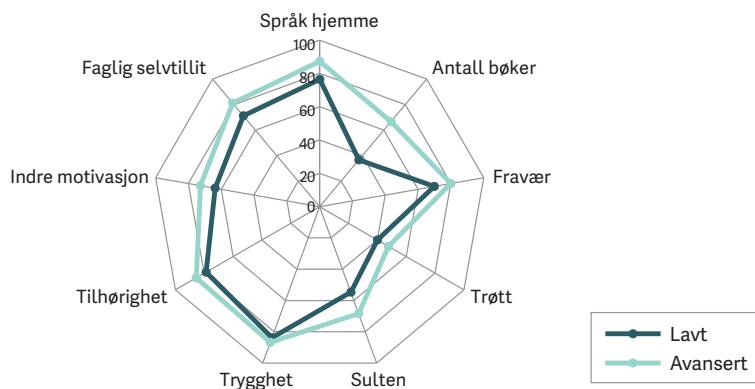
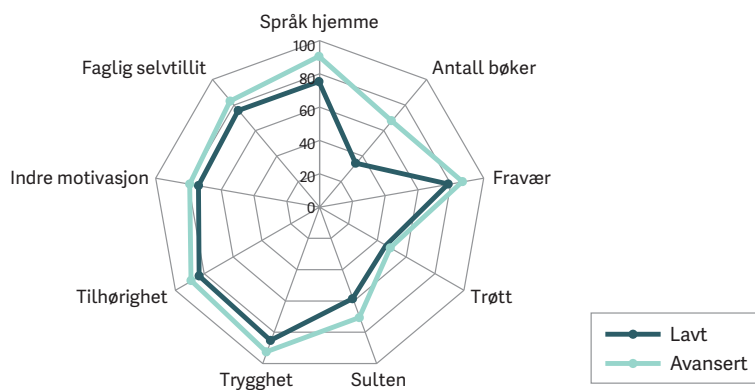
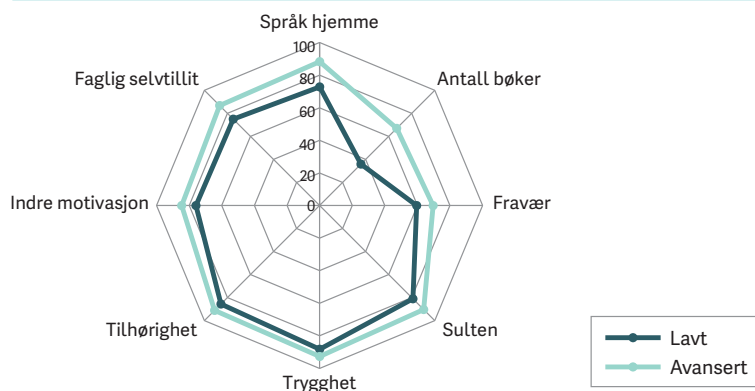
Figur 6.2. Gjennomsnittlige karakteristika av 5. trinnselever med et lavt mestringsnivå i matematikk sammenlignet med elever som presterer på et avansert nivå, TIMSS 2015, TIMSS 2019 og TIMSS 2023.

Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015.

som karakteriserer de to elevgruppene på tvers av sykluser, men i 2015 og 2019 viste SØS større forskjell mellom elevgruppene enn faglig selvtillit. «Sulten» var rangert som nummer 3 i hele perioden. Forskjellene mellom elevgruppene var signifikant forskjellige for alle variabler i 2019 og 2023, men ikke i 2015 (tabell 6.A5).

Figur 6.3 viser gjennomsnittlige karakteristika av 5. trinnselever på et lavt mestringsnivå sammenlignet med de på et avansert mestringsnivå i naturfag. Høye verdier peker på et mer positivt gjennomsnitt enn lave verdier. Den største forskjellen i karakteristika av elevgruppene er «antall bøker» (proxy for SØS), hvor elevgruppen på avansert nivå rapporterer mer positive verdier.

Dersom differansen i gjennomsnittene i 2023 rangeres for de fire gjennomsnittene med størst forskjell mellom elevgruppene, fra størst til minst, viser SØS størst differanse, deretter kommer «sulten» og «språk hjemme», og til slutt «faglig selvtillit». Det er ingen store endringer i hvilke variabler som karakteriserer de to elevgruppene på tvers av sykluser. I 2015 og 2019 viste SØS størst forskjell mellom elevgruppene, og «språk hjemme» den nest største differansen. Forskjellene mellom elevgruppene var signifikant forskjellige for alle variabler i 2019 og 2023, men ikke i 2015 (tabell 6.A5).

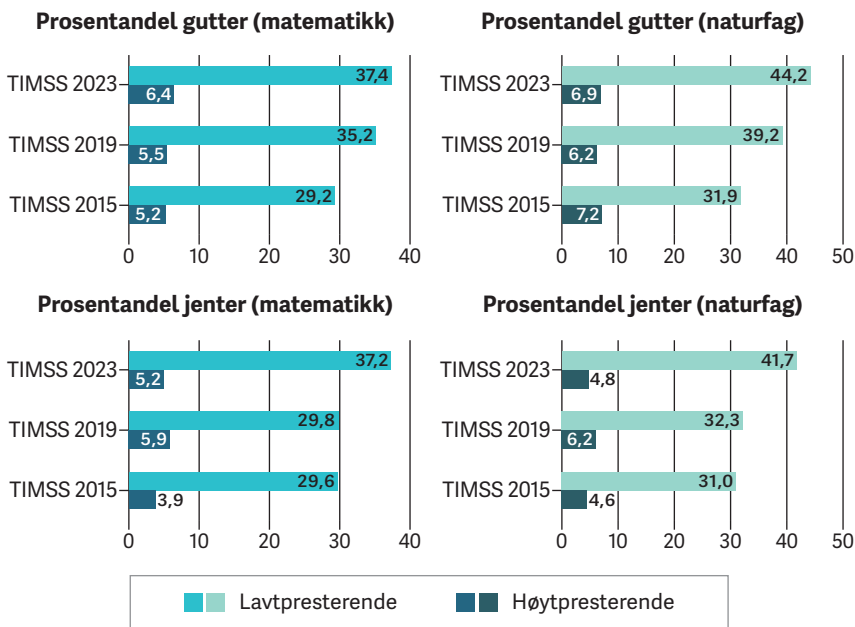
Naturfag, T23 (5. trinn)**Naturfag, T19 (5. trinn)****Naturfag, T15 (5. trinn)**

Figur 6.3. Gjennomsnittlige karakteristika av 5. trinnselever med et lavt mestringsnivå i naturfag sammenlignet med elever som presterer på et avansert nivå i TIMSS 2015, TIMSS 2019 og TIMSS 2023.

Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015.

Figur 6.4 gir en oversikt over hvor stor andel av guttene og jentene som presterte på et lavt eller et avansert mestringsnivå i matematikk og naturfag på 9. trinn. Fra 2015 til 2023 var det en økning i andelen elever som presterte på et lavt mestringsnivå, fra ca. 30 prosent av elevmassen i 2015 til ca. 40 prosent av elevmassen i 2023. Både i 2015 og i 2023 var det en jevn fordeling av gutter og jenter som presterte på et lavt mestringsnivå i begge fagene.

Andelen elever som presterte på et avansert nivå har økt noe i perioden 2015–2023, og omtrent en av tjue elever nådde opp til avansert mestringsnivå i 2023 (figur 6.4). Det er relativt sett flere gutter enn jenter blant de høytpresterende i både matematikk og naturfag. I 2023 utgjorde gutter 55 prosent av de som presterte på et avansert nivå i matematikk og 59 prosent av de som presterte på et avansert nivå i naturfag. I matematikk har spredningen i jentenes prestasjoner steget fra 2015 til 2023 ved at andelen jenter på både et lavt eller et avansert mestringsnivå har økt.



Figur 6.4. Oversikt over hvor stor andel av guttene og jentene som presterte på et lavt eller et avansert mestringsnivå i matematikk og naturfag på 9. trinn i TIMSS 2015, TIMSS 2019 og TIMS 2023.

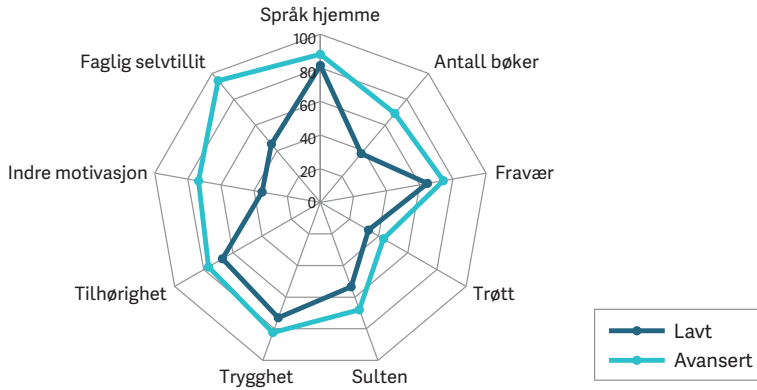
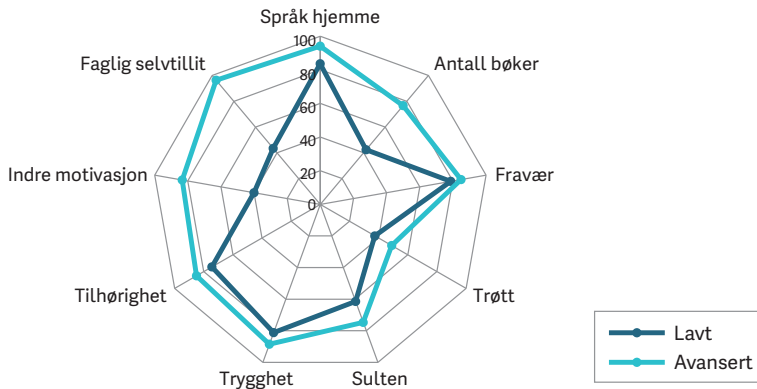
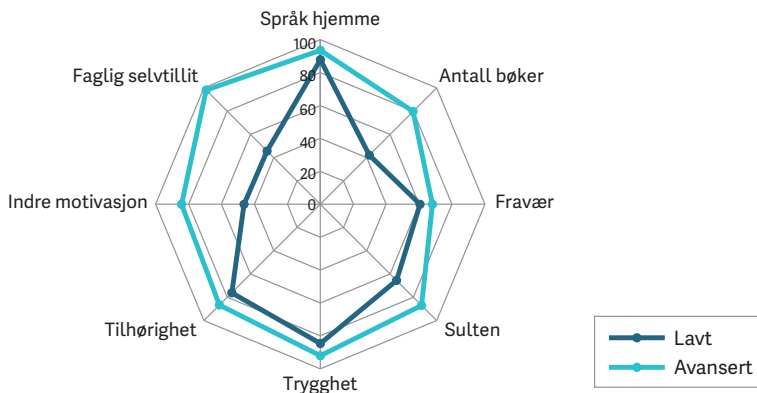
Figur 6.5 viser gjennomsnittlige karakteristika av 9. trinnselever på et lavt mestringsnivå sammenlignet med de på et avansert mestringsnivå i matematikk. Høye verdier peker på et mer positivt gjennomsnitt enn lave

verdier. De største forskjellene i karakteristika av elevgruppene er «faglig selvtillit», «indre motivasjon» og «antall bøker» (proxy for SØS), hvor elevgruppen på avansert nivå rapporterer mer positive verdier. For de andre variablene er det små forskjeller mellom elevgruppene.

Dersom differansen i gjennomsnittene i 2023 rangeres for de fire gjennomsnittene med størst forskjell mellom elevgruppene, fra størst til minst, viser «faglig selvtillit» størst differanse med «indre motivasjon» og SØS rett bak, og deretter «sulten». Det er ingen store endringer i hvilke variabler som karakteriserer de to elevgruppene på tvers av sykluser. Forskjellene mellom elevgruppene var signifikant forskjellige for alle variabler i 2015, 2019 og 2023 (tabell 6.A6).

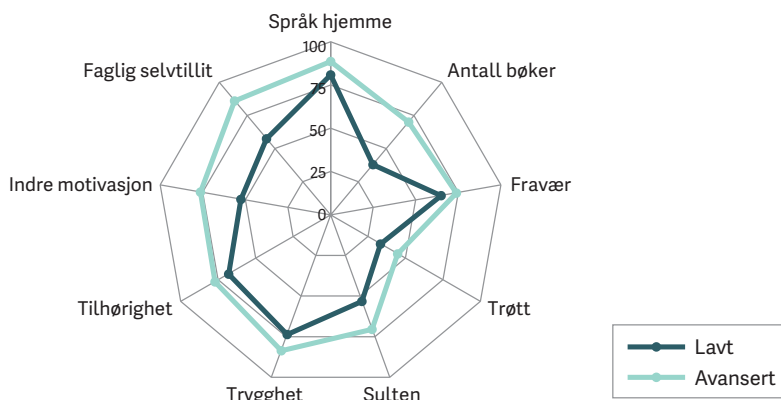
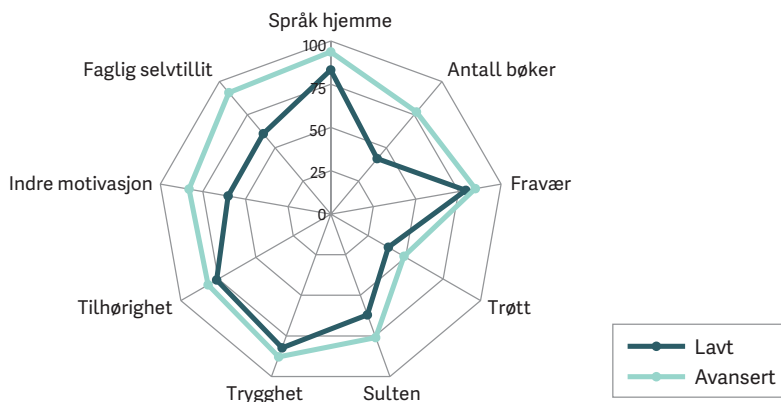
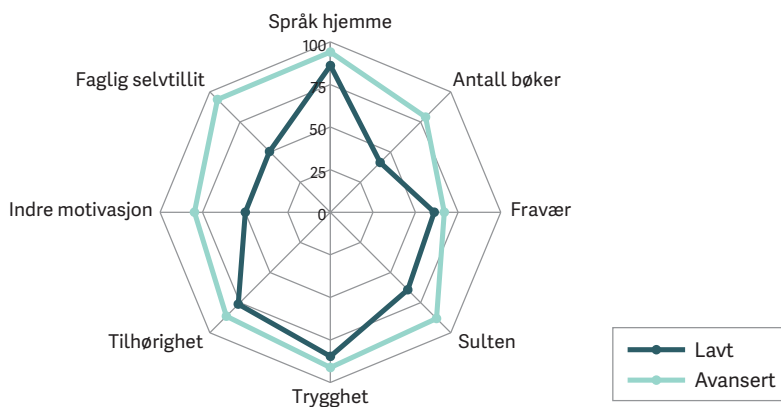
Figur 6.6 viser gjennomsnittlige karakteristika av 9. trinnselever på et lavt mestringsnivå sammenlignet med de på et avansert mestringsnivå i naturfag. Høye verdier peker på et mer positivt gjennomsnitt enn lave verdier. De største forskjellene i karakteristika av elevgruppene er «antall bøker» (proxy for SØS), «faglig selvtillit», «indre motivasjon» og «sulten», hvor elevgruppen på avansert nivå rapporterer mer positive verdier. For de andre variablene er det små forskjeller mellom elevgruppene.

Dersom differansen i gjennomsnittene i 2023 rangeres for de fire gjennomsnittene med størst forskjell mellom elevgruppene, fra størst til minst, viser SØS størst differanse med «faglig selvtillit» og «indre motivasjon» rett bak, og deretter «sulten». Det er ingen store endringer i hvilke variabler som karakteriserer de to elevgruppene på tvers av sykluser, men i 2015 viste «faglig selvtillit» større differanse mellom elevgruppene enn SØS. Forskjellene mellom elevgruppene var signifikant forskjellige for alle variabler i 2015, 2019 og 2023 (tabell 6.A6).

Matematikk, T23 (9. trinn)**Matematikk, T19 (9. trinn)****Matematikk, T15 (9. trinn)**

Figur 6.5. Gjennomsnittlige karakteristika av 9. trinnselever med et lavt mestringsnivå i matematikk sammenlignet med elever som presterer på et avansert nivå, TIMSS 2015, TIMSS 2019 og TIMSS 2023.

Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015.

Naturfag, T23 (9. trinn)**Naturfag, T19 (9. trinn)****Naturfag, T15 (9. trinn)**

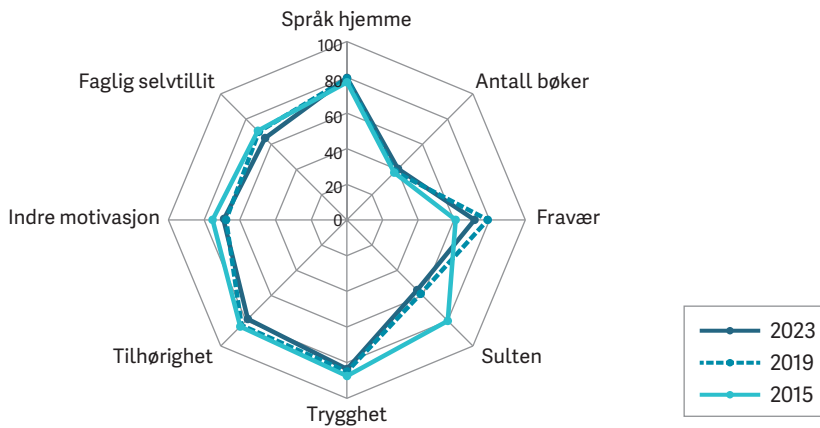
Figur 6.6. Gjennomsnittlige karakteristika av 9. trinnselever med et lavt mestringsnivå i naturfag sammenlignet med elever som presterer på et avansert nivå i TIMSS 2015, TIMSS 2019 og TIMSS 2023.

Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015.

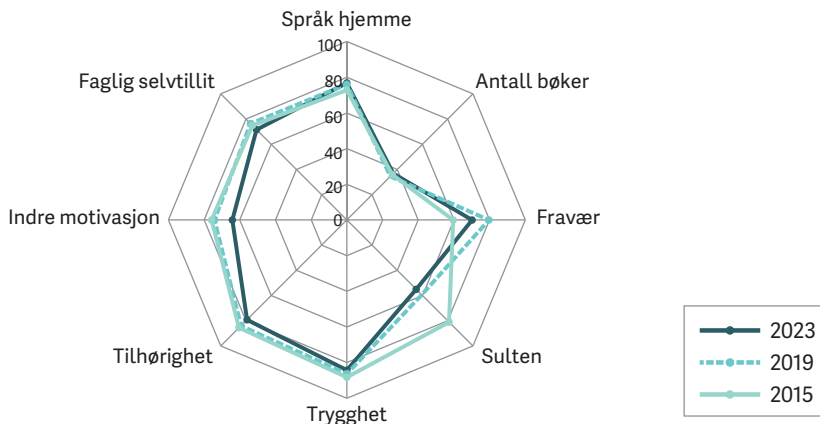
Forskningsspørsmål 2: Endringer i karakteristika av elevgruppen med et lavt mestringsnivå

Figur 6.7 viser gjennomsnittlige karakteristika av 5. trinnselever på et lavt mestringsnivå i matematikk og naturfag i perioden 2015–2023. Det er noen små generelle endringer i elevers gjennomsnittlige karakteristika fra 2015 til 2023 for begge fagene: Elevgruppen er oftere sultne, har mindre fravær og noe redusert motivasjon. Forskjellene mellom elevgruppene var signifikant forskjellige for alle variabler i 2019 og 2023, men ikke i 2015 (tabell 6.A6).

Matematikk, 5. trinn (lavt)



Naturfag, 5. trinn (lavt)

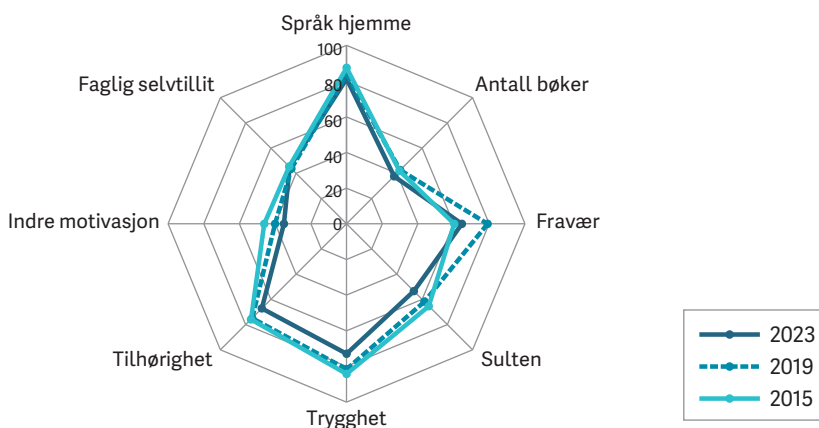


Figur 6.7. Gjennomsnittlige karakteristika av 5. trinnselever med et lavt mestringsnivå i matematikk og i naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023.

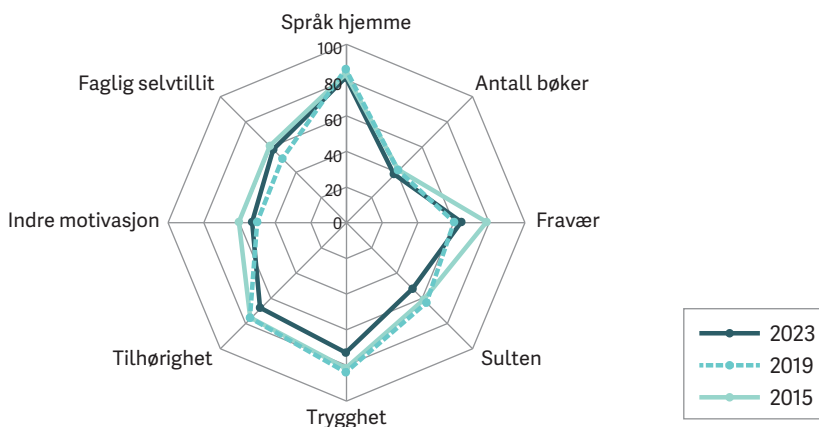
Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015 og er derfor utelatt.

Figur 6.8 viser gjennomsnittlige karakteristika av 9. trinnselever på et lavt mestringsnivå i matematikk og naturfag i perioden 2015–2023. Det er noen små generelle endringer i elevers gjennomsnittlige karakteristika fra 2015 til 2023 for begge fagene: Elevgruppen rapporterer noe lavere tilhørighet, lavere faglig motivasjon, noe redusert trygghetsfølelse og oftere sultne. Forskjellene mellom elevgruppene var signifikant forskjellige for alle variabler i 2015, 2019 og 2023 (tabell 6.A6).

Matematikk, 9. trinn (lavt)



Naturfag, 9. trinn (lavt)



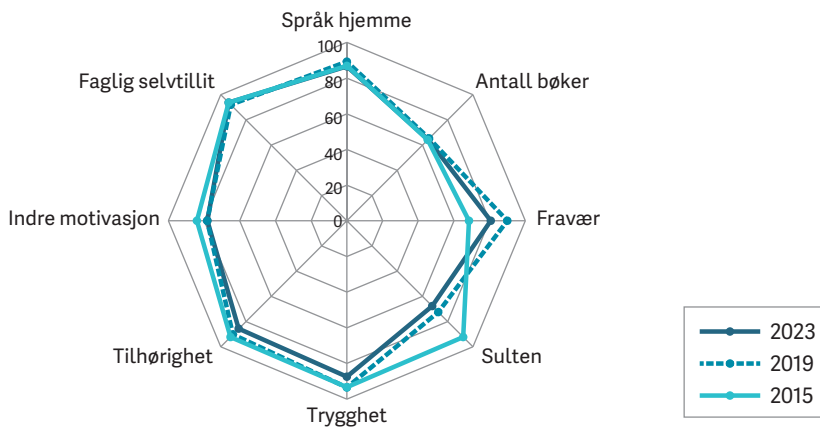
Figur 6.8. Gjennomsnittlige karakteristika av 9. trinnselever med et lavt mestringsnivå i matematikk og i naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023.

Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015 og er derfor utelatt.

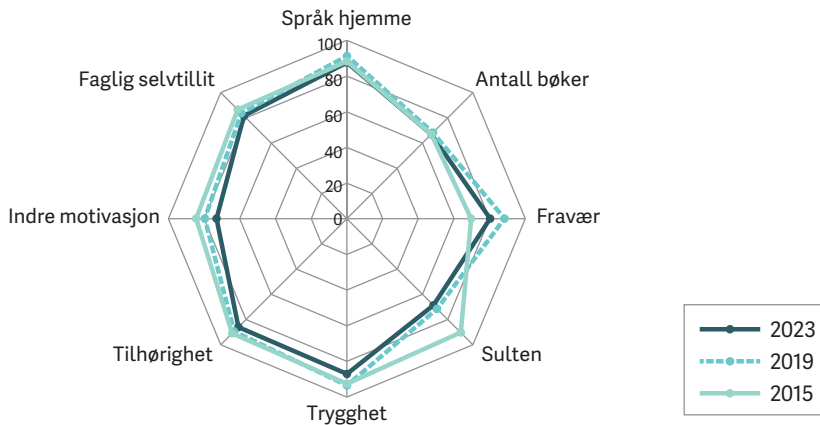
Forskningsspørsmål 3: Endringer i karakteristika av elevgruppen med et avansert mestringsnivå

Figur 6.9 viser gjennomsnittlige karakteristika av 5. trinnselever på et avansert mestringsnivå i matematikk og naturfag i perioden 2015–2023. Det er noen små generelle endringer i elevers gjennomsnittlige karakteristika fra 2015 til 2023 for begge fagene: Elevgruppen er oftere sultne, har mindre fravær og noe redusert motivasjon (spesielt i naturfag).

Matematikk, 5. trinn (avansert)



Naturfag, 5. trinn (avansert)

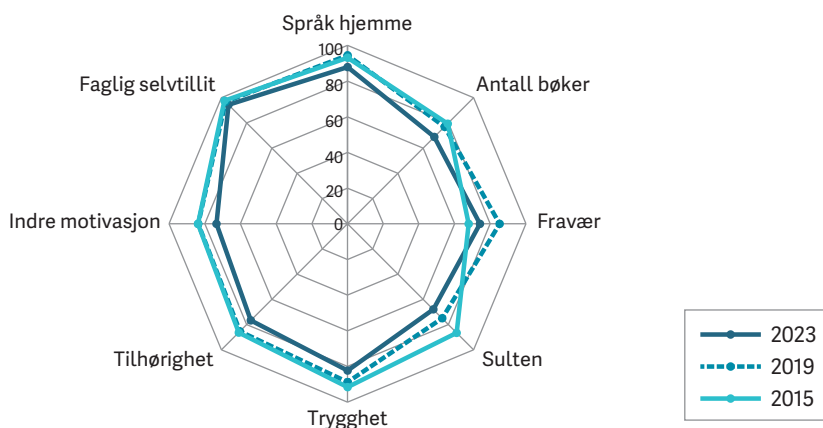


Figur 6.9. Gjennomsnittlige karakteristika av 5. trinnselever med et avansert mestringsnivå i matematikk og i naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023. Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015 og er derfor utelatt.

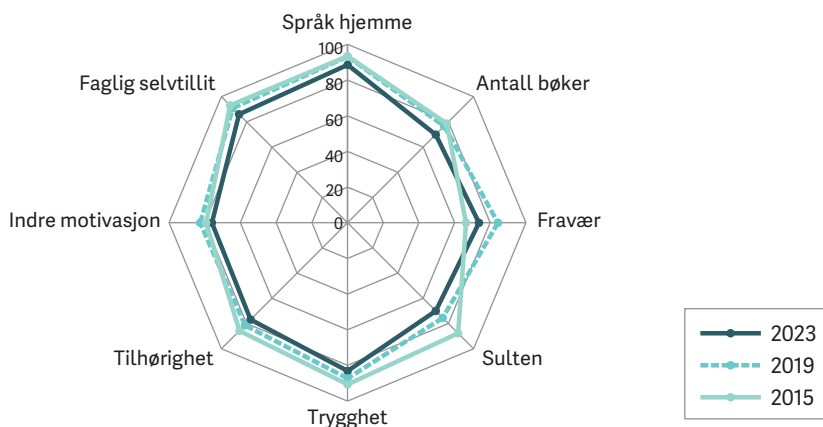
Elevgruppen rapporterer også noe lavere tilhørighet og noe redusert trygghetsfølelse.

Figur 6.10 viser gjennomsnittlige karakteristika av 9. trinnselever på et avansert mestringsnivå i matematikk og naturfag i perioden 2015–2023. Det er noen små generelle endringer i elevers gjennomsnittlige karakteristika fra 2015 til 2023 for begge fagene: Elevgruppen er oftere sultne og har mindre fravær. Elevgruppen rapporterer også noe lavere tilhørighet, noe redusert trygghetsfølelse og litt lavere gjennomsnittlig SØS.

Matematikk, 9. trinn (avansert)



Naturfag, 9. trinn (avansert)



Figur 6.10. Gjennomsnittlige karakteristika av 9. trinnselever med et avansert mestringsnivå i matematikk og i naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023. Note. Utsagnet «Når du kommer på skolen, hvor ofte føler du deg trøtt?» ble ikke stilt i 2015 og er derfor utelatt.

Diskusjon

Sammendrag av resultater

Hovedmålet med dette kapittelet var å karakterisere elevgruppene som presterer på et lavt eller et avansert mestringsnivå i TIMSS. Analysene viser at elevgruppen som presterer på et lavt mestringsnivå kjennetegnes ved lavere sosioøkonomisk status, lavere faglig selvtillit, lavere indre motivasjon (9. trinn), er oftere sultne (5. trinn) og snakker mindre norsk hjemme (naturfag 5. trinn). Med unntak av matematikk på 5. trinn, hvor det er en større andel jenter blant de lavtpresterende (56 %) og en lik kjønnsfordeling på avansert nivå, er andelen gutter og jenter lik blant de lavtpresterende og andelen gutter noe høyere enn andelen jenter på avansert nivå (55–61 %). Det er noen små generelle endringer i elevers gjennomsnittlige karakteristika fra 2015 til 2023 (oftere sultne, mindre fravær, noe redusert indre motivasjon), og disse endringene er ganske like på tvers av mestringsnivå, fag og trinn.

Karakteristika for lavt og avansert mestringsnivå

At analysene her viser at elevgruppen på et lavt mestringsnivå kjennetegnes ved gjennomsnittlig lavere sosioøkonomisk status (figurene 6.2, 6.3, 6.5 og 6.6) enn elevgruppen som presterer på et avansert nivå, er ikke uventet. Dette samsvarer med internasjonal forskning som viser at det er en sammenheng mellom elevers prestasjoner og SØS (Mullis et al., 2020; OECD, 2023; Sirin, 2005; White et al., 1993). Videre samsvarer dette med analyser av PISA 2022-prestasjonene til norske elever på et lavt mestringsnivå i matematikk, hvor den lavtpresterende elevgruppen hadde lavere gjennomsnittlig sosioøkonomisk status (Jensen & Nilsen, 2024). Sistnevnte studie viste også at elevgruppen med et lavt mestringsnivå i matematikk inneholdt en større andel elever med innvandrerbakgrunn (Jensen & Nilsen, 2024). Hvilket språk elevene snakker hjemme har betydning for prestasjonene, og spesielt i naturfag (Kaarstein et al., 2024; Lehre & Nilsen, 2021), hvilket også er tilfelle for elevgruppen på 5. trinn naturfag her (figur 6.3).

Tidligere forskning har vist at motiverte elever ofte presterer bedre akademisk (Bandura, 1997; Eccles & Wigfield, 2002; Eklöf, 2020), og at faglig selvtillit har sterkere positiv sammenheng med elevers prestasjoner i matematikk og naturfag enn hva indre og ytre motivasjon viser (Ivanova & Michaelides, 2022; Pavešić et al., 2022). Analysene her viser at det er en

større differanse i gjennomsnittlig faglig selvtillit enn gjennomsnittlig indre motivasjon mellom elevgruppene på lavt og avansert mestringsnivå, og da spesielt på 9. trinn. Dette er ikke overraskende, da faglig selvtillit måles ved utsagnet «Jeg gjør det vanligvis bra i matematikk/naturfag» (tabell 6.3), og en ungdomsskoleelev som strever i et fag vil dermed sannsynligvis ikke rapportere at hen gjør det bra i faget.

Studier viser at læringsmiljøet til elevene har en stor betydning for elev-ers prestasjoner (Nilsen & Teig, 2022). I dette kapitlet er det kun sett på elevers følelse av tilhørighet og trygghet på skolen, og disse analysene viser kun små forskjeller mellom elevgruppene på lavt og avansert mestringsnivå (figurene 6.2, 6.3, 6.5 og 6.6). Alle elevgruppene rapporterer om noe redusert følelse av tilhørighet og trygghet på skolen, da særlig i perioden 2019–2023, men gjennomsnittlig trygghet på skolen i Norge er høy sett i et internasjonalt perspektiv (Kaarstein et al., 2024).

For å kunne jobbe godt med skolefagene, må en elev være klar for undervisningen. Dette er her undersøkt ved å se på hvor ofte elevene kommer sultne eller trøtte på skolen, og for begge elevgruppene er det en negativ trend fra 2015 til 2023. Forskning har vist at å få dekket grunnleggende fysiske behov som søvn og mat kan påvirke læringssituasjoner (Alfonsi et al., 2020; Vik et al., 2022b, 2022a; Vrapi et al., 2022), og at det er en direkte sammenheng mellom elevers rapporterte antall timers søvn og deres prestasjoner i matematikk (Lin et al., 2020) og i naturfag (Vik et al., 2022b). Det kunne vært interessant å undersøke søvnmønsteret til de lavtpresterende elevene i matematikk på 5. trinn da det er en overvekt av jenter på dette nivået (56 prosent jenter) og variabelen «trøtt» kommer opp på fjerdeplass over de største forskjeller i differansen mellom elevgruppene på lavt og avansert mestringsnivå i matematikk på 5. trinn.

Kjønn

Analyser av internasjonale storskalaundersøkelser i perioden 1995–2012 viste en økende andel jenter blant de lavtpresterende i matematikk for både barne- og ungdomsskoleelevene, mens det var flest gutter blant de høytpresterende i matematikk og naturfag (Baye & Monseur, 2016). PIRLS 2011 og TIMSS 2011 viste at gutter var overrepresentert blant de lavtpresterende og de høytpresterende i lesing, matematikk og naturfag (Bergold et al., 2017). Data fra PISA 2022 viste at blant norske 15-åringer er det flere lavtpresterende gutter enn jenter i matematikk (Jensen & Nilsen, 2024) samt flere gutter på de øverste mestringsnivåene (Jensen et al., 2023). Dette samsvarer

med resultatene her, som viser at andelen gutter på et avansert mestringsnivå er større enn andelen jenter for 5. trinn matematikk (61 prosent gutter) og 9. trinn matematikk (55–59 % gutter). I naturfag på 5. trinn er kjønnsfordelingen lik på avansert mestringsnivå. For elevgruppene på lavt mestringsnivå, er det lik kjønnsfordeling for matematikk 9. trinn og naturfag på begge trinn. I matematikk på 5. trinn har det vært en økende andel jenter blant de lavtpresterende, og de utgjør 56 prosent av elevgruppen i TIMSS 2023. Dette samsvarer med trenden som har blitt observert helt fra 1995 med en økende andel jenter blant de lavtpresterende i matematikk (Baye & Monseur, 2016). I tillegg har andelen jenter på et avansert nivå i matematikk blitt redusert fra 2015 til 2023 (figur 6.1). Dette er en trend som bør analyseres nærmere for å kunne sette inn tiltak for å hindre ytterligere reduksjon i jenters matematikprestasjoner. Grunnleggende matematikkunnskaper er blant annet viktige byggesteiner for andre fag høyere opp i utdanningssystemet.

Begrensninger ved studien

TIMSS-undersøkelsen er en tverrsnittsundersøkelse, og det er derfor vanskelig å si noe om hvilken retning årsakssammenhenger har da en sammenheng kan gå begge veier (f.eks. holdninger og prestasjoner; Seaton et al., 2010). For å kunne si noe om årsakssammenhenger, er randomiserte kontrollerte studiedesign det beste. Styrken ved TIMSS-undersøkelsen er at den har representative utvalg av elevene på 5. og 9. trinn, og kan gi et godt øyeblikksbilde av forholdene på det tidspunktet den gjennomføres.

Elevgruppene som er definert som lavtpresterende eller høytpresterende, viser ikke til en generell egenskap ved den enkelte elev. Elevgruppene er definert ut fra beskrivelsene av de ulike mestringsnivåene i TIMSS-undersøkelsen (Mullis et al., 2020). Videre har elevgruppen på et lavt mestringsnivå en høyere prosentandel manglende svar på spørreskjemaene enn elevgruppen på et avansert mestringsnivå (tabellene 6.A1–6.A4).

Denne studien har ikke inkludert variabler fra spørreskjemaene til faglærere eller rektorer, men sett på spørreskjemaene besvart av elevene. Videre forskning kan validere resultatene ved å inkludere overlappende spørsmål fra ulike spørreskjema besvart av elever, foresatte, faglærere og rektorer.

Sluttkommentar

Kuttpunktene for mestringsnivåene varierer mellom fag, og klassifiseringen av å ha en lav mestringsgrad i et fag påvirkes av hvor kuttpunktet et

satt. Generelt er ferdigheter normalfordelt, og innenfor matematikk er det beregnet at ca. 7 prosent av elevene har matematikkvansker og at ytterligere 10–15 prosent har milde, men varige, vansker i faget (Geary et al., 2007). Det er dermed forventet at det vil være en til to elever med matematikkvansker i hver skoleklasse (Azemi & Vestby, 2024). Resultatene i dette kapittelet viser at andelen med elever på et lavt mestringsnivå i matematikk (5. trinn: 23 %; 9. trinn: 37 %) er noe høyere enn hva som er forventet (17–22 %). Videre viser resultatene at andelen elever på et lavt mestringsnivå er økende både i naturfag og matematikk, uten at kuttpunktet for mestringsnivåene er endret fra 2015 til 2023. Den samme utviklingen finner man igjen i matematikk i PISA-undersøkelsen, hvor andelen 15-åringer på et lavt mestringsnivå (under nivå 2) har økt fra 19 prosent i 2018 til 31 prosent i 2022 (OECD, 2023). Nasjonale prøver i regning viste en svakt økende andel 5. trinnselever på laveste mestringsnivå fra 2022 til 2023, fra 25,2 til 26,3 prosent (Utdanningsdirektoratet, 2023). At ulike typer prøver viser at andelen elever på et lavt mestringsnivå er økende, er foruroligende for et skolesystem som både bygger på en egalitær ideologi med forventning om kun små forskjeller mellom elevers prestasjoner (Blossing et al., 2014) og som har likeverd som en grunntanke (Nilsen et al., 2018; Opplæringslova, 1998). I et likeverdsperspektiv har «alle individer en rett til å oppnå det minimumsnivået av elementære kunnskaper som er nødvendige for å klare seg i samfunnet» (Nilsen et al., 2018, s. 154).

Forfatterbiografier

Anne-Catherine Lehre er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en bakgrunn innenfor språk og realfag, og har en doktorgrad i biostatistikk. Lehre har jobbet i TIMSS-prosjektet siden 2020.

Manvi Aurora Rohatgi er rådgiver ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en mastergrad i samfunns-sikkerhet og praktisk pedagogisk utdanning, og har jobbet med storskalaundersøkelser ved UiO siden 2021.

Referanser

- Alfonsi, V., Scarpelli, S., D'Atri, A., Stella, G. & De Gennaro, L. (2020). Later school start time: The impact of sleep on academic performance and health in the adolescent population. *International Journal of Environmental Research and Public Health*, 17(7), Artikkel 2574. <https://doi.org/10.3390/ijerph17072574>

- Azemi, F. & Vestby, M. H. M. (2024). *Oppfølging av lavt-presterende elever i matematikk*. Universitetet i Oslo.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Baye, A. & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-Scale Assessments in Education*, 4(1), Artikkel 1. <https://doi.org/10.1186/s40536-015-0015-x>
- Bergem, O. K. (2016a). Hovedresultater i matematikk. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag* (s. 22–44). Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016-03>
- Bergem, O. K. (2016b). Hovedresultater i naturfag. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag* (s. 44–63). Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016-04>
- Bergem, O. K., Kaarstein, H. & Nilsen, T. (Red.). (2016). *Vi kan lykkes i realfag: Resultater og analyser fra TIMSS 2015*. Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016>
- Bergold, S., Wendt, H., Kasper, D. & Steinmayr, R. (2017). Academic competencies: Their interrelatedness and gender differences at their high end. *Journal of Educational Psychology*, 109(3), 439–449. <https://doi.org/10.1037/edu0000140>
- Blossing, U., Imsen, G. & Moos, L. (2014). Nordic schools in a time of change. I U. Blossing, G. Imsen & L. Moos (Red.), *The Nordic education model: «A school for all» encounters neo-liberal policy* (s. 1–14). Springer Netherlands. https://doi.org/10.1007/978-94-007-7125-3_1
- Creemers, B. & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. <https://doi.org/10.4324/9780203939185>
- Creemers, B. & Kyriakides, L. (2010). School factors explaining achievement on cognitive and affective outcomes: Establishing a dynamic model of educational effectiveness. *Scandinavian Journal of Educational Research*, 54(3), 263–294. <https://doi.org/10.1080/00313831003764529>
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., Iver, D. M. & Feldlaufer, H. (1993). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal*, 93(5), 553–574. <https://doi.org/10.1086/461740>
- Eklöf, H. (2020). Student motivation and self-beliefs. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (s. 1–24). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_44-1
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L. & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, 78(4), 1343–1359. <https://doi.org/10.1111/j.1467-8624.2007.01069.x>
- Hascher, T. & Hagenauer, G. (2010). Alienation from school. *International Journal of Educational Research*, 49(6), 220–232. <https://doi.org/10.1016/j.ijer.2011.03.002>
- Hedges, L. V. & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269(5220), 41–45.
- Ivanova, M. & Michaelides, M. P. (2022). Motivational components in TIMSS 2015 and their effects on engaging teaching practices and mathematics performance. *Studies in Educational Evaluation*, 74, Artikkel 101173. <https://doi.org/10.1016/j.stueduc.2022.101173>
- Jensen, F. & Nilsen, T. (2024). Lavtpresterende elever i matematikk. I A. Pettersen & F. Jensen (Red.), *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (s. 187–211). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.222.ch8>
- Jensen, F., Pettersen, A., Frønes, T. S., Eriksen, A., Løvgren, M. & Narvhus, E. K. (2023). *PISA 2022. Norske elevers kompetanse i matematikk, naturfag og lesing*. Cappelen Damm Forskning. <https://doi.org/10.23865/noasp.205>
- Kaarstein, H. (2017). Elevers motivasjon – TIMSS 2015. *Tangenten*, 2, 17–22.
- Kaarstein, H., Lehre, A.-C., Radišić, J. & Rohatgi, M. A. (2024). *TIMSS 2023. Kortrapport*. Universitetet i Oslo.

- Kaarstein, H., Radišić, J., Lehre, A.-C., Nilsen, T. & Bergen, O. K. (2020). *TIMSS 2019. Kortrapport*. Universitetet i Oslo.
- Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M. & de Boer, H. (2020). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Research Papers in Education*, 35(6), 641–680. <https://doi.org/10.1080/02671522.2019.1615116>
- Lee, J.-S. (2012). The effects of the teacher–student relationship and academic press on student engagement and academic performance. *International Journal of Educational Research*, 53, 330–340. <https://doi.org/10.1016/j.ijer.2012.04.006>
- Lehre, A., Lehre, K. P., Laake, P. & Danbolt, N. C. (2009). Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Developmental Psychobiology*, 51(2), 198–206. <https://doi.org/10.1002/dev.20358>
- Lehre, A.-C. & Nilsen, T. (2021). Språk i hjemmet og naturfagprestasjoner fra TIMSS 2015 til TIMSS 2019. I T. Nilsen & H. Kaarstein (Red.), *Med blikket mot naturfag* (s. 165–182). Universitetsforlaget. <https://doi.org/10.18261/9788215045108-2021-07>
- Lin, L., Somerville, G., Boursier, J., Santisteban, J. A. & Gruber, R. (2020). Sleep duration is associated with academic achievement of adolescent girls in mathematics. *Nature and Science of Sleep*, 12, 173–182. <https://doi.org/10.2147/NSS.S237267>
- Lorah, J. (2022). Analyzing large-scale assessment data with multilevel analyses: Demonstration using the Programme for International Student Assessment (PISA) 2018 data. I M. S. Khine (Red.), *Methodology for multilevel modeling in educational research: Concepts and applications* (s. 121–139). Springer. https://doi.org/10.1007/978-981-16-9142-3_7
- Meinck, S. & Brese, F. (2019). Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large-Scale Assessments in Education*, 7(1), Artikkel 8. <https://doi.org/10.1186/s40536-019-0076-3>
- Meld. St. 34 (2023–2024). *En mer praktisk skole. Bedre læring, motivasjon og trivsel på 5.–10. trinn*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-34-20232024/id3052898/>
- Mullis, I. V. S. & Martin, M. O. (Red.). (2017). *TIMSS 2019 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <http://timssandpirls.bc.edu/timss2019/frameworks/>
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L. & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center at Boston College website. <https://timss2019.org/reports>
- Nilsen, T., Björnsson, J. K. & Olsen, R. V. (2018). Hvordan har likeverd i norsk skole endret seg de siste 20 årene? I J. K. Björnsson & R. V. Olsen (Red.), *Tjue år med TIMSS og PISA i Norge: Trender og nye analyser* (s. 150–172). Universitetsforlaget. <https://doi.org/10.18261/9788215030067-2018>
- Nilsen, T. & Pettersen, A. (2024). Matematikkundervisning i norske og nordiske klasserom – klasseledelse, støttende undervisning og faglige utfordringer. I A. Pettersen & F. Jensen (Red.), *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (s. 139–165). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.222.ch6>
- Nilsen, T. & Teig, N. (2022). A systematic review of studies investigating the relationships between school climate and student outcomes in TIMSS, PISA, and PIRLS. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (s. 1053–1086). Springer International Publishing. https://doi.org/10.1007/978-3-030-88178-8_34
- Norges forskningsråd. (2012). *Realfagsrekruttering langs nye spor. Sluttrapport*. <https://kudos.dfo.no/documents/1878/files/1828.pdf>
- NOU 2019: 3. (2019). *Nye sjanser – bedre læring. Kjønnforskjeller i skoleprestasjoner og utdanningsløp*. Kunnskapsdepartementet.
- OECD. (2023). *PISA 2022 Results (Vol. I)*. PISA, OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Opplæringslova. (1998). *Lov om grunnskolen og den vidaregåande opplæringa (LOV-1998-07-17-61)* [Opphevet]. Lovdata. <https://lovdata.no/LTI/lov/1998-07-17-61>
- Pavešić, B. J., Radović, M. & Brese, F. (2022). Students' interests, motivation, and self-beliefs. I B. J. Pavešić, P. Koršňáková & S. Meinck (Red.), *Dinaric perspectives on TIMSS 2019*:

- Teaching and learning mathematics and science in South-Eastern Europe* (s. 65–100). Springer International Publishing. https://doi.org/10.1007/978-3-030-85802-5_4
- Reynolds, K. A., Mullis, I. V. S. & Martin, M. O. (2021). TIMSS 2023 context questionnaire framework. I I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Rosén, M., Steinmann, I. & Wernersson, I. (2020). Gender differences in school achievement. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (s. 1–48). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_46-1
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Seaton, M., Marsh, H. W. & Craven, R. G. (2010). Big-fish-little-pond effect: Generalizability and moderation – two sides of the same coin. *American Educational Research Journal*, 47(2), 390–433. <https://doi.org/10.3102/0002831209350493>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Strand, S., Deary, I. J. & Smith, P. (2006). Sex differences in cognitive abilities test scores: A UK national picture. *British Journal of Educational Psychology*, 76(3), 463–480. <https://doi.org/10.1348/000709905X50906>
- Teddlie, C. & Reynolds, D. (2000). *The international handbook of school effectiveness research*. Routledge. <https://doi.org/10.4324/9780203454404>
- Thapa, A., Cohen, J., Guffey, S. & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, 83(3), 357–385. <https://doi.org/10.3102/0034654313483907>
- Utdanningsdirektoratet. (2023, 9. november). *Resultater fra nasjonale prøver 5. trinn*. <https://www.udir.no/tall-og-forskning/statistikk/statistikk-grunnskole/analyser/2023/analyse-av-nasjonale-prover-5-trinn-2023/>
- Vik, F. N., Nilsen, T. & Øverby, N. C. (2022a). Aspects of nutritional deficits and cognitive outcomes – triangulation across time and subject domains among students and teachers in TIMSS. *International Journal of Educational Development*, 89, Artikkel 102553. <https://doi.org/10.1016/j.ijedudev.2022.102553>
- Vik, F. N., Nilsen, T. & Øverby, N. C. (2022b). Associations between sleep deficit and academic achievement – triangulation across time and subject domains among students and teachers in TIMSS in Norway. *BMC Public Health*, 22(1). <https://doi.org/10.1186/s12889-022-14161-1>
- von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U. & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs6460>
- Vrapi, R., Alia, A. & Brese, F. (2022). Characteristics of high- and low-performing students. I B. J. Pavešić, P. Koršňáková & S. Meinck (Red.), *Dinaric perspectives on TIMSS 2019: Teaching and learning mathematics and science in South-Eastern Europe* (s. 191–212). Springer International Publishing. https://doi.org/10.1007/978-3-030-85802-5_9
- Wang, M.-T. & Degol, J. L. (2016). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, 28(2), 315–352. <https://doi.org/10.1007/s10648-015-9319-1>
- White, S. B., Reynolds, P. D., Thomas, M. M. & Gitzlaff, N. J. (1993). Socioeconomic status and achievement revisited. *Urban Education*, 28(3), 328–343. <https://doi.org/10.1177/0042085993028003007>

Vedlegg

Tabell 6.A1. Prosentandel¹ manglende elevsvar for 5. trinnselever på lavt og avansert mestringsnivå i matematikk for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Variabel	Matematikk: lavt mestringsnivå			Matematikk: avansert mestringsnivå		
	2015	2019	2023	2015	2019	2023
SØS (bøker)	4 %	10 %	7 %	1 %	3 %	2 %
Språk hjemme	2 %	8 %	9 %	0 %	2 %	4 %
Motivasjon	3 %	9 %	26 %	1 %	3 %	8 %
Selvtillit	3 %	11 %	15 %	0 %	2 %	4 %
Trygg	6 %	9 %	15 %	4 %	2 %	3 %
Tilhørighet	3 %	9 %	10 %	1 %	3 %	3 %
Trøtt ²	-	9 %	7 %	-	3 %	2 %
Sulten	2 %	13 %	10 %	0 %	4 %	3 %
Fravær	3 %	9 %	6 %	1 %	3 %	2 %

Note. ¹ Avrundet til nærmeste heltall. ² Variabelen var ikke inkludert i TIMSS 2015.

Tabell 6.A2. Prosentandel¹ manglende elevsvar for 5. trinnselever på lavt og avansert mestringsnivå i naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Variabel	Naturfag: lavt mestringsnivå			Naturfag: avansert mestringsnivå		
	2015	2019	2023	2015	2019	2023
SØS (bøker)	4 %	8 %	8 %	1 %	2 %	2 %
Språk hjemme	2 %	7 %	10 %	0 %	2 %	3 %
Motivasjon	3 %	8 %	13 %	1 %	2 %	4 %
Selvtillit	4 %	10 %	16 %	0 %	3 %	6 %
Trygg	5 %	6 %	17 %	3 %	2 %	4 %
Tilhørighet	3 %	7 %	12 %	1 %	2 %	3 %
Trøtt ²	-	7 %	8 %	-	3 %	2 %
Sulten	2 %	12 %	12 %	0 %	5 %	3 %
Fravær	2 %	7 %	8 %	0 %	3 %	2 %

Note. ¹ Avrundet til nærmeste heltall. ² Variabelen var ikke inkludert i TIMSS 2015.

Tabell 6.A3. Prosentandel¹ manglende elevsvar for 9. trinnselever på lavt og avansert mestringsnivå i matematikk for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Variabel	Matematikk: lavt mestringsnivå			Matematikk: avansert mestringsnivå		
	2015	2019	2023	2015	2019	2023
SØS (bøker)	2 %	4 %	6 %	0 %	2 %	0 %
Språk hjemme	2 %	4 %	5 %	0 %	2 %	0 %
Motivasjon	2 %	4 %	12 %	1 %	2 %	1 %
Selvtillit	3 %	8 %	16 %	0 %	2 %	1 %
Trygg	2 %	3 %	11 %	0 %	2 %	2 %
Tilhørighet	3 %	5 %	12 %	1 %	2 %	1 %
Trøtt ²	-	5 %	7 %	-	2 %	1 %
Sulten	2 %	6 %	9 %	0 %	2 %	1 %
Fravær	4 %	8 %	8 %	2 %	3 %	0 %

Note. ¹ Avrundet til nærmeste heltall. ² Variabelen var ikke inkludert i TIMSS 2015.

Tabell 6.A4. Prosentandel¹ manglende elevsvar for 9. trinnselever på lavt og avansert mestringsnivå i naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Variabel	Naturfag: lavt mestringsnivå			Naturfag: avansert mestringsnivå		
	2015	2019	2023	2015	2019	2023
SØS (bøker)	3 %	4 %	6 %	0 %	2 %	1 %
Språk hjemme	2 %	4 %	5 %	0 %	3 %	1 %
Motivasjon	3 %	7 %	16 %	0 %	2 %	1 %
Selvtillit	3 %	9 %	20 %	1 %	2 %	2 %
Trygg	2 %	4 %	11 %	0 %	2 %	1 %
Tilhørighet	3 %	5 %	11 %	0 %	2 %	1 %
Trøtt ²	-	4 %	7 %	-	2 %	1 %
Sulten	2 %	5 %	9 %	0 %	3 %	1 %
Fravær	5 %	8 %	7 %	2 %	3 %	1 %

Note. ¹ Avrundet til nærmeste heltall. ² Variabelen var ikke inkludert i TIMSS 2015.

Tabell 6.A5. Signifikanstest (Students t-test) for sammenligning av 5. trinnselever på lavt og avansert mestringsnivå i matematikk og naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Variabel	Matematikk			Naturfag		
	2015	2019	2023	2015	2019	2023
SØS (bøker)	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Språk hjemme	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Motivasjon	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Selvtillit	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Trygg	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Tilhørighet	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Trøtt ¹		$p < 0,05$	$p < 0,05$		$p < 0,05$	$p < 0,05$
Sulten	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$
Fravær	$p > 0,05$	$p < 0,05$	$p < 0,05$	$p > 0,05$	$p < 0,05$	$p < 0,05$

Note. ¹ Variabelen var ikke inkludert i TIMSS 2015. Det er kun i 2015 at forskjellene i bakgrunnsvariabler mellom elever på lavt og avansert mestringsnivå ikke er signifikante.

Tabell 6.A6. Signifikanstest (Students t-test) for sammenligning av 9. trinnselever på lavt og avansert mestringsnivå i matematikk og naturfag for TIMSS 2015, TIMSS 2019 og TIMSS 2023 i Norge

Variabel	Matematikk			Naturfag		
	2015	2019	2023	2015	2019	2023
SØS (bøker)	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Språk hjemme	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Motivasjon	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Selvtillit	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Trygg	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Tilhørighet	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Trøtt ¹		$p < 0,05$	$p < 0,05$		$p < 0,05$	$p < 0,05$
Sulten	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$
Fravær	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$	$p < 0,05$

Note. ¹ Variabelen var ikke inkludert i TIMSS 2015. Forskjellene mellom bakgrunnsvariabler for elever på lavt og avansert mestringsnivå er signifikante for begge fag og i alle sykluser.

KAPITTEL 7

Læreplan og faglig innhold i matematikk på 5. trinn i Norge

Mark White, Alexander Jonas Viktor Selling, Andreas Pettersen og Jelena Veletić

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: Hva elevene lærer på skolen, avhenger i stor grad av hva læreren underviser i. For å få en bedre forståelse av sammenhengen mellom undervisning og elevenes læring, er det derfor viktig å utforske både det faglige innholdet i læreplanen og hva som faktisk blir undervist i klasserommet. I dette kapitlet ser vi på det faglige innholdet i matematikkundervisning fra 1. til 5. trinn i både TIMSS 2019 og TIMSS 2023, for å få innsikt i hvilke emner norske lærere rapporterer at de underviser i. Vi undersøker også endringer i innholdet i matematikkundervisning fra 2019 til 2023, noe som er interessant med tanke på innføringen av læreplanverket LK20 i 2020. Vi finner at emnene som ble undervist til både norske og nordiske elever var stort sett like i 2019, selv om det var stor variasjon innenfor landene. Norge skilte seg imidlertid ut i 2023 ved at norske elever blir undervist i færre emner både i inneværende skoleår og i skoleløpet opptil 5. trinn. Denne endringen er mest tydelig innenfor hovedemnene *geometri og måling* og *statistikk*, som blir introdusert etter 5. trinn i den nye nasjonale læreplanen. Denne endringen er mindre tydelig innenfor hovedemnet *tall*, som utgjør hoveddelen av den nasjonale læreplanen før 5. klasse. Implikasjonene av disse endringene blir diskutert videre i kapitlet.

Nøkkelord: mulighet til å lære, matematikk, den implementerte læreplanen, den intenderte læreplanen, geometri, TIMSS

Abstract**Curriculum and Academic Content in Mathematics in Grade 5 in Norway**

What is taught has an important influence on what students learn in school. It is therefore important to study both what school systems intend for students to learn and what actually gets taught to students. Given changes in the national Norwegian curriculum in 2020 (LK20), this chapter looks at the content of mathematics instruction up through the 5th grade from both TIMSS 2019 and TIMSS 2023 to understand what is taught to Norwegian students and how this has changed across time. We find that the content taught to Norwegian students and students from the other Nordic countries was generally similar in 2019, albeit there was large variation within countries. However, Norway distinguishes itself in 2023 with Norwegian students being taught fewer topics both in the current school year and overall (from 1st through 5th grade). This change is most evident in the content areas of geometry and measurement and statistics, which are only present in the new national curriculum after 5th grade, but, to a lesser extent, the change also is evident in the content area of numbers, which comprises the bulk of the national curriculum before 5th grade. We discuss the implications of these changes.

Keywords: opportunity to learn, mathematics, intended curriculum, TIMSS, implemented curriculum, geometry

Introduksjon

Den nasjonale læreplanen er en viktig styringsmekanisme i den norske skolen (Karseth et al., 2020) og setter rammene for grunnopplæringen i Norge. I læreplanene for matematikkfaget er det beskrevet hvilke kunnskaper og ferdigheter (dvs. kompetansemål) elevene skal tilegne seg og når i skoleløpet elevene bør mestre dem. I matematikkfaget er det fastsatt kompetansemål som beskriver hva elevene skal lære på hvert klassetrinn fra 2. til 10. trinn (Kunnskapsdepartementet, 2019).

I forskningslitteraturen brukes ofte begrepet *opportunity to learn* (OTL) som en betegnelse på det faglige innholdet elevene har fått mulighet til å lære (Berliner, 1990). OTL er viktig både ved at det kan bidra til å forklare forskjeller i elevers prestasjoner og at det brukes til å belyse forskjeller mellom land når det kommer til hvilket faglig innhold som inngår i undervisningen. OTL kan dermed støtte beslutningstagere i å forstå i hvilken grad lave prestasjoner kan forklares med at elevene ikke har fått undervisning i noe av det faglige innholdet, snarere enn kvaliteten på undervisningen, og gjør det mulig å sammenligne innholdet i læreplaner på tvers av land (Floden, 2002).

Forskning knyttet til OTL viser ofte til tre forskjellige nivåer av læreplanen (Goodlad et al., 1979): den *intenderte* læreplanen viser til hva det er forventet at elevene skal lære fastsatt på systemnivå. I Norge vil dette si målene og innholdet i fagene slik det er beskrevet i den nasjonale læreplanen. Den *implementerte* læreplanen er det som faktisk blir undervist i klasserommet, mens den *oppnådde* læreplanen viser til hva elevene har lært i undervisningen (McDonnell, 1995). Ved å skille mellom disse tre nivåene er det mulig å finne forskjeller i eller avvik mellom det som er forventet (intendert), det som undervises (implementert) og det elevene lærer (oppnådd). Dette viser i hvilken grad man oppnår målene med utdanningen og kan bidra med verdifull kunnskap i arbeid med læreplanutvikling og undervisningsstrategier.

I Norge har det nylig vært en endring i den nasjonale læreplanen, fra Kunnskapsløftet 2006 (LK06; Kunnskapsdepartementet, 2006) til Kunnskapsløftet 2020 (LK20; Kunnskapsdepartementet, 2017), noe som innebærer endringer i den intenderte læreplanen. Samtidig har lærere som har deltatt i *Trends in International Mathematics and Science Study* (TIMSS) i 2019 og 2023 blitt spurt om hvilke emner de har dekket, eller ikke har dekket, i matematikkundervisningen. Dette gjør det mulig å undersøke

det faglige innholdet i matematikkundervisningen i norske skoler, samt undersøke endringer i den implementerte læreplanen og i hvilken grad disse samsvarer med endringene i den nasjonale læreplanen.

I denne studien utforsker vi mulighetene for læring i matematikk i norsk skole, ut fra lærernes rapportering om egen undervisning, samt hvordan disse har endret seg mellom 2019 og 2023 og sammenhenger mellom endringer i det som blir undervist og endringer innført i overgangen fra LK06 til LK20. For å se resultatene i et større perspektiv sammenligner vi også den implementerte læreplanen i Norge med de andre nordiske landene som deltok i TIMSS 2023. Målet er at disse analysene skal bidra med verdifull kunnskap om matematikkundervisningen i norsk skole og gi innsikt som kan støtte videreutviklingen av skolepraksis.

I dette kapittelet stiller vi følgende forskningsspørsmål:

1. Hvordan har den implementerte læreplanen endret seg mellom 2019 og 2023 i Norge?
2. I hvilken grad er endringene i den implementerte læreplanen i tråd med endringer fra LK06 til LK20 i Norge?
3. Hvordan er den implementerte læreplanen i 5. trinn matematikk i Norge sammenlignet med de andre nordiske landene?

Bakgrunn og teori

OTL-rammeverket

OTL er et bredt begrep som kan omfatte mange ulike sider ved elevenes læring. Carrolls (1963) opprinnelige OTL-rammeverk inkluderte for eksempel faktorer som tid brukt til læring, kvalitet på undervisningen og elevenes intelligens, i et overordnet rammeverk som prøvde å forklare elevenes læring (Floden, 2002). På grunn av kompleksiteten i dette overordnede rammeverket, har det senere blitt utviklet mer moderne versjoner av OTL, hvor man heller vektlegger de tre læreplannivåene: den intenderte, den implementerte og den oppnådde læreplanen.

Det er ment å være en sterk sammenheng mellom de tre nivåene (Floden, 2002; Goodlad et al., 1979; McDonnell, 1995). Kompetansemålene i den intenderte læreplanen definerer læringsmålene for forskjellige klassetrinn, og lærerne har ansvar for å legge til rette for at elevene når disse læringsmålene til rett tid. Dette vil si at den implementerte læreplanen

skal være i samsvar med den intenderte læreplanen, og at den oppnådde læreplanen, altså elevenes læringsutbytte, også reflekterer det innhold som forventes å bli lært i den intenderte læreplanen.

I praksis er det derimot ikke en slik «perfekt» sammenheng mellom disse nivåene. Ofte er det store forskjeller mellom forventet læring, læreres undervisning og elevenes læringsutbytte. Én av årsakene til dette er at lærere tolker innholdet i den intenderte læreplanen ulikt, noe som kan føre til avvik mellom intendert og implementert læreplan, og uønsket variasjon i undervisningen på tvers av elever (Remillard & Heck, 2014).

Disse forskjellene er interessante fra et styringsperspektiv siden forskjellene viser hvor den intenderte læreplanen feiler som en styringsmekanisme. De er også interessante fra et forskningsperspektiv siden de kan vise til utfordringer knyttet til tolkningen av den intenderte læreplanen eller elevenes læringsprosesser. I dette kapittelet skal vi fokusere på 5. trinnlæreres rapporterte dekning av den implementerte læreplanen på 1.–5. trinn i Norge, slik denne er operasjonalisert i TIMSS. Vi sammenligner denne på tvers av land, samt mot den intenderte læreplanen i Norge.

OTL i TIMSS

Via TIMSS-undersøkelsen er det mulig å undersøke hvorvidt det er variasjoner i det faglige innholdet som undervises i en rekke ulike utdannings-systemer rundt omkring i verden (det vil si forskjeller i den implementerte læreplanen). I TIMSS måles den implementerte læreplanen gjennom lærerspørreskjemaer, hvor lærere blir spurt om hvilke faglige emner de har undervist, undervisningstid og vektlegging av spesifikke emner. For å sikre at de faglige emnene som inngår i spørreskjemaet er relevante for alle deltakerlandene brukes internasjonale eksperter og representanter fra de ulike landene i utviklingen av spørsmålene. Denne prosessen omhandler flere runder med evalueringer, revideringer og utprøvinger av spørsmålene i spørreskjemaet, for å sørge for at det blir tatt hensyn til kulturelle forskjeller og forskjeller i læreplanene. Målet er å forsikre seg om at spørreskjemaet er bredt nok til at det kan brukes på tvers av alle deltakerlandene, men også for å legge til rette for lokale tilpasninger (Houang & Schmidt, 2008).

Imidlertid kan det være at lærerspørreskjemaet ikke er like godt tilpasset alle landene, og at det dermed er bedre tilpasset enkelte lands læreplaner, noe som kan føre til *measurement bias*. Dette kan skje om spørsmålene som måler faglig dekningsgrad i TIMSS ikke dekker innhold som er inkludert

i læreplanen eller hvis spørsmålene inkluderer emner som ikke er inkludert i læreplanen. Denne begrensningen er spesielt relevant når det gjelder forskningsspørsmål 2, som undersøker i hvilken grad endringer i den implementerte læreplanen er i tråd med endringer fra LK06 til LK20 i Norge, siden vi bare ser på endringene ut fra det TIMSS faktisk måler. Denne begrensningen er imidlertid mindre problematisk for forskningsspørsmål 3, hvor vi sammenligner Norge med de andre nordiske landene, siden alle landene blir målt etter det samme rammeverket. Det tredje forskningsspørsmålet blir dermed bare begrenset av dette hvis TIMSS-rammeverket har en bias for eller mot læreplanen for spesifikke land, heller enn at det er usystematiske ulikheter.

Tidligere forskning på OTL

Flere studier har undersøkt sammenhengen mellom den implementerte læreplanen (undervisningens innhold og omfang) og den oppnådde læreplanen (elevenes prestasjoner) på tvers av land. Mange av disse studiene bygger på data fra TIMSS, hvor den implementerte læreplanen er målt gjennom spørsmål til lærere om hvorvidt de har undervist i ulike faglige emner. Scheerens (2016) analyserte data fra TIMSS 2011 for 4. og 8. trinn i 23 land, med fokus på koblingen mellom innholdet i undervisningen og prestasjoner i matematikk og naturfag. Scheerens (2016) fant en middels korrelasjon mellom antallet emner som dekkes i undervisningen og elevenes prestasjoner i 12 av de 23 landene i matematikk, mens det ikke ble funnet en signifikant korrelasjon i naturfag. Dette tyder på at sammenhengen mellom dekningsgrad av den implementerte læreplanen, det vil si antall forskjellige emner som undervises innenfor et fag, og prestasjoner kan variere mellom fag og mellom kontekster.

Rolfe et al. (2021) gjorde en tilsvarende analyse med data fra 78 land som deltok i TIMSS. De fant signifikante variasjoner i dekningsgraden av den implementerte læreplanen mellom landene, der elever med høyere sosioøkonomisk status (SØS) hadde generelt blitt eksponert for mer faglig innhold enn elever med lavere SØS. Likevel fant man også noen få eksempler på land der elever med lavere SØS hadde større grad av eksponering for faglig innhold enn elever med høyere SØS, på grunn av kompensatoriske mekanismer i utdanningssystemet. Til tross for disse variasjonene i dekningsgrad av den implementerte læreplanen konkluderte studien med at eksponering for mer faglig innhold ikke var en signifikant medierende

faktor i relasjonen mellom SØS og prestasjoner i de fleste landene, noe som setter spørsmålstegn ved i hvilken grad utdanningsmuligheter alene kan adressere ulikheter basert på SØS.

Rosén og Nilsen (2024) undersøkte sammenhengen mellom endringer i dekningsgrad av emner og endringer i prestasjoner i matematikk og naturfag mellom 2011 og 2019 i de fire nordiske landene som deltok i disse to TIMSS-undersøkelsene. De fant at økt deknning av emner i matematikk over tid var assosiert med en liten positiv effekt på prestasjoner i de fleste landene (Rosén & Nilsen, 2024). For norske elever kunne høyere prestasjoner delvis forklares av at det var mer undervisning i emnene relatert til tallforståelse, mens en nedgang i undervisning av geometri kan ha bidratt til å begrense ytterligere framgang i prestasjoner. I Finland og Danmark, der det var en nedgang i elevprestasjoner, viste resultatene at det også hadde vært en økning i undervisning av tallforståelse, og at dette til en viss grad begrenset nedgangen i prestasjoner. For Sverige ble det derimot ikke funnet noen signifikant sammenheng mellom den implementerte læreplanen og prestasjoner, noe som tyder på at dette ikke kunne forklare framgangen i prestasjoner blant svenske elever. Basert på disse resultatene konkluderte Rosén og Nilsen (2024) med at endringer i den implementerte læreplanen, spesielt innenfor matematikk, hadde en begrenset, men positiv effekt på prestasjoner.

Forskning på den implementerte læreplanen har også vært knyttet til likeverd og hvorvidt forskjellige grupper av elever får tilsvarende implementerte læreplaner. For eksempel avslørte Blömekes (2024) analyse av dataene fra TIMSS 2019 tydelige mønstre i den implementerte læreplanen på tvers av de nordiske landene. Elever med høyere SØS i Danmark og Finland hadde større sannsynlighet for å bli undervist i flere emner i matematikk enn elever med lavere SØS. Til sammenligning hadde elever i Sverige og Norge like stor sannsynlighet for å få undervisning i flere emner uavhengig av SØS.

Kang og Cogan (2022) undersøkte også sammenhengen mellom implementert læreplan og prestasjoner, men brukte en annen operasjonalisering av den implementerte læreplanen der elevene ble spurt om deres kjennskap til ulike typer matematikkoppgaver. Er bredere implementert lærerplan (dvs. en med høyere dekningsgrad) innebar i dette tilfellet at elevene var kjent med flere eksempeloppgaver og at de hadde lignende oppgaver på prøver og i timene. Resultatene viste at en implementert læreplan med høyere dekningsgrad samsvarte med høyere prestasjoner i matematikk i både TIMSS- og PISA-undersøkelsen (*Programme for International Student Assessment*) (Kang & Cogan, 2022).

Samlet sett viser forskningen at dekningsgraden av den implementerte læreplanen, det vil si hvor mange emner som dekkes i undervisningen, er en signifikant faktor for å predikere elevers prestasjoner, spesielt i matematikk. Samtidig er forholdet mellom dekningsgraden av den implementerte læreplanen, elevenes sosioøkonomiske status og prestasjoner komplekst og kontekstavhengig. En viktig begrensning ved forskningen som er diskutert her er hvordan den implementerte læreplanen måles. Disse studiene har tatt utgangspunkt i *antallet* emner lærere rapporterer at de har dekket eller antallet typer eksempeloppgaver elevene er kjent med. Disse metodene tar ikke hensyn til *hva* som faktisk har blitt undervist, noe som ikke gjør det mulig å undersøke hvorvidt elever som blir undervist i et emne faktisk mestrer det emnet, noe som er sentralt i teorien bak OTL. Dette kapitlet skiller seg fra tidligere forskning ved å ta utgangspunkt i de faglige emnene som lærerne rapporterer å ha undervist, og ikke bare antallet emner.

De intenderte læreplanene – fra LK06 til LK20

I Norge har det blitt innført nye læreplanverk med jevne mellomrom: Mønsterplanen 87, Læreplanen 97, Kunnskapsløftet 06, en revisjon til Kunnskapsløftet 06 i 2013 og Kunnskapsløftet 2020. Innføringen av de nasjonale læreplanene er ment å være et virkemiddel i en nasjonal utdanningsreform og har som formål å være både en pedagogisk og faglig plattform for lærere og et styringsverktøy for den implementerte læreplanen (Aasen et al., 2015).

For å kunne tolke endringer i den implementerte læreplanen slik den måles i TIMSS 2019 og 2023, er det nødvendig å se på det faglige innholdet i LK06 og LK20 – de intenderte læreplanene som var gjeldende i henholdsvis 2019 og 2023 (Kunnskapsdepartementet, 2006, 2017, 2019). Nedenfor diskuterer vi innholdet i LK06 og LK20 og endringene mellom dem.

Endringer fra LK06 til LK20

I en evaluering av fagfornyelsen (LK20), undersøkte Karseth et al. (2020) strukturen og innholdet i læreplanen. De så spesielt på innhold og kompetansemål for to fag, norskfaget og KRLE-faget, og fant blant annet at kompetansemålene i LK20 var mer overordnede og generelle, og mindre konkrete og tydelige, enn i LK06. Videre fant de at LK20 var mindre innholdsorientert, og at det var et utydelig forhold mellom kompetanse og ferdigheter, samt at det var en økt vektlegging av prosesser (som for eksempel «å utforske») i LK20 enn i LK06 (Karseth et al., 2020).

Det faglige målet for undervisningen er beskrevet i kompetansemålene. En markant forskjell mellom LK06 og LK20 er struktureringen av kompetansemålene i matematikkfaget. Mens det i LK20 er gitt kompetansemål etter hvert årstrinn, ble det i LK06 gitt kompetansemål etter 2., 4., 7. og 10. årstrinn. Når vi tolker den implementerte læreplanen opp mot den intenderte læreplanen, har vi derfor sett på kompetansemålene fra 1. til og med 7. trinn i de to læreplanene. Spesielt har vi sett på

- om det faglige innholdet i kompetansemålene i LK06 også finnes i kompetansemålene i LK20, og omvendt
- om det er faglig innhold i kompetansemål i LK06 som er flyttet til andre trinn i LK20
- hvorvidt det er endringer i formuleringer som åpner for ulike tolkninger og uklarheter i det faglige innholdet som skal undervises

Forskjeller i kompetansemålene mellom LK06 og LK20

Når vi sammenligner de to læreplanene ser vi at kompetansemålene i LK20 ofte er bredere og mindre presist formulert enn i LK06. Dette kan for eksempel dreie seg om at man ikke konkretiserer hvilke prosesser, metoder eller representasjoner som skal brukes, eller at det ikke gis føringer for eventuelle begrensninger av et mål. Noen eksempler:

- I LK06 var det eksplisitt uttrykt at elevene skulle telle til 100, bygge mengder til 10 og jobbe med tier-grupper. I kompetansemålene i LK20 finner vi at elevene skal «eksperimentere med telling forlengs og baklengs» og «utforske tall».
- De faglige begrepene *doble* og *halvere* var eksplisitt nevnt i LK06, men er ikke nevnt i LK20.
- I LK06 skulle elevene addere og subtrahere med «tosifra tall», i LK20 er det ikke spesifisert størrelse på tall.
- I LK06 var det ett kompetansemål på å kjenne igjen og videreføre tallmønstre, og ett for å lage og utforske geometriske mønstre. I LK20 finner vi et generelt kompetansemål for å kjenne igjen og lage mønstre, ikke spesifisert om det er tallmønstre og/eller geometriske mønstre.
- Multiplikasjonstabellen inngikk i LK06, men ikke er eksplisitt nevnt i LK20.

I hvilken grad kompetansemål fra LK06 samsvarer med målene i LK20 vil derfor avhenge av hvordan målene tolkes. For eksempel kan

kompetansemålet «eksperimentere med telling» i LK20 tolkes som å telle til 100 av noen lærere (slik som kompetansemålet i LK06), mens andre lærere kan vurdere at det er tilstrekkelig å telle forlengs og baklengs til 50. Et annet eksempel er dobling og halvering, som ikke er eksplisitt nevnt i LK20, men som tolkes til å være en del av kompetansemålene knyttet til det å utforske tall og egenskaper ved tall. Vår vurdering av kompetansemålene i matematikk, som bredere og mindre konkrete i LK20 enn i LK06, samsvarer i stor grad med funnene til Karseth et al. (2020) for norskfaget og KRLE-faget. En slik endring medfører at tolkningsrommet for kompetansemålene øker, og gir lærere både større frihet og mer ansvar når det kommer til å bestemme det konkrete innholdet i undervisningen. Det er også grunn til å tro at dette kan føre til større variasjon i innholdet i matematikkfaget mellom lærere og skoler.

Selv om det generelle inntrykket er at kompetansemålene har blitt mindre konkrete, finnes det også eksempler på konkrete matematiske begreper som beskrives i kompetansemål i LK20, men ikke i LK06. For eksempel nevnes den kommutative og assosiative egenskapen eksplisitt i 2. trinn i LK20, men ikke i LK06.

Begreper og emner i LK20

I LK20 brukes begreper fra kjerneelementene mye i kompetansemålene. For eksempel finner vi spor etter kjerneelementet «Utforsking og problemløsning» i mange av kompetansemålene på 2.–4. trinn, hvor verbet «utforske» er hyppig brukt og elevene ofte skal «løse problemer». Mens det er naturlig at utforsking blir synliggjort i kompetansemålene, siden det inngår som et kjerneelement i LK20, kan det også bidra til uklarhet. Begrepet «utforske» kan ha flere ulike betydninger og tolkes på ulike måter (f.eks. Artigue & Blomhøj, 2013), noe som gjør det utfordrende å vite akkurat hva det innebærer i praksis. For eksempel kan det at elevene skal «utforske egenskaper ved de fire regneartene» tolkes på mange måter og dermed resultere i forskjellig læringsoppnåelse mellom klasser og skoler.

I LK20 defineres utforsking som det «at elevene leter etter mønstre, finner sammenhenger og diskuterer seg frem til en felles forståelse» (Kunnskapsdepartementet, 2017). Selv om dette peker på viktige deler av det å forstå matematiske konsepter, kan det være utfordrende å vite nøyaktig hva det er forventet at elevene skal lære. Karseth et al. (2020) poengterer også utfordringene ved verbet «utforske» og viser til at dette er verbet med

den mest omfattende og mangetydige beskrivelsen i utdanningsdirektoratets beskrivelser av verbene i LK20.

Brøk er et av temaene som er blitt mer fremtredende i den nye læreplanen. I LK20 finner vi seks kompetansemål på 5. trinn der brøk spiller en sentral rolle. Elevene skal for eksempel kunne bruke brøk til å diskutere sannsynlighet i spill og som en del av mengder og tall på tallinjen, og utforske og forklare sammenhenger mellom brøker, desimaltall og prosent. I LK06 var brøk en del av et mer sammensatt mål etter 4. trinn der elevene skulle «beskrive plassverdisystemet for dei heile tala, bruke positive og negative heile tal, enkle brøkar og desimaltal i praktiske samanhengar, og uttrykkje talstorleikar på varierte måtar» (Kunnskapsdepartementet, 2006). Brøk inngikk også i to kompetansemål etter 7. trinn. Selv om LK06 ikke hadde egne kompetansemål som kun omhandlet brøk, kan det sammensatte målet ha bidratt til at sammenhengen mellom brøk, desimaltall og prosent likevel ble naturlig å introdusere.

I LK06 var geometri et av hovedområdene med flere kompetansemål etter både 2. og 4. trinn. Disse kompetansemålene omfatter flere geometriske delmål der det vises til spesifikke figurer (f.eks. sirkler, mangekanter, kuler), egenskaper til geometriske figurer (f.eks. hjørner, kanter, flater) og fenomener (f.eks. bruke speilsymmetri og parallell forskyving i praktiske situasjoner). I LK20 finnes det kun ett kompetansemål fra 1. trinn til og med 5. trinn som omhandler geometri: «utforske, beskrive og sammenligne egenskaper ved to- og tredimensjonale figurer ved å bruke vinkler, kanter og hjørner» (på 4. trinn) (Kunnskapsdepartementet, 2017). Dette viser at geometri er mye mindre vektlagt på de tidlige trinnene i dagens læreplan.

Måling var et annet hovedområde i LK06. I LK20 er det noen kompetansemål som omfatter måling på 2., 3. og 4. trinn i LK20, men færre enn tidligere. Noe av det som skiller de to læreplanene innenfor dette området er at myntenheter og salg er fjernet som konkrete mål i LK20, samt at begrepet overslag ikke nevnes eksplisitt, slik det ble i LK06. Her bør det igjen nevnes at overslag kan inkluderes i læreres egne tolkninger i de bredere målene i LK20, selv om det ikke blir spesifisert.

Statistikk er et av hovedområdene innenfor matematikk i TIMSS, men har fått redusert rolle i de trinnene vi ser på i denne studien. I LK06 var statistikk et eget hovedområde for 1. til 4. trinn, representert med ett kompetansemål etter 2. trinn og ett etter 4. trinn. Her ble også spesifikke statistiske representasjoner nevnt (f.eks. søylediagram). I LK20 introduseres ikke statistikk, diagrammer og sentralmål før på 7. trinn.

LK20 er det introdusert kompetansemål knyttet til lek og programmering. Det finnes omtrent ett kompetansemål per trinn som omfatter programmering, og noen mål på 2., 3. og 4. trinn som tar for seg «lek». Lek og programmering var ikke eksplisitt nevnt i kompetansemålene i LK06. Selv om disse målene ikke er relevante for emneomfanget i TIMSS, kan disse områdene ha ført til at andre mål har blitt tatt ut. Dette kan være med på å forklare det reduserte fokuset på geometri og statistikk i de aktuelle trinnene.

Metode

TIMSS-undersøkelsen

Vi henviser til det første kapittelet for mer informasjon om TIMSS-studien generelt (se kapittel 1 i denne antologien). I dette kapittelet fokuserer vi på svar fra lærere som underviste i matematikk på 5. trinn i Norge og de andre nordiske landene i 2019 og 2023.¹ Vi bruker kun data fra lærerspørreskjemaet i 2023 ($N_{\text{Norge}} = 493$; $N_{\text{Sverige}} = 351$; $N_{\text{Danmark}} = 430$; $N_{\text{Finland}} = 423$) og i 2019 ($N_{\text{Norge}} = 375$; $N_{\text{Sverige}} = 301$; $N_{\text{Danmark}} = 309$; $N_{\text{Finland}} = 362$). Utvalget i TIMSS er representativt for elever, ikke lærere. Derfor kan prosentandelene som rapporteres nedenfor tolkes som andelen elever som har fått undervisning i ulike emner på ulike tidspunkter. Selv om det kun er lærere på 5. trinn som har deltatt, gir lærernes svar informasjon hvilke faglige emner elevene har blitt undervist i fram til slutten av 5. trinn.

Operasjonalisering av den implementerte læreplanen i TIMSS

Som nevnt tidligere måler TIMSS den implementerte læreplanen gjennom spørsmål i lærerspørreskjemaet (tabell 1). Disse spørsmålene er utviklet basert på TIMSS-rammeverket, som er ment å være et felles rammeverk for å sammenligne på tvers av land. TIMSS-rammeverket inneholder altså ikke alle emnene som undervises i Norge, men gir et standardisert grunnlag for å forstå den implementerte læreplanen. De matematiske emnene i TIMSS-rammeverket på 5. trinn er fordelt over tre hovedemner: *tall*; *måling og geometri*; *statistikk*. Innenfor hvert hovedemne svarer lærerne på hvorvidt

¹ Deltakende elever fra de andre nordiske landene er i 4. trinn. Norge deltar med 5. trinn for at norske elever har det samme alder som elever fra andre land.

de har undervist spesifikke emner, som for eksempel «Finne omkretsen av polygoner» (*måling og geometri*), eller «dividere tresifret tall med ensifret tall» (*tall*). Svaralternativene i spørreskjemaet er «Hovedsakelig undervist før dette skoleåret», «Hovedsakelig undervist dette skoleåret» eller «Bare så vidt introdusert eller ikke undervist (ennå)».² For å spare plass i figurene, kaller vi disse svarkategoriene «før dette året», «dette året» og «så vidt introdusert» i figurene. Spørreskjemaet fanger derfor opp emner som har blitt undervist i hele skoleløpet opp til omtrent april i 5. trinn.

Tabell 7.1. Operasjonalisering av den implementerte læreplanen i TIMSS

TIMSS 2019	TIMSS 2023
Følgende liste inneholder hovedemnene i matematikktesten i TIMSS. Velg det alternativet som best beskriver når elevene i TIMSS-klassen ble undervist i hvert av emnene.	Følgende liste inneholder hovedemnene i matematikktesten i TIMSS. Velg det alternativet som best beskriver når elevene i TIMSS-klassen ble undervist i disse emnene.
TALL <i>Begrepet hele tall, inkludert posisjonssystemet og ordning av tall</i> Addere, subtrahere, multiplisere og/eller dividere hele tall <i>Begrepene multipler og faktorer; partall og oddetall</i> Regneuttrykk (finne manglende tall, representere situasjoner/problem ved hjelp av regneuttrykk) Tallmønster (utvide tallmønster og å finne manglende ledd) <i>Begrepet brøk, inkludert representere, sammenligne og ordne, addere og subtrahere enkle brøker</i> <i>Begrepet desimaltall, inkludert posisjonssystemet, ordning av desimaltall og addere og subtrahere desimaltall</i>	TALL <i>Kjenne posisjonssystemet og ordne hele tall etter størrelse</i> Addere og subtrahere tall med opptil 4 sifre Multiplisere tresifret tall med ensifret tall, og tosifret tall med tosifret tall Dividere tresifret tall med ensifret tall <i>Løse oppgaver med oddetall og partall, multipler eller faktorer</i> <i>Runde av eller gjøre overslag med hele tall</i> Finne det manglende tallet eller den manglende operasjonen i en ligning <i>Utvide et tallmønster eller finne et manglende tall i mønsteret</i> Representere , sammenligne og ordne brøker etter størrelse Addere og subtrahere enkle brøker <i>Representere, sammenligne og ordne desimaltall med inntil to desimaler</i> Addere og subtrahere desimaltall med inntil to desimaler

(Forts.)

2 I Norge i 2023, var svarkategorien «Bare så vidt introdusert eller ikke undervist (ennå)» delt i to svarkategorier («Bare så vidt introdusert» og «Ikke undervist (ennå)») for å ha fire svarkategorier. Siden det kun var Norge i 2023 som brukte fire svarkategorier, har vi valgt å bruke kun tre svarkategorier slik at data er like på alle land og tidspunkter.

Tabell 7.1. (Forts.)

TIMSS 2019	TIMSS 2023
MÅLING OG GEOMETRI	MÅLING OG GEOMETRI
Løse oppgaver som handler om lengde , inkludert måling og estimering	Måle, anslå, addere og subtrahere lengder
Løse oppgaver som handler om masse, volum og tid	Addere og subtrahere masse, volum eller tid med passende enheter
Finne og beregne omkrets, areal og volum	Finne omkretsen av polygoner
Parallele og perpendikulære linjer	Finne arealet av figurer ved å dekke med kvadrater, eller volum ved å fylle med
Sammenligne og tegne vinkler	Gjenkjenne og tegne parallelle linjer og linjer som står vinkelrett på hverandre
Elementære egenskaper for vanlige geometriske figurer	Sammenligne og tegne vinkler
Tredimensjonale figurer , inkludert sammenhenger mellom disse og deres todimensjonale representasjoner	Beskrive og tegne vanlige todimensjonale figurer (f.eks. sirkler, trekanter, firkanter)
	Beskrive og tegne vanlige tredimensjonale figurer (f.eks. terninger, rettvinklede bokser, kjegler, sylindre, kuler)
STATISTIKK	STATISTIKK
Lese og tolke data fra tabeller, piktogrammer, søylediagrammer, linjediagrammer og sektordiagrammer	Lese og tolke framstillinger av data
Organisere og representere data for å kunne svare på spørsmål (dataframstilling)	Lage eller fullføre dataframstilling
Trekke konklusjoner fra framstillinger av data	Trekke konklusjoner fra to eller flere datakilder

Notat: Fet skrift er brukt som tittel på emner i figurene 1–7. Kursiv skrift er brukt for å vise emner som er vurdert som sammenlignbare på tvers årene.

Spørsmålene i lærerspørreskjemaet ble oppdatert i 2023 i tråd med revideringene av TIMSS sitt faglige rammeverk (se tabell 7.1). Dette betyr at spørsmålene er ulike i 2019 og 2023. I 2023 var det

- tolv emner innenfor *tall*
- åtte emner innenfor *måling og geometri*
- tre emner innenfor *statistikk*

Til sammenligning var det i 2019

- syv emner innenfor *tall*
- syv emner innenfor *måling og geometri*
- tre emner innenfor *statistikk*

Kun ti av emnene var direkte sammenlignbare mellom de to årene (disse er markert i kursiv i tabell 7.1). Dette medfører begrensninger når det kommer til å utforske endringer i den implementerte læreplanen fra 2019 til 2023.

Analysen

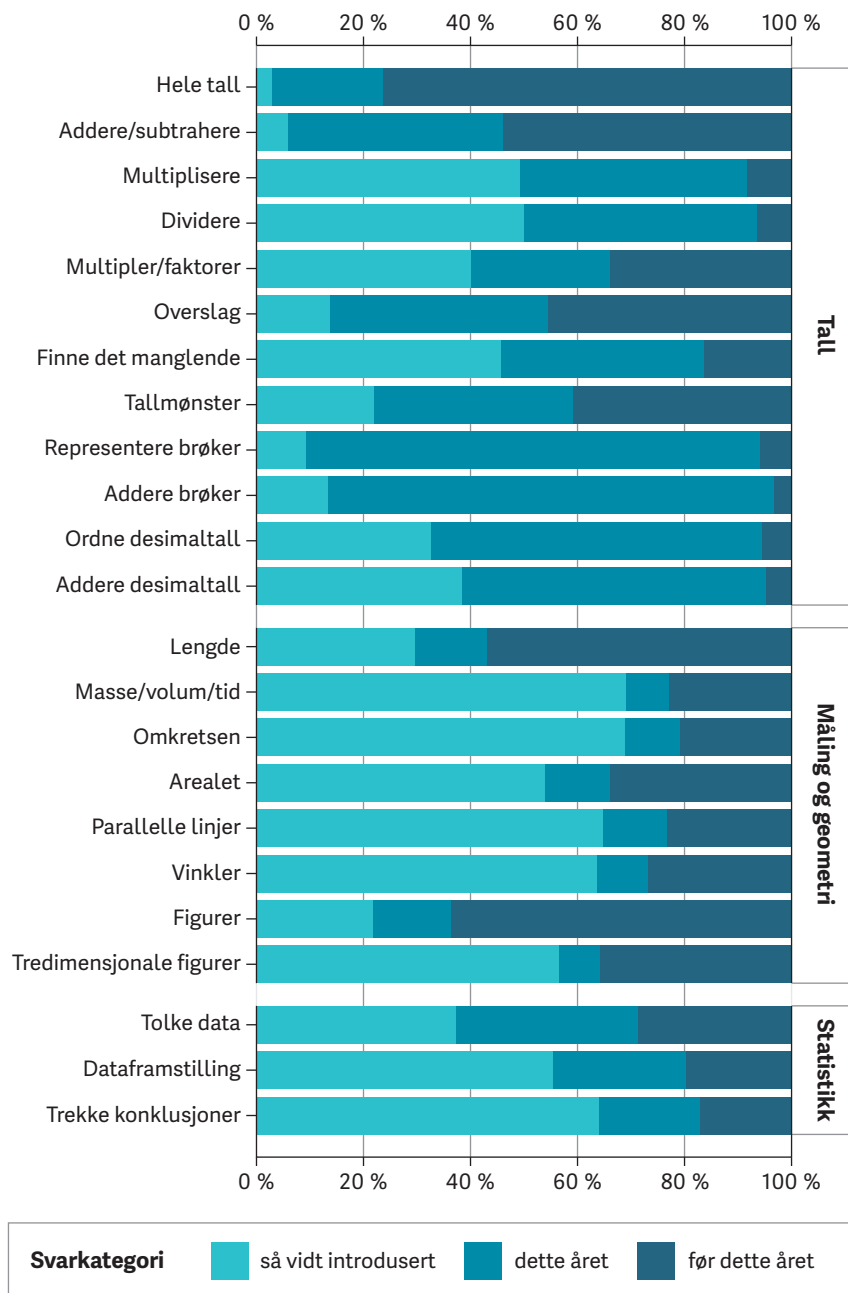
Vi bruker enkle analysemetoder hvor vi fokuserer på å presentere vektete gjennomsnitt for de tre svarkategoriene i spørreskjemaet. Resultatene presenteres i figurer som gjør det mulig å sammenligne Norge med de andre nordiske landene. For å se på endringer over tid utforsker vi om det finnes sammenhenger mellom Norge og de andre nordiske landene innen hvert år og om sammenhengene har endret seg over tid. I tillegg gjør vi en direkte sammenligning av de ti emnene som var like i 2019 og 2023. Gjennomsnittene representerer den gjennomsnittlige elevens undervisning, slik den rapporteres av lærerne.

Når det kommer til vår tolkning av lærernes svar, har vi valgt å fokusere på de to svarkategoriene «bare så vidt introdusert eller ikke undervist (ennå)» og «hovedsakelig undervist dette skoleåret». Når lærere svarer at et emne er «bare så vidt introdusert eller ikke introdusert (ennå)», kan dette gi informasjon om mulige mangler i den implementerte læreplanen. For eksempel, hvis Norge har en høyere andel slike svar for flere emner enn de andre landene, kan dette tyde på at den implementerte læreplanen i Norge er smalere ved at den dekker færre faglige områder. Når lærere svarer at et emne er «hovedsakelig undervist dette skoleåret», kan dette indikere hvor sterkt den intenderte læreplanen styrer undervisningen. Dersom den intenderte læreplanen var den eneste faktoren som påvirket det faglige innholdet i undervisningen, ville vi forventet at enten 100 prosent eller 0 prosent av lærerne ville rapportert at et gitt emne ble undervist dette året avhengig av om emnet var i den intenderte læreplanen for dette trinnet. Avvik fra dette mønsteret sier noe om hvor sterk påvirkning den intenderte læreplanen har på det faglige innholdet i undervisningen.

Resultater

Norske resultater

Figur 1 viser norske læreres rapportering av når elevene har fått undervisning i de ulike emnene i matematikk i TIMSS 2023. Som forventet viser



Figur 7.1. Lærernes rapporterte dekning av faglige emner i Norge i 2023 på 5. trinn.
Note: Fulltekst til svarkategoriene er «Hovedsakelig undervist før dette skoleåret», «Hovedsakelig undervist dette skoleåret» og «Bare så vidt introdusert eller ikke undervist (ennå)». Se fotnote 1 for notat om svarkategorier brukt i Norge i 2023.

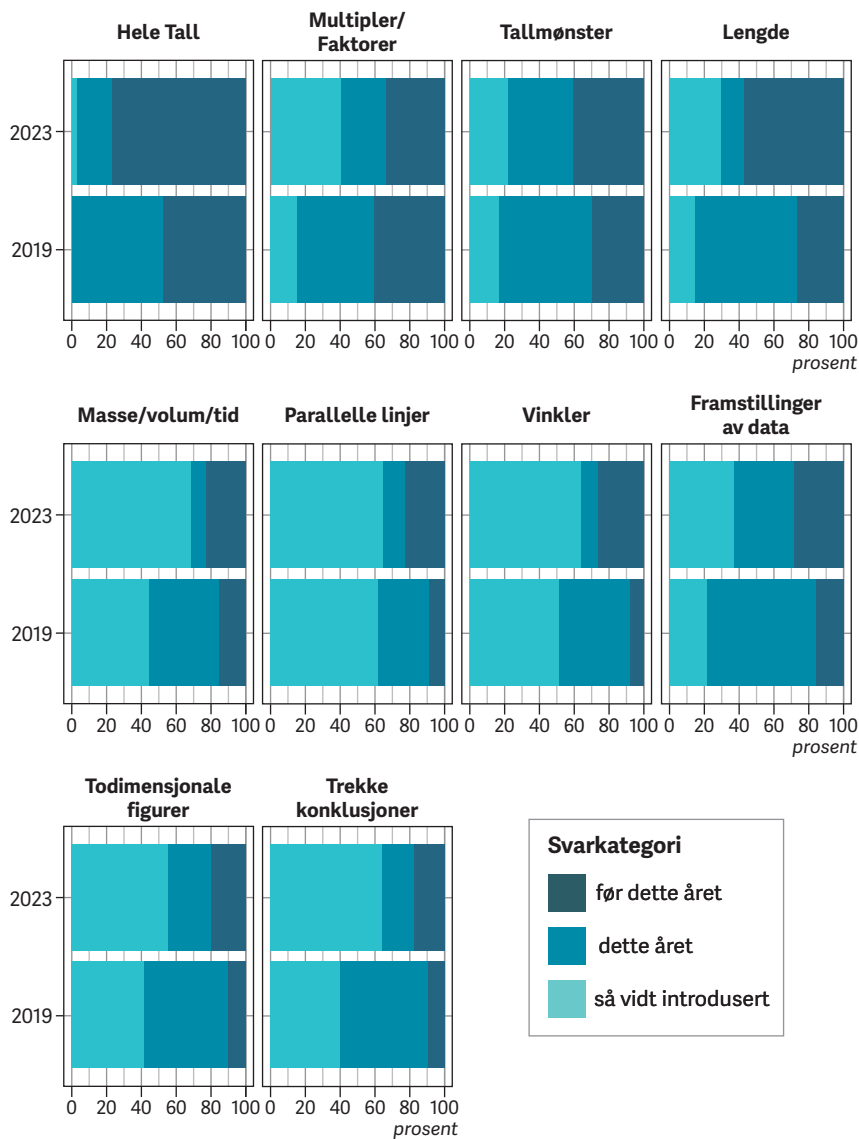
figuren at de ulike faglige emnene undervises på ulike tidspunkter. Vi ser for eksempel at de fleste elevene ser ut til å ha fått undervisning i «Hele tall» og addisjon og subtraksjon i svarkategorien «før dette året» eller «dette året». Derimot har under ti prosent av elevene fått undervisning i multiplikasjon og divisjon før dette skoleåret (figur 7.1).

Generelt ser vi at emnene innenfor *tall* i større grad har blitt undervist enten før eller i løpet av dette skoleåret, sammenlignet med emnene innenfor *måling og geometri* og *statistikk* (figur 7.1). For eksempel rapporterer over 60 prosent av lærerne at egenskaper ved geometriske figurer, som masse, volum, omkrets og vinkler, er i svarkategorien «så vidt introdusert». Vi kan også se at innenfor *tall*, er det en forholdsvis stor andel lærere som rapporterer at elevene blir undervist i mange av disse emnene dette året. Dette kan tyde på at disse emnene er tydeligere koblet til kompetansemålene for 5. trinn enn de andre emnene.

Forskningsspørsmål 1: Hvordan har den implementerte læreplanen endret seg i Norge mellom 2019 og 2023?

Figur 7.2 viser en direkte sammenligning av de norske lærernes svar fra 2019 og 2023 for de emnene som er like i de to undersøkelsene. I 2023 er det en større andel elever med lærere som rapporterte at emnet var «så vidt eller ikke ennå introdusert» eller «undervist før dette året», og en reduksjon i andelen som rapporterer at emnene ble undervist dette året. Vi tolker dette som en innsnevring av den implementerte læreplanen (det vil si færre emner ble undervist på 5. trinn). Samtidig er det en større variasjon mellom lærere, siden det flere rapporterer både at emner er undervist før dette året og at emner ennå ikke er undervist. Dette er en utfordring sett i lys av LK20 som en styringsmekanisme, siden dette kan tyde på mindre enighet om når emner skal undervises. Det er viktig å understreke at denne analysen kun omfatter et begrenset antall emner. Det er mulig at andre emner, som ikke er inkludert i denne analysen, har blitt mer vektlagt i undervisningen.

Sammenligning av 2019 og 2023 i Tall

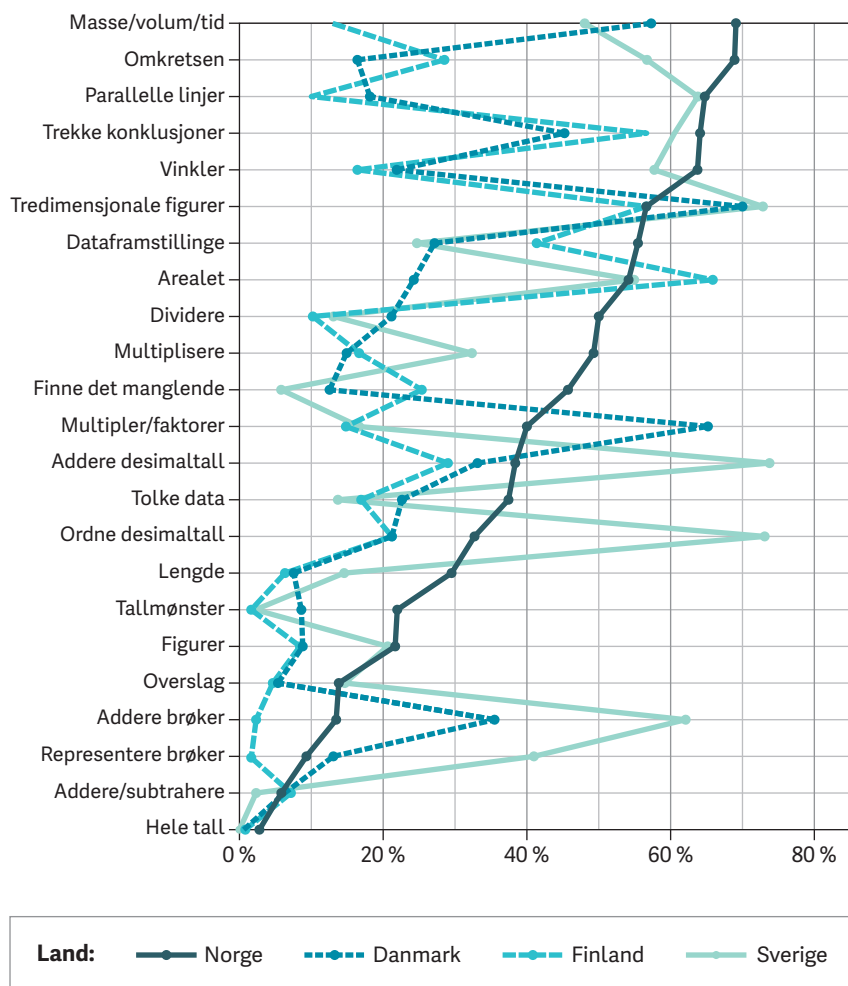


Figur 7.2. Sammenligning av spørsmålene som var like i TIMSS 2019 og 2023 på 5. trinn. Fulltekst til svarkategoriene er «Hovedsakelig undervist før dette skoleåret», «Hovedsakelig undervist dette skoleåret» eller «Bare så vidt introdusert eller ikke undervist (ennå)». Se fotnote 1 for notat om svarkategorier brukt i Norge i 2023.

Forskningsspørsmål 2: I hvilken grad er endringene i den implementerte læreplanen i tråd med endringer fra LK06 til LK20?

Vår tolkning av endringene i læreplanene viste at LK20 er mindre konkret enn LK06 når det gjelder faglig innhold i kompetansemålene og i hvilke trinn det er forventet at elevene skal ha lært dette innholdet. Dette legger opp til mer rom for tolkninger, og større frihet, blant lærere. Samtidig var det også noen konkrete endringer i LK20 når det kom til det faglige innholdet. For eksempel ble brøk tatt inn i kompetansemålene for 5. trinn, noe som gjenspeiles i analysene ved at andelen som har blitt undervist i brøk økte fra 40 prosent i 2019 (figur 4) til over 80 prosent i 2023 (figur 7.3). Når det kommer til emnene innenfor *måling og geometri*, er disse ofte rapportert som så vidt eller ennå ikke introdusert (figur 7.1). Dette kan muligens forklares ved at kompetansemålene i geometri ble flyttet til 6. trinn i LK20. Det vil si at for de tydeligste endringene vi fant i sammenligningen av LK20 og LK06, kan vi se mønster i dataene som støtter påstanden om at lærere har endret undervisningspraksis i tråd med endringen i LK20.

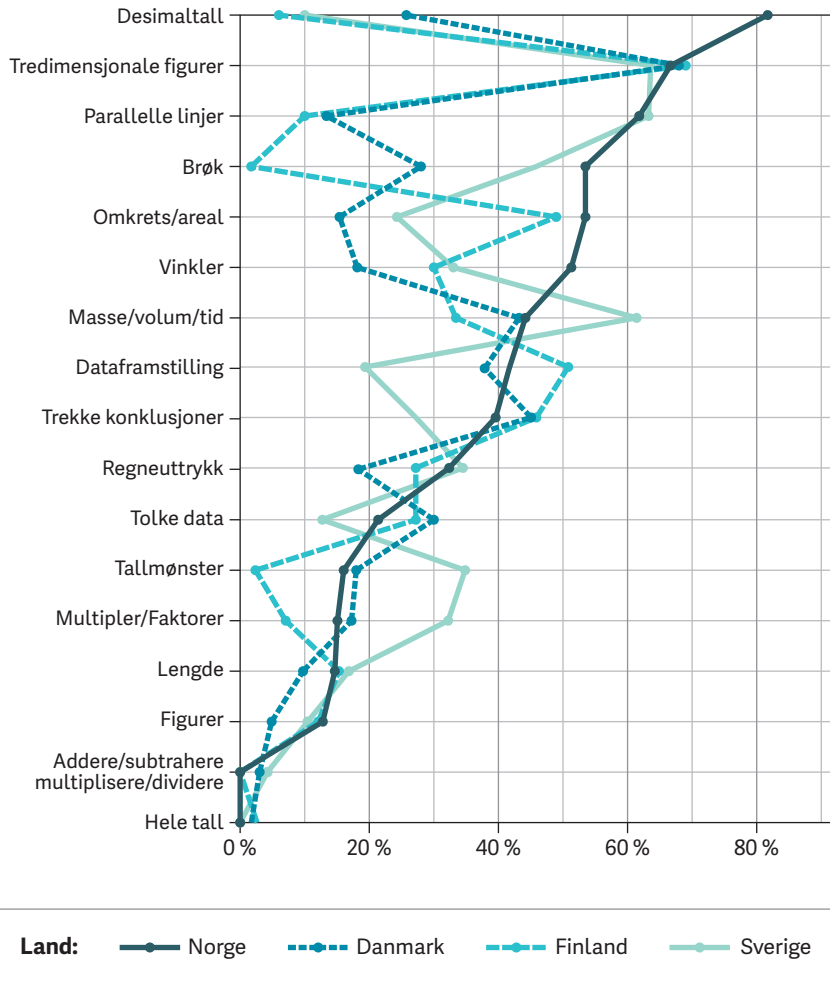
Vi ser også en økt spredning i når emner undervises, der flere lærere rapporterer at emner enten ble undervist før dette året eller ikke ennå. Dette kan tyde på at LK20 gir mindre tydelige føringer for når faglig innhold skal undervises. Dette samsvarer også med vår gjennomgang, der vi fant at kompetansemålene er bredere og mindre konkrete enn tidligere, noe som åpner opp for at lærere i større grad enn tidligere må tolke og bestemme innholdet i undervisningen på hvert trinn.



Figur 7.3. Prosentandel elever i 2023 med en lærer som rapporterte at emnet var «bare så vidt introdusert eller ikke undervist (ennå)», for hvert emne i de fire nordiske landene.

Forskningsspørsmål 3: Hvordan er den implementerte læreplanen i matematikk i Norge sammenlignet med de andre nordiske landene?

Figur 7.3 og 7.4 viser andelen elever som har lærere som rapporterer at hvert emne bare var «så vidt introdusert» eller «ikke introdusert ennå». Figur 7.3 viser data fra 2023 og figur 7.4 fra 2019. Emnene er sortert etter andelen i Norge. I 2023 har Norge den høyeste andelen slike svar i 14 av 23 emner (61 %). Kun 5 av disse 14 emnene er innenfor hovedemnet *tall*, selv om



Figur 74. Prosentandel elever i 2019 med en lærer som rapporterte at emnet var «Bare så vidt introdusert eller ikke undervist (ennå)», for hvert emne i de fire nordiske landene.

halvparten av emnene totalt er innenfor *tall*-kategorien. Til sammenligning hadde Norge høyest andel i 6 av 17 (35 %) emner i 2019.

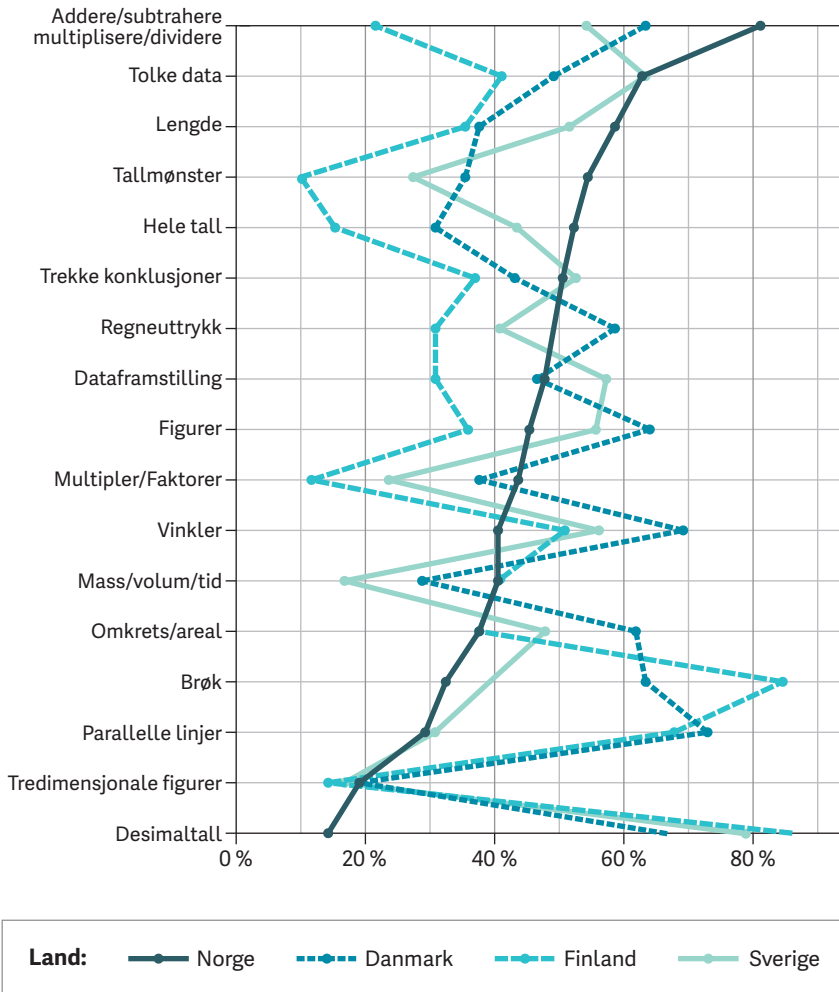
Med en sammenligning av fire land, ville man forvente at hvert land hadde høyest andel i omtrent 25 prosent av emnene dersom landene var like. Dette tyder på at norske elever i gjennomsnitt har fått undervisning i færre av emnene som er inkludert i TIMSS enn elever i de andre nordiske landene i opplæringen fram til det tidspunktet spørreskjema ble svart. Dette gjelder hovedsakelig for emner innenfor *måling og geometri* og *statistikk*. Selv om vi ikke har data for emner som ikke er inkludert i TIMSS, er

det ingen grunn til å tro at TIMSS-rammeverket har svakere sammenheng med det som er undervist i Norge enn det som er undervist i andre land. Det er derfor liten grunn til å tro at disse forskjellene skyldes skjevheter i hvilke emner som er inkludert i TIMSS.

Figur 7.5 og 7.6 viser tilsvarende resultater for lærere som har rapportert at emnet hovedsakelig er undervist dette året. Vi fokuserer på «dette året» siden det gir innsikt i hvor sterkt den intenderte læreplanen styrer undervisningen. Dersom læreplanen fungerer som tydelig styringsmekanisme, med tydelige rammer for når elevene skal lære ulike kunnskaper og ferdigheter, ville vi forvente at emner som er ment å undervises på 5. trinn



Figur 7.5. Prosentandel elever i 2023 med en lærer som rapporterte at emnet var «hovedsakelig undervist dette skoleåret», for hvert emne i de fire nordiske landene.



Figur 7.6. Prosentandel elever i 2019 med en lærer som rapporterte at emnet var «hovedsakelig undervist dette skoleåret», for hvert emne i de fire nordiske landene.

nærmer seg 100 prosent, mens emner som ikke er en del av kompetansemålene for dette trinnet nærmer seg 0 prosent.

Som figurene viser, varierer andelen elever som har fått undervisning i hvert emne dette året fra 20 til 80 prosent i de fleste landene. Dette viser betydelig variasjon både innen og mellom land når det gjelder tidspunktet for når emner blir undervist. Her skiller Norge seg ut fra de andre nordiske landene i 2023 (se figur 7.5), men ikke i 2019 (se figur 7.6). I 2023 får nesten alle elever i Norge undervisning om brøk, og en stor andel får undervisning om desimaltall. Samtidig er det relativt få elever som får undervisning i de andre emnene. Dette gjelder spesielt innenfor *måling og geometri* og

statistikk, som er undervist til mindre enn 20 prosent av elevene på 5. trinn. Dette tyder på en innsnevring av den implementerte læreplanen på 5. trinn i Norge. Det er mulig at norske lærere underviser andre emner som ikke er inkludert i lærerspørreskjemaet i TIMSS. Fra et styringsperspektiv, vil i så fall dette kunne tolkes som et positivt tegn dersom dette er i tråd med intensjonene bak LK20.

Diskusjon

I dette kapittelet har vi undersøkt den implementerte læreplanen i Norge og sammenlignet den med de andre nordiske landene. I norsk skole brukes den nasjonale læreplanen som et styringsverktøy for å legge til rette for at alle elever skal få like muligheter til læring. Et av myndighetenes formål med endringen fra LK06 to LK20 var å fremme dybdelæring og bedre forståelse blant elevene (Meld. St. 28 (2015–2016)). Vi vet imidlertid at læreplaner ikke overføres direkte til klasserommene, men at lærere tolker innholdet og tilpasser undervisningen basert på sin forståelse og erfaring (Spillane et al., 2002). For å få innsikt i hva som faktisk undervises, er det derfor viktig å undersøke den implementerte læreplanen.

Vi fant at norske elever i 2023 i større grad enn elever i de andre nordiske landene har lærere som rapporterer at flere emner bare var så vidt introdusert eller ikke undervist ennå. Samtidig er det færre norske elever som har lærere som rapporterte at emnene har blitt undervist stort sett dette året. Dette gjelder spesielt emner innenfor *måling og geometri*, mens forskjellene er mindre for hovedemnet *tall*.

I 2019 var det få systematiske forskjeller mellom de nordiske landene, noe som tyder på at endringene har skjedd i perioden mellom 2019 og 2023. Dette kan tyde både på en innsnevring av den implementerte læreplanen i Norge, og mer variasjon i hvilke emner som undervises og når de blir undervist. Det er viktig å nevne at det var mye variasjon innad og mellom landene, men at Norge fortsatt skiller seg ut.

Disse funnene kan tyde på at matematikkundervisningen til norske elever er knyttet til færre emner enn matematikkundervisningen til elever i andre nordiske land, noe som kan ha negative konsekvenser for bredden i norske elevers kunnskaper og ferdigheter. Samtidig er det viktig å understreke at dette ikke nødvendigvis er negativt. Dersom norske elever oppnår en dypere forståelse i de emnene som undervises, kan dette anses som en positiv endring som er i tråd med dybdelæring, som er et av målene med LK20 (Kunnskapsdepartement, 2017).

Resultatene fra TIMSS 2023 viser at nedgangen i norske elevers prestasjoner på 5. trinn er størst innenfor hovedemnet *geometri og målinger*. Dette er det hovedemnet som ble mest redusert både i den intenderte og implementerte læreplanen i 1.–5. trinn. Samtidig var nedgangen i prestasjoner minst i hovedemnet *tall*, som også har færrest endringer i den implementerte læreplanen. Likevel er nedgangen i *tall* også betydelig (omtrent 10 skalapoeng; Kaarstein et al., 2024), noe som kan tyde på at undervisning i færre emner ikke nødvendigvis gir bedre læring. Dette peker på at endringer i den implementerte læreplanen kan være én forklaringsfaktor for nedgangen i prestasjoner. For å bekrefte dette, trengs mer presise analyser som direkte kobler emnene som har blitt undervist med det faglige innholdet i oppgavene på prøven. Vi arbeider med å undersøke slike koblinger og funnene vil bli presentert i framtidige artikler.

En viktig begrensning ved denne studien er at emnene som lærerne blir spurt om i TIMSS ikke dekker hele bredden av LK20. Det er derfor mulig at innsnevringen vi fant skyldes at norske elever blir undervist i emner som ikke inngår TIMSS, som for eksempel programmering. De faglige emnene som inngår i lærerspørreskjemaet ble valgt ut for å gi best mulig dekning av læreplanene i deltakerlandene. Emnene som inkluderes i TIMSS bør være ganske omfattende, hvilket kan bekreftes med vår vurdering, men det er mulig (men kanskje ikke så sannsynlig) at de ikke dekker det som undervises i Norge like godt som det som undervises i de andre nordiske landene.

Endringene vi fant i den implementerte læreplanen fra 2019 til 2023 var stort sett i tråd med endringer fra LK06 til LK20. Dette inkluderer mindre undervisning i geometri og måling, og mer undervisning i brøk i 5. trinn. Dette kan tyde på at LK20 har fungert som en styringsmekanisme når det kommer til det faglige innholdet i matematikkfaget. Vi har ikke utforsket andre aspekter av LK20 her, men tidligere studier peker på at læreplaner fungerer dårligere som styringsverktøy når det kommer til undervisningsmetoder og mer komplekse aspekter av undervisning (White et al., 2022). Dette er fordi lærere må tolke læreplanen for å forstå det som må endres av praksis, og denne tolkningen påvirkes av egne erfaringer og kunnskap (Cohen, 1990). Mange av endringene LK20 var ment å føre til, som dybdelæring og tverrfaglige temaer (Kunnskapsdepartement, 2017), er mer komplekse enn endringer i faglig innhold. Det er fortsatt uklart i hvilken grad LK20 har ført til disse endringene, men flere evalueringsprosjekter har pekt på at det har skjedd enkelte endringer i tråd med LK20 (f.eks. Brevik et al., 2023, 2024).

Vår vurdering av LK20 viser også at læreplanen er vagere enn LK06 når det gjelder å beskrive det forventede faglige innholdet i undervisningen for

hvert trinn. Dette punktet har vært gjenstand for offentlig debatt i norsk media (Bjerke & Solomon, 2025; Jambak, 2024; Sunnanå, 2025; Westermoen, 2025). Det er mulig at den økte spredningen i den implementerte læreplanen som ble funnet, det vil si at færre lærere som rapporterer at et emne var dekket i år og flere som rapporterer at det ble dekket før dette året eller ikke ennå, er en konsekvens av denne vagheten. En mindre presis læreplan åpner for flere tolkninger av innholdet, noe som kan føre til uenighet om når visse emner bør undervises. Dette kan ha negative konsekvenser for likeverd i norsk skole, noe som er et sentralt prinsipp i det norske utdanningssystemet (Opplæringslova, 2023, § 1–3), ettersom større variasjon i den implementerte læreplanen kan føre til ulik tilgang til det faglige innholdet i skolen. Denne ulikheten kan ytterligere forsterkes av at en relativt stor andel lærere mangler formell undervisningskompetanse i matematikk, siden kompetansen til lærerne påvirker deres tolkninger av læreplanen (Cohen, 1990).

Overordnet viser funnene i denne studien at det er behov for mer forskning på den implementerte læreplanen i Norge, og hvilke konsekvenser endringene har for elevenes læring og den skolen som helhet. Den mulige innsnevringen av den implementerte læreplanen, altså at færre emner undervises, samt en større variasjon i når ulike emner blir undervist, er bekymringsfull fra et likeverdperspektiv og når det kommer til å gi elever kunnskapen og ferdighetene de trenger for å lykkes senere i livet. Denne innsnevringen var imidlertid delvis forventet, ut fra endringene fra LK06 til LK20, hvor målet var å gi tid og plass for dybdelæring av det faglige innholdet som undervises og elevene innsikt i viktige tverrfaglige temaer (Utdanningsdirektoratet, 2021). Hvorvidt dette fører til en positiv endring eller ikke, avhenger av om elevene faktisk oppnår dypere forståelse av det faglige innholdet og de tverrfaglige temaene. Den tilsynelatende økte variasjonen i den implementerte læreplanen, altså at lærere underviser ulike emner på ulike tidspunkter, var derimot ikke forventet og gir grunn til bekymring. Det er nødvendig med mer forskning for å forstå hvorfor den implementerte læreplanen varierer, og hva det betyr for elevenes læring og for den norske skolen.

Forfatterbiografier

Mark White er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han leder prosjektet *Teaching and Learning International Survey (TALIS) 2024* i Norge. Hans forskningsinteresse er blant annet undervisningskvalitet, lærerprofesjonalitet og kvantitative metoder.

Alexander Jonas Viktor Selling er postdoktor ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han har en doktorgrad i matematikkdiraktikk og bakgrunn som lektor i matematikk og naturfag. Hans forskningsinteresser inkluderer undervisningskvalitet, matematikk-lærerutdanning og problemløsning i matematikk.

Andreas Pettersen er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han har en doktorgrad i matematikkdiraktikk og har jobbet med PISA-undersøkelsen siden 2017. Pettersen har utgitt flere artikler om matematikkdiraktikk og vært redaktør for boka *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (utgitt i 2024).

Jelena Veletić er postdoktor ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har doktorgrad i utdanningsmåling og bakgrunn fra psykologi. Hun arbeider særlig med analyser av storskala utdanningsdata, og forskningsinteressene hennes omfatter skoleledelse, skolemiljø, læreres jobbtilfredshet og andre forhold som påvirker lærere.

Referanser

- Artigue, M., Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM Mathematics Education*, 45(1), 797–810. <https://doi.org/10.1007/s11858-013-0506-6>
- Berliner, D. C. (1990). What's all the fuss about instructional time? I M. Ben-Peretz & R. Bromme (Red.), *The nature of time in school* (s. 3–35). Teachers College Press.
- Bjerke, A. H. & Solomon, Y. (2025, 20. februar). *Nivågrupper i matematikkfaget – en dårlig idé*. Utdanningsnytt. <https://www.utdanningsnytt.no/bedre-skole-fagartikkel-matematikk/nivagruupper-i-matematikkfaget-en-darlig-ide/426047>
- Blömeke, S. (2024). Equality in content coverage in the Nordic countries? I N. Teig, T. Nilsen & K. Y. Hansen (Red.), *Effective and equitable teacher practice in mathematics and science education* (s. 231–250). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-49580-9_8
- Brevik, L. M., Gudmundsdottir, G. B., Barreng, R. L. S., Dodou, K., Doetjes, G., Evertsen, I., Goldschmidt-Gjerløw, B., Hartvigsen, K. M., Hatlevik, O. E., Isaksen, A. R., Magnusson, C. G., Mathé, N. E. H., Siljan, H., Stovner, R. B. & Suhr, M. L. (2023). *Å mestre livet i 8. klasse. Perspektiver på livsmestring i klasserommet i sju fag* (Rapport 2 fra forsknings- og evalueringsprosjektet EDUCATE). Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Brevik, L. M., Gudmundsdottir, G. B., Aashamar, P. N., Barreng, R. L. S., Dodou, K., Doetjes, G., Hartvigsen, K. M., Hatlevik, O. E., Mathé, N. E. H., Roe, A., Siljan, H., Stovner, R. B. & Suhr, M. L. (2024). *Å jobbe utforskende på Vg1 og Vg2. Den enkelte lærers undervisning har mer å si enn fagenes egenart* (Rapport 3 fra forsknings- og evalueringsprosjektet EDUCATE). Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, 64(8), 723–733.
- Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, 12(3), 311–329.
- Floden, R. E. (2002). The measurement of opportunity to learn. I A. C. Porter & A. Gamoran (Red.), *Methodological advances in cross-national surveys of educational achievement* (s. 231–266). National Academy Press.

- Goodlad, J. I. (1979). *Curriculum inquiry. The study of curriculum practice*. McGraw-Hill.
- Houang, R. T. & Schmidt, W. H. (2008). TIMSS international curriculum analysis and measuring educational opportunities. *International Journal of Educational Research*, 47(4), 235–251. <https://doi.org/10.1016/j.ijer.2008.01.003>
- Jambak, T. (2024, 4. juni). Nå trengs det en grundig diskusjon om veien videre for matematikkfaget. *Aftenposten*. <https://www.aftenposten.no/meninger/debatt/i/73KEm4/>
- Kang, H. & Cogan, L. (2022). The differential role of socioeconomic status in the relationship between curriculum-based mathematics and mathematics literacy: The link between TIMSS and PISA. *International Journal of Science and Mathematics Education*, 20(1), 133–148. <https://doi.org/10.1007/s10763-020-10133-2>
- Karseth, B., Kvamme, O. A., & Ottesen, E. (2020). *Fagfornyelsens læreplanverk: Politiske intensjoner, arbeidsprosesser og innhold* (Rapport nr. 1). Universitetet i Oslo. <https://www.uv.uio.no/forskning/prosjekter/fagfornyelsen-evaluering/publikasjoner/eva2020-delrapport-1.pdf>
- Kunnskapsdepartementet. (2006). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Fastsatt som forskrift ved kongelig resolusjon. Læreplanverket for Kunnskapsløftet 2006.
- Kunnskapsdepartementet. (2017). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Fastsatt som forskrift ved kongelig resolusjon. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/overordnet-del/>
- Kunnskapsdepartementet. (2019). *Læreplan i matematikk 1.–10. trinn* (MAT01–05). Fastsatt som forskrift. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/mat01-05>
- McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*, 17(3), 305–322. <https://doi.org/10.2307/1164509>
- Meld. St. 28 (2015–2016). *Fag – fordypning – forståelse: En fornyelse av Kunnskapsløftet*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/>
- Opplæringslova. (2023). *Lov om grunnskolen og den vidaregåande opplæringa* (LOV-1998-07-17-61). Lovdata. <https://lovdata.no/lov/1998-07-17-61>
- Remillard, J. T. & Heck, D. J. (2014) Conceptualizing the curriculum enactment process in mathematics education. *ZDM Mathematics Education*, 46, 705–718. <https://doi.org/10.1007/s11858-014-0600-4>
- Rolfe, V., Strietholt, R. & Hansen, K. Y. (2021). Does inequality in opportunity perpetuate inequality in outcomes? International evidence from four TIMSS cycles. *Studies in Educational Evaluation*, 71, Artikkel 101086 <https://doi.org/10.1016/j.stueduc.2021.101086>
- Rosén, M. & Nilsen, T. (2024). Are changes in content coverage related to changes in achievement over time? I N. Teig, T. Nilsen & K. Y. Hansen (Red.), *Effective and equitable teacher practice in mathematics and science education* (s. 187–205). Springer. https://doi.org/10.1007/978-3-031-49580-9_6
- Scheerens, J. (2016). Meta-analyses of school and instructional effectiveness. I J. Scheerens (Red.), *Educational effectiveness and ineffectiveness: A critical review of the knowledge base* (s. 175–223). Springer Netherlands. https://doi.org/10.1007/978-94-017-7459-8_8
- Spillane, J. P., Reiser, B. J. & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387–431.
- Sunnanå, S. (2025, 2. mai). *Er matematikkfaget blitt for teoretisk?* Khrono. <https://www.khrono.no/er-matematikkfaget-blitt-for-teoretisk/964995>
- Utdanningsdirektoratet. (2021, 24. juni). *Kunnskapsløftet 2020 – hvorfor har vi fått nye læreplaner?* <https://www.udir.no/laring-og-trivsel/lareplanverket/fagfornyelsen/hvorfor-nye-lareplaner/>
- Westermoen, B. N. (2025, 21. juni). Hva skal vi bruke matematikken til? *Stavanger Aftenblad*. <https://www.aftenbladet.no/i/4B4aqa>
- White, M., Maher, B. & Rowan, B. (2022). Common core–related shifts in English language arts teaching from 2010 to 2018. *The Elementary School Journal*, 123(1), 1–27. <https://doi.org/10.1086/720732>
- Aasen, P., Prøitz, T. & Rye, E. (2015). Nasjonal læreplan som utdanningspolitisk dokument. *Norsk pedagogisk tidsskrift*, 99(6), 417–433. <https://doi.org/10.18261/ISSN1504-2987-2015-06-03>

KAPITTEL 8

Undervisningskvalitet og matematikkprestasjoner på 5. trinn fra TIMSS 2019 til TIMSS 2023

Mark White, Hege Kaarstein og Armin Jentsch

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: TIMSS 2023 viste en signifikant nedgang i matematikkprestasjonene på 5. trinn i Norge. Denne nedgangen sammenfaller med flere andre indikatorer som tyder på at norske elevers prestasjoner er på vei ned. Fordi undervisningskvalitet anses som en av de viktigste, formbare faktorene med positiv sammenheng med elevers prestasjoner, utforsker vi i dette kapittelet om elevenes oppfatning av undervisningskvalitet har endret seg. Videre utforskes det i hvilken grad nedgangen i prestasjoner har sammenheng med en eventuell endring i elevenes oppfatning av undervisningskvalitet. Kapittelet bruker data fra TIMSS 2019 og 2023 på 5. trinn. Resultatene viser at det er en liten, men statistisk signifikant nedgang i elevrapportert undervisningskvalitet fra 2019 til 2023, og at 17 prosent av nedgangen i prestasjoner henger sammen med nedgangen i elevenes oppfatning av undervisningskvalitet. Dette tyder på at nedgangen i elevenes oppfatning av undervisningskvalitet kan være en liten, men viktig forklaringsfaktor for nedgangen i prestasjoner selv om det meste av nedgangen i prestasjoner har sammenheng med andre faktorer.

Nøkkelord: undervisningskvalitet, Norge, matematikkprestasjoner, TIMSS, regresjon

Abstract**Instructional Quality and Mathematics Achievement in Grade 5 from TIMSS 2019 to TIMSS 2023**

TIMSS 2023 showed a significant decrease in students' mathematics achievement at 5th grade in Norway. This decrease coincides with several other indicators suggesting that Norwegian students show lower levels of learning. Since teaching quality is considered one of the most important, malleable factors that is associated with student achievement, this chapter seeks to understand if the observed decrease in achievement coincides with decreases in students' perception of teaching quality. The chapter further explores if there is an association between these two downturns. The chapter uses data from 5th grade mathematics in TIMSS 2019 and 2023. The results show that there is a small, but statistically significant decrease in student-reported teaching quality from 2019 to 2023 and that 17 percent of the decrease in achievement is associated with this decrease in teaching quality. This suggests that decrease in teaching quality could be a small, but important, explanation for the decrease in achievement, but most of the decrease in achievement is driven by other factors.

Keywords: teaching quality, Norway, mathematics achievement, TIMSS, regression

Introduksjon

De norske resultatene fra *Trends in International Mathematics and Science Study* (TIMSS) 2023 viste en signifikant nedgang i elevenes matematikkprestasjoner på 5. trinn fra den forrige TIMSS-undersøkelsen i 2019 (Kaarstein et al., 2024). Den gjennomsnittlige nedgangen var på 12 poeng. Dette utgjør rett i overkant av en femtedel av differansen mellom matematikkresultatet for elevene på 4. og 5. trinn fra TIMSS-undersøkelsen i 2015 (Bergem, 2016). Nedgangen kan sies å tilsvare omtrent en femtedel av et skoleår.

Når det rapporteres om endringer, er det naturlig å stille spørsmål om hvorfor de har inntruffet eller hva som har forårsaket endringene. I skole-sammenheng er det spesielt viktig å identifisere mulige årsaker, eller faktorer som skolen, lærerne og/eller utdanningssystemet kan gjøre noe med (i motsetning til for eksempel elevenes hjemmebakgrunn, som skolene ikke kan gjøre noe med). En av flere faktorer som tidligere forskning har pekt ut som mulig å gjøre noe med, som også er veldig viktig for elevenes prestasjoner, er lærernes undervisningskvalitet (se f.eks. Bergem et al., 2016; Klieme & Nilsen, 2022; Seidel & Shavelson, 2007; Senden et al., 2023). Undervisningskvalitet er undervisningspraksis som skaper læringsmuligheter og støtter elevenes læring på en slik måte at den øker sannsynligheten for ønsket læringsutbytte (Charalambous et al., 2021). Derfor er målet med dette kapittelet å undersøke om lærernes undervisningskvalitet har endret seg, og om en endring kan ha sammenheng med den observerte nedgangen i elevenes matematikkprestasjoner på 5. trinn. For å nå dette målet, stilles følgende forskningsspørsmål:

- 1 Hvordan har elevenes oppfatning av undervisningskvalitet i matematikk, sett fra elevenes perspektiv, endret seg fra TIMSS 2019 til TIMSS 2023?
- 2 Dersom det har vært en endring i elevenes oppfatning av undervisningskvalitet, hvor stor andel av denne endringen har sammenheng med nedgangen i matematikkprestasjoner på 5. trinn fra TIMSS 2019 til TIMSS 2023?

Forskningsspørsmål 2 avhenger av resultatet på forskningsspørsmål 1. Dersom elevenes oppfatning av lærernes undervisningskvalitet ikke har endret seg fra TIMSS 2019 til TIMSS 2023, vil det ikke gi mening å gå videre til forskningsspørsmål 2.

Bakgrunn

I det store bildet vil en enkeltregistrering av nedgang mellom to målepunkter normalt sett ikke være nok til å utløse bekymring. Når flere undersøkelser og resultater peker i samme retning, derimot, er det viktigere å identifisere mulige årsaker.

Fra 2015 til 2023 viser trendresultatene fra TIMSS-undersøkelsen en nedgang i matematikkprestasjoner for elevene på både 5. og 9. trinn (Kaarstein et al., 2024). Elevene på 9. trinn hadde en signifikant nedgang fra 2015 til 2019 på 9 poeng (omtrent en tredjedels skoleår, Kaarstein et al., 2020). Også de norske 15-åringenes matematikkresultater i undersøkelsen *Programme for International Student Assessment* (PISA) 2022 gikk ned i forhold til de tidligere stabile PISA-resultatene (Jensen et al., 2023). For PISA 2022 ble dataene riktignok samlet inn rett i etterkant av covid-19-pandemien, men den store nedgangen i matematikkresultatene på 33 poeng (som kan sies å tilsvare over ett års skolegang, se f.eks. Avvisati & Givord, 2023) ble bare delvis tilskrevet pandemien (Jensen et al., 2023). For de nasjonale prøvene i regning, som måler elevenes grunnleggende ferdigheter til regning på 5., 8. og 9. trinn, har det vært mulig å se på trend fra 2022 (NOU 2023: 1), og trenden for de tre trinnene viser også tendenser til nedgang, noe større på ungdomstrinnet enn på 5. trinn (Utdanningsdirektoratet, 2024). Med utgangspunkt i resultatene fra de ulike undersøkelsene nevnt her, kan det tyde på at norske elevers kompetanse i regning og matematikk er på vei nedover.

Samtidig med dårligere kompetanse i matematikk hos norske elever, vil et økende mangfold i befolkningen (NOU 2015: 8) sannsynligvis også by på større utfordringer når undervisningen skal tilpasses en stadig mer heterogen elevmasse. I tillegg rapporteres det om økende stress og vanskeligheter blant elever og foreldre (Meld. St. 34 (2023–2024); Nortvedt et al., 2020; Pascoe et al., 2020). I den norske kortrapporten fra TIMSS 2023, antydes det en negativ utvikling av skolemiljøet (Kaarstein et al., 2024). Flere elever rapporterer om å bli mobbet hver uke nå enn tidligere, det er færre lærere som føler seg trygge på skolen, og rektorene rapporterer om flere problemer. Videre kommer det også fram at rektorene rapporterer om et mindre læringstrykk, og lærerne rapporterer om flere begrensinger i undervisningen. Det er eksempelvis en større andel elever som er sultne/trøtte, mangler forkunnskaper og har liten konsentrasjonsevne (Vik et al., 2022a, 2022b). Ifølge tidligere forskning vil begrensningene igjen kunne

påvirke lærerens undervisningskvalitet (Oliver et al., 2011; Wang & Degol, 2014). Og, som nevnt over, er lærernes undervisningskvalitet knyttet til elevenes prestasjoner.

Med dette som bakgrunn, er det overordnede spørsmålet for dette kapitlet om den norske skolen har klart å opprettholde kvaliteten i undervisningen i et (skole)samfunn som har gjennomgått forandringer. Hvis undervisningskvaliteten ligger på samme nivå som tidligere, til tross for endringene, kan det tyde på at lærerne har klart å tilpasse seg de nye forholdene. På liknende måte, dersom nedgangen i elevenes prestasjoner har liten eller ingen sammenheng med endringer i lærernes undervisningskvalitet, kan prestasjonsnedgangen sannsynligvis tilskrives samfunnsmessige endringer, som for eksempel pandemien.

Undervisningskvalitet

Undervisningskvalitet er en av de viktigste formbare faktorene som påvirker elevprestasjon (Klieme & Nilsen, 2022; Nilsen & Gustafsson, 2024; Seidel & Shavelson, 2007; Senden et al., 2023), elevmotivasjon (Fauth et al., 2014) og læreres tilfredshet med jobben (Lervåg, 2024; Wartenberg et al., 2023). Selv om det generelt sett ser ut til å være enighet blant forskerne i utdanningsfeltet om at undervisningskvalitet har en positiv sammenheng med prestasjoner (se f.eks., Baumert et al., 2010; Kaarstein & Nilsen, 2021; Senden et al., 2023), er det ofte blandede empiriske resultater for ulike dimensjoner og på tvers trinn, land og år (Senden et al., 2023; Teig & Luoto, 2024).

Det finnes flere ulike teorier om hva konseptet undervisningskvalitet innebærer (Charalambous et al., 2021), og med det også ulike forslag til hvordan det kan konseptualiseres (Seidel & Shavelson, 2007). Denne studien har tatt utgangspunkt i et rammeverk for undervisningskvalitet i matematikk som ofte benyttes i Europa (mest i tysktalende land og Norden) og i internasjonale studier som TIMSS. Dette rammeverket deler undervisningskvalitet i de tre grunnleggende dimensjonene (eng. *the three basic dimensions*; Praetorius et al., 2018); *klasseledelse* (*classroom management*), *kognitiv aktivering* (*cognitive activation*) og *støtte til elevene* (*student support*). Innen dimensjonen *støtte til elevene*, legger TIMSS-rammeverket spesielt stor vekt på klar og tydelig instruksjon. Forstyrrende atferd vektlegges i dimensjonen *klasseledelse* (Reynolds et al., 2021). I samsvar med TIMSS, legges det vekt på disse begrepene nedenfor.

I klasserom hvor det utøves god og effektiv klasseledelse settes det klare og konsekvente regler for og forventninger til elevenes oppførsel. Lærerne sørger for stabile rutiner samtidig som de følger med på og justerer elevenes atferd (Klieme et al., 2009). Effektiv klasseledelse kan støtte elevenes læring både sosialt, emosjonelt og faglig. Den kan også øke tilstedeværelsen for både elever og lærere samt forebygge eller avverge alvorlig aggresjon eller atferdsproblemer blant elever (Brouwers & Tomic, 2000; Oliver et al., 2011; Seidel & Shavelson, 2007). TIMSS måler ikke alle aspekter ved klasseledelse, men fokuserer på forstyrrende atferd.

Kognitiv aktivering tar for seg de undervisningspraksisene som engasjerer elevene i tenkning, for eksempel problemløsning eller diskusjon i hele klassen (Lipowsky et al., 2009; Praetorius et al., 2018). Kognitiv aktivering bygger på den konstruktivistiske ideen om at læring ikke kan påtvinges elever. Lærerne må legge til rette for og tilby læringsmuligheter ut fra elevenes behov slik at de kan konstruere sin egen kunnskap og kompetanse. Ifølge tidligere studier kan kognitiv aktivering ha en positiv sammenheng med prestasjoner (Baumert et al., 2010; Lipowsky et al., 2009). Dersom lærerne i tillegg støtter og utfordrer elevene til å jobbe seg gjennom krevende oppgaver, finner både Praetorius et al. (2018) og Senden et al. (2023) en positiv, signifikant sammenheng mellom kognitiv aktivering og elevprestasjoner.

Klar og tydelig instruksjon er en del av dimensjonen *støtte til elevene*, og handler om lærernes evne til å tilrettelegge og tilpasse undervisningen etter elevenes individuelle kognitive evner og behov (kalt *scaffolding* på engelsk, van de Pol et al., 2010). Tidligere forskning har vist at det er sammenheng mellom klar og tydelig instruksjon og elevenes prestasjoner i både matematikk og naturfag (Bergem et al., 2016; Kleickmann et al., 2020; Kaarstein & Nilsen, 2021; Senden et al., 2023)

Metode

Data og utvalg

For å svare på forskningsspørsmålene benyttes norske data for elever på 5. trinn fra TIMSS-undersøkelsene som ble gjennomført i 2019 og 2023. Data om elevers oppfatning av matematikklærernes undervisningskvalitet hentes fra den delen av elevenes spørreskjema som handler om hvordan de oppfatter matematikklærerens undervisning. I tillegg benyttes elevenes samlede

matematikkprestasjon. For mer informasjon om TIMSS-undersøkelsens design og gjennomføring, se kapittel 1.

På 5. trinn deltok henholdsvis 3 951 og 5 301 elever i 2019 og 2023. I begge undersøkelsene deltok det like mange jenter som gutter. Elevenes gjennomsnittsalder da undersøkelsen ble gjennomført var 10,7 år i 2019 og 10,8 i 2023 (se almanakk for elever, Fishbein et al., 2021, 2025).

Selv om det er et mål at spørreskjemaene i TIMSS-undersøkelsen er så konsistente som mulig fra en undersøkelse til den neste for å spore trender, må de også oppdateres for å reflektere samfunnsutviklingen (Reynolds et al., 2021). Dette resulterer noen ganger i at enkelte variabler eller utsagn fjernes og nye legges til. Når målet er å undersøke endringer over tid, er det avgjørende at utsagnene forblir uendret over tid for å sikre sammenliknbarhet. Derfor har bare de utsagnene som er identiske i TIMSS 2019 og TIMSS 2023 blitt inkludert i analysene til dette kapittelet. Selv om den norske prosjektgruppen har lagt inn tre nasjonale tilleggsutsagn om kognitiv aktivering i elevenes spørreskjemaer i 2019 og 2023, dekkes ikke hele begrepet slik det vanligvis er beskrevet i de tre grunnleggende dimensjonene (se f.eks. Praetorius et al., 2018). Det samme gjelder for klar og tydelig instruksjon, hvor seks av ni utsagn er videreført fra 2019 til 2023, og for forstyrrende atferd, hvor tre av seks utsagn er videreført.

I samsvar med tidligere forskning benyttes ikke lærernes, men elevenes rapportering om lærernes undervisningskvalitet i matematikk her (se f.eks. Fauth et al., 2014; Kyriakides et al., 2014; Wagner et al., 2016). I spørreskjemaet blir elevene bedt om å tenke på matematikkundervisningen de har fått og angi enten hvor enige de er i utsagnene eller hvor ofte utsagnene forekommer. Tabell 8.1 viser de inkluderte utsagnene, utsagnenes ID for de to syklusene av TIMSS-undersøkelsen og hvilken dimensjon av undervisningskvalitet utsagnene tilhører. Det er viktig å nevne her at vår måling av undervisningskvalitet, på lik linje med nesten all annen forskning på undervisningskvalitet, ikke dekker begrepet fullstendig. Vi måler kun noen aspekter av begrepet.

Elevenes matematikkprestasjoner måles ved hjelp av omtrent 180 oppgaver laget ut fra rammeverket for matematikk i TIMSS (se f.eks. Aldrich et al., 2024). Rammeverket definerer hva som skal måles (se f.eks. Philpot et al., 2021). Oppgavene blir fordelt i mindre oppgavesett, og hver enkelt elev får 36 minutter til å løse 20–26 oppgaver. Fordi hver elev kun får et utvalg av de 180 oppgavene, beregnes det i etterkant fem plausible verdier for hver elev (se f.eks. Bezirhan & von Davier, 2024).

Tabell 8.1. Utsagn om lærerens undervisningskvalitet i matematikk, inkludert i elevenes spørreskjema både i TIMSS 2019 og TIMSS 2023

ID i 2023	ID i 2019	Utsagn	Dimensjon	Skala
ASBM03B	ASBM03B	Læreren er lett å forstå	Klar og tydelig ²	Svært enig, Litt enig, Litt uenig, Svært uenig
ASBM03C	ASBM03C	Læreren har klare svar på spørsmålene mine	Klar og tydelig	
ASBM03D	ASBM03D	Læreren er flink til å forklare matematikk	Klar og tydelig	
ASBM03E	ASBM03E	Læreren gjør mange forskjellige ting for å hjelpe oss å lære	Klar og tydelig	
ASBM03F	ASBM03F	Læreren forklarer et emne om igjen når vi ikke forstår det	Klar og tydelig	
ASBM03L ¹	ASXM03J ¹	Læreren forteller hvordan nye ting henger sammen med det jeg allerede kan	Klar og tydelig	
ASBM03I	ASXM04A ¹	Læreren ber meg om å forklare svarene mine	Kognitiv aktivering	Svært enig, Litt enig, Litt uenig, Svært uenig
ASBM03M ¹	ASXM04B ¹	Elevene diskuterer løsninger på oppgaver med hverandre	Kognitiv aktivering	
ASBM03N ¹	ASXM04E ¹	Læreren oppmuntrer oss til å spørre om ting vi lurte på i matematikk	Kognitiv aktivering	
ASBM04A	ASBM04A	Elevene hører ikke på hva læreren sier	Forstyrrende atferd	Hver eller nesten hver time, Omtrent halvparten av timene, Noen timer, Aldri
ASBM04C	ASBM04D	Læreren må vente lenge på at elevene roer seg	Forstyrrende atferd	
ASBM04D	ASBM04E	Elevene avbryter læreren	Forstyrrende atferd	

Note. ¹ Nasjonalt tilleggsutsagn. ² Klar og tydelig = Klar og tydelig instruksjon.

Analysemetode

For å svare på forskningsspørsmålene benyttes regresjonsanalyse (Lumley, 2024b). Alle analysene er utført i analyseprogrammet *R* (R Core Team, 2021). For å ta hensyn til det komplekse designet til TIMSS, hvor elevene tilhører eller er gruppert i klasser og klassene tilhører skoler (se f.eks. Siegel & Foy, 2024), er *survey*- og *svylme*-pakkene i *R* tatt i bruk (Lumley, 2024b, 2024a). Disse pakkene inkluderer også alle vekter på elevnivå i analysene. For å ta høyde for manglende data (*missings*) benyttes *R*-pakken *mice* til *multiple imputation* (van Buuren, 2018), og elevenes vekter inkluderes i imputasjonsmodellene (Kim et al., 2006). *Multiple imputation* er en statistisk teknikk for å håndtere manglende data. Teknikken anbefales for håndtering av manglende data i det fleste tilfeller (van Buuren, 2018).

For å svare på det første forskningsspørsmålet ble elevenes oppfatning av lærernes undervisningskvalitet satt som avhengig variabel. De uavhengige variablene i denne regresjonsmodellen inkluderte målingstidspunkt (T_n , dvs. $T_1 = \text{TIMSS 2019}$ og $T_2 = \text{TIMSS 2023}$), elevenes kjønn og en interaksjon av målingstidspunkt og kjønn. Elevenes kjønn og interaksjonen mellom kjønn og målingstidspunkt er brukt som kontrollvariabler. For det andre forskningsspørsmålet ble ny regresjonsmodell kjørt med elevenes matematikkprestasjoner som den avhengige variabelen. Uavhengige variabler i denne regresjonsmodellen inkluderte målingstidspunkt og elevrapportert undervisningskvalitet. Elevenes sosioøkonomiske status (SØS)¹, alder, kjønn samt en interaksjon av (elevenes) kjønn og målingstidspunkt er brukt som kontrollvariabler. Variabelen som måler elevenes SØS er en samlevariabel som inkluderer elevenes angivelse av antall bøker i hjemmet, om elevene har egne rom, om de har nettforbindelsen hjemme og lengden på foreldres utdanning (Martin et al., 2020). I alle analysene ble de fem tilgjengelige plausible verdiene for elevprestasjonene (Bezirhan & von Davier, 2024) brukt og kombinert i samsvar med etablerte regler for regresjonsmodeller som bruker prestasjon som avhengig variabel («Rubin's rules», se f.eks. Laukaityte & Wiberg, 2017; Lorah, 2020; Rutkowski et al., 2013).

På bakgrunn av rammeverket fra Stapleton et al. (2016), defineres undervisningskvalitet på elevnivå siden det antas at hver elev har sin egen opplevelse av undervisningen. Videre, på grunn av nylige diskusjoner om validiteten til vanlige målingsmodeller (White, 2025), testes sammenhengen mellom hvert utsagn i tabell 1 med elevprestasjon i matematikk. De tre valgte dimensjonene (kognitiv aktivering, klar og tydelig instruksjon og forstyrrende atferd) ble sjekket på samme måte. Da alle sammenhengene var like, ble det definert én samlevariabel som inneholdt gjennomsnittet av alle utsagnene i tabell 8.1. Denne sammenslåingen av variabler med ulike svarkategorier kan gjøre det litt vanskeligere å tolke regresjonskoeffisientene, men kan i tillegg øke den eksterne validiteten av en skala (Marfeo et al., 2014). I dette tilfellet, fordi undervisningskvalitet handler om både frekvens og kvalitet av de ulike aktivitetene som foregår i undervisningen, kan samlevariabelen sies å representere elevens oppfatning av undervisningskvalitet på et mer overordnet nivå. For å unngå vanskeligheter ved å tolke regresjonskoeffisienter, presenterer vi resultater fra regresjonsmodeller ved bruk av simulasjon.

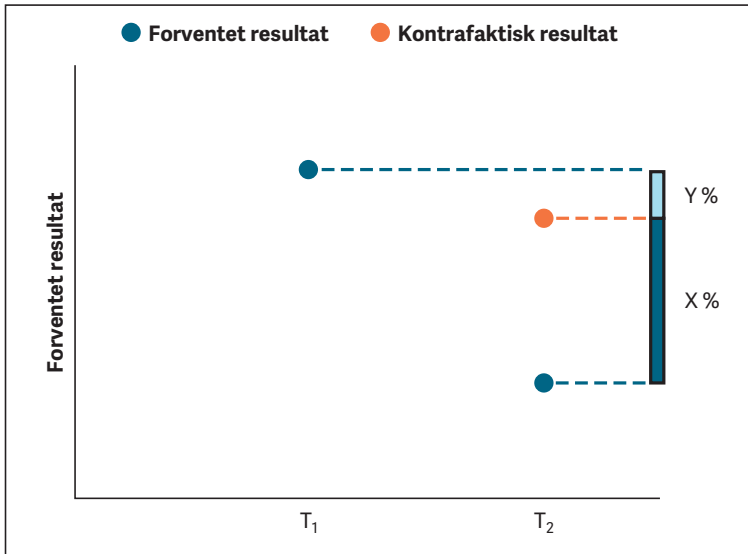
1 Navn på skala/variabel i datasettet er ASBGHRL.

For å presentere resultatene av regresjonsanalysene, er det gjennomført simuleringer slik King et al. (2000) anbefaler. Denne typen presentasjon tar utgangspunkt i at regresjonsmodellen er en «svart boks» hvor verdiene for de uavhengige variablene puttes inn og et estimat for den avhengige variabelen kommer ut. Estimatet for den avhengige variabelen kalles «forventet resultat».

I analysene for å svare på det første forskningsspørsmålet i dette kapitlet, estimeres to forventede resultater av elevenes oppfatning av undervisningskvalitet; ett for TIMSS 2019 og ett for TIMSS 2023. Deretter gjøres en enkel sammenlikning mellom de forventede resultatene av variabelen, for å se om verdien for de forventede resultatene er ulike i 2019 og 2023. Gjennom denne prosedyren sammenliknes gjennomsnittene av de forventede resultatene på tvers av år etter å ha justert for kjønn. Dersom de forventede resultatene er signifikant ulike, har elevenes oppfatning av undervisningskvaliteten endret seg. For tekniske detaljer om simuleringen og informasjon om målingsusikkerhet, se King et al. (2000).

For å svare på det andre forskningsspørsmålet, om hvor mye av nedgangen i prestasjoner som er relatert til en eventuell endring i undervisningskvalitet, ble det laget to scenarioer. For det første scenarioet ble det estimert et forventet resultat for prestasjonene i 2023 gitt den observerte endringen i undervisningskvalitet. Det samme er gjort for 2019. For det andre scenarioet ble det estimert et forventet resultat for prestasjonene i 2023 gitt at undervisningskvaliteten *ikke* har endret seg, altså at den var den samme som i 2019. Det siste estimatet gir verdien til et såkalt kontrafaktisk resultat fordi det viser resultatet av en simulering for regresjonsmodellen for en ikke-observert verdi. Figur 1 viser et eksempel på en presentasjon av en slik analyse. De blå punktene viser de forventede resultatene i målingstidspunkt 1 og 2 (T_1 og T_2). Det oransje punktet viser det forventede resultatet av prestasjoner i T_2 dersom undervisningskvaliteten *ikke* har endret seg fra T_1 til T_2 . Det oransje punktet viser altså verdien til det kontrafaktiske resultatet. Fordi den eneste forskjellen mellom disse to forventede resultatene er knyttet til undervisningskvalitet, må en eventuell forskjell mellom de forventede resultatene ha sammenheng med en endring i undervisningskvalitet.

Ut fra eksempelet illustrert i figur 8.1 ville resultatet av analysene antydnet at Y prosent av nedgangen (differansen mellom de forventede resultatene for prestasjonene i målingstidspunkt T_1 og T_2) i elevenes prestasjoner *ikke vil ha* sammenheng med endringer i lærernes undervisningskvalitet, mens X prosent av nedgangen i prestasjoner *vil ha* sammenheng med nedgangen



Figur 8.1. Eksempel på forventede resultater for målingstidspunkt T_1 og T_2 og kontrafaktisk resultat for T_2 .

i lærernes undervisningskvalitet. Det er derfor verdien av X som gir svaret på forskningsspørsmål 2.

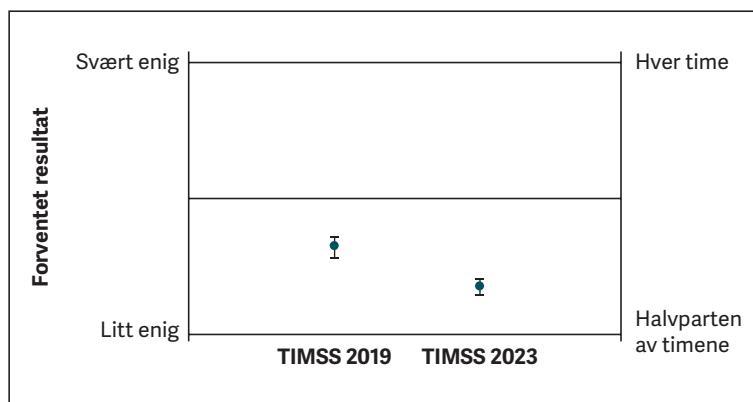
Resultater

I denne seksjonen viser vi resultatene fra de beskrevne simuleringene. Regresjonskoeffisientene fra analysene er vist i vedlegget til dette kapitlet. Regresjonskoeffisientene tilhørende analysene utført for forskningsspørsmål 1 vises i tabell 8.V1, mens regresjonskoeffisientene tilhørende forskningsspørsmål 2 vises i tabell 8.V2.

Endring i undervisningskvalitet fra TIMSS 2019 til TIMSS 2023

Resultatet av analysene utført for å svare på spørsmålet om lærernes undervisningskvalitet har endret seg fra TIMSS 2019 til TIMSS 2023, er vist i figur 2. Y-aksene i figur 2 viser de skalaene som er benyttet for utsagnene. Svaralternativene for ni av de tolv utsagnene som er analysert samlet i konstruktet undervisningskvalitet er Svært uenig, Litt uenig, Litt enig og Svært enig (Y-akse venstre side), mens de tre siste utsagnene besvares på

en skala inkluderende Aldri, Noen timer, Omtrent halvparten av timene og Hver eller nesten hver time (vist på høyre side i figuren). Figuren viser kun de to øverste nivåene i begge skalaene. Punktene viser den gjennomsnittlig verdien av undervisningskvalitet for de forventede resultatene med 95 prosent konfidensintervaller.



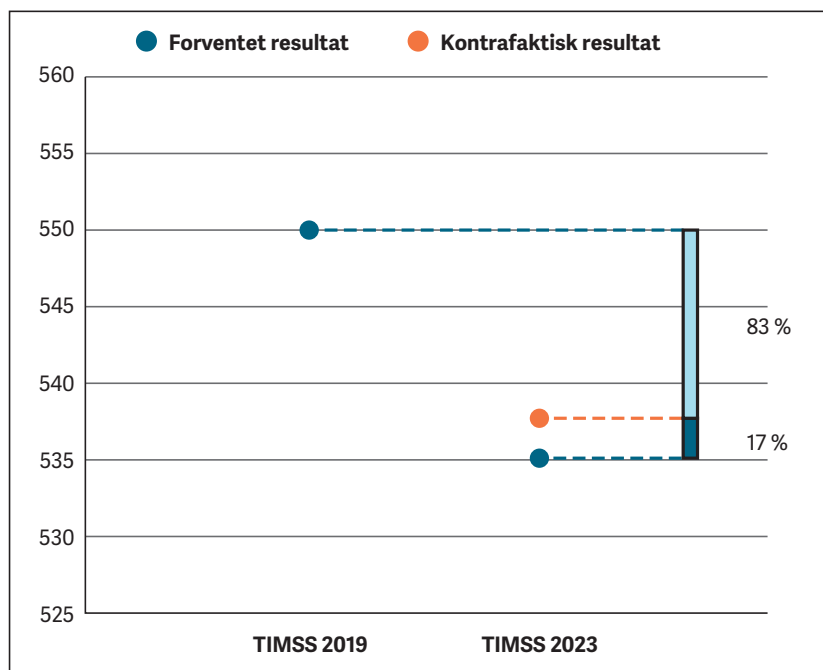
Figur 8.2. Forventede resultater for gjennomsnittet av variabelen lærernes undervisningskvalitet i TIMSS 2019 og TIMSS 2023. Y-aksene inkluderer kun det to høyeste svarkategoriene for å synliggjøre forskjell og usikkerhet.

I figur 8.2 vises endringen i lærernes undervisningskvalitet fra TIMSS 2019 til 2023. Nedgangen er på 0,15 poeng. Med tanke på at undervisningskvalitet er målt på en skala fra 0 til 3 poeng, er ikke nedgangen på 0,15 poeng stor. For å få en litt bedre forståelse av størrelsen på nedgangen, tenk deg at vi spør en klasse med tjue elever om undervisningskvaliteten til læreren, og vi gjør det ved to anledninger. En nedgang på 0,15 poeng oppstår dersom for eksempel tre av elevene gir lærerens undervisningskvalitet ett poeng mindre den andre gangen de blir spurt mens de resterende sytten elevene opprettholder sin første vurdering. Sett fra dette perspektivet er ikke endringen stor. Men ved å estimere et gjennomsnitt for hver klasse (summen av alle elevene) som deltar og sammenlikne disse, utgjør nedgangen på 0,15 omtrent ett standardavvik fra det totale gjennomsnittet for variabelen. Objektivt sett er nedgangen liten, men sett i relasjon til standardavviket er nedgangen relativt stor tatt i betraktning at det ikke er stor variasjon på tvers av klasserommene.

Selv om nedgangen i lærernes undervisningskvalitet bare er på 0,15 poeng, er det interessant å gå videre med det andre forskningsspørsmålet, hvor sammenhengen mellom endringen i undervisningskvalitet og elevenes nedgang i matematikkprestasjon utforskes.

Sammenheng mellom nedgangen i undervisningskvalitet og elevenes matematikkprestasjon

I figur 8.3 vises resultatet av analysene utført for å svare på det andre forskningsspørsmålet. Figuren inneholder de forventede (simulerte) matematikkresultatene for TIMSS 2019 og TIMSS 2023, samt det kontrafaktiske matematikkresultatet for TIMSS 2023. Den forventede poengverdien (549,7) for matematikkprestasjoner i TIMSS 2019 ligger, med 95 prosent sikkerhet, i intervallet fra 543,9 poeng til 555,8 poeng. For TIMSS 2023 er forventet verdi 535,2 poeng (529,9–542,2) og kontrafaktisk verdi på 537,7 poeng (532,7–543,4).



Figur 8.3. Forventede resultater for elevenes matematikkprestasjoner i TIMSS 2019 og TIMSS 2023 og det kontrafaktiske resultatet for elevenes matematikkprestasjoner i TIMSS 2023.

Ut fra avstanden mellom de to forventede verdiene i figur 8.3, er det beregnet at 83 prosent av elevenes nedgang i matematikkprestasjoner ikke har sammenheng med endringen i undervisningskvalitet, mens 17 prosent er relatert til endringen.

Diskusjon

Fordi undervisningskvalitet er viktig for elevenes læring samtidig som det er en faktor som skoler kan gjøre noe med eller kontrollere (Nilsen & Gustafsson, 2024; Seidel & Shavelson, 2007; Senden et al., 2023), har det vært et mål i dette kapittelet å undersøke om skolene har klart å opprettholde kvaliteten på undervisningen til tross for de endringene samfunnet har vært gjennom i de siste årene. Vi har derfor utforsket endringen i elevens oppfatning av undervisningskvalitet fra TIMSS 2019 til TIMSS 2023 og hvorvidt endringen har sammenheng med nedgangen i elevprestasjoner.

Som nevnt over, står den norske skolen overfor mange utfordringer, fra økende mobbing til flere og flere begrensninger for undervisningen (f.eks., elever som er sultne/trøtte; Kaarstein et al., 2024; Meld. St. 34 (2023–2024); NOU 2015: 8). Derfor er det ikke overraskende at resultatene fra denne studien viser at elevene oppfatter kvaliteten på undervisningen som noe lavere i 2023 enn i 2019. Men fordi nedgangen er liten, kan én tolkning av resultatene være at skolene og lærerne håndterer utfordringene relativt godt. Denne tolkningen støttes av resultatet fra regresjonsanalysene hvor sammenhengen mellom nedgang i undervisningskvalitet og nedgang i prestasjoner ble undersøkt. Disse analysene viste at 17 prosent av nedgangen i elevenes matematikkprestasjoner hang sammen med nedgangen i elevenes oppfatning av undervisningskvalitet, mens 83 prosent av prestasjonsnedgangen ikke hang sammen med nedgangen i undervisningskvalitet.

Den største andelen av elevenes nedgang i matematikkprestasjoner hadde ikke sammenheng med endringen i elevrapportert undervisningskvalitet. Det er fremdeles mulig at denne delen av nedgangen er relatert til undervisningskvalitet, men da ikke til elevens oppfatning av undervisningskvalitet som målt i TIMSS. Det er også mulig at de faktorene som har sammenheng med den største delen av nedgangen, er faktorer som skolene kan kontrollere, som for eksempel skoleledelse. Å finne den eller disse faktorene er en svært vanskelig, om ikke umulig, oppgave. Når det gjelder faktorer som skolene ikke kan gjøre noe med, for eksempel økende heterogenitet og stress, som nevnt tidligere (Nortvedt et al., 2020; Pascoe et al., 2020), er det viktig å vurdere hvordan andre deler av samfunnet kan bidra.

Med utgangspunkt i sammenhengen mellom nedgangen i undervisningskvalitet og prestasjoner, peker resultatene i denne studien på at det

vil være viktig for skoler å bidra med støtte og ressurser til lærere for å jobbe med undervisningskvaliteten. Dersom kausalitet antas, vil en høyere undervisningskvalitet kunne føre til høyere elevprestasjoner. Dersom denne antagelsen ikke holder mål, vil det fortsatt være viktig å vurdere tiltak fordi undervisningskvalitet har sterk sammenheng med lærernes tilfredshet med jobben (Wartenberg et al., 2023) og elevers motivasjon (Fauth et al., 2014). Derfor kan det å støtte lærere med å oppnå høy undervisningskvalitet, eller trives med jobben også hjelpe til med å beholde lærere i yrket – hvilket er et viktig politisk mål gitt problemer med rekruttering til læreryrket i Norge (Lervåg, 2024).

En begrensning for denne studien er knyttet til målingen av undervisningskvalitet i TIMSS. Fra TIMSS 2019 til TIMSS 2023 er det en forskjell på tjué utsagn i elevspørreskjemaet, hvilket er en stor endring. Dermed kan det påstås at TIMSS ikke har en stabil måling av undervisningskvalitet eller måler hele begrepet. I tillegg er TIMSS en såkalt tverrsnittsundersøkelse hvor det er mulig å si noe om sammenhenger mellom ulike faktorer, men ikke mulig å si noe om årsak og virkning.

Avsluttende kommentarer

Med utgangspunkt i nedgangen i de norske matematikkprestasjonene på 5. trinn fra TIMSS 2019 til TIMSS 2023 (Kaarstein et al., 2024), har vi i dette kapittelet sett spesielt på endring i elevenes oppfatning av undervisningskvalitet og sammenhengen mellom denne oppfatningen og prestasjonsendring. Resultatene av de gjennomførte analysene viste at elevenes oppfatning av undervisningskvalitet hadde endret seg, og at endringen hang sammen med en liten, men betydelig andel av nedgangen i prestasjoner.

På grunnlag av disse resultatene ser det ut til at skolene har et større behov for støtte og ressurser til å møte utfordringer i (skole)hverdagen enn de har for å jobbe med undervisningskvalitet. Denne studien kan ikke gi informasjon om hva slags støtte eller hvilke ressurser skoler og lærere trenger for å håndtere for eksempel det økende mangfoldet (Nortvedt et al., 2020), økende krav til digitale ferdigheter (Fernández-Batanero et al., 2021) og/eller økende stress og utrygghet for alle aktørene (Pascoe et al., 2020; Ramberg et al., 2020). Det trengs mer forskning for å bedre forstå hva skoler og lærere trenger for å øke/opprettholde høy undervisningskvalitet, hva lærerutdanningsinstitusjoner kan gjøre for å bedre støtte lærere og hvordan de ulike utfordringene bør håndteres og prioriteres.

Forfatterbiografier

Mark White er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han leder prosjektet *Teaching and Learning International Survey (TALIS) 2024* i Norge. Hans forskningsinteresse er blant annet undervisningskvalitet, lærerprofesjonalitet og kvantitative metoder.

Hege Kaarstein er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad i matematikkdidaktikk og har jobbet med TIMSS-undersøkelsen siden 2009, og vært prosjektleder fra 2018. I tillegg deltar hun i den internasjonale ekspertgruppen for matematikk i TIMSS. Hennes interesser er blant annet lærernes undervisningskompetanse i matematikk, elevenes motivasjon for faget og svar på åpne oppgaver.

Armin Jentsch er forsker ved det humanvitenskapelige fakultet ved Universitetet i Köln. Han har doktorgrad i utdanningsvitenskap og har jobbet med TIMSS-undersøkelsen siden 2019. Hans interesser er knyttet til undervisningskvalitet, måling av kompetanse og psykometriske metoder.

Referanser

- Aldrich, C. E. A., Bookbinder, A. & Khorramdel, L. (2024). Developing the TIMSS mathematics and science achievement instruments. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report (methods and procedures)* (s. 1.1–1.23). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs3063>
- Avvisati, F. & Givord, P. (2023). The learning gain over one school year among 15-year-olds: An international comparison based on PISA. *Labour Economics*, 84, Artikkel 102365. <https://doi.org/10.1016/j.labeco.2023.102365>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180. <https://doi.org/10.3102/0002831209345157>
- Bergem, O. K. (2016). Hovedresultater i matematikk. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015* (s. 22–43). Universitetsforlaget.
- Bergem, O. K., Nilsen, T. & Scherer, R. (2016). Undervisningskvalitet i matematikk. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015* (s. 120–136). Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016-08>
- Bezirhan, U. & von Davier, M. (2024). TIMSS achievement scaling methodology: Item response theory and population models. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report* (s. 11.1–11.15). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs8288>
- Brouwers, A. & Tomic, W. (2000). *Disruptive student behavior, perceived self-efficacy, and teacher burnout* [Konferansepresentasjon]. APA 2000 Convention, Washington, DC, United States.

- Charalambous, C. Y., Praetorius, A.-K., Sammons, P., Walkowiak, T., Jentsch, A., & Kyriakides, L. (2021). Working more collaboratively to better understand teaching and its quality: Challenges faced and possible solutions. *Studies in Educational Evaluation*, 71, 101092. <https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1016%2Fj.stueduc.2021.101092&data=05%7C02%7C%7C84678dd0e40341ac8a3008de073cd2ac%7C64393617747f4e8391bbb318e7eaa668%7C0%7C0%7C638956154983608982%7CUnknown%7CTWFpbGZsb3d8eyJFbXB0eU1hcGkiOnRydWUsIlYiOiIwLjAuMDAwMCIsIlAiOiJXaW4zMtIsIkFOljoifQ%3D%3D%7C0%7C%7C%7C%7C&sdata=kOqWhhKPAydxHV%2BXihh4M5G2oxO%2FjPLNjRqQGE8omkI%3D&reserved=0>
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Fernández-Batanero, J.-M., Román-Graván, P., Reyes-Rebollo, M.-M. & Montenegro-Rueda, M. (2021). Impact of educational technology on teacher stress and anxiety: A literature review. *International Journal of Environmental Research and Public Health*, 18(2), Artikkel 2. <https://doi.org/10.3390/ijerph18020548>
- Fishbein, B., Foy, P. & Yin, L. (2021). *TIMSS 2019 user guide for the international database* (2. utg.). Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/international-database/>
- Fishbein, B., Taneva, M. & Kowolik, K. (2025). *TIMSS 2023 user guide for the international database*. Boston College, TIMSS & PIRLS International Study Center. <https://timss2023.org/data>
- Jensen, F., Pettersen, A., Frønes, T. S., Eriksen, A., Løvgren, M. & Narvhus, E. K. (2023). *PISA 2022. Norske elevers kompetanse i matematikk, naturfag og lesing*. Cappelen Damm Akademisk. <https://cdforskning.no/cdf/catalog/book/205>
- Kim, J. K., Brick, J. M., Fuller, W. A. & Kalton, G. (2006). On the bias of the multiple-imputation variance estimator in survey sampling. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(3), 509–521. <https://doi.org/10.1111/j.1467-9868.2006.00546.x>
- King, G., Tomz, M. & Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science*, 44(2), 347–361. <https://doi.org/10.2307/2669316>
- Kleickmann, T., Steffensky, M. & Praetorius, A.-K. (2020). Quality of teaching in science education. more than three basic dimensions? I A.-K. Praetorius, J. Grünkorn & E. Klieme (Red.), *Empirische Forschung zu Unterrichtsqualität – Theoretische Grundfragen und quantitative Modellierungen* (s. 37–55). Beltz Juventa. <https://doi.org/10.25656/01:25862>
- Klieme, E. & Nilsen, T. (2022). Teaching quality and student outcomes in TIMSS and PISA. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education: Perspectives, methods and findings* (s. 1089–1134). Springer International Publishing. https://doi.org/10.1007/978-3-030-88178-8_37
- Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras study. I T. Janik & T. Seidel (Red.), *The power of video studies in investigating teaching and learning in the classroom* (s. 137–160). Waxmann.
- Kyriakides, L., Creemers, B. P. M., Panayiotou, A., Vanlaar, G., Pfeifer, M., Cankar, G. & McMahon, L. (2014). Using student ratings to measure quality of teaching in six European countries. *European Journal of Teacher Education*, 37(2), 125–143. <https://doi.org/10.1080/02619768.2014.882311>
- Kaarstein, H., Lehre, A.-C., Rohatgi, M. & Radišić, J. (2024). *TIMSS 2023. Kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Kaarstein, H. & Nilsen, T. (2021). Lærerkompetanse, undervisningskvalitet og naturfagprestasjoner fra TIMSS 2015 til TIMSS 2019. I T. Nilsen & H. Kaarstein (Red.), *Med blikket mot naturfag* (s. 183–206). Universitetsforlaget. <https://doi.org/10.18261/9788215045108-2021-08>
- Kaarstein, H., Radišić, J., Lehre, A.-C., Nilsen, T. & Bergem, O. K. (2020). *TIMSS 2019. Kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.

- Laukaityte, I. & Wiberg, M. (2017). Using plausible values in secondary analysis in large-scale assessments. *Communications in Statistics – Theory and Methods*, 46(22), 11341–11357. <https://doi.org/10.1080/03610926.2016.1267764>
- Lervåg, M.-L. (2024, 5. april). *Flere studenter – færre tar lærerutdanning*. Statistisk sentralbyrå. <https://www.ssb.no/utdanning/hoyere-utdanning/statistikk/studenter-i-universitets-og-hogskoleutdanning/artikler/flere-studenter-faerre-tar-laererutdanning>
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E. & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>
- Lorah, J. (2020). Estimating a multilevel model with complex survey data: Demonstration using TIMSS. *Journal of Modern Applied Statistical Methods*, 18(2), 2–20. <https://doi.org/10.22237/jmasm/1604190360>
- Lumley, T. (2024a). *Survey: Analysis of complex survey samples* (Versjon 4.4) [Programvare]. <https://CRAN.R-project.org/package=survey>
- Lumley, T. (2024b). *svylme: Linear mixed models for complex survey data* (Versjon 1.5-1) [Programvare]. <https://CRAN.R-project.org/package=svylme>
- Marfeo, E. E., Ni, P., Chan, L., Rasch, E. K. & Jette, A. M. (2014). Combining agreement and frequency rating scales to optimize psychometrics in measuring behavioral health functioning. *Journal of Clinical Epidemiology*, 67(7), 781–784. <https://doi.org/10.1016/j.jclinepi.2014.02.005>
- Martin, M. O., von Davier, M. & Mullis, I. V. S. (Red.). (2020). *Methods and procedures: TIMSS 2019 technical report*. TIMSS & PIRLS International Study Center.
- Meld. St. 34 (2023–2024). (2024). *En mer praktisk skole Bedre læring, motivasjon og trivsel på 5.–10. Trinn*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-34-20232024/id3052898/?ch=1>
- Nilsen, T. & Gustafsson, J.-E. (2024). Changes in teacher practices related to changes in student achievement. I N. Teig, T. Nilsen & K. Y. Hansen (Red.), *Effective and equitable teacher practice in mathematics and science education*. Springer.
- Nortvedt, G. A., Wiese, E., Brown, M., Burns, D., McNamara, G., O'Hara, J., Altrichter, H., Fellner, M., Herzog-Punzenberger, B., Nayir, F. & Taneri, P. O. (2020). Aiding culturally responsive assessment in schools in a globalising world. *Educational Assessment, Evaluation and Accountability*, 32(1), 5–27. <https://doi.org/10.1007/s11092-020-09316-w>
- NOU 2015: 8. (2015). *Fremtidens skole. Fornyelse av fag og kompetanser*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2015-8/id2417001/>
- NOU 2023: 1. (2023). *Kvalitetsvurdering og kvalitetsutvikling i skolen*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2023-1/id2961070/>
- Oliver, R. M., Wehby, J. H. & Reschly, D. J. (2011). Teacher classroom management practices: Effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews*, 7(1), 1–55. <https://doi.org/10.4073/csr.2011.4>
- Pascoe, M. C., Hettrick, S. E. & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth*, 25(1), 104–112. <https://doi.org/10.1080/02673843.2019.1596823>
- Philpot, R., Lindquist, M. & Aldrich, C. E. A. (2021). TIMSS 2023 mathematics framework. I V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM*, 50(3), 407–426. <https://doi.org/ggq4g4>
- R Core Team. (2021). *R: A language and environment for statistical computing* [Programvare]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramberg, J., Låftman, S. B., Åkerstedt, T. & Modin, B. (2020). Teacher stress and students' school well-being: The case of upper secondary schools in Stockholm. *Scandinavian Journal of Educational Research*, 64(6), 816–830. <https://doi.org/10.1080/00313831.2019.1623308>

- Reynolds, K. A., Mullis, I. V. S. & Martin, M. O. (2021). TIMSS 2023 context questionnaire framework. I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Rutkowski, L., von Davier, M. & Rutkowski, D. (Red.). (2013). *Handbook of international large-scale assessment background, technical issues, and methods of data analysis*. CRC Press. <https://doi.org/10.1201/b16061>
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Senden, B., Nilsen, T. & Teig, N. (2023). The validity of student ratings of teaching quality: Factorial structure, comparability, and the relation to achievement. *Studies in Educational Evaluation*, 78, Artikkel 101274. <https://doi.org/10.1016/j.stueduc.2023.101274>
- Siegel, P. & Foy, P. (2024). TIMSS sample design. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report* (s. 3.1–3.30). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs7952>
- Stapleton, L. M., Yang, J. S. & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics*, 41(5), 481–520. <https://doi.org/gfzcxg>
- Teig, N. & Luoto, J. M. (2024). Teaching quality and assessment practice: Trends over time and correlation with achievement. I N. Teig, T. Nilsen & K. Y. Hansen (Red.), *Effective and equitable teacher practice in mathematics and science education: A Nordic perspective across time and groups of students* (s. 155–186). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-49580-9_5
- Utdanningsdirektoratet. (2024). *Resultater fra nasjonale prøver for ungdomstrinnet 2024. Klar nedgang i regneferdighetene*. <https://www.udir.no/tall-og-forskning/statistikk/statistikk-grunnskole/analyser/2024/resultater-fra-nasjonale-prover-for-ungdomstrinnet-2024/>
- van Buuren, S. (2018). *Flexible imputation of missing data* (2. utg.). CRC Press. <https://stefvanbuuren.name/fimd/sec-evaluation.html>
- van de Pol, J., Volman, M. & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296. <https://doi.org/10.1007/s10648-010-9127-6>
- Vik, F. N., Nilsen, T. & Øverby, N. C. (2022a). Aspects of nutritional deficits and cognitive outcomes – triangulation across time and subject domains among students and teachers in TIMSS. *International Journal of Educational Development*, 89, Artikkel 102553. <https://doi.org/10.1016/j.ijedudev.2022.102553>
- Vik, F. N., Nilsen, T. & Øverby, N. C. (2022b). Associations between sleep deficit and academic achievement – triangulation across time and subject domains among students and teachers in TIMSS in Norway. *BMC Public Health*, 22(1), Artikkel 1790. <https://doi.org/10.1186/s12889-022-14161-1>
- Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B. & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108(5), 705–721. <https://doi.org/10.1037/edu0000075>
- Wang, M.-T. & Degol, J. (2014). Staying engaged: Knowledge and research needs in student engagement. *Child Development Perspectives*, 8(3), 137–143. <https://doi.org/10.1111/cdep.12073>
- Wartenberg, G., Aldrup, K., Grund, S. & Klusmann, U. (2023). Satisfied and high performing? A meta-analysis and systematic review of the correlates of teachers' job satisfaction. *Educational Psychology Review*, 35(4), Artikkel 114. <https://doi.org/10.1007/s10648-023-09831-4>
- White, M. (2025). A peculiarity in psychological measurement practices. *Psychological Methods*. <https://doi.org/10.1037/met0000731>

Vedlegg

Tabell 8.V1. Regresjonskoeffisienter (Beta) for regresjon med undervisningskvalitet som avhengig variabel

Variabel	Beta	sf	t	p
Skjæringspunkt (Intercept)	2,32	0,02	100,04	<0,001
År er 2023	-0,16	0,03	-5,87	<0,001
Elev er gutt	0,00	0,02	-0,15	0,45
Interaksjon av År er 2023 og Elev er gutt	0,04	0,02	1,75	0,09

Note. sf = standardfeil for Beta; t = t-statistikk; p = p-verdi

Tabell 8.V2. Regresjonskoeffisienter (Beta) for regresjon med elevprestasjon som avhengig variabel

Variabel	Beta	sf	t	p
Skjæringspunkt (Intercept)	510,01	18,08	28,21	<0,001
Læringsressurser i hjemmet (ASBGHRL)	14,63	2,46	5,94	0,005
Alder (ASDAGE)	20,93	3,69	5,68	0,005
Elev er gutt	6,07	3,85	1,57	0,11
År er 2023	-15,95	4,2	-3,77	0,02
Interaksjon av År er 2023 og Elev er gutt	7,77	4,58	1,70	0,09
Undervisningskvalitet	15,83	6,87	2,30	0,05

Note. sf = standardfeil for Beta; t = t-statistikk; p = p-verdi

KAPITTEL 9

Motivasjon, responstid og prestasjoner i matematikk blant norske elever på 9. trinn

Jelena Radišić og Andreas Pettersen

Institutt for Lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: I dette kapittelet undersøker vi sammenhengen mellom elevers responstid, motivasjon og prestasjoner. Studien tar utgangspunkt i data fra TIMSS 2023, med fokus på matematikk blant elever på 9. trinn i Norge. Ved å bruke en personsentrert tilnærming kombinerer vi profil- og mønsteranalyse for å finne ut om tiden elevene bruker på å svare på oppgaver – spesielt utfordrende oppgaver – kan si noe om deres motivasjon og prestasjoner i faget. Elever med høy selvtillit, eller som har en kombinasjon av selvtillit i og interesse for faget, viser mer optimale prestasjonsmønstre. Samtidig viser det seg at gjennomsnittlig responstid og svarmønstre blir mer komplekse når man ser på motivasjon og prestasjoner sammen. Når man tar hensyn til oppgavenes vanskegrad, blir mønstrene mindre tydelige.

Nøkkelord: motivasjon, prestasjoner, innsats, rask responsadferd, TIMSS 2023

Abstract

Motivation, Response Time, and Mathematics Achievement – A Latent Profile Analysis of Grade 9 Students in Norway

Building on prior research, this chapter explores the intricate relationship between students' response times, motivation, and performance. It utilises data from TIMSS 2023, concentrating on ninth-grade mathematics in Norway. Adopting a person-centred approach, it integrates profile and pattern analysis to ascertain whether

students' response times, especially for challenging items, provide valuable insights into their motivation and performance. Profiles with high self-confidence or combined interest and confidence demonstrated more optimal performance patterns. However, the average reaction time and response patterns become more complex and vary depending on motivation and achievement considered together, while task difficulty contributes to more unclear patterns.

Keywords: motivation, performance, effort, rapid response behaviour, TIMSS 2023

Introduksjon

Innenfor utdanningsforskning har data om elevers responstid fått mye oppmerksomhet de siste årene (Goldhammer et al., 2014; Lundgren & Eklöf, 2023; Tan & Hew, 2019). Begrepet responstid viser til hvor lang tid elever bruker på å svare på en oppgave. Den økte oppmerksomheten skyldes i stor grad at mange av prøvene og undersøkelsene som gjennomføres i skolen har blitt digitale, blant annet internasjonale storskalaundersøkelser som *Trends in International Mathematics and Science Study* (TIMSS) og *Programme for International Student Assessment* (PISA). Med digitale prøver er det mulig å automatisk samle inn prosessdata fra blant annet elevenes svarmønstre. Elevenes responstid på oppgaver kan gi innsikt i ulike aspekter av deres kompetanse (Hoogland & Tout, 2018; Stacey & Wiliam, 2014), deres atferd når de gjennomfører prøven, samt kognitive (Castro et al., 2019; Goldhammer et al., 2014; Ratcliff et al., 2015; Tan & Hew, 2019) og ikke-kognitive prosesser (Hofverberg et al., 2022; Lundgren & Eklöf, 2023; Soland et al., 2019). Dette gjør det mulig å utforske elevenes kunnskaper og forståelse og hvordan de engasjerer seg i prøvesituasjonen (Ivanova et al., 2020; Stacey & Wiliam, 2014).

Forskning på responstid har ofte blitt brukt for å få mer kunnskap om elevers faglige prestasjoner, men resultatene fra denne forskningen har vært varierende (Castro et al., 2019; Goldhammer et al., 2014; Ratcliff et al., 2015). De siste årene har det også vært noe forskning som har undersøkt hvorvidt responstid kan gi verdifull innsikt i elevers sosio-emosjonelle begreper, som for eksempel motivasjon (Hofverberg et al., 2022; Soland et al., 2019). I dette kapittelet vil vi utforske sammenhengen mellom elevenes responstid på matematikkoppgaver, tolket som en indikator på innsats, deres motivasjon for matematikk og deres matematikkprestasjoner. Vi bruker data fra TIMSS 2023 for 9. trinn i Norge. I analysene benytter vi en personorientert tilnærming (Ferguson et al., 2020), kombinert med mønsteranalyse (på engelsk *pattern analysis*) for å utforske om, og i hvilken grad, tiden elevene bruker på å besvare matematikkoppgaver henger sammen med deres motivasjon og prestasjoner.¹

1 KI ble benyttet til språkhjelp i artikkelen.

Bakgrunn

Motivasjon spiller en avgjørende rolle

Tidligere forskning viser at elevenes faglige motivasjon har en positiv sammenheng med deres prestasjoner i ulike fag, blant annet i matematikkfaget (Eccles & Wigfield, 2023; Marsh et al., 2012; Prast et al., 2018; Schöber et al., 2018; Skaalvik et al., 2015; Stankov & Lee, 2017; Wang, 2012; Watt et al., 2012). Studier har også vist at mange elever blir mindre motiverte etter hvert som de kommer lenger i utdanningsløpet (Lepper et al., 2005; Muenks et al., 2018) og at dette påvirker deres fremtidige utdannings- og yrkesvalg (Dowker et al., 2016; Toh & Watt, 2022; Wang, 2012; Watt et al., 2012). I tillegg er motivasjonsnivået også knyttet til nåværende engasjement og innsats, for eksempel under prøver (Wise & Kong, 2005).

En av de mest brukte teoriene for å forstå elevers motivasjon er *situated expectancy-value theory* (SEVT) (Eccles & Wigfield, 2023, 2024; Wigfield & Eccles, 2000). SEVT tar utgangspunkt i hvordan elevers valg påvirkes av en rekke faktorer, som for eksempel tidligere erfaringer, kulturell kontekst og oppfatninger. Disse faktorene er knyttet til elevens mestringsforventninger og de subjektive verdiene de tillegger arbeidsoppgaver, noe som igjen påvirker både prestasjoner og beslutninger, særlig når det gjelder karriereveier (Eccles & Wigfield, 2020, 2023, 2024; Wigfield & Eccles, 2000). Innenfor SEVT stammer mestringsforventning fra en persons oppfatning, som er formet av individets erfaringer, eller syn på egne evner til å lykkes med fremtidige arbeidsoppgaver, vanligvis målt gjennom begreper som selvtillit og selvoppfatning (Lee & Stankov, 2018). I vår studie benyttes selvtillit som en indikator for mestringsforventninger. I SEVT defineres fire subjektive oppgaveverdier: interesseverdi, nytteverdi, måloppnåelsesverdi og kostnad (Wigfield & Eccles, 2000; Eccles & Wigfield, 2020, 2023, 2024), der de to første er relevante for vår studie. Interesseverdi handler om den forventede gleden ved å engasjere seg i en oppgave, og ligger tett opp til ideene om interesse (Hidi & Renninger, 2006) og indre motivasjon (Ryan & Deci, 2016). Nytteverdi representerer oppgavens relevans for andre mål og er ofte knyttet til ytre motivasjon (Ryan & Deci, 2016). I denne studien holder vi oss innenfor rammen av SEVT-konseptualiseringen.

Resultater fra tidligere studier danner et noe sammensatt bilde når det kommer til sammenhenger mellom interesse, innsats, mestringsforventning og prestasjoner. Interesseverdi er mer gjennomgående knyttet til engasjement og innsats i bestemte situasjoner, mens både oppgaveverdier, spesielt interesseverdi, og mestringsforventning er knyttet til utdannings- og

yrkesvalg, for eksempel matematikk (Wang, 2012; Watt et al., 2012). Mestringsforventning er først og fremst knyttet til akademiske prestasjoner, hovedsakelig gjennom variabler som selvtillit og mestringsstro (Marsh et al., 2017; Wang, 2012). Studier har også vist at gutter generelt rapporterer om høyere selvtillit enn jenter (Nagy et al., 2010), og det er rapportert om forskjellige resultater for gutter og jenter når det gjelder ulike oppgaveverdier (Toh & Watt, 2022). Videre viser forskning at forholdet mellom motivasjonsvariabler og prestasjoner kan bli direkte, indirekte og gjensidig (Habók et al., 2020; Schöber et al., 2018).

I Norge har tidligere forskning vist en positiv sammenheng mellom prestasjoner i matematikk, interesseverdi, selvtillit og nytteverdi (Kaarstein et al., 2020, 2024; Kaarstein & Nilsen, 2016; Skaalvik et al., 2015). I tillegg er det observert kjønnsforskjeller i guttenes favør blant elever på 9. trinn i de to siste TIMSS-syklusene, spesielt når det gjelder interesseverdi og selvtillit (Kaarstein et al., 2020, 2024).

Prøveatferd, motivasjon og innsats

Motivasjon er et komplekst fenomen som ofte kommer til uttrykk i form av innsats, engasjement eller iherdighet for å påvirke læring og livsvalg. I denne studien er vi hovedsakelig opptatt av innsats, og dette analyseres innenfor rammen av oppgave- eller prøvespesifikk innsats. Tidligere forskning tyder på at elevers manglende motivasjon for å gjennomføre prøver, og konsekvensen dette har for elevenes innsats, utgjør et betydelig problem i situasjoner hvor resultatene ikke har noen direkte konsekvenser for elevene (Akhtar & Firdiyanti, 2023). Internasjonale storskalaundersøkelser, som for eksempel TIMSS, faller inn under denne kategorien. Lav motivasjon i prøvesituasjoner kan føre til manglende innsats blant elevene, noe som igjen kan svekke påliteligheten av resultatene fra undersøkelsene. Flere studier har vist at motiverte testdeltakere presterer bedre enn sine mindre motiverte medelever, selv når det er tatt høyde for deltakernes evnenivå (Baumert & Demmrich, 2001; Duckworth et al., 2011; Silm et al., 2019; Wise & DeMars, 2005). Wise og DeMars (2005) definerer innsats på prøver som det engasjementet og den energien elevene investerer i prøven for å oppnå best mulig resultat. Elevenes innsats måles vanligvis gjennom selvrappotering eller gjennom observasjon av atferd under prøver. Etter hvert som mange prøver har blitt digitale, har det blitt vanlig å loggføre og analysere elevenes interaksjoner med prøveoppgaver (Stacey & Wiliam, 2014). Dette har ført til en økning i forskning på elevenes synlige innsats på prøver (Hofverberg et al., 2022; Ivanova et al., 2020; Valdivia et al., 2023).

Responstid brukes ofte som et mål på elevenes atferd på prøven. Responstid refererer vanligvis til perioden fra elevene får tildelt en oppgave til de har fullført den – eller har gitt opp og gått videre. Dette begrepet er utbredt i mange fagområder, spesielt i yrkesvitenskap, hvor det har blitt grundig undersøkt for eksempel i studier av radarstimulusdeteksjon. Selv om digitaliseringen av prøver og undersøkelser har ført til mer forskning på responstid de senere årene, har dette fenomenet vært undersøkt innen utdanningsforskningen i lang tid. For eksempel undersøkte Blommers og Lindquist i 1944 elevers leseforståelse ved å måle tiden elevene brukte på å svare på oppgaver og hvilken sammenheng dette hadde med svarets nøyaktighet.

Når man bruker responstid for å undersøke elevers prøveatferd, er det flere faktorer som må tas i betraktning. Et veldig raskt svar på en oppgave kan indikere manglende innsats og engasjement, mens lengre tidsbruk kan tolkes som et tegn på større innsats og såkalt «løsningsatferd», det vil si at eleven nøye har lest, vurdert og svart på oppgaven (Goldhammer et al., 2014; Hofverberg et al., 2022; Schnipke & Scrams, 1997; Wise & DeMars, 2005). Samtidig viser forskning at høytpresterende elever ofte løser oppgaver raskere enn lavtpresterende, og at lang responstid dermed er en indikator på svakere ferdigheter. Goldhammer et al. (2014) fant også at sammenhengen mellom tidsbruk og prestasjoner varierer med både oppgavetype og elevenes ferdighetsnivå.

Responstidsinnsats (*response time effort* – RTE; Wise & Kong, 2005) er et mye brukt mål på hvor mye innsats elever legger i en prøve. Responstidsinnsats beregnes som andelen oppgaver der eleven viser løsningsatferd, altså har brukt tilstrekkelig tid på oppgaven til å lese, vurdere og forsøke å løse den. Løsningsatferd avgjøres ved å sammenligne elevens tidsbruk med en forhåndsdefinert tidsgrense for hver oppgave. Hvis eleven bruker mer tid enn tidsgrensen, antas det at oppgaven er forsøkt løst (løsningsatferd). Bruker eleven mindre tid, tolkes det som at eleven ikke har gjort et reelt forsøk på å løse oppgaven (ikke løsningsatferd, ofte kalt rask gjetting).

Tidstersklene kan fastsettes på ulike måter. En vanlig metode er visuell undersøkelse av fordelingen av tidsbruken på svarene på oppgaven (se for eksempel Wise, 2006), eller å bruke ti prosent av gjennomsnittlig responstid blant elevene (se for eksempel Wise & Ma, 2012). De ulike metodene for å fastsette tidsterskler kan ha forskjellige styrker og svakheter (Sahin & Colvin, 2020). I tillegg bør konteksten for prøven tas i betraktning når tidsgrensen fastsettes (Hofverberg et al., 2022). Feilklassifisering kan oppstå dersom tidsgrensen settes for lavt, for eksempel ved at elever som gjetter

raskt kan bli feilaktig vurdert til å ha forsøkt å løse oppgaven (løsningsatferd). Dersom tidsgrensen settes for høyt, risikerer man at elever som har løst oppgaven raskt og effektivt blir feilklassifisert som å ikke ha gjort et reelt forsøk (ikke løsningsatferd).

Til tross for slike svakheter, har målinger av responstidsinnsats vist seg å være gode indikatorer på elevers innsats på prøver (Wise & Kong, 2005). Tidligere forskning har generelt funnet en positiv sammenheng mellom innsats og prestasjon på prøver (Silm et al., 2019, 2020; Wise & DeMars, 2005). Samtidig finnes det få studier som undersøker hvordan ulike aspekter av motivasjon henger sammen med innsats, og resultatene fra disse studiene er ikke alltid entydige. For eksempel fant Akhtar og Firdiyanti (2023) at innsats var sterkere knyttet til elevenes forventninger om å lykkes med å løse oppgavene enn hvor verdifullt de mente det ville være å lykkes. Pools og Monseur (2021) analyserte data fra PISA 2015 og fant at interesse og glede for naturfag hadde en positiv sammenheng med innsats på prøven, mens elevenes selvtillit i naturfag hadde en svakere sammenheng når man kontrollerte for bakgrunnsfaktorer. Hofverberg et al. (2022) kom til lignende funn og fant at elever med profiler som indikerte mest positiv motivasjon og avanserte epistemiske oppfatninger, både presterte bedre og anstrengte seg mest. Samtidig tyder enkelte rapporter på at elever generelt viser høyt engasjement, selv på prøver uten direkte konsekvenser for elevene (*low-stakes*-prøver), selv om det finnes synlige forskjeller mellom demografiske grupper (Valdivia et al., 2023).

Denne studien

I denne studien undersøker vi sammenhengen mellom elevers innsats, motivasjon og prestasjoner i matematikk. Studien bygger på data fra TIMSS 2023 for 9. trinn i Norge. Ved hjelp av en personorientert tilnærming (Ferguson et al., 2020) kombinerer vi profilanalyse med mønsteranalyse av elevenes svaratferd. Hensikten er å undersøke om elevenes responstid – også med hensyn til oppgavenes vanskelighetsgrad – gir nyttig innsikt i deres motivasjon for matematikk og deres faglige prestasjoner. Ut fra dette ønsker vi å undersøke de to følgende forskningsspørsmålene:

- 1 Hvilke motivasjonsprofiler i matematikk finner vi blant norske elever i TIMSS 2023, og i hvilken grad varierer fordelingen av gutter og jenter i de ulike profilene?

Denne delen av analysen bygger på SEVT-rammeverket (Eccles & Wigfield, 2020, 2023, 2024; Wigfield & Eccles, 2000) og lignende arbeid (Hofverberg et al., 2022; Radišić & Jensen, 2021). Ut fra dette forventer vi å finne profiler med høy og lav interesse for matematikk, og at de ulike profilene vil ha elever som i ulik grad ser nytteverdien av matematikk og har en blanding av høy og lav selvtillit. Videre forventer vi at gutter vil være overrepresentert i profiler som viser høyere interesseverdi og faglig selvtillit (Kaarstein et al., 2020, 2024).

- 2 Hva er sammenhengen mellom motivasjonsprofilene og elevenes responstid og prestasjoner i matematikk? Og i hvilken grad påvirkes disse sammenhengene av oppgavenes vanskelighetsgrad?

Basert på resultater fra tidligere forskning (Silm et al., 2019, 2020; Wise & DeMars, 2005), forventer vi at vi vil finne en positiv sammenheng mellom elevenes innsats på prøven og deres prestasjoner. Når det kommer til sammenhengen mellom motivasjonsprofilene, innsats og prestasjoner, finnes det få studier i matematikk (Hofverberg et al., 2022). Resultatene fra studiene som finnes, gir et blandet bilde (Akhtar & Firdiyanti, 2023; Pools & Monseur, 2021). Vi har derfor ingen bestemt hypotese om disse sammenhengene, men forventer å finne en positiv sammenheng mellom høytpresterende elever og motivasjonsprofiler med høyt engasjement (Valdivia et al., 2023).

Metode

I denne delen beskriver vi utvalget i studien, variablene vi har brukt og hvordan analysene er gjennomført for å svare på forskningsspørsmålene.

Utvalg

I denne studien har vi benyttet oss av data fra de norske elevene som deltok i TIMSS 2023 på ungdomstrinnet. Dette datasettet består av besvarelser fra et representativt utvalg elever på 9. trinn ($n = 6\,234$, 47,2 prosent jenter). Utvalgene i TIMSS trekkes tilfeldig i to steg: først trekkes skolene, deretter trekkes opptil to hele klasser fra hver skole (Siegel & Foy, 2024). Mer informasjon om utvalget finnes i kapittel 1 i denne boka.

Variabler og konstrukter

I TIMSS 2023 gjennomførte elevene først en 90-minutters faglig prøve i matematikk og naturfag, og deretter et spørreskjema om deres bakgrunn,

holdninger og oppfatninger om skolen og læringsmiljøet knyttet til naturfag og matematikk. Den faglige prøven benytter et rotert blokkdesign (Aldrich et al., 2024), som vil si at oppgavene er satt sammen i «blokker» og hver elev får noen slike blokker med oppgaver. En slik utforming innebærer at hver elev bare svarer på et utvalg av matematikkoppgavene i matematikkprøven. I det påfølgende spørreskjemaet svarer alle elevene på alle de samme spørsmålene (Reynolds, 2024).

I denne studien har vi kun brukt data fra oppgavene i matematikk og spørsmålene knyttet til matematikk i spørreskjemaet. Tabell 9.1 gir en oversikt over variablene og konstruktene som er benyttet i analysene. En mer detaljert oversikt over spørsmålene innenfor hvert konstrukt finnes i vedlegg 9.1.

Tabell 9.1. Variabler og konstrukter brukt i studien

Variabel/konstrukt	Beskrivelse	Gjennomsnitt (standardfeil) og andel missing (%)	Reliabilitet
	Prestasjoner		
Prestasjoner i matematikk	Elevenes skår på matematikkprøven i TIMSS, beregnet gjennom plausible verdier*	503 (2,4), 0 %	/
	Demografiske kategorier		
Kjønn	Kategorisk, som gutt og jente (47,2% jenter)	/, 0%	/
	Konstrukter knyttet til motivasjon		
Liker å lære matematikk	Samleverdi. Eksempel: «Jeg liker å lære matematikk.»	9,17 (0,24), 6,7 %	0,947
Selvtillit i matematikk	Samleverdi. Eksempel: «Matematikk er lett for meg.»	9,94 (0,28), 9,4 %	0,890
Verdsetter matematikk	Samleverdi. Eksempel: «Jeg vil gjerne ha en jobb som innebærer bruk av matematikk.»	8,81 (0,26), 9,9 %	0,905
	Variabler knyttet til responstid		
Gjennomsnittlig tidsbruk på oppgavene (AveTT)	Avvik fra gjennomsnittlig tidsbruk på oppgavene	0,014 (0,005)	/
Responstidsinnsats (RTE)	Andelen oppgaver som elevene besvarte med løsningsatferd	0,83 (0,002)	/

Merk: * Alle fem plausible verdier ble brukt i alle analysene. Alle konstruktene ble brukt i tråd med TIMSS 2023-rammeverket med rapportert statistikk, med unntak av responstidskonstruktene som vi har utviklet for denne studien.

Elevenes motivasjon for matematikk måles gjennom tre konstrukter (se tabell 9.1). Konstruktet «Liker å lære matematikk» legger vekt på de indre aspektene ved motivasjon, og i hvilken grad elevene trives med og liker matematikk, eller om de oppfatter matematikk som kjedelig. Konstruktet «Verdsetter matematikk» er knyttet til motivasjonens nytteverdi og i hvilken grad elevene opplever fordeler ved å studere matematikk eller gjøre karriere innenfor matematikkfaget. Det siste konstruktet, «Selvtillit i matematikk», er knyttet til elevens egen vurdering av sine evner til å håndtere matematisk innhold (von Davier et al., 2024).

To variabler ble utviklet som en indikator på elevenes innsats, basert på deres responstid på oppgavene: gjennomsnittlig tid på oppgavene (AveTT = *average time on task*) og responstidsinnsats (RTE = *response time effort*).

Gjennomsnittlig tid på oppgavene (AveTT) måler hvordan en elevs tidsbruk skiller seg fra gjennomsnittlig tidsbruk på oppgavene. Beregningen av AveTT ble gjort i tre trinn:

- 1 Først ble gjennomsnittlig tidsbruk og standardavvik beregnet for hver oppgave, basert på alle elevene som hadde fått tildelt oppgaven.
- 2 Deretter ble hver elevs tidsbruk på hver oppgave standardisert ved hjelp av z-skårer, og beregnet som avvik fra den gjennomsnittlige tidsbruken for den aktuelle oppgaven.
- 3 Til slutt ble elevens AveTT beregnet som gjennomsnittet av disse avvikene på tvers av alle oppgavene eleven hadde fått tildelt.

Ved å bruke standardiserte avvik blir elevenes AveTT-verdier sammenlignbare, selv om de har jobbet med ulike oppgaver. En AveTT-verdi på 0,1 betyr at eleven i gjennomsnitt har brukt 0,1 standardavvik mer tid per oppgave enn gjennomsnittet. Tilsvarende vil en verdi på -0,1 bety at eleven har brukt 0,1 standardavvik kortere tid enn gjennomsnittet per oppgave. En positiv verdi på AveTT betyr at eleven brukte lengre tid enn gjennomsnittet, mens en negativ verdi betyr kortere tid. I vår studie varierte AveTT fra -1,01 til 14,39., hvor nesten halvparten av utvalget hadde en negativ verdi i AveTT og bare 4 elever (0,1 %) oppnådde verdier over 6,73.

Responstidsinnsats (RTE) måler hvor ofte en elev bruker nok tid på en oppgave til at det er sannsynlig at hen har forsøkt å lese og løse den, det vil si viser *løsningsatferd*. For å avgjøre dette, satte vi en tidsgrense for hver oppgave på ti prosent av gjennomsnittlig tidsbruk (Hofverberg et al., 2022; Valdivia et al., 2023; Wise, 2015). Det vil si at hvis en elev brukte minst ti

prosent av den gjennomsnittlige tiden alle elever brukte på oppgaven, ble det regnet som et forsøk på å løse den (løsningsatferd). Denne tidsgrensen varierte fra 1,66 til 19,14 sekunder, med et gjennomsnitt på 7,18 sekunder, for de ulike oppgavene. RTE ble beregnet som andelen oppgaver der eleven nådde denne grensen. Siden RTE-målet hadde lav variasjon og var skjevt fordelt (for eksempel hadde 76 prosent av elevene RTE på minst 0,8), valgte vi å gjøre RTE om til en dikotom variabel. Elever med RTE under 0,80, det vil si viste løsningsatferd på mindre enn 80 prosent av oppgavene, ble klassifisert som å ha lav innsats på prøven.

Kjønn ble registrert som en dikotom variabel (gutt/jente).

Analyser

Analysene i studien ble gjennomført i to steg. I det første steget ble det utført en latent profilanalyse (LPA) basert på de tre motivasjonskonstruksjonene «Liker å lære matematikk», «Selvtillit i matematikk» og «Verdsetter matematikk». Latent profilanalyse (LPA) er en metode for blanding av latente variabler som gjør det mulig å teste tilpasning og signifikans av ulike latente profiler blant individer i et datasett (Bergman & Trost, 2006; Collins & Lanza, 2010). Denne metoden bidrar til å avgjøre om det finnes ett enkelt mønster eller om en kombinasjon av flere latente profiler er til stede i en fordeling. I vår studie er disse latente profilene grupper av elever med like verdier for de tre motivasjonskonstruksjonene, som vi omtaler som motivasjonsprofiler. Modeller med to til ni latente klasser ($k = 2-9$) ble evaluert for å avgjøre hvor mange motivasjonsprofiler vi fant blant de norske elevene på 9. trinn. Hver modell ble estimert ved hjelp av 5 000 sett med tilfeldige startverdier med 100 iterasjoner, og de 200 beste løsningene ble beholdt til den endelige optimeringsfasen. Valget av den endelige løsningen (det endelige settet med profiler) tok hensyn til entropiindeksens cut-off (Geiser, 2013) sammen med en kombinasjon av *bootstrapped likelihood ratio test* (BLRT), Vuong-Lo-Mendell-Rubin *likelihood ratio test* (VL-LRT) og Lo-Mendell-Rubin justert LRT-test (LMR), som anbefalt av Asparouhov og Muthén (2012). Til slutt ble den endelige modellen validert en gang til i henhold til Geisers (2013) retningslinjer for bruk av den beste *log-likelihood*-verdien. Kjønnforskjeller mellom profilene ble undersøkt med R3STEP-oppsjonen (Asparouhov & Muthén, 2020) i Mplus og deskriptive analyser i SPSS.

I det neste steget undersøkte vi sammenhengen mellom motivasjonsprofilene og elevenes prestasjoner og innsats (AveTT og RTE). Her brukte

vi BCH-metoden i Mplus (Asparouhov & Muthén, 2020; Bolck et al., 2004), kombinert med deskriptive analyser i SPSS.

Resultater

Resultatene presenteres i to deler. Først beskriver vi hvordan motivasjonsprofilene til elevene ble bestemt og presenterer de seks profilene vi fant blant elevene (forskningsspørsmål 1). Deretter beskriver vi resultatene fra analysen av sammenhengene mellom motivasjonsprofilene, elevenes responsmønstre og elevenes prestasjoner i matematikk (forskningsspørsmål 2).

Elevenes motivasjonsprofiler

I tråd med anbefalte praksis for LPA (Geiser, 2013) undersøkte vi en serie løsninger (fra 2 til 9 profiler). Under denne analysen tok vi hensyn til implikasjonene av det tidligere nevnte teoretiske rammeverket som ligger til grunn for denne undersøkelsen og studier som bruker en lignende tilnærming, hvor vi spesielt har vektlagt den studien som er gjennomført i Norge (Radišić & Jensen, 2021). I tillegg har vi tatt hensyn til profilegenskapene til løsningene og deres tolkbarhet i lys av kommende mønsteranalyser.

Tabell 9.2 viser tilpasningsstatistikken (på engelsk *fit statistics*) for de ulike analyseløsningene. Mens løsningene med åtte og ni profiler ga resultater uten å replikere *log-likelihood*-verdien, pekte løsningen med syv profiler konsekvent på en problematisk BLRT-verdi. Løsningene for tre til seks profiler er statistisk gyldige (se LMR, BLRT og VL-LRT verdier). Blant disse viste alternativet med seks profiler de beste resultatene, med en entropiverdi på 0,853. Denne løsningen ga seks distinkte profiler, som reflekterer nyanserte forskjeller i elevens motivasjon knyttet til interesse, selvtillit og nytteverdi. Basert på de videre analysene anser vi løsningen med seks profiler som den mest detaljerte og passende fremstillingen av elevenes motivasjonsmønstre når det gjelder matematikk. Gjennomsnittlige latente klassesannsynligheter ligger også i tilfredsstillende område (0,836 til 0,960). Den foreslåtte seksklassemodellen ble validert ved å følge Geisers (2013) retningslinjer for optimal repetisjon av *log-likelihood*-verdien.

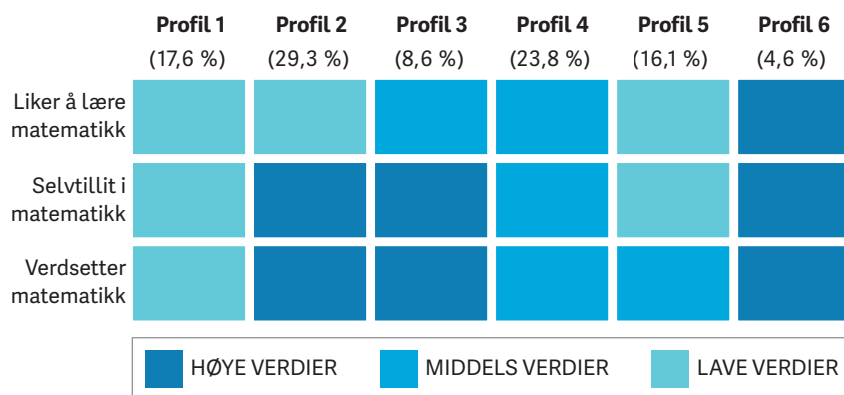
Figur 9.1 viser en oversikt over verdiene for de tre motivasjonskonstruktene for de seks profilene. Profil 1 kjennetegnes ved lave verdier på alle de tre konstruktene (*Liker å lære matematikk*, *Selvtillit i matematikk* og *Verdsetter matematikk*). Denne profilen, som vi anser som den minst

Tabell 9.2. Oversikt over evaluerte løsninger. Løsningen med seks profiler ble vurdert som best og er uthevet med grå bakgrunn

#No.	Log likelihood	#fp	AIC	BIC	SABIC	LMR	BLRT	VL-LRT	Entropi	Minste gruppe-frekvens
2	-34695,969	10	69411,94	69478,870	69447,09	/	/	/	0,637	46,00 %
3	-33909,619	14	67847,24	67940,942	67896,45	0	0	0	0,805	13,90 %
4	-33659,807	18	67355,61	67476,091	67418,89	0	0	0	0,789	8,30 %
5	-33482,926	22	67009,85	67157,102	67087,19	0	0	0	0,781	8,40 %
6	-33306,119	26	66664,24	66838,261	66755,64	0	0	0	0,853	4,60 %
7	-33145,004	30	66350,01	66550,803	66455,47	0	n/a	0	0,840	4,60 %
8	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
9	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note: #No = antall profiler; #fp = frihetsgrader; BIC = Bayesian information criterion; AIC = Akaike's information criterion; SABIC = sample-size adjusted BIC; LMR = Lo-Mendell-Rubin adjusted likelihood ratio test; BLRT = parametric bootstrapped likelihood ratio test; VL-LRT = Vuong-Lo-Mendell-Rubin likelihood ratio test

optimale, gjelder for 17,6 prosent av elevene. Profil 2 er den mest utbredte, med 29,3 prosent av elevene. Dette er elever som har moderat selvtillit og verdsetter matematikk til en viss grad, men som ikke liker å lære matematikk. Profil 3 (8,6 prosent av elevene) har de høyeste verdiene for selvtillit og nytteverdi, og har middels høy interesse for matematikk, mens profil 4 (23,8 prosent av elevene) har middels høye verdier på alle de tre konstruktene. Profil 5, hvor vi finner 16,1 prosent av elevene, kjennetegnes ved svært lav interesse for og selvtillit i matematikk, men hvor faget til en viss grad verdsettes. Profil 6 er den mest optimale profilen ved at den har høye verdier på alle de tre konstruktene. Dette er også den minst utbredte profilen, som vi kun finner blant 4,6 prosent av de norske elevene på 9. trinn.



Figur 9.1. Variasjonen i de isolerte motivasjonsprofilene

Note: Høye, middels og lave verdier på de tre motivasjonskonstruktene tilsvare grenseverdiene som TIMSS har angitt for hver skala. (von Davier et al., 2024).

I noen av profilene finner vi kjønnsforskjeller når det kommer til fordelingen av gutter og jenter (tabell 3). I profil 4 og 6 er det en overvekt av gutter, mens profil 5 domineres av jenter ($\chi^2 = 69,517$, $df = 10$, $p < 0,001$). Det er altså flere gutter enn jenter med høy interesse, selvtillit og verdsettelse av matematikk (profil 6), mens i den minst optimale profilen (profil 1) finner vi ingen signifikante forskjeller i andelen gutter og jenter. I den nest minst optimale profilen (profil 5) finner vi derimot en overvekt av jenter. Interessant nok kan det se ut som relativt tydelige kjønnsforskjeller i profil 3, men disse er ikke statistisk signifikante.

Tabell 9.3. Fordeling av gutter og jenter for de seks motivasjonsprofilene

		Profil 1	Profil 2	Profil 3	Profil 4	Profil 5	Profil 6
Gutter	% kjønn	49,4 %	51,3 %	57,4 %	55,4 %	43,3 %	64,8 %
	% profil	16,0 %	29,4 %	8,9 %	25,7 %	14,0 %	6,0 %
	SR	-1,0	-0,3	1,7	1,9	-3,7	3,1
Jenter	% kjønn	50,5%	48,7 %	42,6 %	44,6 %	56,7 %	35,2 %
	% profil	17,5%	29,9 %	7,1 %	22,2 %	19,8 %	3,5 %
	SR	1,0	0,3	-1,8	-2,0	3,9	-3,2

Merk: Den første %-linjen for kjønn angir prosentandelen innenfor den spesifikke profilen. Den andre %-linjen for kjønn betyr prosentandelen innenfor kjønnskategorien. SR står for standardisert residual; verdier over 1,96 i absolutt verdi er statistisk signifikante.

Motivasjonsprofiler og prestasjoner i matematikk

Resultatene viser at elever i motivasjonsprofil 3 har de høyeste gjennomsnittsprestasjonene i matematikk, etterfulgt av elever i profil 4 og 6 (tabell 9.4). Det er ingen statistisk signifikant forskjell mellom gjennomsnittsprestasjonene for profil 4 og 6. Det er også verdt å merke seg at profil 3 har høyest verdier for både selvtillit og nytteverdi, men noe lavere på interesse. Til sammenligning har elevene i profil 6 jevnt over høye verdier på alle de tre konstruktene, men noe lavere verdi på selvtillit enn profil 3. Vi ser også at de to profilene hvor elevene har lavest prestasjoner er de som kjennetegnes med lav selvtillit (profil 1 og 5). Dette kan tyde på at selvtillit i faget har en sterkere sammenheng med matematikkprestasjoner enn interesse og nytteverdi.

Tabell 9.4. Gjennomsnittsprestasjoner i matematikk for elever i de ulike motivasjonsprofilene

	Profil 1	Profil 2	Profil 3	Profil 4	Profil 5	Profil 6
Prestasjoner i matematikk	452,872 sig ^{2,3,4,5,6}	502,254 sig ^{1,3,4,5,6}	561,915 sig ^{1,2,4,5,6}	533,534 sig ^{1,2,3,5}	482,407 sig ^{1,2,3,4,6}	529,542 sig ^{1,2,3,5}

Merk: Tabellen viser kun verdier for den første plausible verdien. Se vedlegg 9.2 for en sammenligning av gjennomsnittlige matematikkresultater for alle plausible verdier på tvers av profilene. Indeks i nedre rad indikerer statistisk signifikans mellom profilene ($p < 0,01$).

Motivasjonsprofiler, prestasjoner og innsats

Analysene viser tydelige forskjeller mellom profilene når det kommer til hvor lang tid elevene brukte på oppgavene. Elevene i profil 3 brukte i gjennomsnitt mest tid (høyest AveTT), etterfulgt av elevene i profil 4 og profil

2. Når vi ser dette i sammenheng med elevenes prestasjoner, ser vi at elever i *profil 3*, som har høyest AveTT, også er de høyest presterende elevene. Elevene i profil 4, som har nest høyest prestasjoner i matematikk, har også relativt høy tidsbruk (tabell 9.5).

Elevene i profil 2 har også høyere responstid enn gjennomsnittet, som tyder på at de gjør en betydelig innsats, men ikke lykkes like godt som elevene i profil 4 og 3 med å løse oppgavene. Likevel verdsetter disse elevene matematikk og har høy selvtillit i faget. Profil 6, som kjennetegnes ved at elevene både liker å lære matematikk, har selvtillit og verdsetter faget, skiller seg fra de andre profilene når det kommer til innsats og prestasjoner. Disse elevene har relativt lav tidsbruk på oppgavene, men presterer likevel relativt høyt. Dette tyder på at de i stor grad lykkes med å løse matematikkoppgavene raskt og effektivt.

Tabell 9.5. Gjennomsnittlig tidsbruk på oppgavene (AveTT) for elever i de ulike profilene

	Profil 1	Profil 2	Profil 3	Profil 4	Profil 5	Profil 6
AveTT	-0,114 sig ^{2,3,4,5,6}	0,051 sig ^{1,3}	0,105 sig ^{1,2,(4),5,6}	0,055 sig ^{1,(3),5}	0,014 sig ^{1,3,4}	0,004 sig ^{1,3}

Merk: Indeks i nedre rad indikerer statistisk signifikans mellom profilene ($p < 0,01$). Tall i parentes, f.eks. (4), indikerer $p < 0,05$.

For å bedre forstå sammenhengen mellom motivasjonsprofilene, prestasjoner og innsats, brukte vi også responstidsinnsats (RTE) som mål på innsats. RTE ble beregnet som en dikotom variabel, hvor elevene som brukte mer tid enn grenseverdien på minst 80 prosent av oppgavene ble klassifisert til å vise løsningsatferd. Resultatene viser en statistisk signifikant forskjell mellom profilene ($\chi^2 = 101,692$, $df = 5$, $p < 0,001$). I profil 1 er det en signifikant høyere andel elever som viser ingen løsningsatferd enn i de andre profilene ($SR > \mp 1,96$). I profil 2, 3 og 4 er det motsatt, det vil si en lavere andel elever som viser ingen løsningsatferd sammenlignet med de andre profilene. For profilene 5 og 6 er det ingen statistisk signifikante forskjeller i andelen elever med løsningsatferd sammenlignet med de andre profilene.

Resultatene bekrefter i stor grad resultatene fra gjennomsnittlig tidsbruk (tabell 9.6). Vi ser at blant de svært lavt motiverte elevene i profil 1, finner vi den høyeste andelen elever som ikke viser løsningsatferd, det vil si at de går raskt gjennom oppgavene uten å bruke tilstrekkelig tid til å lese, vurdere og løse dem. Denne andelen er betydelig høyere enn for alle de andre profilene. Til sammenligning er det relativt få elever som ikke viser

løsningsatferd i profil 6, hvor elevene også hadde relativt lav tidsbruk. De høyt motiverte og høytpresterende elevene i profil 6 bruker altså lite tid, men de fleste bruker tilstrekkelig tid til å gjøre det som trengs for å kunne løse oppgavene. Dette forsterker inntrykket av elevene i profil 6 som raske og effektive til å løse matematikkoppgavene.

Tabell 9.6. Prosentandel elever som viser og ikke viser løsningsatferd for de seks motivasjonsprofilene

	Profil 1	Profil 2	Profil 3	Profil 4	Profil 5	Profil 6
Ingen løsningsatferd						
% profil	34,8 %	20,5 %	16,6 %	20,3 %	23,6 %	21,1 %
SR	7,7	-2,3	-2,9	-2,2	0,3	-0,7
Løsningsatferd						
% profil	65,2 %	79,5 %	83,4 %	79,7 %	76,4 %	78,9 %
SR	-4,2	1,2	1,6	1,2	-0,2	0,4

Merk: %-linjen for profil angir prosentandelen innenfor den spesifikke profilen. SR står for standardisert residual; verdier over 1,96 i absolutt verdi er statistisk signifikante.

Til slutt undersøkte vi også om oppgavens vanskelighetsgrad ser ut til å påvirke hvorvidt elevene viser løsningsatferd eller ikke. For å undersøke dette, brukte vi oppgavekategoriseringen som er gjort i TIMSS 2023 (von Davier et al., 2024), der oppgavene deles inn i fire kategorier: lav, middels, høy og avansert.

Analysene viste ingen tydelige mønstre når vi tok hensyn til oppgavenes vanskegrad. De bidro heller ikke til å tydeliggjøre resultatene som er presentert tidligere i dette kapitlet. I stedet tyder analysene på at tiden eleven bruker på en oppgave hovedsakelig er knyttet til innholdet og kompleksiteten i oppgaven. For eksempel, dersom oppgaven består av flere deler, eller krever at elevene skriver inn svaret med egne ord i stedet for å trykke på riktig svar (flervalg).

Diskusjon

I dette kapitlet har vi undersøkt om det finnes ulike motivasjonsprofiler knyttet til matematikk blant elever på 9. trinn i Norge, med utgangspunkt i situated-expectancy value theory (Eccles & Wigfield, 2020, 2023, 2024; Wigfield & Eccles, 2000). Gjennom analysen fant vi seks ulike profiler som viser forskjeller i elevenes interesse for matematikk, i hvilken grad de ser matematikkens nytteverdi og selvtillit i faget. Disse funnene bekrefter

hypotesen vår og er i samsvar med tidligere STEM-forskning og funn knyttet til Norge (Hofverberg et al., 2022; Radišić & Jensen, 2021).

Når vi ser på andelen elever i hver motivasjonsprofil, finner vi at under fem prosent av elevene tilhører profilen med høyest verdi på alle de tre motivasjonskonstruktene (profil 6). Også profil 3, som viser de høyeste verdiene for to av tre motivasjonskonstruksjoner, inkluderer under ti prosent av elevene. Til sammenligning finner vi flere elever i profil 1, den minst optimale motivasjonsprofilen, enn i profil 3 og 6 til sammen. Disse resultatene er i tråd med tidligere funn (Radišić & Jensen, 2021) og stemmer også overens med resultatene fra de to siste TIMSS-syklusene, som viser en tydelig nedgang i motivasjonen for matematikk blant norske elever på 9. trinn (Kaarstein et al., 2020, 2024). Samtidig viser resultatene i dette kapitlet at 42 prosent av elevene har høy selvtilitt og ser verdien av matematikk for fremtidig utdanning, jobb eller hverdagsliv (profilene 2, 3 og 6). Dette kan ses som et positivt funn og et område som kan styrkes ytterligere, ved at disse elevene også kan utvikle en interesse for matematikk, som er det motivasjonsaspektet som er minst synlig blant elever på 9. trinn i Norge.

Vi forventet å finne at gutter ville være mer fremtredende enn jenter i profiler med høy interesse verdi og selvtilitt (Kaarstein et al., 2020, 2024), og dette ble delvis bekreftet. Gutter dominerte profilene med størst interesseverdi og selvtilitt, men det er ingen tydelige kjønnsforskjeller i profilene med minst motivasjon. Den eneste profilen der jenter er mer representert, viser elevene med en viss grad av anerkjennelse av matematikkens nytteverdi for fremtiden.

Tidligere forskning viser at elever som har høy interesseverdi for matematikk ofte presterer godt i faget (Prast et al., 2018; Schöber et al., 2018; Skaalvik et al., 2015). Våre resultater tyder imidlertid på at elevenes selvtilitt i matematikk har enda sterkere sammenheng med prestasjoner (Marsh et al., 2017; Wang, 2012). De profilene med høyest selvtilitt, er også de som presterer best. Selv om interesse for faget er det minst fremtredende motivasjonsaspektet i vårt utvalg, viser resultatene våre at elever som liker å lære matematikk, som i profil 6, presterer godt over gjennomsnittet blant norske elever. Dette understreker at interesse fortsatt er viktig. I tillegg finner vi en klar sammenheng mellom innsats på prøven og prestasjoner (Silm et al., 2019, 2020; Wise & DeMars, 2005), der profilene med høyest innsats (profil 4 og 6) også er de med høyest prestasjoner.

Et interessant funn er knyttet til innsatsen til profil 2. Elevene her presterer omtrent på gjennomsnittet, men bruker mer tid på oppgavene enn

gjennomsnittet. Disse elevene rapporterer om svært lav interesse for matematikk, men har høy selvtilitt i og ser nytteverdien av faget. Dette kan være med på å forklare hvorfor de fortsetter å jobbe med oppgavene, selv om de ikke er så interesserte i faget.

For profil 4 og 6 er det et tydelig samsvar mellom elevenes motivasjon, prestasjoner og innsats i prøvesituasjonen, som samsvarer med funnene til Valdivia et al. (2023). For andre profiler er bildet mer sammensatt, noe som stemmer overens med det funn fra tidligere studier (Akhtar & Firdiyanti, 2023; Pools & Monseur, 2021). Dermed støtter våre resultater funnene til Akhtar og Firdiyanti (2023) om betydningen av mestringsforventninger for innsats, og funnene til Pools og Monseur (2021) og Hofverberg et al. (2022).

Når man ser på elevenes innsats på prøver, diskuteres det i mange tilfeller den faktiske situasjonen ved prøven og om prøven oppleves som relevant for eleven (Hofverberg et al., 2023; Valdivia et al., 2023). Et viktig funn er at resultatene tyder på at de fleste norske elevene gjør en innsats, uavhengig av hvor vanskelige oppgavene er eller hvor godt de presterer. Dette styrker validiteten til resultatene fra undersøkelsen (Akhtar & Firdiyanti, 2023), selv om det fortsatt er behov for mer detaljerte analyser for å fullt ut forstå alle elevenes atferd.

Begrensninger og videre forskning

Som med alle studier, er det en begrensning i hvilke begreper vi har kunnet analysere. Studien bygger på tilgjengelige data, og selv om vi har brukt situated expectancy-value theory (Eccles & Wigfield, 2020, 2023, 2024; Wigfield & Eccles, 2000), er ikke alle motivasjonsaspektene som inngår i SEVT inkludert i vår studie. Dette gjør at vi må være forsiktige med hvilke konklusjoner vi trekker. Likevel har den personorienterte tilnærmingen gitt oss en mer nyansert forståelse av hvordan motivasjon varierer blant elever, og hvordan dette henger sammen med prestasjoner og atferd på prøver.

En annen viktig begrensning er at TIMSS-dataene er tverrsnittdata, det vil si data som er samlet inn på ett tidspunkt. Det kan derfor ikke trekkes noen konklusjoner om profilenes stabilitet over tid. Ut fra resultater fra tidligere studier, kan vi imidlertid observere at visse profiler går igjen i flere TIMSS-sykluser (Radišić & Jensen, 2021), noe som kan tyde på en viss stabilitet over tid.

Et annet trekk ved TIMSS-undersøkelsen er at det er en *low-stakes*-prøve, som vil si at den ikke har direkte konsekvenser for elevene. Selv om

vi har funnet mønstre som samsvarer med tidligere funn fra Valdivia et al. (2023), er det fortsatt viktig å observere elevenes prestasjoner, motivasjon og prøvetatferd i situasjoner der prøvesituasjonen oppleves som relevant og viktig for elevene. I videre forskning bør man også se nærmere på innholdet i matematikkoppgavene som elevene arbeider med, spesielt hvordan oppgaver der elevene må konstruere egne svar kraver om en konstruert respons kontra flervalgsoppgaver påvirker deres atferd.

Forfatterbiografi

Jelena Radišić er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Der har hun vært siden 2016. Hennes arbeid er knyttet til forskning på elev-, lærer- og skolekarakteristika som påvirker elevenes resultater. Hun er spesielt opptatt av hvordan elevenes læring påvirkes av motivasjon, selvtillit i matematikk, akademiske følelser, læreres holdninger og undervisningspraksis. I sin forskning benytter Radišić en rekke ulike tilnærminger som inngår i kvantitativ, kvalitativ og blandet metodeforskning.

Andreas Pettersen er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han har en doktorgrad i matematikkdidaktikk og har jobbet med PISA-undersøkelsen siden 2017. Pettersen har utgitt flere artikler om matematikkdidaktikk og vært redaktør for boka *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (utgitt i 2024).

Referanser

- Akhtar, H. & Firdiyanti, R. (2023). Test-taking motivation and performance: Do self-report and time-based measures of effort reflect the same aspects of test-taking motivation? *Learning and Individual Differences*, 105, Artikkel 102323. <https://doi.org/10.1016/j.lindif.2023.102323>
- Aldrich, C. E. A., Bookbinder, A. & Khorramdel, L. (2024). Developing the TIMSS mathematics and science achievement instruments. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report (methods and procedures)* (s. 1.1–1.23). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs3063>
- Asparouhov, T. & Muthén, B. (2020). *Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model*. Mplus. <https://www.statmodel.com/examples/webnotes/webnote21.pdf>
- Baumert, J. & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16(3), 441–462.

- Bergman, L. R. & Trost, K. (2006). The person-oriented versus the variable-oriented approach: Are they complementary, opposites, or exploring different worlds? *Merrill-Palmer Quarterly*, 52, 601–632.
- Blommers, P. & Lindquist, E. F. (1944). Rate of comprehension of reading; its measurement and its relation to comprehension. *Journal of Educational Psychology*, 35(8), 449–473. <https://doi.org/10.1037/h0054306>
- Bolck, A., Croon, M. & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, 12(1), 3–27. <https://doi.org/10.1093/pan/mp001>
- Castro, M. J., López, M., Cao, M. J., Fernández-Castro, M., García, S., Frutos, M. & Jiménez, J. M. (2019). Impact of educational games on academic outcomes of students in the degree in nursing. *PloS One*, 14(7), Artikkel e0220388. <https://doi.org/10.1371/journal.pone.0220388>
- Collins, L. M. & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. John Wiley & Sons.
- Dowker, A., Sarkar, A. & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology*, 7, 1–16.
- Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R. & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences of the United States of America*, 108(19), 7716–7720. <https://doi.org/10.1073/pnas.1018601108>
- Eccles, J. S. & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61. <https://doi.org/10.1016/j.cedpsych.2020.101859>
- Eccles, J. S. & Wigfield, A. (2023). Expectancy-value theory to situated expectancy-value theory: Reflections on the legacy of 40+ years of working together. *Motivation Science*, 9(1), 1–12. <https://doi.org/10.1037/mot0000275>
- Eccles, J. S. & Wigfield, A. (2024). The development, testing, and refinement of Eccles, Wigfield, and colleagues' situated expectancy-value model of achievement performance and choice. *Educational Psychology Review*, 36(2), Artikkel 51. <https://doi.org/10.1007/s10648-024-09888-9>
- Ferguson, S. L., G. Moore, E. W. & Hull, D. M. (2020). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development*, 44(5), 458–468. <https://doi.org/10.1177/0165025419881721>
- Geiser, C. (2013). *Data analyses with Mplus*. Guilford Press.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H. & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, 106(3), 608–626. <https://doi.org/10.1037/a0034716>
- Habók, A., Magyar, A., Németh, M. B. & Csapó, B. (2020). Motivation and self-related beliefs as predictors of academic achievement in reading and mathematics: Structural equation models of longitudinal data. *International Journal of Educational Research*, 103, Artikkel 101634. <https://doi.org/10.1016/j.ijer.2020.101634>
- Hidi, S. & Renninger, K. A. (2006). The four phase model of interest development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
- Hofverberg, A., Eklöf, H. & Lindfors, M. (2022). Who makes an effort? A person-centered examination of motivation and beliefs as predictors of students' effort and performance on the PISA 2015 science assessment. *Frontiers in Education*, 6, 1–17. <https://doi.org/10.3389/educ.2021.791599>
- Hoogland, K. & Tout, D. (2018). Computer-based assessment of mathematics into the twenty-first century: Pressures and tensions. *ZDM Mathematics Education*, 50(4), 675–686. <https://doi.org/10.1007/11858-018-0944-2>
- Ivanova, M., Michaelides, M. & Eklöf, H. (2020). How does the number of actions on constructed-response items relate to test-taking effort and performance? *Educational Research and Evaluation*, 26(5–6), 252–274.

- Kaarstein, H., Lehre, A.- C., Radišić, J. & Rohatgi, M. A. (2024). *TIMSS 2023. Kortrapport*. Universitet i Oslo.
- Kaarstein, H., Radišić, J., Lehre, A. C., Nilsen, T. & Bergem, O. K. (2020). *TIMSS 2019. Kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Kaarstein, H. & Nilsen, T. (2016). Motivasjon. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015* (s. 63–77). Universitetsforlaget.
- Lee, J. & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, 65, 50–64. <https://doi.org/10.1016/j.lindif.2018.05.009>
- Lepper, M. R., Corpus, J. H. & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology*, 97(2), 184–196. <https://doi.org/10.1037/0022-0663.97.2.184>
- Lundgren, E. & Eklöf, H. (2023). Questionnaire-taking motivation: Using response times to assess motivation to optimise on the PISA 2018 student questionnaire. *International Journal of Testing*, 23(4), 231–256. <https://doi.org/10.1080/15305058.2023.2214647>
- Marsh, H. W., Xu, M. & Martin, A. J. (2012). Self-concept: A synergy of theory, method, and application. I K. R. Harris, S. Graham, T. Urda, C. B. McCormick, G. M. Sinatra & J. Sweller (Red.), *APA educational psychology handbook: Vol. 1. Theories, constructs, and critical issues* (s. 427–458). American Psychological Association.
- Marsh, H. W., Martin, A. J., Yeung, A. S. & Craven, R. G. (2017). Competence self-perceptions. I A. J. Elliot, C. S. Dweck & D. S. Yeager (Red.), *Handbook of competence and motivation: Theory and application* (2. utg., s. 85–115). Guilford Press.
- Muenks, K., Wigfield, A. & Eccles, J. S. (2018). I can do this! The development and calibration of children's expectations for success and competence beliefs. *Developmental Review*, 48, 24–39. <https://doi.org/10.1016/j.dr.2018.04.001>
- Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O. & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories? *Journal of Research on Adolescence*, 20(2), 482–506. <https://doi.org/10.1111/j.1532-7795.2010.00644.x>
- Pools, E. & Monseur, C. (2021). Student test-taking effort in low-stakes assessments: Evidence from the English version of the PISA 2015 science test. *Large-scale Assessments in Education*, 9(1), Artikkel 10. <https://doi.org/10.1186/s40536-021-00104-6>
- Prast, E. J., Van de Weijer-Bergsma, E., Miočević, M., Kroesbergen, E. H. & Van Luit, J. E. H. (2018). Relations between mathematics achievement and motivation in students of diverse achievement levels. *Contemporary Educational Psychology*, 55, 84–96. <https://doi.org/10.1016/j.cedpsych.2018.08.002>
- Radišić, J. & Jensen, F., (2021). Norske 9.-trinnslevers motivasjon for naturfag og matematikk – en latent profilanalyse av TIMSS 2019. I T. Nilsen & H. Kaarstein (Red.), *Med blikket mot naturfag. Nye analyser av TIMSS-data og trender 2015–2021* (s. 103–139). Universitetsforlaget.
- Ratcliff, R., Thompson, C. A. & McKoon, G. (2015). Modeling individual differences in response time and accuracy in numeracy. *Cognition*, 137, 115–136. <https://doi.org/10.1016/j.cognition.2014.12.004>
- Reynolds, K. A. (2024). Updating the TIMSS instruments for describing contexts of student learning. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report (methods and procedures)* (s. 2.1–2.8). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs6609>
- Ryan, R. M. & Deci, E. L. (2016). Facilitating and hindering motivation, learning, and well-being in schools: Research and observations from self-determination theory. I K. R. Wentzel & D. B. Miele (Red.), *Handbook of motivation at school* (2. utg., s. 96–119). Erlbaum.
- Sahin, F. & Colvin, K. F. (2020). Enhancing response time thresholds with response behaviors for detecting disengaged examinees. *Large-scale Assessments in Education*, 8(1), Artikkel 5. <https://doi.org/10.1186/s40536-020-00082-1>

- Schnipke, D. L. & Scrams, D. J. (1997). Modeling item response times with a two-state mixture model: A new method of measuring speededness. *Journal of Educational Measurement*, 34(3), 213–232. <https://doi.org/10.1111/j.1745-3984.1997.tb00516.x>
- Schöber, C., Schütte, K., Köller, O., McElvany, N. & Gebauer, M. M. (2018). Reciprocal effects between self-efficacy and achievement in mathematics and reading. *Learning and Individual Differences*, 63, 1–11. <https://doi.org/10.1016/j.lindif.2018.01.008>
- Siegel, P. & Foy, P. (2024). TIMSS sample design. I M. von Davier, B. Fishbein & A. Kennedy (Red.), *TIMSS 2023 technical report (methods and procedures)* (s. 3.1–3.30). Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs7952>
- Silm, G., Must, O. & Täht, K. (2019). Predicting performance in a low-stakes test using self-reported and time-based measures of effort. *Trames*, 23(3), 353–376. <https://doi.org/10.3176/tr.2019.3.06>
- Silm, G., Pedaste, M. & Täht, K. (2020). The relationship between performance and test-taking effort when measured with self-report or time-based instruments: A meta-analytic review. *Educational Research Review*, 31, Artikkel 100335. <https://doi.org/10.1016/j.edurev.2020.100335>
- Skaalvik, E. M., Federici, R. A. & Klassen, R. M. (2015). Mathematics achievement and self-efficacy: Relations with motivation for mathematics. *International Journal of Educational Research*, 72, 129–136. <https://doi.org/10.1016/j.ijer.2015.06.008>
- Soland, J., Wise, S. L. & Gao, L. (2019). Identifying disengaged survey responses: New evidence using response time metadata. *Applied Measurement in Education*, 32(2), 151–165. <https://doi.org/10.1080/08957347.2019.1577244>
- Stacey, K. & Wiliam, D. (2014). Technology and assessment in mathematics. I M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick & F. K. S. Leung (Red.), *Third international handbook of mathematics education* (s. 721–51). Springer.
- Stankov, L. & Lee, J. (2017). Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence*, 61, 11–16. <https://doi.org/10.1016/j.intel.2016.12.001>
- Tan, C. Y. & Hew, K. F. (2019). The impact of digital divides on student mathematics achievement in Confucian heritage cultures: A critical examination using PISA 2012 data. *International Journal of Science and Mathematics Education*, 17(6), 1213–1232. <https://doi.org/10.1007/10763-018-9917-8>
- Toh, L. & Watt, J. M. G. (2022). How do adolescent mathematical self-concept and values explain attainment of different kinds of STEM degrees in adulthood? *Contemporary Educational Psychology*, 69, Artikkel 102057. <https://doi.org/10.1016/j.cedpsych.2022.102057>
- Valdivia, S. D., Rutkowski, L., Rutkowski, D., Canbolat, Y. & Underhill, S. (2023). Test engagement and rapid guessing: Evidence from a large-scale state assessment. *Frontiers in Education*, 8, Artikkel 1127644. <https://doi.org/10.3389/educ.2023.1127644>
- von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U. & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs6460>
- Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology*, 48(6), 1643–1657. <https://doi.org/10.1037/a0027247>
- Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P. & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology*, 48(6), 1594–1611. <https://doi.org/10.1037/a0027838>
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>

- Wise, S. L. (2006). An investigation of the differential effort received by items on a low-stakes computer-based test. *Applied Measurement in Education*, 19(2), 95–114. https://doi.org/10.1207/s15324818ame1902_2
- Wise, S. L. (2015). Effort analysis: Individual score validation of achievement test data. *Applied Measurement in Education*, 28(3), 237–252. <https://doi.org/10.1080/08957347.2015.1042155>
- Wise, S. L. & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1–17. https://doi.org/10.1207/s15326977ea1001_1
- Wise, S. L. & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), 163–183. https://doi.org/10.1207/s15324818ame1802_2
- Wise, S. L. & Ma, L. (2012). *Setting response time thresholds for a CAT item pool: The normative threshold method* [Paperpresentasjon]. Annual meeting of the National Council on Measurement in Education, Vancouver, Canada.

Vedlegg

Vedlegg 9.1. Konstrukter knyttet til motivasjon

Konstrukter	Påstander i spørreskjemaet
Liker å lære matematikk	«Jeg liker å lære matematikk.» «Jeg lærer mye interessant i matematikk.» «Jeg liker matematikk.» «Jeg liker alt skolearbeid som har med tall å gjøre.» «Jeg liker å løse oppgaver i matematikk.» «Jeg liker å løse oppgaver i matematikk.» «Jeg liker å løse oppgaver i matematikk.»
Selvtillit i matematikk	«Jeg liker å løse oppgaver i matematikk.» «Matematikk er vanskeligere for meg enn for mange andre i klassen.» «Jeg er rett og slett ikke flink i matematikk.» «Matematikk er lett for meg.» «Jeg er flink til å løse vanskelige oppgaver i matematikk.» «Jeg er flink til å forklare matematikk for andre.» «Matematikk er vanskeligere for meg enn noe annet fag.» «Matematikk gjør meg forvirret.»
Verdsetter matematikk	«Jeg tror at å lære matematikk kommer til å hjelpe meg i dagliglivet.» «Jeg trenger matematikk for å lære andre skolefag.» «Jeg må gjøre det bra i matematikk for å komme inn på den utdanningen jeg helst vil.» «Jeg må gjøre det bra i matematikk for å få den jobben jeg ønsker meg.» «Jeg kunne tenke meg en jobb hvor jeg får bruk for matematikk.» «Det er viktig å lære om matematikk for å komme seg fram her i verden.» «Å lære matematikk vil gi meg flere muligheter til arbeid når jeg blir voksen.» «Det er viktig å gjøre det bra i matematikk.»

Vedlegg 9.2. Gjennomsnittresultater i matematikk for de seks motivasjonsprofilene for alle plausible verdier

	Profil 1	Profil 2	Profil 3	Profil 4	Profil 5	Profil 6
Plausibel verdi (1)	452,872	502,254	561,915	533,534	482,407	529,542
Plausibel verdi (2)	454,629	503,042	564,939	532,776	482,672	529,461
Plausibel verdi (3)	456,755	502,173	562,085	533,117	484,155	530,966
Plausibel verdi (4)	454,111	502,219	564,340	533,659	484,566	531,064
Plausibel verdi (5)	454,914	502,529	564,254	533,586	485,483	529,090

KAPITTEL 10

Norske elevers kunnskaper og ferdigheter i algebra – den vanskelige oppgaven med å lage matematiske uttrykk

Andreas Pettersen og Jelena Radišić

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: Kunnskaper og ferdigheter i algebra utgjør en viktig del av elevers matematiske kompetanse. Tidligere forskning har vist at norske elever strever med å lære seg denne delen av matematikkfaget, noe som har fått mye oppmerksomhet i både forskningslitteraturen og utdanningspolitikk. For å få dypere innsikt i norske elevers kunnskaper og ferdigheter i algebra, har vi analysert elevsvarene på tre utvalgte oppgaver fra TIMSS 2023. Oppgavene er rettet mot en spesifikk del av algebra: å generalisere et mønster og formulere et matematisk uttrykk. Gjennom en analyse av elevsvarene på disse oppgavene, og se disse i sammenheng med elevenes prestasjoner i matematikk, undersøker vi hva slags typer svar, både riktige og feilaktige, elevene gir og diskuterer hva hvilken forståelse som kan ligge til grunn for ulike typer svar. Studien gir dermed en innsikt i norske elevers evne til å generalisere mønstre og beskrive dette ved bruk av formell matematisk notasjon.

Nøkkelord: algebra, matematisk kompetanse, matematiske uttrykk, TIMSS

Abstract**Norwegian Students' Knowledge and Skills in Algebra – The Challenging Task of Creating Algebraic Expressions**

Knowledge and skills in algebra constitute an important part of students' overall mathematical competence. Previous research has shown that Norwegian students often struggle with learning this aspect of mathematics, an issue that has received considerable attention in both educational research and policy. To gain deeper insight into Norwegian students' understanding of algebra, we have analysed student responses to three selected tasks from TIMSS 2023. These tasks target a specific aspect of algebra: generalising a pattern and formulating an algebraic expression. Through an analysis of the students' responses and by examining these in relation to their overall mathematics performance, we investigate the types of answers, both correct and incorrect, that students provide, and discuss the understandings that may underlie these responses. The study thus offers insight into Norwegian students' ability to generalise patterns and express them using formal mathematical notation.

Keywords: algebra, mathematical competence, mathematical expressions, TIMSS

Innledning

Algebra er et komplekst begrep uten en entydig definisjon, og hvordan begrepet forstås og brukes varierer fra person til person (Mason, 1996). Mange ser på algebra som matematikkens språk (Grønmo, 2018; Mason 1996), der bokstaver og matematiske symboler kombineres og manipuleres for å beskrive situasjoner og løse problemer. Kunnskaper og ferdigheter innenfor algebra ansees også som en sentral del av matematisk kompetanse (Niss & Højgaard, 2019) og blir gjerne sett på som en «portvokter» i matematikkfaget, siden det å mestre algebra er avgjørende for å kunne gå videre med matematikk i studier og arbeidslivet (Knuth et al., 2005). Algebra spiller også en sentral rolle i matematikkfaget i alle land (Leung et al., 2014).

Tradisjonelt har algebra vært knyttet til det å løse ligninger, formulere og håndtere algebraiske uttrykk og oversette tekstoppgaver til algebraiske uttrykk med variabler og ukjente (Kieran, 2020). Innenfor forskningslitteraturen finnes det derimot ulike syn på hva det innebærer å lære algebra i skolen (Kieran, 2020) og det er store forskjeller på tvers av land når det kommer til hva som undervises og hvordan. Til tross for disse forskjellene, er det bred enighet om at overgangen fra aritmetikk til algebra er svært vanskelig for mange elever (Cai & Knuth, 2011; Kieran, 2020). Tidligere studier kan også tyde på at norske elever har større vansker med algebra enn elever i mange andre land (Bergem, 2016; Grønmo & Olsen, 2006). Dette er trolig årsaken til at algebraens plass i skolematematikken er og har vært et hett tema for politikk, praksis og forskning i Norge de siste årene (se for eksempel Grønmo & Helgesen, 2018; Vinje, 2023). Sammenlignet med tidligere undersøkelser viser resultatene fra TIMSS 2023 (*Trends in International Mathematics and Science Study*) en oppgang i norske elevers resultater innenfor algebra på 9. trinn (Kaarstein et al., 2024). Likevel presterer norske elever fortsatt svakere i algebra enn elever i mange andre land (von Davier et al., 2024).

Mens gjennomsnittresultater fra TIMSS og andre internasjonale undersøkelser gir et overordnet bilde av elevers prestasjoner, sier de lite om hva som ligger bak resultatene, for eksempel hva slags feil elevene gjør og hvordan dette henger sammen med elevenes kunnskaper og forståelse. Det å gjøre feil er en svært viktig del av elevenes læringsprosess (Käfer et al., 2019; Tulis, 2013), og ved å undersøke elevers feilsvar kan vi få verdifull innsikt i deres kunnskaper, ferdigheter og grunnleggende begrepsmessige

forståelser. I dette kapittelet undersøker vi norske 9. trinnselevers (feil)svar på tre algebra-oppgaver i TIMSS 2023. De tre oppgavene er valgt fordi de måler den samme typen algebraisk kompetanse: evnen til å generalisere et matematisk mønster og formulere dette som et matematisk uttrykk. Gjennom analysene av elevsvarene ønsker vi å finne ut hva slags typer (feil)svar norske elever gir på denne typen for å få innsikt elevenes forståelse av algebra og deres bruk av formell matematisk notasjon.

Hva er algebra?

Begrepet algebra har røtter tilbake til 800-tallet og var opprinnelig knyttet til det å løse ligninger (Kieran, 2020). Det var dette synet på algebra, som et verktøy for å manipulere symboler og løse ligninger, som var gjeldende frem til andre halvdel av 1900-tallet. Dette synet har også preget undervisningen i algebra, som i stor grad har handlet om det å løse ligninger, håndtere algebraiske uttrykk og oversette tekstoppgaver til uttrykk med variabler og ukjente (Kieran, 2020). Tradisjonelt har algebra blitt introdusert på ungdomstrinnet og senere i utdanningsløpet, mens undervisningen på barnetrinnet har handlet om aritmetikk (Cai & Knuth, 2011; Kaput, 2008). Etter hvert har det imidlertid vokst frem et bredere syn på hva algebra er. Dette inkluderer ikke bare bruk av variabler og symboler, men også representasjoner som funksjoner, grafer og tabeller. I tillegg har forskning på algebraisk tenkning og algebraiske ideer fått mye oppmerksomhet, hvor det argumenteres for at dette er noe som bør introduseres tidligere i skoleløpet enn den tradisjonelle algebraen (Blanton et al., 2017; Cai & Knuth, 2011; Kieran, 2022).

I forskningslitteraturen finnes det ulike forståelser og definisjoner av algebra. Ifølge Kongelf (2015) er det synet på algebra som generalisert tallære som i dag dominerer undervisning. Wu (2001), for eksempel, ser på algebra som generalisert tallære og at dette handler om generalisering og abstraksjon uttrykt med symboler.

For mange elever er overgangen fra aritmetikk, det vil si læren om tall og deres egenskaper, til algebra svært krevende. En tidlig introduksjon av algebraisk tenkning handler ikke bare om at denne overgangen skal bli enklere, men handler også om å gi elevene en dypere forståelse for underliggende matematiske strukturer og sammenhenger (Cai & Knuth, 2011). Algebraisk tenkning bygger på to sentrale elementer: variabler som generelle og abstrakte enheter, og en strukturell forståelse av algebra. Disse

elementene er involvert i alle algebraaktiviteter, som for eksempel generalisering, ligninger og funksjoner (Warren et al., 2016).

Algebra som en del av matematisk kompetanse

Algebra blir sett på som en sentral del av matematisk kompetanse og omtales som en «portvokter» fordi den er avgjørende for å mestre mer avansert matematikk (Filer & Chang, 2008; Hunter & Miller, 2022; Lee & Mao, 2021; Star et al., 2015). Algebra danner et grunnlag for komplekse matematiske ideer, og gjennom arbeid med symboler og generaliserte begreper utvikler elevene evnen til å analysere og løse komplekse matematiske problemer som ikke lar seg håndtere med mer grunnleggende matematiske ferdigheter (Cai et al., 2011; Carpenter & Levi, 2000). Å beherske algebra gjør det også mulig å løse problemer raskere og mer effektivt enn med grunnleggende aritmetikk (Grønmo, 2018). Algebra fungerer også som et universelt språk i matematikken, som gjør det mulig å uttrykke relasjoner mellom størrelser og variabler på en konsis og presis måte (Grønmo, 2018; Kvasz, 2006). Dette gjør det lettere å kommunisere matematiske ideer og løse problemer på tvers av ulike fagområder.

Algebra er tett integrert i andre deler av matematikken, som geometri og trigonometri (Wasserman, 2016). Forståelse av algebraiske begreper er dermed avgjørende for å lykkes også innenfor andre matematiske områder. I tillegg styrker algebraen evnen til logisk tenkning og resonnering (Kieran, 1992). Elevene lærer å tilnærme seg problemer systematisk, dele dem opp i mindre deler og løse dem trinn for trinn, noe som er en verdifull ferdighet i alle deler av matematikken. Med tanke på den viktige rollen algebra spiller i matematikkfaget, og betydningen av kunnskaper og ferdigheter algebra for en generell matematisk kompetanse, er det ikke overraskende at den utgjør en viktig del av matematikkfaget i skolesystemer over hele verden (Leung et al., 2014).

Algebra i læreplanen i Norge

I læreplanverket LK20 utgjør «Abstraksjon og generalisering» ett av seks kjerneelementer som skal beskrive det viktigste faglige innholdet som gjelder på tvers av trinnene 1 til 13 (Kunnskapsdepartementet, 2019). «Abstraksjon og generalisering» handler om at elevene gradvis skal kunne formalisere tanker, strategier og matematisk språk, håndtere formelt symbolspråk, og å kunne formalisere sammenhenger med algebra

(Kunnskapsdepartementet, 2019). Videre finner vi generalisering og å lage algebraiske uttrykk med variabler, som de tre oppgavene i vår studie handler om, i ulike kompetansemål i læreplanen:

- «bruke variabler og formler til å uttrykke sammenhenger i praktiske situasjoner» (6. trinn)
- «lage og forklare regneuttrykk med tall, variabler og konstanter knyttet til praktiske situasjoner» (8. trinn)
- «beskrive og generalisere mønstre med egne ord og algebraisk» (8. trinn)

Algebra hatt en sentral rolle også i tidligere læreplaner. Matematikkfaget har tidligere vært strukturert i ulike matematiske temaer eller hovedområder, der algebra har utgjort et av disse. I LK06 var for eksempel «Tall og algebra» et av hovedområdene fra og med 5. trinn, mens «Tall» var hovedområdet fra 1. til 4. trinn (Kunnskapsdepartementet, 2006). I L97 ble algebra først introdusert som et såkalt *målområde* («Tall og algebra») på 8. trinn (Kirke-, utdannings- og forskningsdepartementet, 1996). I den foreløpig siste realfagsstrategien, *Tett på realfag* (2015), fremheves det at algebra har blitt styrket i LK06 i på 4., 7. og 10. trinn (Kunnskapsdepartementet, 2015). Her henvises det blant annet til resultater fra TIMSS 2011, som viser at norske elever har svak kompetanse i algebra og at norske elever introduseres for algebra senere enn elever i andre land (Kunnskapsdepartementet, 2015). Innføringen av «Abstraksjon og generalisering» som et kjerneelement fra 1. trinn, kan dermed tolkes som et uttrykk for et ønske om å styrke algebraundervisningen tidligere i utdanningsløpet enn før. I tillegg innebar innføringen av kompetansemål på hvert trinn i LK20 en innstramming i lærernes handlingsom, siden strukturen i LK06 med kompetansmål for 5.–7. trinn ga lærerne mulighet til å utsette algebraundervisningen til 7. trinn.

Elevers kunnskaper og ferdigheter i algebra

Forskningen på elevers kunnskaper og ferdigheter i algebra har ikke vært like omfattende som forskningen på generell matematisk kompetanse. Likevel finnes det en rekke studier og noen kunnskapsoppsummeringer som gir innsikt i elevers forståelse for og utfordringer med algebra (se for eksempel Kieran, 2006; Warren et al., 2016). Resultatene fra disse studiene viser til ulike aspekter ved algebra som er krevende for mange elever og

som, sett under ett, tegner et bilde av algebra som et vanskelig tema for elevene å lære. For eksempel finner vi at mange elever strever med å gjenkjenne og uttrykke mønsterstrukturer, tolke, manipulere og gjenkjenne strukturen i algebraiske uttrykk, forstå ulike sider ved konseptet variabel og bruke algebra til å løse tekstoppgaver (Kieran, 2006; Warren et al., 2016). Warren et al. (2016) viser til at mange av de vanskelighetene elevene opplever med algebra, kan forklares med den tradisjonelle oppbygningen av matematikkundervisningen hvor elevene først arbeidet med aritmetikk og deretter algebra. Dette kan blant annet medføre en mangelfull forståelse av likhetstegnet, hvor likhetstegnet blir forstått som noe som skiller en prosedyre fra et resultat (Kieran, 1992) eller som et symbol for at elevene skal «gjøre noe» (Knuth et al., 2005). Flere studier har funnet at en slik mangelfull forståelse av likhetstegnet er svært utbredt, ikke bare på barnetrinnet, men også på ungdomstrinnet (se for eksempel Knuth et al., 2005). Dette sees på som svært problematisk siden forståelse for likhet og likhetstegnet er svært viktig for å kunne lære algebra (Hornburg et al., 2022; Knuth et al., 2005).

Kieran (2020) viser til annen forskning på elevers læring i algebra, som finner at mange elever har problemer med å «godta» at matematiske uttrykk uten likhetstegn (som for eksempel $x + 3$) kan være et gyldig svar på en oppgave, at mange elever kun ser på uttrykk som utregningsprosesser og ikke som matematiske objekter, og at mange elever ikke ser hvordan algebra fungerer som et verktøy til å representere strukturer i en situasjon. I tillegg viser forskning at elever generaliserer sekvenser av figurer (mønstre), slik som de tre oppgavene vi har undersøkt i vår studie, på ulike måter (Kieran, 2020).

Videre viser Bush og Karp (2013) til en rekke tidligere studier som har vist at variabler er et konsept som er krevende å forstå, og at elever har mange ulike misoppfatninger av variabler. Dette handler blant annet om at mange elever ser på variabler som merkelapper (Stacey & MacGregor, 1997) og ikke forstår at variabler kan ha ulike verdier og ikke er en verdi som mangler (Asquith et al., 2007). Phillipp (1992) påpeker at det ikke er rart elever har problemer med å forstå konseptet med variabler, med tanke på at det forstås og brukes på såpass mange ulike måter i matematikken. Dette støttes av MacGregor og Stacey (1997), som fant at elever gjerne gjetter når de skal tolke bokstaver og algebraiske uttrykk eller at de tolker det ut fra andre symbolsystemer de kjenner. Knuth et al. (2005) hadde noe mer oppløftende resultater i sin studie av elevers forståelse av variable. De fant at andelen som tolket bokstav-symboler til å kunne representere mer enn én verdi økte fra 50 prosent på 6. trinn til 75 prosent på 8. trinn (Knuth

et al., 2005). Også Kieran (2020) trekker frem noen oppløftende resultater når hun viser til at resultatene fra TIMSS 2015 viste en fremgang i algebra-resultatene på 8. trinn blant 14 av deltakerlandene. Hun mener dette kan være et tegn på en forbedring når det kommer til elevenes læring i algebra rundt omkring i verden (Kieran, 2020).

Det tidligere nevnte algebra-problemet som beskrives i den internasjonale forskningslitteraturen, hvor algebra undervises som et eget fag isolert fra andre matematiske områder, kan vi også finne spor av i Norge. Kongelf (2015) undersøkte seks norske lærebøker i matematikk og fant at det sjelden legges opp til at algebra-opplæringen bygger videre på kunnskaper fra tallære, men at det i stor grad fremstår som et isolert emne. Resultater fra tidligere TIMSS-undersøkelser har også vist at utfordringer med å lære algebra ser ut til å være et like stort, eller til og med et større problem, for norske elever som for elever i mange andre land (Bergem, 2016). I de to tidligere TIMSS-gjennomføringene, i 2019 og 2023, var Norge blant deltakerlandene med størst forskjell mellom de generelle matematikkresultatene og resultatene i algebra (Mullis et al., 2020; von Davier et al., 2024). Pedersen (2015) undersøkte elevsvar på algebraoppgaver fra TIMSS Advanced 2008 og fant at norske elever strevde med å manipulere symboler og uttrykk. Naalsund (2012) kom frem til lignende resultater i sin undersøkelse av norske elever, hvor hun fant at mange av vanskelighetene med formell algebra handlet om elevenes manglende kunnskap om aritmetiske notasjoner og konvensjoner, både for elever på 8. trinn og på 10. trinn.

Elevers feilsvar som kilde til kunnskap

Feil er en integrert del av læringsprosessen (Käfer et al., 2019; Tulis, 2013; Wischgoll et al., 2015) og kan gi viktig innsikt som er relevant for undervisning, læring og vurdering (Heemsoth & Heinze, 2014; O'Brien et al., 2017). Analyse av disse feilene kan avdekke ulike nivåer av matematisk kompetanse (Neidorf et al., 2020) og skille mellom feil som skyldes prosessuelle ferdigheter og grunnleggende konseptuelle misforståelser (Kingsdorf & Krawec, 2014). Selv om dette gjelder for matematikkfeltet generelt, kan det også overføres til fagområder innenfor matematikk som for eksempel algebra (Booth et al., 2017). Analyser av elevers besvarelser har vist seg å gi en innsikt i elevers forståelse og ferdigheter som kan brukes til å tilpasse undervisningen (Heemsoth & Heinze, 2014; Riccomini, 2005). Videre tyder forskning på at elever med svake prestasjoner ofte benytter seg av uhensiktsmessige operasjonelle strategier, sannsynligvis på grunn

av mangelfull konseptuell forståelse, mens elever med høye prestasjoner viser større fleksibilitet og nøyaktighet i strategier i arbeid med algoritmer (Zhang et al., 2013). Effekten varierer imidlertid mellom grupper, hvor faglig sterkere elever har større utbytte av å analysere feilaktige eksempler, mens elever med lite forkunnskap har større utbytte av riktige eksempler (Heemsoth & Heinze, 2014).

Vår studie

Denne studien tar utgangspunkt i elevsvarene tre oppgaver i TIMSS 2023 hvor elevene skal generalisere et matematisk mønster og formulere dette som et matematisk uttrykk. Ved å analysere norske elevers svar på disse oppgavene, ønsker vi å undersøke kunnskaper og ferdigheter i denne delen av algebra blant elever på 9. trinn i Norge. Dette kan bidra til å belyse hvorfor det å generalisere mønstre og formulere matematiske uttrykk er vanskelig for elevene.

Vi ønsker å besvare følgende forskningsspørsmål:

1. Hva slags typer (feil)svar finner vi i de tre TIMSS-oppgavene hvor elevene skal generalisere mønstre og formulere matematiske uttrykk?
2. Hva er sammenhengen mellom ulike typer elevsvar og prestasjoner i matematikk og i algebra?

Metode

I denne delen beskriver vi utvalget av elever og elevsvarene, samt algebra-oppgavene som analysen bygger på. Videre gjør vi rede for kodeskjemaet som er utviklet for å kategorisere elevsvarene, og for metodene som er brukt i analysen.

Utvalg av elevsvar og oppgaver

TIMSS er en internasjonal undersøkelse som måler elevenes kompetanse i matematikk og naturfag på 5. og 9. trinn. Undersøkelsen gjennomføres hvert fjerde år og inkluderer representative utvalg av elever fra deltakerlandene. Dette kapittelet bygger på data fra TIMSS 2023 i matematikk på 9. trinn, hvor 6 234 norske elever deltok (kapittel 1 i denne antologien).

I TIMSS 2023 var det omtrent 200 oppgaver i matematikk på 9. trinn, hvor hver elev fikk tildelt et utvalg av disse (Kaarstein et al., 2024). Oppgavene var fordelt på fire emneområder (*tall; algebra; geometri;*

statistikk og sannsynlighet) og tre kognitive områder (*å kunne; å anvende; å resonnere*). Omtrent 30 prosent av disse oppgavene var innenfor emneområdet algebra (Kaarstein et al., 2024), og disse var knyttet til to underområder: (1) uttrykk, operasjoner og ligninger og (2) sammenhenger og funksjoner (Philpot et al., 2021).

I denne studien undersøker vi elevsvarene på tre algebraoppgaver som kan knyttes til det første underområdet, nærmere bestemt det å formulere matematiske uttrykk for å beskrive problemstillinger. I disse oppgavene presenteres elevene for figurer som utvikler seg i et bestemt mønster, der hver figur er bygget opp av mindre bestanddeler (for eksempel trekanter). For å løse oppgaven må elevene formulere et uttrykk som beskriver antall deler i figur n . Figur 10.1 viser én av disse tre oppgavene, «Små trekanter i figur (ME72108)». De to andre oppgavene er konfidensielle, og verken oppgavene eller elevenes svar på disse oppgavene kan derfor vises. I oppgaven (figur 10.1) får elevene se tre figurer i et mønster bestående av små trekanter og de må gjenkjenne hvordan mønsteret utvikler seg. Oppgaven består av to deloppgaver. I oppgave A skal elevene oppgi antallet små trekanter i figur 10. I oppgave B skal elevene formulere et uttrykk for antall små trekanter i figur n . I vår studie er det elevsvarene på oppgave B som undersøkes.

Ut fra kompetansemålene i LK20, som er omtalt tidligere i kapitlet, er målet for matematikkopplæringen at elevene skal ha tilegnet seg de nødvendige kunnskapene og ferdighetene for å kunne løse denne typen oppgaver etter 8. trinn.

For hver oppgave der elevene skal formulere egne svar, benytter TIMSS-undersøkelsen en kodeguide for å vurdere om svarene er riktige eller gale. For oppgaven «ME72108B» (se figur 10.1) vil uttrykket « n^2 », samt alle ekvivalente matematiske uttrykk (som for eksempel « $n \cdot n$ ») vurderes som riktige svar. I tillegg fremgår det i kodeguiden at også riktig matematisk uttrykk formulert med ord (for eksempel «multiplisere figurtallet med seg selv») skal godkjennes.

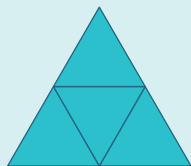
Tabell 10.1 viser resultatene for de tre oppgavene blant norske elever. Vi ser at rundt 900 norske elever fikk tildelt hver av oppgavene, og at andelen riktige svar varierer mellom 12,8 og 17,4 prosent. Dette tyder på at de tre oppgavene var vanskelig for mange norske elever. I tillegg ser vi en forholdsvis høy andel elever som ikke har svart på oppgavene (tabell 10.1), noe som er vanlig både for vanskelige oppgaver og oppgaver med åpne svarfelt. Vi antar at for mange av elevene som ikke har svart på oppgaven skyldes dette at de manglet nødvendig matematisk kunnskap eller ferdigheter for å løse dem.

Dette er de tre første figurene i et mønster.

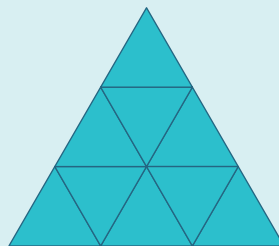
Den neste figuren lages ved å legge til en rad med trekanter på bunnen av figuren før.



Figur 1



Figur 2



Figur 3

A. Fullfør tabellen for figur 10.

Figur	1	2	3	4	10
Antall små trekanter	1	4	9	16	

B. Skriv et uttrykk for antall små trekanter i figur n .

Svar:

Figur 10.1. Oppgaven «Små trekanter i figur (ME72108B)»¹. Dette er en av de tre oppgavene kapittelet bygger på.

Tabell 10.1. Resultater for norske elever på de tre algebraoppgavene. N er antall elever som fikk tildelt oppgaven. «Ikke nådd» vil si andelen elever som ikke rakk å komme frem til denne oppgaven i prøvesettet sitt (hentet fra Fishbein et al., 2025)

Oppgave	N	Riktig svar	Feil svar	Ikke svart	Ikke nådd
ME72110B	903	14,8 %	45,8 %	36,7 %	2,6 %
ME72106C	912	12,8 %	56,9 %	22,4 %	7,9 %
ME72108B	899	17,4 %	48,9 %	31,2 %	2,4 %

Det er ikke kun blant norske elever vi finner en lav andel riktige svar på de tre oppgavene. Tabell 10.2 gir en oversikt over resultatene for norske elever og for elever i Sverige og Finland, de tre nordiske landene som deltok i TIMSS 2023.

Selv om det er noen variasjoner mellom oppgavene og landene, er andelen riktige svar gjennomgående lav på alle de tre oppgavene i de nordiske landene (tabell 10.2). Samtidig finnes det land hvor andelen riktige svar er betydelig høyere. For eksempel viser resultatene for Japan, et land

¹ Oppgave ME72108B, TIMSS 2023 Copyright © 2025. Gjengitt med tillatelse fra TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College.

Tabell 10.2. Resultater for elever i Norge, Sverige og Finland på de tre algebraoppgavene i TIMSS 2023

Oppgaver	Riktig svar	Feil svar	Ikke svart	Ikke nådd
ME72110B				
Norge	14,8 %	45,8 %	36,7 %	2,6 %
Sverige	8,6 %	58,2 %	31,2 %	2,0 %
Finland	11,7 %	62,1 %	25,8 %	0,4 %
ME72106C				
Norge	12,8 %	56,9 %	22,4 %	7,9 %
Sverige	25,6 %	48,3 %	17,5 %	8,6 %
Finland	4,9 %	70,8 %	21,4 %	3,0 %
ME72108B				
Norge	17,4 %	48,9 %	31,2 %	2,4 %
Sverige	18,4 %	52,6 %	27,5 %	1,6 %
Finland	16,0 %	58,9 %	23,9 %	1,2 %

hvor elevene presterer høyt på internasjonale undersøkelser i matematikk (Kaarstein et al., 2024; OECD, 2023), at nærmere 50 prosent av elevene svarte riktig svar på disse oppgavene (Fishbein et al., 2025).

Utvikling av kodeskjema og koding av elevsvar

For å undersøke de norske elevsvarene på de tre oppgavene ble det utviklet et kodeskjema. Hensikten med kodeskjemaet var å identifisere og kategorisere ulike typer elevsvar som kunne knyttes til ulike matematiske ferdigheter og kunnskaper. Noen kategorier ble utviklet deduktivt basert på kunnskap fra forskningslitteraturen, som for eksempel manglende forståelse for likhetstegnet (kategori 21). Andre kategorier ble utviklet med en mer induktiv tilnærming, ut fra en gjennomgang av elevsvarene i forkant av selve kodingen, som for eksempel svar der eleven oppgir kun et tall (kategori 23).

Kodeskjemaet inneholder kategorinummer, beskrivelser, forklaringer og eksempelsvar. Kategorinummene, beskrivelsene og forklaringene (av feilsvar) er formulert slik at de er felles på tvers av de tre oppgavene, mens forklaringene av riktige svar og eksempelsvarene er tilpasset hver enkelt oppgave. Forfatterne diskuterte seg først frem til en kodeguide med åtte kategorier: to for riktig svar og seks for ulike typer feilsvar. Denne ble prøvd ut på et tilfeldig utvalg elevsvar (rundt 10 prosent av svarene per oppgave), der alle «blanke» og manglende svar var ekskludert. I denne utprøvingen kategorisert de to forfatterne de samme svarene uavhengig av hverandre og samsvaret i kodingen (interkoder-reliabilitet) ble beregnet ved hjelp av Cohens kappa (Cohen, 1960). Tabell 10.3 viser samsvaret mellom de to forfatterne i utprøvingen.

Elevsvarene der forfatterne var uenige om kategoriseringen ble diskutert, og det ble gjort en felles vurdering om hvordan svaret skulle

Tabell 10.3. Samsvaret i koding mellom de to forfatterne for 10 prosent av elevsvarene. Samsvaret (interkoder-reliabilitet) er beregnet med Cohens kappa

Oppgave	Antall svar	Reliabilitet (SE)
ME72110B	57	0,888 (0,47)
ME72110B	61	0,785 (0,61)
ME72110B	58	0,871 (0,49)

kategoriseres. Ettersom samsvaret i kodingen var relativt høyt for alle oppgavene, ble det kun gjort mindre endringer i kodeguiden før den ble brukt til å kategorisere resten av elevsvarene. Ut fra uenighetene i utprøvingen, prøvde vi å tydeliggjøre forskjellene mellom kategoriene og legge det ble lagt til flere eksempelsvar. Det ble blant annet bestemt at elevsvar som var formulert med « $f(n)=$ » eller « $fn=$ », og deretter et korrekt matematisk uttrykk, skulle vurderes som riktig, da vi antar at mange elever lærer å skrive funksjonsuttrykk på denne måten. Vi ble også enige om at «.» skulle vurderes som en gyldig måte å skrive multiplikasjonstegnet på, slik at for eksempel « $n.n$ » er vurdert som riktig svar på ME72108B. Formuleringen « nn » ble derimot bestemt å være ugyldig. Videre ble to av kategoriene slått sammen, slik at det endelig kodeskjemaet besto av syv kategorier: to for riktige svar og fem ulike typer feilsvar. Tabell 10.4 viser den endelige versjonen av kodeskjemaet.

Etter oppdateringen av kodeskjemaet ble de resterende 90 prosent av elevsvarene kategorisert. Svarene ble fordelt likt mellom forfatterne, og hver av besvarelsene ble kategorisert av én av forfatterne. Elevsvar der forfatteren var usikker på kategoriseringen, ble diskutert og kategorisert i fellesskap.

Beregning av prestasjoner

Elevenes prestasjoner i TIMSS beregnes på en standardisert skala der 500 poeng representerer skalamidtpunktet. Dette tilsvarer det internasjonale gjennomsnittet som ble beregnet i TIMSS 1995, hvor standardavviket ble satt til 100 poeng. I TIMSS 2023 var gjennomsnittet for de norske elevene på 9. trinn på 501 poeng (Kaarstein et al., 2024).

I TIMSS 2023 er elevenes prestasjoner representert ved fem plausible verdier (se kapittel 1 i denne boka for nærmere forklaring av plausible verdier). I analysene i dette kapittelet ble alle de fem plausible verdiene benyttet for å beregne gjennomsnitt. Det er også brukt vektning på elevnivå (*total student weight*) for å sikre representativitet i utvalget. Alle analyser med gjennomsnittsresultatene ble gjennomført ved hjelp av programvaren IDB Analyzer 5.0.40.

Tabell 10.4. Kodeskjema for å kategorisere elevsvar på de tre TIMSS-oppgavene. Forklaring på riktig svar og eksempelsvar er gitt for den frigitte oppgaven ME72108B (figur 1).

Riktige svar	Beskrivelse	Forklaring	Eksempelsvar
10	Riktig matematisk uttrykk med kun symboler og variabler. Uttrykk som starter med «f(n)=» eller «fn=», der resten er korrekt, kategoriseres som 10	Eleven har skrevet n^* n eller et tilsvarende algebraisk uttrykk. n eller N må være brukt som variabel i uttrykket.	« n^2 » « n^*n » « $f(n) = n^*n$ »
11	Riktig matematisk uttrykk skrevet kun med ord, eller en blanding av ord og symboler.	Eleven har formulert uttrykket n^*n med ord, eller en blanding av symboler og ord.	«figur nummer gange figur nummer» « $n^*n = \text{svar}$ »
Feilsvar			
20	Feil matematisk uttrykk, med symboler, men hvor egnede variabler er brukt (n eller N).	Eleven har skrevet et uttrykk med symboler og egnede variabler (n eller N , eller «f(n)=»), men hvor uttrykket ikke er korrekt i henhold til oppgaven.	« $21n$ » « $n3+2$ » « $5n^2$ »
21	Feil matematisk uttrykk med ekstra eller irrelevante variabler. Eller at svaret er gitt som en ligning (med likhetstegn), dvs. to uttrykk som er satt lik hverandre.	Eleven har skrevet et uttrykk med symboler og variabler, men hvor det er brukt ekstra og irrelevante variabler. Dette gjelder for eksempel hvis det ikke er brukt n eller N i uttrykket, men en annen variabel (f.eks. x), eller hvis det er brukt likhetstegn (=) og svaret dermed er en ligning	« $Fn = n + 3+2$ » « $N+x$ » « $n = f+2$ » « $n = n^*n$ »
22	Feil matematisk uttrykk formulert med ord (uten matematiske symboler) eller en blanding av ord og symboler.	Eleven har skrevet et matematisk uttrykk, eller et regnestykke, som er formulert med en blanding av ord og matematiske symboler/variabler (og uttrykket er feil svar på oppgaven).	«figur gange figur erlik n » « n ganger er lik n i andre » « $n = \text{figur} + \text{neste oddetall}$ »
23	Svaret er gitt som et tall/verdi eller et regnestykke, uten variabel.	Eleven har skrevet et svar på oppgaven som er en verdi og ikke et uttrykk med en variabel. For eksempel et tall eller et regnestykke som blir et tall.	«100» «5-4»
24	Andre svar og «useriøse» svar.	Eleven har gitt et svar som ikke fremstår som et seriøst forsøk på å gi et riktig svar, eller skriver at hen ikke forstår.	«hæ?» «vet ikke hva figur n står for»

Resultater

Denne delen er strukturert etter de to forskningsspørsmålene. Første del tar for seg forskningsspørsmål 1 og presenterer de ulike typene elevsvar vi finner blant norske 9. trinnselever på de tre algebraoppgavene. Andre del er knyttet til forskningsspørsmål 2 og sammenhengen mellom ulike typer elevsvar og prestasjoner i matematikk.

Ulike typer elevsvar

Tabell 10.5 viser antall og fordeling av elevsvar i de ulike kategoriene for hver av de tre algebraoppgavene. Resultatene viser til dels stor variasjon i hvordan svarene fordeler seg i de ulike kategoriene, men også at det i flere av kategoriene er relativt mange elevsvar for alle oppgavene. For alle de tre oppgavene finner vi mer enn 10 prosent av svarene i kategoriene 10, 20, 21 og 23 (tabell 10.5). Videre ser vi at kategori 23, der elevene har gitt et svar som i form av et tall eller et regnestykke uten å bruke en variabel, er svært fremtredende på tvers av oppgavene. Når det gjelder riktige svar, ser vi at kategori 10 utgjør en betydelig større andel enn kategori 11. Blant elevene som har formulert riktig uttrykk, er det altså veldig mange flere som har brukt symboler enn de som har brukt ord eller en blanding av ord og symboler.

Tabell 10.5. Fordelingen av elevsvar i de ulike kategoriene for de tre oppgavene

Svar-kategori	Beskrivelse	Elevsvar ME72110B	Elevsvar ME72106	Elevsvar ME72108B
10	Riktig uttrykk med symboler	91 (16,3 %)	77 (12,4 %)	90 (15,3 %)
11	Riktig uttrykk med ord/blanding	12 (2,2 %)	6 (1,0 %)	21 (3,6 %)
20	Feil uttrykk med riktige symboler/variabler	118 (21,2 %)	124 (19,9 %)	59 (10,0 %)
21	Feil uttrykk ekstra/irrelevante variabler eller ligning	77 (13,8 %)	141 (22,7 %)	128 (21,8 %)
22	Feil uttrykk formulert med ord/blanding	55 (9,9 %)	47 (7,6 %)	52 (8,8 %)
23	Svar gitt som tall/regnestykke	120 (21,5 %)	130 (20,9 %)	142 (24,1 %)
24	Andre/useriøse svar	84 (15,1 %)	97 (15,6 %)	96 (16,3 %)
	Totalt	557 (100 %)	622 (100 %)	588 (100 %)

Det er et lite avvik mellom andelen riktige svar i vår koding (kategori 10 og 11) og de offisielle resultatene fra TIMSS 2023 (tabell 10.1). Dette skyldes trolig at vurderingskriteriene, og hva som godkjennes som riktig svar, er noe annerledes i vårt kodeskjema enn i den offisielle kodeguiden som ble brukt i TIMSS 2023.

Sammenheng mellom type svar og prestasjoner

Tabell 10.6 viser gjennomsnittlig poengsum i matematikk og innenfor emneområdet algebra for elever ulike typer svar på de tre algebraoppgavene. På tross av de relativt høye standardfeilene for enkelte svarkategorier, noe som skyldes et lavt antall elever i disse kategoriene, er det tydelige forskjeller i gjennomsnittsresultater mellom de ulike kategorier. For elever som har gitt riktig svar, er resultatene i matematikk nær eller over 600 og 550 skalapoeng for henholdsvis kategori 10 og 11 (tabell 10.6). For kategoriene hvor elevene har formulert et feil uttrykk, med symboler, variabler eller ord (kategoriene 20, 21 og 22), er gjennomsnittsresultatene betydelig lavere og i overkant av 500 poeng. For kategoriene hvor elevene har gitt et tall, et regnestykke uten variabel, eller andre/useriøse svar (23 og 24) er resultatene fra 460 poeng og nedover for de tre oppgavene (tabell 10.6). Vi ser det samme mønsteret for gjennomsnittsresultatet i matematikk som for resultatet i algebra for de ulike kategoriene.

Tabell 10.6. Gjennomsnittlig poengsum i matematikk og algebra for elever med elevsvar i de ulike kategoriene (med standardfeil SE i parentes)

Svar-kategori	Beskrivelse	Gjennomsnittlig poengsum i matematikk (SE) / algebra (SE)		
		ME72110B	ME72106C	ME72108B
10	Riktig matematisk uttrykk formulert med symboler	596,53 (5,35) / 586,47 (7,78)	596,57 (7,75) / 592,43 (8,82)	588,99 (7,24) / 579,96 (8,01)
11	Riktig matematisk uttrykk formulert med ord eller en blanding av ord og symboler	615,2 (17,21) / 609,97 (17,69)	556,28 (19,78) / 540,51 (24,64)	561,63 (11,91) / 545,21 (15,21)
20	Feil matematisk uttrykk (med egnede symboler/variabler)	528,76 (6,99) / 512,08 (7,49)	531,96 (6,58) / 514,66 (7,06)	510,56 (13,15) / 497,36 (14,71)
21	Feil matematisk uttrykk med ekstra og/eller irrelevante variabler eller en ligning (likhetstegn)	525,10 (10,80) / 504,76 (11,43)	538,70 (6,15) / 527,30 (6,96)	527,58 (6,84) / 510,12 (7,14)
22	Feil matematisk uttrykk formulert som inneholder ord	523,02 (9,37) / 510,94 (8,88)	524,37 (9,67) / 512,37 (11,79)	524,63 (10,50) / 507,73 (12,52)
23	Svar gitt som et tall eller et regnestykke	457,51 (7,89) / 440,30 (8,06)	450,91 (7,14) / 431,60 (7,88)	460,57 (6,91) / 442,83 (6,57)
24	Andre svar / useriøse svar	447,22 (7,46) / 428,76 (8,91)	455,60 (7,95) / 438,75 (9,48)	459,91 (10,38) / 440,26 (9,24)

Note. Øvre rad viser resultater i matematikk og nedre rad viser resultater i algebra.

De syv kategoriene kan skilles i tre hovedgrupper:

- A) Uttrykk som generaliserer mønsteret på korrekt måte (kategori 10 og 11)
- B) Feil uttrykk, formulert med variabler og/eller ord (kategori 20, 21 og 22)
- C) Elevsvar uten generalisering (for eksempel tall), eller andre/useriøse svar (kategori 23 og 24)

Når vi slår sammen kategoriene i disse tre hovedgruppene blir mønsteret for resultatene i tabell 10.6 enda tydeligere. Elevene med riktige svar (A) har et gjennomsnittresultat på det som er omtalt som høyt kompetansenivå i TIMSS. For elever som har formulert et uttrykk (B), er gjennomsnittresultatet på middels høyt kompetansenivå, mens gjennomsnittresultatet for elever i hovedgruppe C er på lavt kompetansenivå (tabell 10.7).

Tabell 10.7. Gjennomsnittlig (Gj.) poengsum i matematikk for de tre hovedgruppene av elevsvar (med standardfeil (s.f.) i parentes og antall elever (N))

	Oppgaver					
	ME72110B		ME72106C		ME72108B	
	N	Gj.(s.f.)	N	Gj.(s.f.)	N	Gj.(s.f.)
Hovedgruppe						
A (svarkategorier 10 og 11)	n = 103	598,70 (5,44)	n = 83	593,74 (7,49)	n = 111	583,63 (6,40)
B (svarkategorier 20, 21 og 22)	n = 250	526,21 (5,19)	n = 312	534,11 (4,46)	n = 239	522,44 (5,44)
C (svarkategorier 23 og 24)	n = 204	453,17 (5,85)	n = 227	452,99 (5,43)	n = 238	460,30 (6,55)

Diskusjon og konklusjon

Algebra er krevende for elever i alle land, men resultater fra tidligere TIMSS-undersøkelser kan tyde på at dette har vært spesielt krevende for norske elever (Bergem, 2016; Mullis et al., 2020; von Davier et al., 2024). Dette kan se ut til å særlig gjelde det å håndtere symboler og formell matematisk notasjon (Naalsund, 2012; Pedersen, 2015). Resultatene fra de tre algebraoppgavene som er denne studien bygger på, kan også tyde på dette: Relativt få norske elever på 9. trinn klarer å formulere et korrekt matematisk uttrykk for å generalisere et mønster.

I læreplanen LK20 finner vi at det å kunne beskrive og generalisere mønstre algebraisk, er et av kompetansemålene på 8. trinn (Kunnskapsdepartementet,

2019). På bakgrunn av dette kunne man forvente at mer enn 15 prosent av de norske elevene på 9. trinn mestrer slike oppgaver. Samtidig er det ikke bare norske elever som strever med denne typen oppgaven, resultatene fra TIMSS 2023 viser at elever i Sverige og Finland også har tilsvarende lave løsningsprosenten. Og selv om man skal være forsiktig med å sammenligne de norske resultatene med land som er mer ulike i kultur og skolesystem enn de nordiske landene, er det interessant å merke seg at det finnes land hvor betydelig flere elever mestrer å løse denne typen oppgaver, for eksempel Japan, hvor halvparten av 8. trinnselevne løser de samme oppgavene korrekt. Med endringene som er gjort fra LK06 til LK20 med kjerneelementet «Abstraksjon og generalisering», og med å beskrive kompetansemål for hvert årstrinn, er det spennende å se hvorvidt dette kan gjøre overgangen fra aritmetikk til algebra enklere for norske elever.

Fordelingen av elevsvar i de ulike kategoriene kan tyde på at det er stor variasjon i elevenes forståelse for matematiske uttrykk. Dette støttes av forskjellene i gjennomsnittsresultater for elevene i de ulike kategoriene. Når elever gir svaret som en tallverdi (kategori 23) eller som en ligning (kategori 21), kan det være fordi eleven ikke ser på et matematisk uttrykk som et gyldig svarformat (Kieran, 2020). Samtidig er det interessant at elever som har prøvd å formulere et uttrykk med et formelt språk, selv om uttrykket er feil, presterer betydelig høyere enn de som gir svaret som en tallverdi (kategori 23). Dette tyder på at også relativt høytpresterende elever har mangelfull kunnskap om matematiske uttrykk og variabler. Feil bruk av symboler og variabler i matematiske uttrykk kan ha flere forklaringer. Det kan komme fra intuisjon eller gjetting, det kan komme av overføring fra andre symbolsystemer elevene kjenner til, eller det kan komme fra misoppfatninger som bygger på misvisende undervisningsmaterieell (MacGregor & Stacey, 1997). Ifølge MacGregor og Stacey (1997) er ikke slike feil nødvendigvis uttrykk for lav kognitiv utvikling, men heller gjennomtenkte forsøk på å forstå ny notasjon eller overføre mening fra andre kontekster. Dette understreker viktigheten av å forstå hvilke tanker og kunnskaper som ligger bak elevenes feil og å bruke dette som utgangspunkt for videre læring.

Våre funn tyder på at elever som tar i bruk variabler og symboler for å formulere matematiske uttrykk, selv om dette involverer feilaktig bruk, har relativt gode matematiske kunnskaper og ferdigheter. Dette kan tolkes som et tegn på at de er i ferd med å lære seg det som for dem er et nytt og formelt (matematisk) språk. Et viktig steg i denne læringsprosessen vil

være å gjøre elevene bevisste på de rollene bokstavsymboler kan ha i matematikk, for eksempel som merkelapper, konstanter, ukjente og parametere (Philipp, 1992).

Det er også verdt å merke seg at svært få av de norske elevene formulerer uttrykk med en blanding av ord og symboler (kategori 11), til tross for at forskning viser at elever ofte generaliserer mønstre på ulike måter (Kieran, 2020). Svarformatet på de tre oppgavene er derimot utformet slik at det legger opp til relativt korte svar, noe som kan tenkes å «oppfordre» elevene til skrive uttrykket med formell notasjon. Det er mulig flere elever ville lykkes med å generalisere mønsteret dersom de hadde fått mulighet til uttrykke seg mer fritt, for eksempel om svarformatet ga mulighet til mer tekst eller tegninger. Samtidig er det å kunne ta i bruk korrekt formell notasjon en viktig del av kunnskaper og ferdigheter i algebra og av matematisk kompetanse generelt (Niss & Højgaard, 2019). De elevene som har uttrykt et korrekt matematisk uttrykk på de tre oppgavene vi har analysert, og som dermed viser tegn til denne kompetansen, presterer også betydelig høyere enn gjennomsnittet i TIMSS 2023.

Kongelf (2015) har tidligere vist at algebra i norske lærebøker i stor grad fremstår som et isolert emne, og at det i liten grad legges opp til at algebraopplæringen bygger videre på kunnskaper fra tallære. Med LK20 legges det opp til at algebra skal innføres tidligere i skoleløpet enn i de tidligere læreplaner i Norge. Samtidig påpeker Kaput (2008) at det å innføre algebra på et tidligere stadium i læreplanen ikke alene vil løse algebra-problemet, men at dette også krever endringer i klasseromsundervisningen, vurderingsformer og i lærerutdanningene. Forhåpentligvis kan resultatene fra denne studien bidra til å belyse noen av utfordringene norske elever møter når de skal uttrykke seg med formelt matematisk språk. Vi håper resultatene vi ser i TIMSS 2023 innenfor emneområdet algebra, hvor de norske elevene viser fremgang sammenlignet med tidligere undersøkelser, er et tegn på at innføringen av algebra tidligere i utdanningsløpet gjør at flere elever lykkes i matematikkfaget.

Begrensninger og videre forskning

I denne studien har vi kun undersøkt elevsvarene på tre oppgaver, som er knyttet til en liten del av algebra-området. Dette gjør at vi i liten grad kan trekke konklusjoner ut over hva disse tre oppgavene måler. Samtidig styrker observasjon av tre lignende oppgaver gyldigheten av vår konklusjon og

gjør det mulig for oss å generalisere noen funn for denne typen oppgaver – å generalisere mønstre og lage matematiske uttrykk.

Vi har kun undersøkt elevenes skriftlige besvarelser på oppgavene, og vet derfor ingenting om tankeprosessen bak svarene. Dette begrenser våre funn til hva vi kan observere i de tilgjengelige svarene. Samtidig har utvalget av svar gjort det mulig for oss å trekke konklusjoner om norske 9. trinnselever generelt, ettersom utvalget i datamaterialet er tilstrekkelig stort og representativt. For å få mer innsikt i hva som ligger bak de ulike elevsvarene og de ulike kategoriene, og hva som gjør denne typen oppgaver krevende for elever, ville det vært interessant å inkludere et mindre utvalg av elever og gjøre en datainnsamling hvor vi undersøker elevenes atferd og tankeprosesser mens de løser denne typen oppgaver.

Forfatterbiografier

Andreas Pettersen er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han har en doktorgrad i matematikkdidaktikk og har jobbet med PISA-undersøkelsen siden 2017. Pettersen har utgitt flere artikler om matematikdidaktikk og vært redaktør for boka *Matematisk kompetanse. I dybden på resultater fra PISA 2022* (utgitt i 2024).

Jelena Radišić er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Der har hun vært siden 2016. Hennes arbeid er knyttet til forskning på elev-, lærer- og skolekarakteristika som påvirker elevenes resultater. Hun er spesielt opptatt av hvordan elevenes læring påvirkes av motivasjon, selvtillit i matematikk, akademiske følelser, læreres holdninger og undervisningspraksis. I sin forskning benytter Radišić en rekke ulike tilnærminger som inngår i kvantitativ, kvalitativ og blandet metodeforskning.

Referanser

- Asquith, P., Stephens, A. C., Knuth, E. J. & Alibali, M. W. (2007). Middle school mathematics teachers' knowledge of students' understanding of core algebraic concepts: Equal sign and variable. *Mathematical Thinking and Learning*, 9(3), 249–272. <https://doi.org/10.1080/10986060701360910>
- Bergem, O. K. (2016). Hovedresultater i matematikk. I T. Nilsen, O. K. Bergem & H. Kaarstein (Red.), *Vi kan lykkes i realfag* (s. 22–44). Universitetsforlaget. <https://doi.org/10.18261/97882150279999-2016-03>

- Blanton, M., Brizuela, B. M., Gardiner, A. M., Sawrey, K. & Newman-Owens, A. (2017). A progression in first-grade children's thinking about variable and variable notation in functional relationships. *Educational Studies in Mathematics*, 95(2), 181–202. <https://doi.org/10.1007/s10649-016-9745-0>
- Booth, J. L., McGinn, K. M., Barbieri, C. & Young, L. K. (2017). Misconceptions and learning algebra. I S. Stewart (Red.), *And the rest is just algebra* (s. 63–78). Springer. https://doi.org/10.1007/978-3-319-45053-7_4
- Bush, S. B. & Karp, K. S. (2013). Prerequisite algebra skills and associated misconceptions of middle grade students: A review. *The Journal of Mathematical Behavior*, 32(3), 613–632. <https://doi.org/10.1016/j.jmathb.2013.07.002>
- Cai, J. & Knuth, E. (2011). A global dialogue about early algebraization from multiple perspectives. I J. Cai & E. Knuth (Red.), *Early algebraization: A global dialogue from multiple perspectives* (s. vii–vxi). Springer.
- Cai, J., Ng, S. F., Moyer, J. C. & Knuth, E. (2011). Developing students' algebraic thinking in earlier grades: Lessons from China and Singapore. I J. Cai & E. Knuth (Red.), *Early algebraization* (s. 25–41). Springer. https://doi.org/10.1007/978-3-642-17735-4_3
- Carpenter, T. P. & Levi, L. (2000). *Developing conceptions of algebraic reasoning in the primary grades* (Rapport NCISLA-RR-00-2). University of Wisconsin.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Filer, K. L. & Chang, M. (2008). Peer and parent encouragement of early algebra enrollment & mathematics achievement. *Middle Grades Research Journal*, 3(1), Artikkel 23.
- Fishbein, B., Taneva, M. & Kowolik, K. (2025). *TIMSS 2023 user guide for international database*. Boston College, TIMSS & PIRLS International Study Center. <https://timss2023.org/data>
- Grønmo, L. S. (2018). The role of algebra in school mathematics. I G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt & B. Xu (Red.), *Invited lectures from the 13th International Congress on Mathematical Education* (s. 175–193). Springer.
- Grønmo, L. S. & Helgesen, R. (2018, 11. mai). Norge trenger algebra! *Aftenposten*. <https://www.aftenposten.no/meninger/debatt/i/qnMnz0/norge-trenger-algebra-liv-sissel-groenmo-og-rita-helgesen>
- Grønmo, L. S. & Olsen, R. V. (2006). TIMSS versus PISA: The case of pure and applied mathematics. I *Proceedings of the 2nd IEA International Research Conference*, Washington, DC. The International Association for the Evaluation of Educational Achievement.
- Heemsoth, T. & Heinze, A. (2014). The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instructional Science*, 42, 639–657. <https://doi.org/10.1007/s11251-013-9302-5>
- Hornburg, C. B., Devlin, B. L. & McNeil, N. M. (2022). Earlier understanding of mathematical equivalence in elementary school predicts greater algebra readiness in middle school. *Journal of Educational Psychology*, 114(3), 540–559. <https://doi.org/10.1037/edu0000683>
- Hunter, J. & Miller, J. (2022). Using a culturally responsive approach to develop early algebraic reasoning with young diverse learners. *International Journal of Science and Mathematics Education*, 20(1), 111–131. <https://doi.org/10.1007/s10763-020-10135-0>
- Käfer, J., Kuger, S., Klieme, E. & Kunter, M. (2019). The significance of dealing with mistakes for student achievement and motivation: Results of doubly latent multilevel analyses. *European Journal of Psychology of Education*, 34, 731–753. <https://doi.org/10.1007/s10212-018-0408-7>
- Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? I J. J. Kaput, D. W. Carraher & M. L. Blanton (Red.), *Algebra in the early grades* (s. 5–17). Routledge
- Kieran, C. (1992). The learning and teaching of school algebra. I D. Grouws (Red.), *Handbook of research on mathematics teaching and learning* (s. 390–419). Simon & Schuster.
- Kieran, C. (2006). Research on the learning and teaching of algebra. I A. Gutiérrez & P. Boero (Red.), *Handbook of research on the psychology of mathematics education: Past, present and future* (s. 23–49). Sense Publishers.

- Kieran, C. (2020). Algebra teaching and learning. I S. Lerman (Red.), *Encyclopedia of mathematics education* (s. 36–44). Springer.
- Kieran, C. (2022). The multi-dimensionality of early algebraic thinking: Background, overarching dimensions, and new directions. *ZDM–Mathematics Education*, 54(6), 1131–1150.
- Kingsdorf, S. & Krawec, J. (2014). Error analysis of mathematical word problem solving across students with and without learning disabilities. *Learning Disabilities Research & Practice*, 29(2), 66–74.
- Kirke-, utdannings- og forskningsdepartementet. (1996). *Læreplanverket for den 10-årige grunnskolen*. Nasjonalt læremiddelsenter.
- Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A. & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equivalence & Variable. *Zentralblatt für Didaktik der Mathematik*, 37(1), 68–76.
- Kongelf, T. R. (2015). Introduksjon av algebra i matematikkbøker for ungdomstrinnet i Norge. *NOMAD Nordic Studies in Mathematics Education*, 20(3–4), 83–109. <https://doi.org/10.7146/nomad.v20i3-4.148693>
- Kunnskapsdepartementet. (2006). *Læreplan i matematikk*. Fastsett som forskrift. Læreplanverket for Kunnskapsløftet 2006. <https://www.udir.no/kl06/MAT1-01#>
- Kunnskapsdepartementet. (2015). *Tett på realfag. Nasjonal strategi for realfag i barnehagen og grunnsopplæringen (2015–2019)*. <https://www.regjeringen.no/no/dokumenter/tett-pa-realfag/id2435042/>
- Kunnskapsdepartementet. (2019). *Læreplan i matematikk 1.–10. trinn (MAT01-05)*. Fastsett som forskrift. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/mat01-05>
- Kvasz, L. (2006). The history of algebra and the development of the form of its language. *Philosophia Mathematica*, 14(3), 287–317. <https://doi.org/10.1093/philmat/nkj017>
- Kaarstein, H., Lehre, A.-C., Radišić, J. & Rohatgi, M. A. (2024). *TIMSS 2023. Kortrapport*. Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo.
- Lee, S. W. & Mao, X. (2021). Algebra by the eighth grade: The association between early study of algebra I and students' academic success. *International Journal of Science and Mathematics Education*, 19(6), 1271–1289. <https://doi.org/10.1007/s10763-020-10116-3>
- Leung, F. K. S., Clarke, D., Holton, D. & Park, K. (2014). How is algebra taught around the world? I F. K. S. Leung, D. Clarke, D. Holton & K. Park (Red.), *Algebra teaching around the world* (s. 1–15). Sense Publishers. https://doi.org/10.1007/978-94-6209-707-0_1
- Mason, J. (1996). Expressing generality and roots of algebra. I J. Mason, L. Lee, C. Kieran & N. Bednarz (Red.), *Approaches to algebra* (s. 65–86). Springer Netherlands. https://doi.org/10.1007/978-94-009-1732-3_5
- MacGregor, M. & Stacey, K. (1997). Students' understanding of algebraic notation: 11–15. *Educational Studies in Mathematics*, 33(1), 1–19. <https://doi.org/10.1023/A:1002970913563>
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/international-results/>
- Niss, M. & Højgaard, T. (2019). Mathematical competencies revisited. *Educational Studies in Mathematics*, 102, 9–28.
- Naalsund, M. (2012). *Why is algebra so difficult? A study of Norwegian lower secondary students' algebraic proficiency* [Doktorgradsavhandling]. Universitet i Oslo.
- O'Brien, R., Pan, X., Courville, T., Bray, M. A., Breaux, K., Avitia, M. & Choi, D. (2017). Exploratory factor analysis of reading, spelling, and math errors. *Journal of Psychoeducational Assessment*, 35(1–2), 7–23.
- OECD. (2023). *PISA 2022 results: Vol. I. The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Pedersen, I. F. (2015). What characterizes the algebraic competence of norwegian upper secondary school students? Evidence from timss advanced. *International Journal of Science and Mathematics Education*, 13(Suppl 1), 71–96. <https://doi.org/10.1007/s10763-013-9468-y>

- Philipp, R. A. (1992). The many uses of algebraic variables. *The Mathematics Teacher*, 85(7), 557–561. <https://doi.org/10.5951/MT.85.7.0557>
- Philpot, R., Lindquist, M., Mullis, I. V. & Aldrich, C. E. (2021). TIMSS 2023 mathematics framework. I I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 Assessment Frameworks* (kap. 1). Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- Riccomini, P. J. (2005). Identification and remediation of systematic error patterns in subtraction. *Learning Disability Quarterly*, 28(3), 233–242. <https://doi.org/10.2307/1593661>
- Stacey, K. & MacGregor, M. (1997). Ideas about symbolism that students bring to algebra. *The Mathematics Teacher*, 90(2), 110–113. <https://doi.org/10.5951/MT.90.2.0110>
- Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K. & Gogolen, C. (2015). Learning from comparison in algebra. *Contemporary Educational Psychology*, 40, 41–54.
- Tulis, M. (2013). Error management behavior in classrooms: Teachers' responses to student mistakes. *Teaching and Teacher Education*, 33, 56–68. <https://doi.org/10.1016/j.tate.2013.02.003>
- Vinje, K. (2023, 25. juli). *Det norske matteproblemet*. NOKUT. <https://www.nokut.no/nokut-bloggen/det-norske-matteproblemet/>
- von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U. & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs6460>
- Wasserman, N. H. (2016). Abstract algebra for algebra teaching: Influencing school mathematics instruction. *Canadian Journal of Science, Mathematics and Technology Education*, 16(1), 28–47.
- Warren, E., Trigueros, M. & Ursini, S. (2016). Research on the learning and teaching of algebra. I Á. Gutiérrez, G. C. Leder & P. Boero (Red.), *The second handbook of research on the psychology of mathematics education*. Sense Publishers. https://doi.org/10.1007/978-94-6300-561-6_3
- Wischgoll, A., Pauli, C. & Reusser, K. (2015). Scaffolding – how can contingency lead to successful learning when dealing with errors? *ZDM Mathematics Education*, 47, 1147–1159. <https://doi.org/10.1007/s11858-015-0714-3>
- Wu, H. (2001). How to prepare students for algebra. *American Educator*, 25(2), Artikkel 10.
- Zhang, D., Ding, Y., Barrett, D. E., Xin, J. P. & Liu, R. (2014). A comparison of strategic development for multiplication problem solving in low-, average-, and high-achieving students. *European Journal of Psychology of Education*, 29, 195–214. <https://doi.org/10.1007/s10212-013-0194-1>

KAPITTEL 11

Hva viser grafen? Elevers statistiske kompetanse og kunnskap om vær og klima

Cecilie Weyergang, Hege Kaarstein, Fredrik Jensen og Anne-Catherine Lehre

Institutt for Lærerutdanning og skoleforskning, Universitetet i Oslo

Sammendrag: I denne artikkelen undersøker vi elevers svar på naturfagoppgaver som krever forståelse av statistikk. Mye av den naturfaglige informasjonen vi i dag må forholde oss til, formidles gjennom visuelle representasjoner og statistiske framstillinger. Dette gjelder for eksempel når vi står overfor problemstillinger og tekster knyttet til miljø og klima. Vi tar utgangspunkt i teoretiske perspektiver fra statistisk literacy-feltet og undersøker 9. trinnelevers svar på to oppgaver i TIMSS, som begge omhandler en grafisk framstilling med flere variabler, beskrevet som en klimaopplysning. I studien kombinerer vi ulike metoder. I steg 1 har vi gjennomført en innholdsanalyse av 896 skriftlige elevsvar, og kategorisert disse. I steg 2 har vi gjort høyttenkning med 12 elever på 9. trinn som løser de samme oppgavene. Høyttenkningen er gjennomført både for å få dypere innsikt i elevenes svar og for å validere kategoriene vi etablerer i steg 1 av analysen. Funnene viser at flere elever ser ut til å ha utfordringer med å tolke informasjon og framstillinger som strider mot deres forventninger og forhold i den konteksten de selv befinner seg i. Det kan også se ut til at flere elever har utfordringer med å skille mellom begrepene vær og klima. I både de skriftlige elevsvarene og i høyttenkningen ser det ut til at flere elever har problemer med å forstå en grafisk framstilling med flere variabler.

Nøkkelord: statistisk literacy, vær, klima, 9. trinn, TIMSS

Abstract**What Does the Graph Show? Insights into Students' Statistical Competence and Knowledge of Weather and Climate from TIMSS 2023**

In this article, we investigate students' response patterns to science tasks that require an understanding of statistics. Much of the scientific information we have to deal with today is disseminated through statistics and visual representations. This applies, for example, when we are faced with issues and texts related to the environment and climate. Based on theoretical perspectives from the field of statistical literacy, we investigate 9th grade students' responses to two tasks in TIMSS, both of which deal with a graphical representation, described as climate information. In the study, we combine different methods. In step 1, we conducted a content analysis of 896 written student responses and categorized them. In step 2, we conducted think-alouds with 12 students in 9th grade who solved the same tasks. This was done both to gain deeper insight into the students' response patterns and to validate the categories we established in step 1 of the analysis. The findings show that several students seem to have challenges interpreting information and representations that contradict their expectations and conditions in the context in which they find themselves. It also appears that many students have difficulty distinguishing between the terms weather and climate. In both the written student responses and the thinking aloud, it seems that several students have difficulty understanding a graphical representation with several variables.

Keywords: statistical literacy, weather, climate, grade 9, TIMSS

Innledning

I dagens teknologiske og datadrevne samfunn må vi kunne forholde oss til statistiske framstillinger. Dette er avgjørende for å kunne delta som aktive samfunnsborgere (Gal, 2004; Sharma, 2017; Tishkovskaya & Lancaster, 2012; Weiland, 2017). Mange av tekstene vi leser daglig, tar utgangspunkt i tallmateriale. I nyhetsartikler fra redaktørstyrte aviser, i innlegg på sosiale medier, i populærvitenskapelige artikler og i tekster med kommersielle avsendere brukes tall og statistikk til å underbygge argumentasjon, til å informere oss og til å overbevise (Sharma, 2017). Tall brukes aktivt innenfor politikk for å gi ulike påstander tyngde og troverdighet, men det hender at politiske ledere kaster ut tall og statistikk som ikke tåler å bli ettergått (Bendavid, 2025). Noen ganger kan statistiske framstillinger også vilde oss (Irving et al., 2022; Johannssen et al., 2021). Å kunne forstå og analysere hvordan ulike data genereres og framstilles er en sentral del av naturfaglig kompetanse (Osborne, 2023), hvilket også framheves i den norske læreplanen. Grafiske framstillinger brukes ofte til å integrere og representere kompleks informasjon (Mathai et al., 2025), og kan bidra til å gi et enklere bilde av sammensatte naturfaglige temaer som hvordan fenomener i naturen henger sammen, eller på hvilken måte klimaet endrer seg. Samtidig er slike framstillinger aldri en direkte gjengivelse av virkeligheten, og de må tolkes av leseren (Harold et al., 2016). For å forstå en grafisk framstilling må leseren både ha statistisk kunnskap og kunnskap om temaet som framstillingen omhandler. I denne artikkelen undersøker vi hvordan elever svarer på to oppgaver fra den internasjonale undersøkelsen *Trends in International Mathematics and Science Study* (TIMSS) 2023. Oppgavene krever at elevene har statistisk kompetanse i tillegg til grunnleggende forståelse av de naturfaglige begrepene vær og klima.

Bakgrunn og forskningsspørsmål

Studien vi presenterer i denne artikkelen, tar utgangspunkt i tidligere forskning om elevers forståelse av statistikk og deres kunnskap om begrepene vær og klima. I tillegg må studien forstås i lys av føringer i den norske læreplanen og internasjonale rammeverk for naturfaglig og matematisk kompetanse, som rammeverkene til TIMSS og PISA.

Tidligere forskning

Forståelse av statistikk innebærer både å kunne lese av grafer og diagrammer, men også å kunne reflektere kritisk over den informasjonen slike representasjoner formidler og forstå hva slags slutninger og konklusjoner vi kan trekke ut fra dem (Gal, 2004). Flere studier har vist at det er krevende for elever å forstå statistiske framstillinger (Glazer, 2011). Det kan se ut til at elever har større utfordringer med å lese og forstå ikke-kontinuerlige tekster som er satt sammen av flere elementer, som verbaltekst i kombinasjon med grafiske framstillinger, enn rene kontinuerlige tekster (Pérez-Echeverría et al., 2018). Det viser seg at elever på både mellomtrinn og ungdomstrinn har lett for å forstå et stolpediagram med nominale data, men har utfordringer med å tolke linjediagrammer og integrere informasjon fra flere variabler (Mathai et al., 2025). Det kan også se ut til at elever stoler på argumentasjon som tar utgangspunkt i tall, og at de i liten grad stiller spørsmål ved statistiske framstillinger (McGrew et al., 2018). Forskning har i tillegg vist at elever har utfordringer med å knytte statistiske framstillinger sammen med naturfaglige konsepter, særlig når de blir bedt om å stille seg kritiske til innholdet som grafen formidler (Lai et al., 2016).

«Klimaendringer» er eksempel på et sentralt konsept som ofte formidles gjennom statistikk og visuelle representasjoner (Harold et al., 2016; Schauss et al., 2024). Flere internasjonale utdanningsstudier innenfor feltet klima-literacy tar for seg elevers forståelse av klimaendringer og utforsker egnede undervisningsstrategier for å styrke elevers kunnskap om temaet (Wildbichler et al., 2025). Elever vil møte visuelle representasjoner som informerer om klimaendringer både i og utenfor skolen, og for å kunne engasjere seg i miljøutfordringer verden i dag står overfor, er det avgjørende at de forstår slike framstillinger. Denne forståelsen krever igjen at de forstår kjernebegreper som vær og klima (Da Rosa, 2022). Klima kan beskrives som værstatistikk og inkluderer både gjennomsnittlig vær og typiske variasjoner. Mens været sier noe om forhold i atmosfæren over korte tidsspenn som timer, dager, uker og måneder, beskriver klimaet de samme forholdene, men som et gjennomsnitt basert på observasjoner over mange år (Benestad et al., 2018; Cartwright et al., 2021). Internasjonal forskning har vist at elever på både ungdomsskole- og videregående nivå har utfordringer med å forstå forskjellen på vær og klima, og at de bruker begrepene om hverandre (Cartwright et al., 2021; Dawson, 2015; Wildbichler et al., 2025). I en undersøkelse gjennomført blant elever på videregående nivå, viste det seg at elevene hadde en forståelse av at klima endrer seg saktere

enn været. Likevel var det kun en liten andel som forsto at klima refererer til værforholdene i et område, og i hvilken grad disse forholdene endrer seg over flere tiår (Bodzin et al., 2014). Flere andre studier har vist at også studenter og voksne kan ha vansker med å skille mellom begrepene vær og klima (Da Rosa, 2022; Lombardi & Sinatra, 2012).

Den norske læreplanen og internasjonale rammeverk for matematisk og naturfaglig kompetanse

Forståelse av statistikk trekkes fram som en sentral kompetanse i de norske læreplanene i ulike fag. I naturfag heter det at elevene skal identifisere, tolke og bruke informasjon fra ulike teksttyper og vurdere kritisk hvordan naturvitenskapelig informasjon framstilles. Videre innebærer lesing i naturfag å bruke informasjon fra ulike tekster og forstå tekster med stadig mer avanserte fagbegreper, figurer og tabeller. I kompetansemål etter 10. trinn heter det at elevene skal «analysere og bruke innsamlede data til å lage forklaringer, drøfte forklaringene i lys av relevant teori og vurdere kvaliteten på egne og andres utforskinger» (Kunnskapsdepartementet, 2019b). Også i matematikk framheves statistisk kompetanse. Under det tverrfaglige temaet «Demokrati og medborgerskap» heter det at elevene skal kunne analysere funn fra reelle datasett og tallmateriale fra natur, samfunn, arbeidsliv og hverdagsliv. Det poengteres at dette er grunnleggende for at de skal kunne delta i ulike samfunnsdebatter (Kunnskapsdepartementet, 2019a). Statistisk kompetanse kommer også til uttrykk i andre fag enn realfagene. I samfunnsfag heter det at elevene selv skal kunne samle inn data gjennom ulike metoder, presentere funn og se styrker og svakheter ved eget forskningsdesign (Kunnskapsdepartementet, 2019c).

Siden oppgavene fra TIMSS som denne studien baserer seg på handler om vær og klima, er det relevant å se på hvordan kunnskap om begrepene kommer til uttrykk i læreplanen. Begrepet klima nevnes få ganger i læreplanen for grunnskolen. Under det tverrfaglige temaet «Bærekraftig utvikling» heter det i læreplanen for naturfag at elevene skal kunne ta miljøbevisste valg og se disse opp mot *klimautfordringer*, videre skal de kunne kjenne til løsninger for å begrense *klimautfordringer*. Etter kompetansemål for 10. trinn skal elevene kunne gjøre rede for faktorer som kan forårsake globale *klimaendringer* (Kunnskapsdepartementet, 2019b). Det er ingen formuleringer som presiserer at elevene skal forstå hva som ligger i klimabegrepet. Det er i denne sammenheng også interessant å se til læreplanen

i samfunnsfag, for her er også klimabegrepet nevnt. I kompetansemål etter 2. trinn heter det at elevene skal «kunne utforske og gi døme på korleis menneska påverkar klimaet og miljøet, og dokumentere korleis påverkningane kjem til syne i nærmiljøet» (Kunnskapsdepartementet, 2019c). Men heller ikke her vektlegges forståelse av hva klima faktisk er. Dette krever at lærere selv forstår at de må arbeide med denne typen grunnleggende fagbegreper. Begrepet vær er kun nevnt én gang i kompetansemålene i læreplanen i naturfag. Etter 2. trinn heter det at elevene «skal kunne planlegge og gjennomføre undersøkelser av vær og himmelfenomener og sammenligne målinger, observasjoner og værtegn gjennom året».

Det er også relevant å se vår studie i lys av internasjonale rammeverk for matematikk og naturfag. I rammeverket til TIMSS poengteres det at elevene lever i et samfunn hvor de må kunne forholde seg til store mengder kvantitativ informasjon. Elevene trenger å kunne organisere og representere data, men også å kunne tolke og trekke slutninger fra representasjoner som tabeller og grafer. Videre heter det at elevene skal kunne bruke kunnskap om naturfaglige konsepter til å tolke både rene verbaltekster, tabeller, bilder og grafisk informasjon (Centurino & Kelly, 2021).

I faglige rammeverk for andre internasjonale storskalaundersøkelser som Norge deltar i, framheves også forståelse av statistikk i sammenheng med naturfaglig kompetanse. Et av hovedområdene i naturfagsrammeverket til PISA-undersøkelsen handler om å evaluere naturvitenskapelige forskningsdesign og tolke data som blir innhentet, på en kritisk måte. I tillegg framheves evnen til å vurdere hvordan observasjoner leder til ulike modeller som kan forutsi hvordan naturfaglige fenomener kan utvikle seg (OECD, 2023).

Studiens forskningsspørsmål

På bakgrunn av tidligere forskning og føringer i den norske læreplanen og internasjonale rammeverk for naturfag og matematikk, er det behov for mer kunnskap om hvordan elever tolker og forstår statistikk og grafiske framstillinger, særlig i tilknytning til kompleks naturfaglig tematikk som vær og klima. Denne studien tar derfor utgangspunkt i følgende forskningsspørsmål:

Hva kjennetegner elevers svar på oppgaver der klimaopplysninger er gitt i en grafisk framstilling?

Vi utforsker 9. trinnselevers svar på to oppgaver fra TIMSS-undersøkelsen. Vi undersøker både elevers skriftlige svar ($n = 896$) og vi

har gjennomført høyttenkning med et mindre utvalg elever ($n = 12$) som besvarer de samme oppgavene. Videre presenterer vi teoretiske perspektiver som ligger til grunn for studien.

Teoretiske perspektiver

Funnene i denne studien belyses fra det teoretiske feltet statistisk literacy. I tillegg knytter vi elevenes kunnskap om vær og klima til forskningsfeltet *conceptual change*, som handler om forståelse av naturfaglige fenomener og begreper.

Statistisk literacy

Det er ikke fullstendig enighet i forskningsfellesskapet om hvordan begrepet statistisk literacy, eller på norsk statistisk kompetanse, kan defineres og operasjonaliseres. Wallmann (1993) skriver at statistisk literacy handler om å forstå og reflektere over statistiske resultater, og samtidig kunne anerkjenne verdien slike framstillinger kan ha, både privat og i offentlige sammenhenger. Gal (2004) beskriver kompetansen som evnen til å tolke og kritisk evaluere statistisk informasjon og datarelaterte argumenter som opptrer i ulike kontekster. I tillegg handler det om å kunne diskutere egne oppfatninger av framstillingene, og reflektere over hva slags implikasjoner de kan ha. Ifølge Gals rammeverk (2004) består statistisk literacy av fem kognitive elementer som henger tett sammen: statistisk kunnskap, matematisk kunnskap, kjennskap til konteksten for framstillingen, leseferdigheter og evnen til å kunne stille kritiske spørsmål. Rammeverket knytter seg til det nivået av statistisk literacy som er forventet for å kunne forstå framstillinger i ulike samfunnskontekster. Videre gir vi en mer utfyllende beskrivelse av elementene i rammeverket.

For å kunne forstå statistiske framstillinger kreves det for det første en *statistisk kunnskapsbase* der man har kjennskap til begreper relatert til deskriptiv statistikk som gjennomsnitt, median og prosent (Gal, 2004). Videre innebærer en statistisk kunnskapsbase at man forstår hvorfor data trengs i ulike sammenhenger og hvordan data genereres. Det forutsettes også en bevissthet om grunnleggende begreper knyttet til grafiske framstillinger som x-akse og y-akse, og hvordan ulike data i et diagram eller en tabell er organisert. I tillegg forutsettes en grunnleggende forståelse av hva som menes med *sannsynlighet*, og av hvordan man kommer fram til statistiske konklusjoner.

Til grunn for den statistiske kompetansen ligger det videre *matematisk kunnskap* og en grunnleggende forståelse av termer og begreper innenfor matematikk (Gal, 2004). Dette innebærer blant annet kjennskap til hvordan man kommer fram til ulike prosentandeler av et utvalg.

Statistisk literacy innebærer videre *kjennskap til konteksten* som den statistiske framstillingen, for eksempel et diagram, opptrer innenfor (Gal, 2004). Også innenfor leseforskningen er det en etablert oppfatning om at forkunnskaper er avgjørende i møte med ulike tekster (Cervetti & Wright, 2020). For at elevene skal forstå grafen om klimaopplysninger i denne studien, må de faktisk vite hva klimaopplysninger er, som igjen krever at de vet hva klima er. Dette er avgjørende for at de skal forstå hva slags slutninger som kan trekkes ut fra dataene som presenteres.

Videre vil så å si alle statistiske framstillinger kreve *leseferdigheter* (Gal, 2004). I noen tilfeller opptrer statistiske framstillinger i sammenheng med kortere tekstpassasjer. Andre ganger kan den statistiske framstillingen være et element i lengre resonnementer. Å forstå grafer handler ofte om å lese og tolke lengre tekstpassasjer og reflektere over overordnede aspekter ved teksten. Man må blant annet kunne identifisere hensikten med teksten som den statistiske framstillingen er en del av. De som har laget teksten, er ikke nødvendigvis interessert i å gi en objektiv eller balansert framstilling, selv om de viser til statistikk og tallmateriale.

Selv om det er ulikheter i rammeverk for statistisk literacy, er det likevel enighet om at det innebærer mer enn å lese av en graf eller et diagram riktig, og at *evnen til å stille kritiske spørsmål* er en sentral del av kompetansen (Gal, 2004; Koga, 2022). Flere rammeverk beskriver statistisk literacy taksonomisk (Sharma, 2017). På de grunnleggende nivåene i taksonomien kan man lese av informasjon i enkle grafer, men ha utfordringer med å forstå komplekse framstillinger og sammenligne ulike datasett. På høyere nivåer kan man reflektere kritisk over hvordan dataene er samlet inn, utvalget som ligger til grunn for framstillingen og hvilke slutninger som kan trekkes ut fra dataene. Avansert kompetanse innebærer også å identifisere og forklare usikkerheten som ligger i dataframstillingen og komme med alternative tolkninger av resultatene som presenteres.

Begrepsforståelse i naturfag

Elevers forståelse av statistiske framstillinger er påvirket av hvilke forkunnskaper de har om det temaet framstillingen skal si noe om, og i møte

med oppgaven fra TIMSS vil det være relevant å ha en naturfaglig forståelse av klimabegrepet. Elevers forestillinger om naturfaglige begreper og fenomener er et sentralt tema innen naturfagdidaktikk med lange tradisjoner (Krajcik & Shin, 2023). Driver (1989) skriver at elever kommer til naturfagundervisningen med en forforståelse som styrer hvordan de tolker konsepter og fenomener de skal lære. Ofte kan forforståelsene avvike og stå i motsetning til hvordan konseptene beskrives og forklares innenfor den naturfaglige tradisjonen. Forståelsene elevene har med seg, kan være seiglivet og vanskelige å endre. Dette kan forstås i lys av forskningsfeltet *conceptual change* (Da Rosa, 2022). Haug og Mork (2021) framhever viktigheten av å arbeide spesifikt med fagbegreper eller det de betegner som *nøkkelbegreper* i naturfag, for at elever skal forstå hvordan begrepene benyttes i ulike faglige sammenhenger. Man kan argumentere for at vær og klima er nøkkelbegreper innenfor temaer som miljø og bærekraft. Elever kjenner til begrepene vær og klima, men kommer ikke nødvendigvis til naturfagundervisningen med en faglig forståelse av hva begrepene betyr. Det er avgjørende at de får justert sine forforståelser av konseptene for at de skal kunne reflektere over ulike faglige problemstillinger (Da Rosa, 2022). Slik vi viste til tidligere, har flere studier pekt mot at elever har utfordringer med å skille vær og klima fra hverandre. Det kan være flere årsaker til dette, men en årsak kan være at elevene møter begrepene i ulike sammenhenger. Informasjon og diskusjoner om klima opptrer i ulike kilder som elever har tilgang til både i og utenfor skolen: i læreverk skrevet av naturfagdidaktikere, i avisartikler skrevet av journalister og i innlegg fra influensere i sosiale medier (Wildbichler et al., 2025). Disse kildene kan variere mye i innhold og framstillingsform, og flere av dem kan ligge langt fra den naturfaglige tradisjonen.

Metode

For å besvare forskningsspørsmålet i dette kapitlet benyttes ulike metoder og analyser, og studien er gjennomført i to steg. Begge stegene tar utgangspunkt i to oppgaver for 9. trinn fra TIMSS 2023.

Beskrivelse av studiens to steg

I det første steget har vi gjort en induktiv innholdsanalyse og kategorisert norske elevsvar på de to oppgavene. I det andre steget gjennomføres

høyttenkning med utgangspunkt i de samme oppgavene med et mindre utvalg elever på 9. trinn. Steg 1 gir oss et bredt bilde av elevenes svar, mens vi i steg 2 går dypere inn i kategoriene vi etablerer i steg 1 og forsøker å forstå hvorfor elevene svarer som de gjør. Steg 2 bidrar i tillegg til å validere kategoriene vi etablerer i innholdsanalysen fra steg 1. I denne sammenheng er det viktig å presisere at utvalget i steg 2 på ingen måte er representativt for hvordan elever som har deltatt i TIMSS 2023 tenker når de svarer, men høyttenkningen kan gi oss noen indikasjoner på hvorfor elever svarer som de gjør. For mer informasjon om TIMSS-undersøkelsen, se innledningskapittelet (Kaarstein og Lehre).

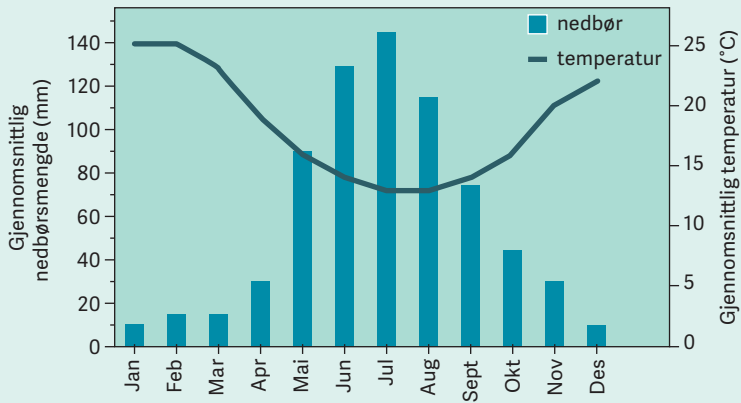
Materiale

Vi har tatt utgangspunkt i to oppgaver med samme kontekst.

Alle oppgaver som inkluderes i TIMSS-undersøkelsen for å måle elevenes kompetanse i matematikk og naturfag, er laget ut fra rammeverkene for undersøkelsen (Mullis et al., 2021). De to oppgavene vist i figur 11.1, er kategorisert som naturfagoppgaver innen emneområdet geofag. Videre er oppgave A kategorisert som en oppgave som krever at elevene anvender kunnskapene sine, mens oppgave B er kategorisert som en resonneringsoppgave der elevene blir bedt om å begrunne hvorfor de krysser av for «Ja» eller «Nei». Oppgavenes kontekst er den samme: Sebastian og familien skal på ferie, og de bruker klimaopplysninger, gitt i form av et diagram, som utgangspunkt for valg av tidspunkt for reisen. Det er ikke oppgitt i oppgaven hvilket land klimaopplysningen gjelder for.

Selv om oppgavene er laget for å måle elevenes kompetanse i naturfag, krever de i tillegg det Gal (2004) beskriver som *statistical literacy*. Som tidligere nevnt presenteres mye naturfaglig informasjon gjennom visuelle representasjoner, slik som diagrammet i figur 11.1. Vi gjenfinner flere av de kognitive elementene fra Gals (2004) rammeverk i oppgavene. For å lese og tolke diagrammet og komme fram til januar, som er riktig svar på oppgave A, må elevene ha både matematisk og statistisk kunnskap. Elevene må kombinere grafene i diagrammet både med riktig nøkkel og y-akse. Det er avgjørende at de leser av grafenes verdier på riktig y-akse. Videre må de finne den måneden som oppfyller begge kravene til familien. Måneden skal være både den varmeste og den tørreste. For å få rett svar på oppgave B, må elevene ha kunnskap om statistiske begreper.

Sebastian og familien hans ville dra på ferie på den varmeste og tørreste tiden av året. For å planlegge turen brukte de klimaopplysningene vist under.



A. Hvilken måned dro Sebastian og familien hans på ferie?

B. Men været var ikke varmt og tørt den uken de var på ferie, det regnet mye og var kaldt. Sebastian konkluderte med at klimaopplysningene han brukte, var gale.

Er Sebastians konklusjon riktig?

(Velg ett svar.)

☐ Ja

☐ Nei

Forklar svaret ditt.

Figur 11.1. Oppgave A og B som ligger til grunn for analysen i dette kapittelet.¹

¹ Oppgave SE62211A og SE62211B, TIMSS 2023 Copyright © 2025. Gjengitt med tillatelse fra TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College.

At begge y-aksenes enheter i diagrammet handler om gjennomsnitt per måned (for temperatur og nedbør), vil kunne gi en elev med kunnskap om begrepet gjennomsnitt, et hint om løsningen. Videre må de ha kjennskap til konteksten som den grafiske framstillingen opptrer innenfor. De må vite hva klima og klimaopplysninger er for å kunne velge «nei» og gi en forklaring som viser at de har forstått at været er det vi møter fra dag til dag, mens klima handler om gjennomsnitt av været for et område over en lang periode. Ut ifra kunnskaper om klima, må de altså innta en kritisk holdning til den konklusjonen som Sebastian trekker.

Prosedyre og analyse steg 1 (innholdsanalyse av elevsvar)

I steg 1 av studien analyserte vi elevsvar fra gjennomføringen av TIMSS 2023. I TIMSS 2023 er det til sammen 218 naturfagoppgaver, og oppgavene fordeles på 14 mindre sett. Elevene får ett oppgavesett hver (tilsvarende for matematikk, Yin & Foy, 2021). Av de 6 324 elevene som deltok i TIMSS 2023, er det derfor 896 elever som har fått de to utvalgte oppgavene, vist i figur 11.1, i sitt oppgavesett i naturfag. Av de 896 elevene, er det 820 som har svart på oppgave A og 843 som har svart på oppgave B. Av de 843 elevene som har svart på oppgave B, har 436 formulert en begrunnelse i tillegg til å krysse av for «ja» eller «nei».

Analysen av elevenes besvarer på oppgave B ble gjennomført med en induktiv tilnærming. I metodelitteraturen betegnes dette også som *konvensjonell innholdsanalyse* der det empiriske materialet styrer kategoriseringen, i motsetning til at man benytter et på forhånd definert analyserammeverk (Fauskanger & Mosvold, 2014; Hsieh & Shannon, 2005). I innholdsanalyse kan man kategorisere tekstmaterialet etter det bokstavelige innholdet med liten grad av abstraksjon, eller man kan etablere videre kategorier der man i større grad gjør tolkninger (Lindgren et al., 2020). I denne analysen har vi etablert kategorier på bakgrunn av våre tolkninger av elevsvarene. Hver forfatter leste gjennom elevsvarene og kom med forslag til svarkategorier før disse ble diskutert, og det første utkastet til analyserammeverk ble satt opp. Vi ble i starten av arbeidet enige om å bruke kodeguiden fra TIMSS for å identifisere de riktige svarene i materialet. Kategorier for feilsvar har vi utviklet selv. Som en test av analyserammeverket, kategoriserte alle forfatterne 60 utvalgte elevsvar. Uenigheter i kategoriseringen for de 60 svarene ble oppklart, og kategoriene i analyserammeverket justert. De 60 elevsvarene fungerte videre som eksempler på de valgte kategoriene. Det endelige

analyserammeverket er vist i tabell 11.1. De resterende elevbesvarelsene ble fordelt likt mellom alle forfatterne, med en overlappt på ti prosent av svarene. Med nominale svarkategorier og flere enn to kodere ble enighet i kategorisering mellom forfatterne sjekket med Fleiss-Kappa (Fleiss, 1971). Uenigheter i kategoriseringen ble diskutert og løst. De elevsvarene hver av forfatterne fremdeles var usikre på, ble også diskutert i fellesskap før endelige kategorier ble tildelt. I kodingen av elevsvarene nådde vi et samsvar på Fleiss-Kappa 0,68. I Landis og Kochs (1977) beskrivelse av Fleiss Kappa-verdier, antydes det at en verdi over 0,2 er «fair», 0,4 er «moderate», 0,6 er «substantial» og over 0,8 er nesten perfekt.

Tabell 11.1. Analyserammeverk for kategorisering av elevbesvarelser på oppgave B

Kategori	Eksempler
Riktig svar (basert på kodeguiden i TIMSS)	«[Nei] Siden det kun er gjennomsnittlig høy varme og lite nedbør, betyr det ikke at det alltid er sånn.» «[Nei] Tabellen viser gjennomsnitt, så det kan være litt forskjellig fra år til år. De var bare uheldig med når de dro.»
Forveksler vær og klima	«Ja, siden det ble helt motsatt vær enn værmeldingen hadde sakt.» «[Ja] det er riktig fordi det er ikke alltid værmeldingene viser riktig.» «[Nei] du vet aldri nøyaktig når det skal bli dårlig vær»
Norsk kontekst	«[Ja] Det er lite sannsynlig at det er 25 grader om vinteren og at sommeren er den kaldeste årstiden.» «[Ja] Det er ikke logisk at det skal være varmest og tørrest midt på vinteren.»
Eleven mener det er en feil i grafen: Temperatur og nedbør er byttet om	«[Ja] Fordi jeg tror grafene har byttet om på navnene og det som står i grafen er veldig usannsynlig.» «[Ja] nedbør og temperatur burde bytte plass»
Eleven leser grafen feil: forveksler temperatur og nedbør	«[Ja] det sto at det skulle være drit varmt»
Vage svar	«[Ja] han burde byttet måned» «[Ja] da har han regnet feil» «[Ja] de dro på feil ferie sted»
Annet	Tull og tøys, eller eleven skriver «vet ikke»

Igjenn er det viktig å presisere at kategoriene baserer seg på våre tolkninger. I steg 1 av analysen har vi eksempelvis kommet fram til en kategori som har fått betegnelsen «forveksler vær og klima». Her har vår egen tolkning påvirket kategoriseringen, og det er viktig å presisere at vi ikke kan vite med sikkerhet at svarene er et uttrykk for at elevene ikke vet forskjell på vær og klima.

Prosedyre og analyse steg 2 (høyttenkning)

I studiens andre steg gjennomførte vi høyttenkning med elever, og vi rekrutterte 12 elever på 9. trinn ved å benytte kontakter i eget nettverk. Alle elevene går på samme skole i det sentrale østlandsområdet. Elevene hadde ikke deltatt i TIMSS-undersøkelsen og hadde ikke sett oppgavene tidligere. Vi ba vår lærerkontakt om å rekruttere så mange elever som mulig uten å ta hensyn til elevenes kjønn, karakter eller kompetansenivå i naturfag. Utvalget kan derfor betegnes som et bekvemmelighetsutvalg (Cohen et al., 2017). Denne delen av studien er godkjent av Sikt, og vi innhentet skriftlig samtykke fra både elever og foreldre siden elevene var under 16 år.

Data som genereres ut fra metoden høyttenkning kan betegnes som *prosessdata*. Prosessdata gir informasjon om tankeprosesser, strategier og tilnærminger som deltakerne bruker når de skal komme fram til og formulere svar på oppgaver. Mens de skriftlige elevsvarene sier noe om produktet av elevenes tolkning, gir høyttenkningen informasjon om prosessene bak tolkningen (Ercikan & Pellegrino, 2017). Overordnet skiller metodelitteraturen mellom to måter å gjennomføre høyttenkning på. Enten kan deltakerne si høyt hva de tenker samtidig som de løser oppgavene, eller så kan forskeren stille deltakerne spørsmål etter at de har løst oppgaven. Sistnevnte betegnes også som *kognitivt intervju* i metodelitteraturen (Padilla & Leighton, 2017). I denne studien har vi valgt å gjennomføre kognitive intervjuer. En slik metode har både begrensninger og fordeler. Når deltakerne får spørsmål i etterkant av oppgaveløsningen, er det fare for at de glemmer hvordan de har tenkt og at spørsmål fra forskeren leder dem til å svare hvordan de tenker at de burde ha løst oppgaven, framfor at de svarer hva de faktisk har gjort. Samtidig kan interaksjonen med forskeren hjelpe deltakeren til å sette ord på egne tankeprosesser (Padilla & Leighton, 2017).

Elevene ble delt inn i to grupper, A og B. To av forfatterne møtte elevene i gruppe A, og to møtte elevene i gruppe B. I forkant utarbeidet vi en intervjuguide, men under intervjuene la vi inn oppfølgingsspørsmål der det var naturlig. Intervjuet kan altså karakteriseres som semistrukturert (Kvale & Brinkmann, 2015). Elevene fikk utdelt oppgavene på ark. De fikk beskjed om å forsøke å løse oppgavene, og deretter spurte vi hvordan de hadde tenkt. Elevene fikk også spørsmål om de hadde arbeidet med denne type grafiske framstillinger i andre fag enn matematikk. For å få bedre innsikt i elevsvar som går under kategorien «forveksler vær og klima», stilte vi i tillegg spørsmål om elevene hadde hørt begrepet klimaopplysning tidligere, og om de visste hva det betydde.

I steg 2 av studien har vi en deduktiv tilnærming i analysen. Analysen nærmer seg her en *direkte innholdsanalyse* (Fauskanger & Mosvold, 2014; Hsieh & Shannon, 2005) der vi bruker kategoriene vi utviklet i steg 1 som et utgangspunkt og undersøker om vi gjenfinner dem i elevers refleksjoner over hvordan de gikk fram for å svare på oppgaven.

Resultater

Innholdsanalyse av elevsvar

Resultater for oppgave A

Fordelingen av elevenes svar på oppgave A vises i tabell 11.2. Elever som ikke svarte på spørsmålet, er ikke inkludert i denne fordelingen.

Tabell 11.2. Fordeling av svar på oppgave A

Valg av måned	Antall elevbesvarelser	Andel elevbesvarelser (%)
Januar	385	47,0
Februar	14	1,7
Mars	7	0,8
April	4	0,5
Mai	40	4,9
Juni	14	1,7
Juli	214	26,1
August	14	1,7
September	31	3,8
Oktober	2	0,2
November	0	0,0
Desember	33	4,0
Annet (tull og tøys, vet ikke)	35	4,3
Eleven svarer med mer enn én måned, enten hele perioden eller minst to av månedene:		
Mai–september	15	1,8
Desember–februar	12	1,5
Sum antall elever	820	100

Mens omtrent halvparten av elevene som besvarte oppgaven skrev det riktige svaret (januar), svarte en tredjedel av elevene at den varmeste og tørreste måneden ligger i perioden mellom mai og september. Dette

sammenfaller med den perioden i året hvor norske elever oftest opplever slike forhold. Når elevene i tillegg får vite at Sebastian skal på ferie, kan muligens denne kombinasjonen lede elevene til å tenke at Sebastian skal på sommerferie. Og for norske skoleelever avvikles sommerferien i løpet av månedene fra juni til august. Det er ikke oppgitt i oppgaven hvilket land klimaopplysningen stammer fra.

Resultater for oppgave B

I tabell 11.3 vises resultatet av kategoriseringen av elevsvarene på oppgave B. Fordelingen inkluderer ikke antall blanke elevbesvarelser.

Tabell 11.3. Fordeling av svar på oppgave B

Kategori	Antall elevbesvarelser	Andel elevbesvarelser (%)
Riktig svar	96	20,7
Forveksler vær og klima	135	29,2
Norsk kontekst	43	9,3
Eleven mener grafen er feil: Temperatur og nedbør er byttet om	7	1,5
Eleven leser grafen feil: forveksler temperatur og nedbør	15	3,2
Vage svar	116	25,1
Annet	51	11,0
Totalt	463	100,0

Ifølge resultatene presentert i tabell 11.3, inneholder kategoriene «riktige svar», «forveksler vær og klima» og «vage svar» over 70 prosent av elevenes besvarelser totalt. Av disse er det flest besvarelser som gir uttrykk for at eleven forveksler vær og klima. Vi ser også at rundt ti prosent av elevene formulerer svar som antyder at de leser grafen ut fra en norsk kontekst.

Høyttenkning

Hensikten med steg 2 av studien var å undersøke om vi kunne gjenfinne kategoriene fra den induktive analysen av elevsvarene (steg 1). Vi finner det mest interessant å se på kategoriene «Tolker grafen ut fra norske forhold», «Forveksler vær og klima» og «Leser grafen feil / tror at det er noe feil med grafen».

Tolker grafen ut fra norske forhold

Av de tolv elevene som deltok i steg 2, var det ni elever som svarte «januar» på det første spørsmålet. Én elev svarte «april», én elev svarte «mai» og én elev svarte «juli». Selv om flertallet av elevene svarte riktig på oppgaven, uttrykte noen av dem usikkerhet i den påfølgende samtalen der de skulle beskrive hvordan de hadde tenkt. Her er et utdrag fra samtalen med elev B1 som eksemplifiserer dette:

B1: Sier den at det er varmest i januar? For jo mer grått det er [peker på stolpene], jo mer regn er det, og jo høyere opp denne linjen er [peker på linjen], jo varmere er det. Ja, det var litt rart at det var veldig varmt i januar.

I: Hvorfor synes du det var rart?

B1: Jeg mener, i Norge er det ganske kaldt i januar ... Eller dette kan være på en asiatick plass? En plass hvor det er varmere?

Elev B1 leser av grafen riktig, men stusser over at den viser at det er varmest i januar. Hen registrerer at dette bryter med norske forhold, men etter å ha tenkt seg om, kommer hen fram til at grafen kan være fra et annet land. Også elev B4 viser lignende refleksjoner når hen skal forklare hvordan hen tenkte for å svare på oppgaven

B4: Ok, så jeg burde kanskje se på månedene først. Og tenke på hvilke måneder som er varme og tørre. Og da er det jo ofte i august, eller fra mai, eller april til august kanskje? Men spørsmålet var jo hvilken måned de dro på ferie.

Elev B4 endte opp med å svare januar, men refleksjonene hans kan være et uttrykk for det flere elever tenker når de møter oppgaven, de leser den opp mot den norske konteksten som de er kjent med. Vi ser i elevsvarene i steg 1 av analysen at flere elever ser ut til å lese grafen ut fra norske forhold, elev B4 klarer derimot å lese mot det hen forventer at grafen skal vise, og hen bruker konteksten som er gitt i oppgaven.

Forveksler vær og klima

Flere av elevene nærmer seg kategorien «forveksler vær og klima» når de skal svare på den andre oppgaven, og uttrykker at Sebastian sin konklusjon om at klimaopplysningene er gale, må være riktig fordi han fikk et annet vær enn det som stod i grafen.

- A5: Jeg tror at konklusjonen hans er riktig, fordi hvis det var kaldt og masse regn når de var på ferie, så kan det ikke stemme, fordi det skal være veldig tørt og veldig varmt i januar. Så da konkluderer jeg med at han hadde rett da, med konklusjonen sin.

Også elev A4 svarer «ja» på spørsmålet om Sebastian sin konklusjon er riktig: «Ja, jeg tenkte kanskje at siden det ikke var varmt og sånt, at det bare sto feil eller noe.»

Elev A6 uttrykker enda mer eksplisitt at hen leser grafen som en værmelding når hen skal svare, og påpeker at en værmelding ikke alltid stemmer overens med været man får: «Det kan jo endre seg i de værmeldingene da, så det kan jo ha vært riktig, og så har de blitt feil.»

Når elevene får spørsmål om de vet hva en klimaopplysning er, er det ingen av dem som kan gi et presist svar.

- I: Og vil du si at klimaopplysning er det samme som en værmelding?

- A5: Kanskje. Begge viser jo temperatur og nedbør. Så, ja, kanskje jeg skal si det da, at det er litt det samme.

Elevene ser også ut til å bruke erfaringer de har med tekster som ligner grafen i TIMSS-oppgaven, når de skal forklare hva en klimaopplysning er, og flere er inne på at klimaopplysningen ligner tekster som finnes på yr.no.

- A3: Klimaopplysningen her er vel ganske lik det man ser på yr, hvis man går inn på værmeldingen, så kan man jo ofte se sånn: Å, det skal bli sol og 10 grader, men hvis man trykker inn så får man hele grafen, akkurat som den her, som da er klimaopplysningen da. Så jeg tror det er nesten det samme.

- B3: Jeg kjenner begrepet, men jeg har ikke hatt så mye om det på skolen. Jeg er jo litt på yr og sånn da.

Både ut fra hvordan elevene svarer på den andre oppgaven og hvordan de beskriver begrepene vær og klima når de får eksplisitt spørsmål om det, kan det se ut til at flere bruker begrepene om hverandre. I denne sammenheng er det også viktig å nevne at vi i samtale med læreren etter intervjuene, fikk vite at elevene skulle lære om klimaendringer og vær fenomener først i 10. klasse.

Leser grafen feil / tror at det er noe feil med grafen

Videre leser flere av elevene grafen feil når de skal forsøke å svare på oppgavene. Noen av elevene blander stolpene og linjen i diagrammet, og forstår det slik at stolpene viser temperatur og linjen viser nedbør.

I: Hva svarte du på oppgaven?

B5: Jeg svarte mai.

I: Har du lyst til å si litt om hvordan du tenkte for å løse oppgaven?

B5: Den streken [peker på en av stolpene i diagrammet] som viser temperatur er på det høyeste, hvor det også er lavest nedbør.

I: Hvilken tenker du er nedbør og hvilken tenker du er temperatur?

B5: Nedbør er den svarte streken, og temperatur er stolpen.

Vi ser her at eleven er inne på kategorien vi har kalt «Blander temperatur og nedbør» i steg 1. I tillegg til at eleven blander hva som er temperatur og hva som er nedbør, ser det ut til at hen oppfatter det slik den tørreste og varmeste tiden er der hvor linjen og én av stolpene møtes i diagrammet. I flere av samtalen framgår det at elevene ikke leser forklaringen av hva linjen og stolpene betyr, men bare antar at stolpene viser temperatur og grafelinjen viser nedbør. I datamaterialet finnes det også eksempler på at elevene gjør andre feillesninger, blant annet elev A6, som mistolker temperaturgraf.

I: Da svarte du april. Og kan du prøve å forklare hvordan du gikk frem for å løse oppgaven?

A6: Ja, jeg så på de her. At den streken der er temperatur. Og de her nedbør. Først tenkte jeg at den [peker på temperaturgraf] måtte være høy for at det skulle være varmt, men så skjønnte jeg at den måtte være lav, fordi det er jo på sommeren, og da var streken ganske lav.

Her ser det ut til at eleven tolker grafen ut fra norske forhold når hen sier at «det var jo på sommeren», og hen antar at jo lavere temperaturlinjen går, jo varmere er temperaturen. Istedenfor å ta utgangspunkt i tallene på y-aksen når hen svarer, bruker hen den norske konteksten som et utgangspunkt.

Enkelte av elevene er også inne på at det er noe galt med grafen, og at de som har laget den, har byttet om på temperatur og nedbør. Elev B1 var den eneste som i intervjuet var inne på at klimaopplysningen kanskje stammet fra et annet land enn Norge. Men når hen skal svare på om Sebastian kan konkludere med at klimaopplysningen er feil, svarer hen at det er godt mulig at det er noe galt med grafen.

B1: Jeg tenkte at det godt kunne være motsatt, at dette her var nedbør [peker på linjen]. At det her var nedbør, og dette her var temperatur. Det ville sett mer riktig ut, men det er jo motsatt ... Var grafen feil da kanskje?

Da vi spurte de tolv elevene om de hadde sett lignende grafer i andre fag, fortalte de at de først og fremst hadde arbeidet med denne type framstillinger i matematikkfaget.

Generelt opplevde vi under høyttenkningen at flere av elevene syntes det var vanskelig å forstå oppgave B, og enkelte trengte en grundigere forklaring etter at de hadde forsøkt å lese oppgaven egen hånd. Ser vi på oppgaveformuleringene, kan det være en utfordring at oppgave B inneholder doble nektelser. Når elevene møter den grafiske framstillingen i TIMSS-undersøkelsen, har de ikke mulighet til å få veiledning i hvordan de skal forstå oppgaven.

Diskusjon

I dette kapittelet har vi undersøkt hva som kjennetegner elevers svar på oppgaver som krever forståelse av en grafisk framstilling som omhandler klimaopplysninger. Gjennom data fra TIMSS-undersøkelsen finner vi at oppgaven er krevende for mange elever, og vår innholdsanalyse av skriftlige elevsvar gir noen indikasjoner på hvordan elevene går fram for å besvare oppgaven. Videre får vi en dypere innsikt i elevenes svar gjennom høyttenkningen der tolv elever forsøker å besvare de samme oppgavene. Overordnet ser vi at oppgavene er krevende for mange elever, noe som samsvarer med annen forskning som er gjort av elevers forståelse av statistikk (Glazer, 2011; Lai et al., 2016; Pérez-Echeverría et al., 2018). Ut fra svarene ser det også ut til elevene har utfordringer med å forstå hva en klimaopplysning er, noe også tidligere forskning har vist (Wildbichler et al., 2025).

Det ser ut til at flere elever har utfordringer med å registrere informasjon som bryter med det de forventer og gå ut av konteksten de befinner

seg i. Slik analysen av elevsvar på oppgave A viste, er det en stor andel av elevene som svarer «juli». I analysearbeidet antok vi at en mulig grunn til dette kunne være at elevene leser grafen opp mot norske forhold og tenker at det er varmest og tørrest i juli. Analysen av elevsvarene til oppgave B antydte også dette. I tillegg så vi i intervjuene at flere av elevenes refleksjoner gikk i denne retningen. I utgangspunktet kan man si at dette ikke er så overraskende, for det er bred enighet innenfor leseforskningen om at vi tolker tekster ut fra de erfaringer og den bakgrunnskunnskapen vi har (Cervetti & Wright, 2020), men her skaper elevenes erfaringer og forforståelse utfordringer for dem. Slik Gal (2004) beskriver det kritiske elementet i sin modell for statistisk literacy, handler det om å stille spørsmål til framstillingene og konklusjonene vi kan trekke ut fra dem. Innenfor leseforskningen beskrives også evnen til å kunne forstå at informasjon bryter med det vi forventer, og på bakgrunn av dette justere oppfatningene og tolkningene våre, som en del av den kritiske tilnærmingen (List et al., 2024; Veum & Skovholt, 2020). Ifølge funn fra vår studie, har flere elever utfordringer med dette.

Videre kan våre funn belyses fra Gals (2004) rammeverk. Ser vi elevenes svar opp mot de kognitive elementene i rammeverket for statistisk literacy, kan det se ut til at flere elever mangler en grunnleggende kunnskapsbase knyttet til statistikk og matematikk. I tidligere forskning har man sett tendenser til at elever særlig har problemer med å tolke linjediagrammer og komplekse framstillinger med flere variabler (Mathai et al., 2025). I oppgaven fra TIMSS møter elever en graf med to y-akser, og dette ser ut til å bli for komplekst for mange. Vi så i høyttenkningen at flere elever blandet temperatur og nedbør, og at de ikke brukte grafinformasjonen da de skulle svare. Vi vil også hevde at analysene våre viser at flere av elevene mangler kunnskap om hva som er forskjellen på vær og klima, og at dette kan være en årsak til at oppgaven er krevende. Tidligere forskning har også vist at elever har utfordringer med å bruke kunnskap om naturfaglige konsepter når de skal forstå statistiske framstillinger (Lai et al., 2016). Slik det framheves i Gals (2004) rammeverk, vil statistisk literacy alltid være avhengig av kjennskap til konteksten en framstilling opptre innenfor, og for at elevene skal kunne stille seg kritiske til den konklusjonen som Sebastian trekker ut fra grafen, må de vite hva en klimaopplysning er. Både elevsvarene fra TIMSS-gjennomføringen og intervjuene indikerer at elevene blander vær og klima. Dette funnet samsvarer også med tidligere forskning som er gjort internasjonalt (Bodzin et al., 2014; Da Rosa, 2022; Dawson, 2015; Lombardi

& Sinatra, 2012). I denne sammenheng er det viktig å påpeke at læreren til elevene vi intervjuet, fortalte at de ikke hadde hatt om tematikk knyttet til klimaspørsmål ennå, og at dette lå først på 10. trinn. Vi vet ikke om det samme er tilfellet for elevene på 9. trinn som gjennomførte TIMSS i 2023. Det kan samtidig stilles spørsmål ved at begrepet klima kun er nevnt tre ganger i læreplanen i naturfag for grunnskolen, og da først og fremst i forbindelse med klimautfordringer. Vi viste også tidligere i kapittelet til et kompetansemål for 2. trinn i samfunnsfag der det heter at elevene skal utforske menneskelig påvirkning på klimaet og dokumentere hvordan påvirkningen viser seg i nærmiljøet. Dette er svært ambisiøse mål som krever omfattende faglig kunnskap. Til grunn for å reflektere over årsaker til klimautfordringer, klimaendringer og miljøfarene vi i dag står overfor, ligger det en forståelse av hva begrepet klima refererer til (Da Rosa, 2022). Arbeid med nøkkelbegreper i naturfag framheves i forskningslitteraturen (Haug & Mork, 2021) og det kan være krevende å arbeide med begreper elevene har kjennskap til fra før (Krajcik & Shin, 2023). Klima og vær er eksempler på begreper elevene har en forståelse av før de kommer inn i klasserommet (Wildbichler et al., 2025). Det er avgjørende at de får hjelp av læreren til å justere disse forståelsene for å kunne reflektere godt og faglig fundert rundt informasjon knyttet til klimaspørsmål.

Det er også relevant å trekke fram at alle elevene vi intervjuet, oppga at de først og fremst møtte grafiske framstillinger i matematikk. Dette kan tyde på at de stort sett er vant til å arbeide med å lese av grafer, og ikke nødvendigvis reflektere over konteksten som grafen presenteres i og stille kritiske spørsmål til framstillingen. Her må det presiseres at dataene våre kun sier noe om elevenes oppfatninger og ikke hva som *faktisk* skjer i klasserommet. Det er mulig at elevene har jobbet med statistiske representasjoner i for eksempel samfunnsfag, selv om de ikke rapporterer om dette i intervjuene. Likevel kan man stille spørsmål ved om det er slik at elever har de matematiske kunnskapene og det statistiske begrepsapparatet på plass, men mangler trening i å lese grafiske framstillinger i mer autentiske sammenhenger, der de må bruke fagkunnskaper for å vurdere innholdet. Det trengs imidlertid mer forskning for å konkludere her.

Til sist vil peke på hvilke muligheter som ligger i prosessdata, og hvordan data fra for eksempel høyttenkning kan brukes til å validere oppgaver i standardiserte prøver og benyttes som en del av oppgaveutviklingen (Ercikan & Pellegrino, 2017). Under intervjuene erfarte vi at flere elever strevde med formuleringen i oppgave B, og det er mulig å tenke seg at

oppgavens vanskegrad har sammenheng med formuleringen i oppgaveteksten, for eksempel dobbel nektelse. Samtidig må det presiseres at vi ikke har belegg for å kunne si noe sikkert om hvorvidt elevene som gjennomførte TIMSS 2023 syntes oppgaveformuleringen var krevende.

Avsluttende betraktninger

Mye av informasjonen elever må forholde seg til i dag, formidles gjennom statistiske framstillinger. Og for å kunne forstå disse framstillingene må de både ha et matematisk og statistisk begrepsapparat, relevant forkunnskap om temaet grafen omhandler og kjenne til konteksten den er en del av, leseferdigheter og evne til å stille kritiske spørsmål til framstillingen. Vår studie tyder på at elever har utfordringer i møte med statistiske framstillinger der de må bruke naturfaglig kunnskap for å forstå. Vi vil understreke viktigheten av å arbeide med statistikk og tallmateriale i ulike fag og ulike kontekster, og trene elevene i å tolke mer komplekse grafer. Dette er avgjørende kompetanse for å kunne tolke og ta stilling til informasjon om verden rundt oss og for å delta aktivt i samfunnet. I videre forskning vil det være aktuelt å se nærmere på hva slags statistiske framstillinger elevene *faktisk* møter i undervisningen og hvilke typer oppgaver de får til disse. Får elevene først og fremst i oppgave å lese av informasjon fra grafer og diagrammer, eller blir de utfordret til å stille seg kritiske til det grafen formidler og bruke faglig kunnskap til å vurdere hva slags slutninger som kan trekkes fra den? En inngang for å besvare et slikt forskningsspørsmål kan være å studere læreverk og oppgavesett i naturfag, men også i andre fag. Det er i tillegg mulig å gjøre observasjonsstudier av læreres undervisning. Både innholdsanalysen av elevsvar og høyttenkningen med elever i denne studien åpner for å studere nærmere hvordan elever presterer på oppgaver som krever statistisk kompetanse. I framtidig forskning vil det være svært relevant å undersøke dette med utgangspunkt i andre typer oppgaver og større utvalg.

Forfatterbiografier

Cecilie Weyergang er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad innenfor leseforskning og arbeider med nasjonale prøver i lesing og med PISA-undersøkelsen. Hennes forskningsinteresser er kritisk lesing og kildekritisk tenkning og leseopplæring.

Hege Kaarstein er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad i matematikkdidaktikk og har jobbet med TIMSS-undersøkelsen siden 2009, og vært prosjektleder fra 2018. I tillegg deltar hun i den internasjonale ekspertgruppen for matematikk i TIMSS. Hennes interesser er blant annet lærernes undervisningskompetanse i matematikk, elevenes motivasjon for faget og svar på åpne oppgaver.

Fredrik Jensen er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Han har en doktorgrad i naturfagdidaktikk og har vært involvert i PISA-undersøkelsen siden 2013, med ansvar som prosjektleder fra og med PISA 2018. Hans forskningsinteresser omfatter elevers kompetanse og holdninger til naturfag. Jensen var medredaktør for boken *Matematisk kompetanse. I dybden på resultater fra PISA 2022*, utgitt i 2024.

Anne-Catherine Lehre er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en bakgrunn innenfor språk og realfag, og har en doktorgrad i biostatistikk. Lehre har jobbet i TIMSS-prosjektet siden 2020.

Referanser

- Bendavid, N. (2025, 30. juli). Tenk på et tall: Donald Trumps magiske matte. *Aftenposten*. <https://www.aftenposten.no/verden/i/KMVlm5/donald-trump-har-et-uanstrengt-forhold-til-tall>
- Benestad, R., Hygen, H. O., & Nordli, Ø. (2018). Er det noen tvil om menneskeskapte klimaendringer?. *Naturen*, 142(4), 136–143. <https://doi.org/10.18261/issn.1504-3118-2018-04-02>
- Bodzin, A. M., Anastasio, D., Sahagian, D., Peffer, T., Dempsey, C. & Steelman, R. (2014). Investigating climate change understandings of urban middle-level students. *Journal of Geoscience Education*, 62(3), 417–430. <https://doi.org/10.5408/13-042.1>
- Cartwright, T. J., Deb, H. & Magee, P. A. (2021). Investigating weather, climate, and climate change understanding of Appalachian middle-level students. *Electronic Journal for Research in Science & Mathematics Education*, 25(2), 6–29. <https://ejrsme.icrsmc.com/article/view/20984>
- Centurino, V. A. S. & Kelly, D. L. (2021). TIMSS 2023 science framework. I I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023/frameworks/chapter-2.html>
- Cervetti, G. N. & Wright, T. S. (2020). The role of knowledge in understanding and learning from text. I E. B. Moje, P. P. Afflerbach, P. Enciso & N. K. Lesaux (Red.), *Handbook of reading research: Volume 5* (kap. 13). Routledge.
- Cohen, L., Manion, L. & Morrison, K. (2017). *Research methods in education* (8. utg.). Routledge. <https://doi.org/10.4324/9781315456539>

- Da Rosa, J. A. (2022). Comparing climate science misconceptions with worldview and cognitive reflection suggests poor understanding and motivated reasoning among undergraduates. *Journal of Geoscience Education*, 70(4), 501–516. <https://doi.org/10.1080/10899995.2021.2006549>
- Dawson, V. (2015). Western Australian high school students' understandings about the socioscientific issue of climate change. *International Journal of Science Education*, 37(7), 1024–1043. <https://doi.org/10.1080/09500693.2015.1015181>
- Driver, R. (1989). Students' conceptions and the learning of science. *International Journal of Science Education*, 11(5), 481–490. <https://doi.org/10.1080/0950069890110501>
- Ercikan, K. & Pellegrino, J. W. (2017). Validation of score meaning using example response processes for the next generation of assessments. I K. Erickson & J. W. Pellegrino (Red.), *Validation of Score Meaning for the Next Generation of Assessments: The Use of Response Processes* (s. 1–8). Routledge. <https://doi.org/10.4324/9781315708591>
- Fauskanger, J. & Mosvold, R. (2014). Innholdsanalysens muligheter i utdanningsforskning. *Norsk pedagogisk tidsskrift*, 98(2), 127–139. <https://doi.org/10.18261/ISSN1504-2987-2014-02-07>
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382. <https://doi.org/10.1037/h0031619>
- Gal, I. (2004). Statistical literacy. I D. Ben-Zvi & J. Garfield (Red.), *The challenge of developing statistical literacy, reasoning and thinking* (s. 47–78). Springer Netherlands. https://doi.org/10.1007/1-4020-2278-6_3
- Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education*, 47(2), 183–210. <https://doi.org/10.1080/03057267.2011.605307>
- Harold, J., Lorenzoni, I., Shipley, T. F., & Coventry, K. R. (2016). Cognitive and psychological science insights to improve climate change data visualization. *Nature Climate Change*, 6(12), 1080–1089. <https://doi.org/10.1038/nclimate3162>
- Haug, B. S. & Mork, S. M. (2021). *Nøkkelbegreper i utforskende arbeid*. Universitetsforlaget.
- Hsieh, H.-F. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Irving, D., Clark, R. W. A., Lewandowsky, S. & Allen, P. J. (2022). Correcting statistical misinformation about scientific findings in the media: Causation versus correlation. *Journal of Experimental Psychology: Applied*, 28(1), 1–9. <https://doi.org/10.1037/xap0000408>
- Johannssen, A., Chukhrova, N., Schmal, F. & Stabenow, K. (2021). Statistical literacy – misuse of statistics and its consequences. *Journal of Statistics and Data Science Education*, 29(1), 54–62. <https://doi.org/10.1080/10691898.2020.1860727>
- Koga, S. (2022). Characteristics of statistical literacy skills from the perspective of critical thinking. *Teaching Statistics*, 44(2), 59–67. <https://doi.org/10.1111/test.12302>
- Krajcik, J. & Shin, N. (2023). Student conceptions, conceptual change, and learning progressions. I N. G. Lederman, D. L. Zeidler & J. S. Lederman (Red.), *Handbook of research on science education* (s. 121–157). Routledge. <https://doi.org/10.4324/9780367855758-7>
- Kunnskapsdepartementet. (2019a). *Læreplan i matematikk* (MAT01-05). Fastsett som forskrift. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/mat01-05>
- Kunnskapsdepartementet. (2019b). *Læreplan i naturfag* (NAT01-04). Fastsett som forskrift. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/nat01-04>
- Kunnskapsdepartementet. (2019c). *Læreplan i samfunnsfag* (SAF01-04). Fastsett som forskrift. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/saf01-04>
- Kvale, S. & Brinkmann, S. (2015). *Det kvalitative forskningsintervju* (T. M. Anderssen & J. Rygge, Overs.; 3. utg.). Gyldendal Akademisk.
- Lai, K., Cabrera, J., Vitale, J. M., Madhok, J., Tinker, R. & Linn, M. C. (2016). Measuring graph comprehension, critique, and construction in science. *Journal of Science Education and Technology*, 25(4), 665–681. <https://doi.org/10.1007/s10956-016-9621-9>
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.

- Lindgren, B.-M., Lundman, B. & Graneheim, U. H. (2020). Abstraction and interpretation during the qualitative content analysis process. *International Journal of Nursing Studies*, 108, Artikel 103632. <https://doi.org/10.1016/j.ijnurstu.2020.103632>
- List, A., Russell, L. A., Yao, E. Z., Campos Oaxaca, G. S. & Du, H. (2024). Critique generation promotes the critical reading of multiple texts. *Learning and Instruction*, 93, Artikel 101927. <https://doi.org/10.1016/j.learninstruc.2024.101927>
- Lombardi, D. & Sinatra, G. M. (2012). College students' perceptions about the plausibility of human-induced climate change. *Research in Science Education*, 42(2), 201–217. <https://doi.org/10.1007/s11165-010-9196-z>
- McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & research in social education*, 46(2), 165–193. <https://doi.org/10.1080/00933104.2017.1416320>
- Mathai, S., Krishnan, P. & Sreevalsan-Nair, J. (2025). Understanding graphical literacy using school students' comprehension strategies. *Contemporary Education Dialogue*, 22(1), 40–74. <https://doi.org/10.1177/09731849241242855>
- Mullis, I. V. S., Martin, M. O. & von Davier, M. (2021). *TIMSS 2023 assessment framework*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023>
- OECD. (2023). *PISA 2025 science framework (draft)*. https://pisa-framework.oecd.org/science-2025/assets/docs/PISA_2025_Science_Framework.pdf
- Osborne, J. (2023). Science, scientific literacy, and science education. I N. G. Lederman, D. L. Zeidler & J. S. Lederman (Red.), *Handbook of research on science education* (s. 785–816). Routledge. <https://doi.org/10.4324/9780367855758-30>
- Padilla, J.-L. & Leighton, J. P. (2017). Cognitive interviewing and think aloud methods. I B. Zumbo & A. Hubley (Red.), *Understanding and investigating response processes in validation research* (s. 211–228). Springer International Publishing. https://doi.org/10.1007/978-3-319-56129-5_12
- Pérez-Echeverría, M. D. P., Postigo, Y. & Marín, C. (2018). Understanding of graphs in social science undergraduate students: Selection and interpretation of graphs. *Irish Educational Studies*, 37(1), 89–111. <https://doi.org/10.1080/03323315.2018.1440248>
- Schauss, M., Nöthen, E., Ottosander, M. P., & Sprenger, S. (2024). Visuals of climate change in school textbooks. *International Research in Geographical and Environmental Education*, 33(3), 245–260. <https://doi.org/10.1080/10382046.2023.2298557>
- Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, 4(1), 118–133. <https://doi.org/10.1080/23265507.2017.1354313>
- Tishkovskaya, S. & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, 20(2), Artikel 4. <https://doi.org/10.1080/10691898.2012.11889641>
- Veum, A. & Skovholt, K. (2020). *Kritisk literacy i klasserommet*. Universitetsforlaget.
- Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, 88(421), 1–8. <https://doi.org/10.1080/01621459.1993.10594283>
- Weiland, T. (2017). Problematising statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96(1), 33–47. <https://doi.org/10.1007/s10649-017-9764-5>
- Wildbichler, S., Haagen-Schützenhöfer, C. & Schubatzky, T. (2025). Students' ideas about the scientific underpinnings of climate change: A systematic review of the literature. *Studies in Science Education*, 61(1), 117–169. <https://doi.org/10.1080/03057267.2024.2395206>
- Yin, L. & Foy, P. (2021). TIMSS 2023 assessment design. I I. V. S. Mullis, M. O. Martin & M. von Davier (Red.), *TIMSS 2023 assessment frameworks*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2023/frameworks/chapter-4.html>

KAPITTEL 12

Økende ulikhet og endringer i norsk skole

Anne-Catherine Lehre og Hege Kaarstein

Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo

I dette kapittelet gis korte oppsummeringer av resultatene fra kapittel 2–11. Samlet sett gir oppsummeringene en fremstilling av utviklingstrekk i norske elevers prestasjoner og læringsforutsetninger i matematikk og naturfag. Gjennom analyser av faglige resultater, motivasjon, undervisningskvalitet og læreplanimplementering belyses både utfordringer og muligheter for å styrke elevenes læring og for å fremme likeverd i skolen. Studiene er basert på TIMSS-data, i hovedsak fra gjennomføringene i 2015, 2019 og 2023.

I denne antologien settes eleven i sentrum, og forfatterne undersøker hvordan ulike faktorer/aspekter ved skolen, undervisningen og fagene henger sammen med elevenes læring og faglige utvikling. Kapitlene gir et forskningsbasert innblikk i hva data fra *Trends in International Mathematics and Science Study* (TIMSS) kan fortelle om norsk skole, og hvordan ulike faktorer på elev-, klasse- og skolenivå henger sammen med elevenes faglige prestasjoner. TIMSS-undersøkelsen, som har blitt gjennomført hvert fjerde år siden 1995, gir et unikt datagrunnlag for å studere utviklingstrekk i elevenes kompetanse i matematikk og naturfag.

TIMSS-undersøkelsen dekker verken alle fagene i skolen eller alle sider ved fagene matematikk og naturfag (som for eksempel faglig kommunikasjon og praktisk arbeid), men undersøkelsen gir likevel informasjon av høy kvalitet for et representativt, nasjonalt utvalg av elever på 5. og 9. trinn. Dette gjør det mulig å undersøke elevenes kompetanse i lys av rammeverkene som definerer hva som skal måles i fagene, og å identifisere faktorer som har sammenheng med læring og prestasjoner. TIMSS-undersøkelsen

kombinerer faglige tester med omfattende kontekstdata fra elever, lærere, skoleledere og foreldre, og gir dermed et mer helhetlig bilde av skolehverdagen.

Selv om TIMSS-undersøkelsen har samlet inn data hvert fjerde år i perioden fra 1995 til 2023, baserer analysene i de fleste av antologiens kapitler seg på data fra undersøkelsene som ble gjennomført i 2015, 2019 og 2023. Et unntak er kapittel 2, skrevet av Nilsen, Ye og Teig, som er det eneste bidraget som inkluderer data helt tilbake til 1995. Deres analyser viser en tydelig utvikling: at sammenhengen mellom elevers sosioøkonomiske status (SØS) og deres matematikkprestasjoner har blitt gradvis sterkere over tid. Denne utviklingen indikerer at sammenhengen mellom elevenes hjemmebakgrunn og deres faglige resultater blir sterkere og sterkere ettersom tiden går, hvilket kan bety at skolene ikke helt får til å kompensere for sosiale ulikheter.

Ifølge Nilsen, Ye og Teig kunne SØS forklare kun 2,4 prosent av variasjonen i elevenes prestasjoner i 1995, mens i 2023 forklarer SØS 12 prosent av variasjonen. At det er variasjon i elevenes prestasjoner er forventet, ettersom elever har ulik hjemmebakgrunn, bor i ulike deler av landet, går på forskjellige skoler og har varierende grad av faglig motivasjon, selvtillit og opplevelse av skolemiljø og undervisning. Samtidig som sammenhengen mellom SØS og prestasjoner har økt, har andelen såkalte resiliente elever – de som presterer godt til tross for lav SØS – sunket fra 18 til 12 prosent. Dette kan tyde på at skolens evne til å kompensere for sosiale forskjeller er noe svekket.

Parallelt med den økende betydningen av elevenes sosioøkonomiske bakgrunn for faglige prestasjoner, tyder funn fra kapittel 3 på at skolemiljøet også har endret seg. Kaarstein, White og Lehre viser at både elever og lærere i 2023 rapporterer lavere opplevelse av trygghet i skolen sammenlignet med 2015. I samme periode har forekomsten av mobbing økt, og lærere melder om mer utfordrende elevatferd og redusert respekt i klasserommet. Rektorer rapporterer samtidig om en økning i problemer knyttet til hærverk og trusler, selv om disse fortsatt vurderes som moderate. Studien peker også på at skoler med en høy andel førstegenerasjonsinnvandrere opplever mer mobbing og lavere grad av tilhørighet blant elevene, noe som understreker behovet for inkluderende tiltak som tar hensyn til kulturelt og språklig mangfold.

Til tross for disse utfordringene viser internasjonale sammenligninger at norske elever, lærere og rektorer fortsatt rapporterer om et relativt trygt

skolemiljø i 2023. Selv om skolemiljøet i norske skoler fortsatt oppleves som relativt trygt, og det finnes strukturer og tiltak som bidrar til å opprettholde trivsel og orden, aktualiserer utviklingen spørsmålet om hvilke faktorer som kan bidra til å styrke skolens evne til å møte disse utfordringene og videre hvilke faktorer som har en positiv sammenheng med elevers læring og prestasjoner.

I kapittel 4 retter Veletić, Radišić og Madsø Jacobsen søkelyset mot skoleledelse som en sentral komponent i arbeidet med gode elevprestasjoner. Gjennom analyser av TIMSS-data undersøker de hvordan rektors fokus på læring, deres utdanningsbakgrunn og erfaring henger sammen med elevenes prestasjoner i matematikk og naturfag. Funnene viser at skoleledelse er en nøkkelfaktor da rektorer med lang erfaring og mastergrad i utdanningsledelse er positivt assosiert med høyere elevprestasjoner. Dette understreker betydningen av profesjonell utvikling og kontinuitet i skoleledelse, og peker på behovet for langsiktige investeringer i skolelederkompetanse. I tillegg viser studien at foreldrenes læringsfokus har en positiv innvirkning, særlig på matematikkprestasjoner, noe som fremhever viktigheten av et godt samarbeid mellom hjem og skole for å støtte elevenes faglige utvikling.

Mens forfatterne av kapittel 4 fremhever betydningen av kvalifisert og stabil skoleledelse for elevenes faglige utvikling, retter forfatterne av kapittel 5 oppmerksomheten mot selve prestasjonsfordelingen blant norske elever. Lehre og Kaarstein undersøker hvordan spredningen i elevenes prestasjoner i matematikk og naturfag har utviklet seg fra 2015 til 2023, med søkelys på forskjeller mellom kjønn og trinn. Analysene viser at spredningen har økt betydelig i begge fag, spesielt på 9. trinn. I matematikk på 5. trinn økte spredningen fra 2015 til 2019, særlig innen emnet tall og blant jenter, mens naturfag på samme trinn viser økt variasjon for begge kjønn og alle emneområder. På 9. trinn er spredningen i matematikkprestasjoner størst i 2023, med tydelige økninger innen tall og geometri, mens naturfag viser størst spredning i 2019, særlig i biologi og fysikk. Gutter på 9. trinn viser generelt større variasjon i prestasjoner enn jenter, noe som gir støtte til den såkalte *greater male variability*-hypotesen. Resultatene peker derfor på viktigheten av tilpasset opplæring – et sentralt prinsipp i norsk skole – og behovet for undervisningsformer som møter elevenes ulike forutsetninger.

Mens forfatterne av kapittel 5 dokumenterer økt spredning i elevenes faglige prestasjoner, retter forfatterne av kapittel 6 oppmerksomheten mot

de elevene som befinner seg i ytterkantene av fordelingen. Lehre og Rohatgi undersøker hva som kjennetegner elever med lave prestasjoner i matematikk og naturfag, og sammenligner dem med elever som presterer på høyt nivå, basert på data fra TIMSS 2015, 2019 og 2023. Analysene viser at andelen elever som presterer på et lavt nivå har økt i perioden, mens andelen elever som presterer på et avansert nivå har vært relativt stabil. Lavtpresterende elever har ofte lavere sosioøkonomisk status, svak faglig selvtilitt og motivasjon, og møter oftere sultne på skolen – særlig på 5. trinn. Det er også en høyere andel som sjelden eller aldri snakker norsk hjemme. Høytpresterende elever har motsatt profil, med høyere selvtilitt, sterkere motivasjon og bedre fysisk velvære. Selv om kjønnsfordelingen generelt er jevn, er det en overvekt av jenter blant de lavtpresterende i matematikk på 5. trinn. Samtidig rapporterer både lavt- og høytpresterende elever i økende grad om at de møter sultne på skolen og om lavere motivasjon, noe som kan ha implikasjoner for læring. I likhet med funnene i kapittel 3, viser også Lehre og Rohatgi at elevenes opplevelse av trygghet og tilhørighet har blitt svekket over tid. Denne utviklingen utfordrer skolens mål om likeverd og sosial utjevning, og understreker behovet for helhetlige tiltak som favner både faglige og sosiale dimensjoner av læring, noe forfatterne av kapittel 2 også påpeker.

Når forskjeller i elevenes faglige prestasjoner blir tydeligere, kan det også reises spørsmål om hvilke muligheter elevene faktisk har hatt til å tilegne seg faglig kunnskap. Dette er utgangspunktet for kapittel 7, der White, Selling, Pettersen og Veletić undersøker hva norske elever på 5. trinn har fått mulighet til å lære i matematikk. Studien ser TIMSS-resultatene i lys av overgangen fra læreplanen LK06 til LK20, som har introdusert bredere og mer prosessorienterte kompetansemål. Forfatterne viser en tydelig sammenheng mellom endringer i læreplanen og den undervisningen elevene faktisk mottar. Fra 2019 til 2023 har det skjedd en endring i undervisningsinnholdet som samsvarer med overgangen fra LK06 til LK20. Denne endringen kan tolkes som et forsøk på å styrke dybdelæring, ved å gi elevene mer tid til å fordype seg i færre emner. Analysene viser at norske elever eksponeres for færre faglige områder enn elever i de øvrige nordiske landene som deltar i TIMSS, noe som kan ha bidratt til den norske nedgangen i prestasjoner – særlig innen geometri og måling. Kapitelforfatterne peker på viktigheten av å sikre at alle elever får tilgang til et bredt og relevant faglig innhold, og at lærere får nødvendig kompetanse og ressurser til å realisere intensjonene i læreplanen.

I kapittel 7 viser White et al. hvordan endringer i læreplanen og undervisningsinnhold kan påvirke elevenes muligheter til å lære, men det er også avgjørende å undersøke hvordan undervisningen faktisk oppleves av elevene. Dette er utgangspunktet for kapittel 8, der White, Kaarstein og Jentsch undersøker om nedgangen i matematikkprestasjoner på 5. trinn fra TIMSS 2019 til 2023 kan knyttes til elevenes oppfatning av lærernes undervisningskvalitet. I denne perioden rapporterer elevene om en liten, men statistisk signifikant nedgang i opplevd undervisningskvalitet, og analysene viser at 17 prosent av prestasjonsnedgangen kan forklares med denne endringen. Forfatterne konkluderer med at norske skoler i stor grad har klart å opprettholde undervisningskvaliteten til tross for økende utfordringer som større elevmangfold, stress og et noe dårligere skolemiljø enn tidligere (jf. kapittel 3 og 6). Likevel understrekes behovet for mer støtte og ressurser til lærere, både for å møte disse utfordringene og for å sikre høy undervisningskvalitet over tid. Dette er viktig ikke bare for elevenes faglige utvikling, men også for lærernes trivsel og profesjonelle utvikling.

I kapittel 8 viser forfatterne hvordan elevenes opplevelse av undervisningskvalitet henger sammen med faglige prestasjoner, og understreker betydningen av lærernes innsats og støtte i en stadig mer kompleks skolehverdag. For å forstå elevenes læring mer inngående, er det imidlertid også nødvendig å se på hvordan de faktisk engasjerer seg i selve prøvesituasjonen. Dette er utgangspunktet for kapittel 9, der Radišić og Pettersen benytter loggfildata fra TIMSS 2023 til å undersøke sammenhengen mellom motivasjon, responstid og matematikkprestasjoner blant norske elever på 9. trinn. Overgangen fra papir- til digitale prøver i TIMSS 2019 har gjort det mulig å analysere elevenes interaksjon med oppgavene – ikke bare resultatene – og gir dermed innsikt i innsats, engasjement og strategier. Ved hjelp av en personorientert tilnærming identifiserer forskerne seks ulike motivasjonsprofiler basert på elevenes selvtillit, interesse og opplevd nytteverdi av matematikk. Analysene viser at høy selvtillit har sterkest sammenheng med høye prestasjoner, mens elever med lav selvtillit presterer svakest. Elevene som hører til profilen som kombinerer høy interesse, selvtillit og nytteverdi, er også de elevene som løser oppgavene raskt og effektivt, og samtidig oppnår høye resultater. Studien avdekker også tydelige kjønnsforskjeller i motivasjonsprofilene. Gutter er overrepresentert i de profilene hvor elevene er mest motiverte, mens jenter dominerer i profilen hvor elevene har lav selvtillit, men synes at faget har høy nytteverdi. Videre peker studien på at innsats i prøvesituasjonen i stor grad påvirkes av oppgavens kompleksitet

og format – ikke nødvendigvis av vanskelighetsgrad alene. Til tross for at TIMSS er en *low-stakes*-prøve, viser majoriteten av elevene tydelig løsningsatferd, noe som styrker validiteten i resultatene. Sammenhengen mellom innsats og prestasjon er tydelig, da elever som bruker mer tid og viser aktiv løsningsatferd presterer bedre. Kapittelet konkluderer med at selv om mange elever viser innsats, er det bekymringsfullt at så få har høy motivasjon. Forfatterne anbefaler videre forskning på hvordan oppgaver og prøvesituasjoner påvirker elevers innsats og motivasjon, samt tiltak for å styrke interesse og selvtillit i matematikk.

I kapittel 10 ser Pettersen og Radišić videre på matematikkresultatene på 9. trinn ved å undersøke norske elevers ferdigheter i algebra, med fokus på evnen til å generalisere mønstre og formulere matematiske uttrykk. Selv om resultatene viser en viss fremgang sammenlignet med tidligere TIMSS-undersøkelser, presterer norske elever fortsatt svakere enn elever i mange andre land, og kun en liten andel klarer å formulere korrekte algebraiske uttrykk. Mange elever gir svar i form av tall eller regnestykker uten variabler, noe som tyder på manglende forståelse for algebraisk notasjon. Samtidig viser analysene at elever som forsøker å bruke symboler og variabler – selv når svaret er feil – har høyere matematikkprestasjoner enn de som ikke gjør det, noe som tolkes som et tegn på begynnende mestring av det formelle matematiske språket. Studien peker også på at algebra fortsatt undervises som et isolert tema i norske lærebøker, og at læreplanendringer alene ikke er tilstrekkelige. For å styrke elevenes algebraiske kompetanse kreves det endringer i undervisningspraksis og lærerutdanning.

I kapittel 10 viser forfatterne hvordan elevers forståelse av algebra og formell notasjon kan gi verdifull innsikt i deres faglige utvikling og behov for støtte i matematikkopplæringen. I kapittel 11 utvider Weyergang, Kaarstein, Lehre og Jensen dette perspektivet ved å undersøke hvordan norske elever på 9. trinn tolker grafiske fremstillinger som formidler klimaopplysninger, basert på åpne svar fra TIMSS 2023. Studien viser at mange elever har vansker med å forstå komplekse grafer med flere variabler, og at de ofte tolker informasjonen ut fra egne erfaringer og norsk kontekst – noe som i flere tilfeller fører til feil svar. I tillegg forveksler en betydelig andel av elevene begrepet klima med vær, og flere misforstår grafens innhold eller tror den inneholder faktiske feil. Disse funnene peker på et behov for å styrke elevenes naturfaglige begrepsforståelse og det forskerne betegner som statistisk literacy. Kapittelet understreker viktigheten av å arbeide med komplekse grafiske fremstillinger i flere fag og kontekster, for å utvikle elevenes evne til

kritisk tolkning og meningsfull bruk av data i møte med samfunnsaktuelle problemstillinger som klimaendringer.

Denne antologien gir et forskningsbasert bilde av utviklingen i norske elevers prestasjoner og læringsforutsetninger i matematikk og naturfag over en periode preget av endringer i skole og samfunn. Gjennom analyser av TIMSS-data belyses det hvordan faktorer som sosioøkonomisk bakgrunn, skolemiljø, ledelse, undervisningskvalitet, motivasjon og læreplaninnhold har sammenheng med elevenes læring. Samlet sett peker funnene på et behov for målrettede tiltak – både pedagogiske og strukturelle – for å sikre likeverdige og inkluderende læringsmiljøer. Samtidig viser studiene at norske elever har potensial til faglig utvikling når de får støtte, utfordringer og muligheter til å engasjere seg i egen læring. Antologien gir dermed et viktig kunnskapsgrunnlag for videre forskning, skoleutvikling og politikkutforming.

Forfatterbiografi

Anne-Catherine Lehre er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en bakgrunn innenfor språk og realfag, og har en doktorgrad i biostatistikk. Lehre har jobbet i TIMSS-prosjektet siden 2020.

Hege Kaarstein er forsker ved Institutt for lærerutdanning og skoleforskning ved Universitetet i Oslo. Hun har en doktorgrad i matematikkdidaktikk og har jobbet med TIMSS-undersøkelsen siden 2009, og vært prosjektleder fra 2018. I tillegg deltar hun i den internasjonale ekspertgruppen for matematikk i TIMSS. Hennes interesser er blant annet lærernes undervisningskompetanse i matematikk, elevenes motivasjon for faget og svar på åpne oppgaver.

Tretti år med TIMSS i Norge: Kan resultater fra 1995 sammenliknes med dagens og framtidige TIMSS-undersøkelser?

Rolf Vegar Olsen

Institutt for spesialpedagogikk, Universitetet i Oslo

Sammendrag: En av de store styrkene til TIMSS, og andre internasjonale undersøkelser, er at de gjentas med faste intervaller. TIMSS gir nå en trendlinje for norske 4./5.- og 8./9. trinnselever for de siste tretti årene. Muligheten for å trekke slutninger om utvikling over tid er i dag en av de sentrale begrunnelsene for at så mange land velger å delta i TIMSS og de andre internasjonale undersøkelsene. Dette kapittelet er en betraktning rundt gyldigheten av å bruke data fra TIMSS for å si noe om utvikling over tid. I kapittelet omtales en rekke generelle årsaker til at det å utvikle lange trendlinjer er vanskelig – og kanskje egentlig umulig?

Nøkkelord: Trender, internasjonale skoleundersøkelser, TIMSS,

Abstract

30 Years of TIMSS in Norway: Can Results from 1995 Be Compared with Current and Future TIMSS Studies?

One of the significant strengths of TIMSS and other international assessments is their regular repetition at set intervals. TIMSS now provides a trend line for Norwegian students in grades 4/5 and 8/9 spanning the last thirty years. The ability to draw conclusions about development over time is currently one of the primary

justifications for why so many countries choose to participate in TIMSS and other international studies. This chapter offers a reflection on the validity of using TIMSS data to make inferences about trends over time. It discusses a number of general reasons why developing long trend lines is challenging—and perhaps fundamentally impossible.

Keywords: Trends, international large-scale assessments, TIMSS

Innledning

Denne boka går inn i en stor samling av norskspråklige publikasjoner med analyser fra en sammenknyttet serie med undersøkelser som startet for tretti år siden. Startpunktet var altså i 1995. Da ble *The Third International Mathematics and Science Survey* gjennomført i 45 land, deriblant også i Norge.¹ Den første nasjonale rapporten ble publisert et par år senere (Lie et al., 1997). I store trekk gir den siste gjennomførte undersøkelsen (TIMSS 2023) tilgang til den samme type data som den første undersøkelsen i 1995.

Ideen med de internasjonale undersøkelsene var opprinnelig å bruke den internasjonale variasjonen for bedre å kunne forstå resultater for sitt eget land. Dette er fortsatt et av de viktigste formålene, men internasjonale sammenlikninger er også notorisk utfordrende fordi man sammenlikner veldig ulike utdanningssystemer. Grad av sentralisering, andel privatskoler, alder ved skolestart og andel av alderskohorter som fortsatt er i utdanningløpet er kun noen eksempler på karakteristikk som varierer på tvers av utdanningssystemer. Slike kilder vil, sammen med mer overordnede språklige, kulturelle, strukturelle og økonomiske forhold, kunne komplisere tolkninger av forskjeller mellom resultater på tvers av land.

Sammenlikninger av norske elever, lærere og skoler over tid er imidlertid i langt mindre grad beheftet med denne typen analytiske utfordringer. De fleste kapitlene i denne boka utnytter nettopp dette tidsdesignet som finnes i TIMSS til å presentere analyser som vektlegger endringer over tid for norsk skole. Nilsen, Teig og Ye (kapittel 2) viser at ulikheten i norsk skole har økt ved å bruke data fra hele rekken av TIMSS-undersøkelser. Kaarstein, White og Lehre (kapittel 3) presenterer analyser av data som konkluderer med at deler av skolemiljøet har blitt dårligere i den samme perioden. Lehre og Kaarstein (kapittel 5) viser at spredningen i de norske elevenes prestasjoner øker fra 2015 til 2023. Lehre og Rohatgi (kapittel 6) presenterer en analyse for å beskrive kjennetegn ved elever på det laveste mestringsnivået i de siste tre TIMSS-undersøkelsene. White, Selling, Pettersen og Veletič (kapittel 7) bruker data fra 2019 og 2023 for å vise at innføringen av LK2020 ikke har ført til betydelige endringer i det innholdet som dekkes i undervisningen. White, Kaarstein og Jentsch (kapittel 8)

1 Som navnet «The Third...» indikerer, finnes det tidlige forløpere til TIMSS-studien, men TIMSS 1995 medførte så store endringer at muligheten for lenking bakover i tid til disse forløperne er svært begrenset.

viser at nedgangen i norske elevers prestasjoner i noen grad kan være knyttet til en nedgang i undervisningskvalitet. Denne muligheten til å kunne gjøre sammenlikninger over tid for læringsresultater, undervisningens innhold og kvalitet og skolemiljø, gir utvilsomt en kraftig analytisk linse for informert diskusjon og drøfting om styrker og svakheter ved det norske utdanningssystemet.

For bedre å kunne forstå de endringene som rapporteres i disse kapitlene, ønsker jeg her å beskrive hvordan TIMSS-undersøkelsen har endret seg gradvis over tid – og ikke minst hvordan de fenomenene vi forsøker å studere også endrer seg.

Fra TIMSS 1995 til TIMSS 2023

Helt siden 1995 har TIMSS beholdt en kjerne i rammeverket som definerer undersøkelsene. Formålet har hele tiden vært å framskaffe data som kan brukes for å sammenlikne og kontrastere læreplaner (intendert læreplan), undervisningspraksiser (implementert læreplan) og elevenes læringsutbytte (oppnådd læreplan) på tvers av mange utdanningssystemer (Mullis & Martin, 2022; Robitaille et al., 1993). Tanken er at dette kan gi de deltakende landene ny og unik innsikt for å forstå og evaluere kvaliteten på utdanningen i sitt eget land. Ideen om at undersøkelsen skulle gjentas med jevne mellomrom var ikke inkludert i rammeverket til undersøkelsen i 1995. Det viste seg imidlertid at det var stor interesse fra flere av de deltakende landene om å gjenta deler av undersøkelsen noen år etterpå, og derfor ble undersøkelsen TIMSS Repeat (TIMSS-R) gjennomført i 1999. Erfaringen med å gjennomføre studien noen år etterpå ble opplagt positivt mottatt. Med TIMSS-R fikk landene tilgang til data om hvordan en rekke elev-, lærer- og skolefaktorer utviklet seg over tid. Dette gir potensielt svært nyttig informasjon for å kunne evaluere virkninger av reformer eller andre større endringer i skolen. Kort tid etter ble TIMSS reetablert som en studie som skulle gjennomføres fast hvert fjerde år, og fra TIMSS 2003 har studien endret navn til *Trends in International Mathematics and Science Study* (Hastedt, 2020; Hastedt & Sibberns, 2022). I etterkant av dette ble også de to første studiene omdøpt fra TIMSS og TIMSS-R til henholdsvis TIMSS 1995 og TIMSS 1999.

Undersøkelsen i 1995 inkluderte det meste av det vi kjenner igjen fra de senere studiene. Dataene som ble samlet inn, besto av elevers besvarer på prøver i matematikk og naturfag og fra spørreskjemaer

til elever, lærere og skoleledere. Studiedesign og metodikk har også vært rimelig stabilt siden 1995: bruk av to-steps utvalgsprosedyrer, utvalgsvekter, IRT-modeller, plausible verdier og rotert heftedesign, for å nevne noen sentrale elementer ved metodikken som fortsatt den dag i dag er gullstandarden for denne typen undersøkelser. Men i tillegg gjennomførte man i 1995 også praktiske prøver i naturfag og matematikk. Dette var en stor logistisk operasjon hvor egne testadministratorer besøkte femti tilfeldig utvalgte skoler (fra Trøndelag og sørover). I koffertene hadde de med seg pakker med utstyr for å gjennomføre et sett med tolv forskjellige standardiserte praktiske prøver (Kind et al., 1999). Som en forløper til TIMSS, ble det også gjennomført en svært omfattende læreplananalyse i noen land, inkludert Norge (det såkalte SMSO-prosjektet). I denne studien ble formelle læreplandokumenter og et utvalg av de mest brukte lærebøkene i matematikk og naturfag(ene) kodet og analysert i detalj. Det ble ikke gitt ut en egen norsk analyse av disse dataene, men resultater for Norge er inkludert i de internasjonale rapportene som ble publisert (Schmidt, McKnight et al., 1997; Schmidt, Raizen et al., 1997). Poenget med å nevne dette er at TIMSS 1995 uten tvil er den mest omfattende og ambisiøse internasjonale studien som noensinne er gjennomført til dags dato.

Til tross for at det er store likheter mellom studiene i 1995 og 2023, er det likevel endringer over tid: Det er endringer i testdesign og elevpopulasjoner, den konkrete definisjonen av de fagområdene som testes oppdateres mellom hver syklus, og innholdet i spørreskjemaene som besvares av elever, foreldre, lærere og skoleledere endres også over tid. I tillegg har antall land som deltar omtrent doblet seg. To av de største endringene i perioden – og disse skal vi komme tilbake til etter hvert – er:

- *Endringer i testdesign:* I 2023-studien ble prøvene og spørreskjemaene for første gang besvart kun elektronisk i Norge og i de fleste andre deltagende landene.
- *Endringer i elevpopulasjonen:* I 1995 var Norge representert med utvalg fra datidens 2. og 3. trinn (den såkalte populasjon 1) og 6. og 7. trinn (populasjon 2). I 2023 var disse to populasjonene representert av henholdsvis elever fra 5. og 9. trinn.

I tillegg er det jo slik at det man måler kan påvirkes av utenforliggende endringer i samfunnet som påvirker hva eller hvordan elever og andre svarer

på undersøkelsene. Et opplagt eksempel på at verden vi ønsker å si noe om har endret seg, er teknologibruk i skolen. Mens man i TIMSS 1995 hadde spørsmål om bruk av kalkulator i spørreskjemaet, har man nå spørsmål om bruk av datamaskiner og nettbrett. I neste syklus bør man vel ikke bli overrasket om spørreskjemaene også vil inneholde spørsmål om bruk av kunstig intelligens (KI). Et slikt eksempel viser tydelig at undersøkelsene må endre seg over tid for at de fortsatt skal gi relevant informasjon.

Det er også verdt å merke seg at hvordan resultater blir rapportert har endret seg betydelig fra 1995 til 2023. I 1995 ble resultatene publisert i fire ulike nasjonale rapporter. I tillegg til hovedrapporten og rapporten knyttet til de praktiske prøvene som er nevnt ovenfor, utga den norske prosjektgruppen to rapporter, en for matematikk og en for naturfag, som inneholdt detaljerte analyser av enkeltoppgavene som ble gjort offentlig tilgjengelige (Brekke et al., 1998; Kjærnsli et al., 1999). I alle disse rapportene ser vi at undersøkelsene i stor grad er knyttet til konstruktivistisk læringsteori og spesifikk litteratur fra matematikk- og naturfagdidaktikk – med et stort fokus på såkalte hverdagsforestillinger og misoppfatninger. De norske forskerne som var involvert den gangen publiserte også relativt lite i vitenskapelige tidsskrifter. I dag er dette veldig annerledes. Nå publiseres gjerne hovedresultatene i en svært kortfattet rapport som følges opp med vitenskapelig publisering i fagfellevurderte antologier (som den du nå leser et kapittel fra) og i artikler i internasjonale tidsskrifter. Vi ser også at de norske forskerne som er involvert i studien i større grad publiserer sammen med kollegaer fra andre land. Til sammen dekker også dagens publisering et langt bredere spekter av teoretiske perspektiver og langt mer avanserte analysemetoder benyttes. Rapporteringen fra 1995 besto i hovedsak av presentasjon av deskriptiv statistikk og korrelasjoner. I denne antologien ser vi at regresjonsbaserte metoder, inkludert bruk av latente variabler, dominerer.

Dette viser for det første at de internasjonale undersøkelsene langt på vei reflekterer de dominerende perspektivene og forskningsfeltet i sin samtid (Blömeke et al., 2022). Samtidig viser disse forskjellene at tretti år med internasjonale undersøkelser i Norge (og mange andre land) har bidratt til å bygge opp nasjonal kompetanse i avansert kvantitativ metodologi innen utdanningsforskningen (Lockheed, 2015). Majoriteten av forfatterne i denne boka har eksempelvis tatt doktorgrader og/eller hatt postdoktorprosjekter tidlig i sine karrierer hvor forskningsspørsmål ble besvart ved hjelp av data fra TIMSS eller en annen internasjonal undersøkelse.

Å måle utvikling over tid med data fra TIMSS: trender og longitudinelle tilleggsstudier

Det fundamentale prinsippet for å måle endringer i elevers læringsutbytte, deres mestringsopplevelse, motivasjon, beskrivelser av hva som skjer i klasserommet – og alt annet som TIMSS inkluderer målinger av – er å *ikke endre målingen!* Metoden som brukes i TIMSS, og i de aller fleste tilfeller hvor man ønsker å måle endringer over tid, er å sørge for at man har utviklet en god test eller et spørreskjema som kan brukes ved flere anledninger. I både de internasjonale undersøkelsene og i de nasjonale prøvene inkluderer man deler av testen som ble gjennomført sist gang. Dette utgjør den såkalte ankertesten. Resultatene på ankertesten benyttes for å kalibrere resten av testen slik at den totale testskåren blir plassert på den samme skalaen i to (eller flere) år. I TIMSS er prøven ankret og lenket på denne måten helt tilbake til 1995. Man kan altså ved hjelp av dette i prinsippet trekke slutninger om hvordan norske elevers prestasjoner i matematikk og naturfag har endret seg i løpet av de siste nesten tretti årene. Ved å sammenstille data fra studier over flere år, kan man framstille og modellere det som gjerne betegnes som trender.

I tillegg har TIMSS et annet designelement for å kunne beskrive utvikling over tid. Den populasjonen som det trekkes elever fra på 5. trinn (populasjon 1) befinner seg på 9. trinn (populasjon 2) når neste syklus av TIMSS gjennomføres. Med andre ord har man i prinsippet et longitudinelt design hvor data fra to studier gjennomført med fire års mellomrom kan brukes for å si noe utviklingen over tid for noen bestemte alderskohorter av elever.

I noen land kan også resultater fra TIMSS koples til andre data for de samme elevene. I noen tilfeller kan man kople TIMSS-data til andre nasjonale datakilder for de samme elevene – både fortidige og framtidige data. Alternativt kan man gjennomføre spesifikke studier hvor man følger hele eller deler av et utvalg som deltok i TIMSS med egne datainnsamlinger over tid. Ved hjelp av slike koplinger til registerdata eller andre data kan man si noe om hvordan eksempelvis matematikkprestasjoner utvikler seg over tid, og man kan bedre modellere hvordan tidlige prestasjoner og andre karakteristikk ved elevene predikerer framtidige utfall. Å etablere slike tilleggsstudier med koplinger til andre datakilder er imidlertid ganske komplekst. Det er svært ressurskrevende, og muligheten begrenses ofte av personvernforordningen (GDPR). Så langt har man ikke lyktes med

å etablere slike koplinger i Norge, men det bør i prinsippet være mulig, siden man har lyktes med dette i andre land som i bunn og grunn har implementert det samme lovverket. I Tsjekkia har man for eksempel koplet 4. trinnselevne fra TIMSS 2011 til en nasjonal longitudinell studie som gjorde at de samme elevene kunne følges gjennom 9. trinn (Greger et al., 2022). I Sverige ble data fra TIMSS 2015 koplet til et begrenset utvalg av variabler fra nasjonale registre som inkluderte informasjon om elevers resultater på nasjonale prøver, karakterer i skolen og hjemmebakgrunn (Wiberg & Rolfman, 2021). Det bør også nevnes at en rekke land deltok i en utvidelse av TIMSS 2023 hvor undersøkelsen ble gjentatt for de samme elevene ett år senere (se <https://timss2023.org/longitudinal-study/>). Ved å gjennomføre slike tilleggsstudier kan man få enda bedre informasjon om endringer over tid. På denne måten får man informasjon om typiske elevprogresjoner i utdanningsløpet. I tillegg får man også genuint longitudinelle data for enkeltindivider som gir langt større muligheter for å trekke slutninger om kausale forhold.

Trege, små og ofte usynlige endringer

Som allerede antydnet er det ikke helt rett fram å lage studier som skal kunne brukes til å si noe om hvordan prosesser, strukturer og resultater i et utdanningssystem endrer seg over tid. I de internasjonale undersøkelsene er det flere eksempler på at endringer i definisjonen av populasjonen, i det innholdet som prøver og spørreskjemaer dekker og i studiedesign og metodikk fører til at sammenlikninger er utfordrende – og TIMSS er ikke et unntak.

Dersom man kun prioriterer hensynet til å kunne si noe om utvikling over tid, er idealet at man beholder alle instrumentene som brukes i undersøkelsen helt uendret over tid. Det er imidlertid mange konkurrerende hensyn som må veies opp mot hverandre når man utvikler spørreskjemaer og prøver i en undersøkelse som TIMSS (Hooper, 2023). Ikke minst må de som er ansvarlige for å utvikle instrumentene til studien, sørge for at de til enhver tid gir relevant og gyldig informasjon. Og for å oppnå dette må man altså avvike fra idealet om å beholde alt uendret ved å oppdatere og endre instrumentene i undersøkelsen slik at de til enhver tid også gir informasjon som er tidsriktig og relevant (Mullis & Martin, 2022).

De av oss som har vært inne i norske klasserom både i 1995 og i dag, vil nok enes om at det er mye som er gjenkjennbart, men det er også mye

som er forskjellig. Elevene sitter fortsatt som regel bak hver sin pult, det er vanligvis én lærer i klasserommet, langt de fleste lærerne har en godkjent lærerutdanning, elevene arbeider litt for seg selv og litt i grupper, undervisningen er organisert i fag og tidsperioder – og slik kunne vi fortsette. Men når vi sammenlikner disse to klasserommene, ser vi også at elevene i dag ikke sitter med en kladdebok, men ofte heller med en bærbar pc eller et nettbrett og løser oppgaver. Oppgavene de løser oppleves også innimellom som litt annerledes (og noen av oss som er opp i årene vil kanskje til og med tenke at «dette er da ikke en oppgave i matematikk!»). Kanskje legger vi også merke til at det er flere elever og lærere i klasserommene som har navn som vi ikke oppfatter å være «typisk norske». Norsk skole er med andre ord ikke den samme nå som 1995 – og data fra TIMSS kan bidra til å beskrive disse endringene.

Som allerede nevnt, må imidlertid TIMSS og andre undersøkelser endre seg for å kunne bidra med relevant informasjon om utdanningssystemene som deltar. TIMSS-undersøkelsen er for eksempel læreplanbasert. Det vil si at undersøkelsens formål enkelt sagt er å måle hvor godt elevene mestrer det faglige innholdet som læreplanene definerer at de skal lære. I en internasjonal undersøkelse er dette strengt tatt ikke mulig: Læreplanene er jo ulike på tvers av land. Men man forsøker altså å lage en «internasjonal læreplan» som på en best mulig måte dekker inn det som typisk undervises i de landene som deltar. Når en slik «internasjonal læreplan» endrer seg, må også TIMSS-undersøkelsen endre seg for å forbli relevant. Men konsekvensen av dette er at hvis 2023 undersøkelsen hadde blitt gjennomført med akkurat den samme prøven som ble brukt i 1995, ville elevene i de fleste land trolig ha prestert dårligere enn elevene i 1995 – og vice versa – fordi de i noen grad vil bli bedt om å løse oppgaver knyttet til innhold som ikke er dekt i deres undervisning. For en del av de andre endringene, eksempelvis endringer i elevenes og lærernes bakgrunn, er strategien at undersøkelsene prøver å fange inn dette i dataene. Dette er kontekstuell informasjon som er viktig for å kunne tolke de endringene man ser. På denne måten vet vi for eksempel at norsk skole i dag har en litt annen sammensetning av både elever og lærere.

Det som er felles for endringene i eksemplene ovenfor, er at de som regel er relativt små og gradvise. Slike små og gradvise endringer skjer hele tiden – og ofte er de ikke så lett å få øye på mens de skjer, og det er vanskelig å vite nøyaktig hvordan de eventuelt påvirker resultatene på prøvene og svarene som gis i spørreskjemaene. Ett eksempel på dette kan

være spørsmålet til elevene om «antall bøker hjemme». Dette er et standardspørsmål som er med i de fleste internasjonale skoleundersøkelsene, i en eller annen form. Hensikten med spørsmålet er at det gir en indikasjon på elevens sosioøkonomiske status (SØS). Denne variabelen har faktisk vært brukt i forskning i over hundre år (Holley, 1916). Hovedgrunnen til at denne variabelen er så populær, er at de aller fleste elevene klarer å gi et svar. Spørsmål for eksempel knyttet til foreldres utdanningsnivå (som jo er et mer direkte mål på SØS) har vist seg å være langt vanskeligere for elever å svare på, noe som fører til problematiske «huller» i datasettene for denne variabelen. Det er uenighet i forskningen om i hvilken grad spørsmålet om antall bøker hjemme er en god indikator for elevens sosioøkonomiske hjemmebakgrunn. Engzell (2019) viser blant annet at det er svak sammenheng mellom elevenes og foreldrenes svar på dette spørsmålet, og konkluderer med at man heller bør bruke andre indikatorer for SØS, mens Wiberg og Rolfman (2021) bruker koplinger til registerdata for å vise at denne variabelen er en treffsikker indikator for SØS i Sverige. Styrken på sammenhengen mellom resultatene på den faglige prøven og denne indikatoren for hjemmebakgrunn blir gjerne tolket som et mål på forekomsten av sosial ulikhet i de deltagende landene. Uavhengig av diskusjonen om hvor godt denne variabelen faktisk reflekterer SØS, er det betimelig å spørre om den underliggende relasjonen mellom antall bøker i hjemmet og familiens SØS er den samme i dag som for tretti år siden. Og som et oppfølgingsspørsmål: Er det mulig å bruke data fra dette spørsmålet for å si noe om hvordan sosial ulikhet i skoleprestasjoner har endret seg over tid? Det er kanskje rimelig å sammenlikne 2019 med 2023, men kan man studere endringen helt fra 1995 for svarene på et slikt spørsmål? Andre, og enda mer usynlige endringer over tid, kan for eksempel være hvordan elevene forholder seg til det å svare på denne typen undersøkelser. Kan vi vite at elevene i dag har samme holdninger og motivasjon til å gjøre sitt beste på prøvene og svare sannferdig på spørreskjemaene? Kan det være små endringer i holdninger og andre verdier over tid som også har en effekt på de dataene som samles? Spørsmålene ovenfor har jeg ikke entydige svar på. De er kun ment retorisk for å understøtte et hovedpoeng: Det er ikke usannsynlig at mange slike små og ofte usynlige endringer vil påvirke og forstyrre tolkninger som gjøres om endringer over tid, og trusselen fra slike snikende endringer er større for analyser som ser på endringer over større tidsrom.

Raske, store og synlige endringer

Andre endringer er større, veldig synlige og raske. Vi har allerede nevnt at TIMSS nå gjennomføres elektronisk, mens den tidligere i all hovedsak har vært gjennomført på papir. Når man gjør slike store endringer i hvordan dataene samles inn, vil de som står bak undersøkelsene inkludere et element i designet som gjør det mulig å studere hvordan endringen påvirker resultatene. I dette tilfellet måtte landene som gjennomførte elektronisk testing i 2019, også inkludere et tilleggsutvalg med elever som besvarte mange av de samme oppgavene på papir. Slike studier refereres gjerne til som brostudier, og hensikten er å undersøke i hvilken grad oppgavene har de samme psykometriske egenskapene i elektronisk og papirformat. I gjennomsnitt var forskjellene små, og man kunne justere for dette i den endelige modellen ved å inkludere en konstant som korrigerer for at oppgaver i papirformat er marginalt enklere (Martin et al., 2020). Med de elektroniske prøvene ble det også innført et sett med nye interaktive oppgaveformater, og slike endringer er det ikke like enkelt å studere og justere for. Dette er oppgaver som introduserer nye aspekter i den faglige kompetansen som prøveresultatene fanger inn.

Fra 2003 til 2007 var det en stor endring i hvordan de ulike versjonene av prøvene ble satt sammen. Kort og godt besto dette i å ha færre oppgaver per elev fordi man hadde observert at det internasjonalt var en økende andel elever som ikke rakk å gjøre seg ferdig med prøven på den gitte tiden i 2003. Konsekvensen av denne endringen ble også eksplisitt studert, og nødvendige justeringer ble gjort (Mullis & Martin, 2022). I perioden TIMSS har eksistert, har det altså vært gjennomført minst to store endringer som har ført til at modellen som brukes for å lenke elevenes skårer justeres. Og nok en gang, for å stille et åpent spørsmål hvor det ikke finnes et entydig svar: Hvor mange slike justeringer kan man egentlig tillate seg å gjøre uten at det til slutt medfører feilkilder i trendmålingene?

I tillegg er det en rekke andre endringer som fra et trendperspektiv er uheldig. I spørreskjemaene kan innholdet som inkluderes variere betydelig over tid. Man har en ordning i TIMSS hvor de deltakende landene kan prioritere hva de ønsker å inkludere i spørreundersøkelsene. Som regel vil endringene fra en undersøkelse til den neste være moderate, men over tid endrer innholdet i spørsmålene seg såpass mye at det kun er noen få målinger hvor man har tilgang til data for over en lengre tidsperiode. I tillegg finnes det også eksempler på at enkeltspørsmål endrer seg på et tidspunkt som

gjør at tidsserien endres. Et illustrerende eksempel her er at spørsmålet til elevene om foreldrenes utdanningsnivå i noen år har inkludert alternativet «Vet ikke». I 2007 valgte derfor ca. halvparten av elevene å rapportere at de faktisk ikke visste hva de foresattes høyeste utdanningsnivå var. Dette er en viktig variabel i beregningen av samlevariabelen for elevenes SØS. Dette gjør det vanskelig å bruke denne enkeltvariabelen eller samlevariabelen i analyser som har som mål å si noe om hvordan sammenhengen med SØS er i endring fra 1995 til 2023.

En siste stor endring som også har blitt omtalt kort ovenfor, er endringer i elevpopulasjonene som deltar i undersøkelsen. I 1995 hadde undersøkelsen en streng aldersbasert definisjon av klasstrinn som skulle delta. Dette ble opplagt gjort fordi sammenhengen mellom klasstrinn og alder var, og er fortsatt, veldig ulik i de deltakende landene. Den gangen var definisjonen for populasjon 1 at man skulle ha utvalg som representerte de to klasstrinnene med størst andeler av 9-åringer på en gitt dato. Og en tilsvarende definisjon eksisterte for populasjon 2, hvor de to klasstrinnene med flest 13-åringer skulle inkluderes. Norge deltok dermed den gangen med det som da var henholdsvis 2./3. trinn og 6./7. trinn (Martin & Kelly, 1997). I tiden siden den gangen har populasjonsdefinisjonen endret seg litt flere ganger, og i dag refereres det til at elevene som deltar skal være fra henholdsvis 4. og 8. trinn, gitt at gjennomsnittsalderen er minst 9,5 år ved det tidspunktet testen gjennomføres. Siden de norske elevene begynner på skolen ett år tidligere enn de landene vi ønsker å sammenlikne oss med, spesielt de andre nordiske landene, fikk Norge framforhandlet et unntak fra denne definisjonen i 2015, og vi fikk fra dette tidspunktet delta med det som for oss er 5. og 9. trinn (Bergem et al., 2016). De norske elevene har derfor samme gjennomsnittsalder som elevene i de andre deltakende nordiske landene. Disse endringene i definisjonene medfører imidlertid at elevene som deltok i TIMSS 2023 i de fleste deltakende land i gjennomsnitt har gått lenger på skolen og er eldre enn de elevene som deltok i 1995.

I 2015 deltok Norge med elever på henholdsvis 4./5. trinn og 8./9. for de to populasjonene. Resultatene viser at i gjennomsnitt skårer norske 5. trinnselever om lag 50 poeng bedre enn elever på 4. trinn, mens forskjellen mellom 8. og 9. trinn var om lag 25 poeng (begge forskjellene er for matematikk) (Bergem et al., 2016). Olsen og Bjørnsson (2018) gjorde en analyse av de samme dataene som dekomponerte disse forskjellene i to: en del som kan knyttes til at elevene på 5. trinn har gått ett år lenger på skolen, og en del som reflekterer at elevene på 5. trinn i gjennomsnitt er ett

år eldre. Omtrent 60 prosent av disse forskjellene for begge populasjonene kan knyttes til at man har gått ett år lenger i skolen, mens de resterende 40 prosent kan knyttes til at elevene har blitt ett år eldre. I sum må vi derfor nok en gang slå fast at den ikke-konstante definisjonen av populasjonene i TIMSS skaper utfordringer for den som ønsker å bruke data fra undersøkelsen for å se på endringer over tid, og nok en gang er problemet størst for tolkninger av endringer over lengre tidsrom. Denne store endringen i 2015 er også hovedgrunnen til at de fleste kapitlene i denne boka som ser på endringer over tid, kun forholder seg til data fra de tre siste syklusene i TIMSS.

Det finnes også andre raske og store endringer som påvirker tolkninger av trender i de internasjonale undersøkelsene. Her er covid-19-pandemien et godt eksempel. Med studier i henholdsvis 2019 og 2023 har nok resultatene for TIMSS blitt relativt moderat påvirket av dette, men det er et faktum at elevene som deltok i 2023 har gått glipp av et betydelig antall undervisningstimer sammenliknet med elevene som har deltatt i de tidligere undersøkelsene. Dette var jo en ekstrem hendelse, men fungerer likevel som et eksempel på at det kan være forhold som ikke er knyttet til undervisning, skole og læreplaner som i noen grad kan virke inn på resultatene i undersøkelsen på en måte som skaper ustabilitet i trendlinjene.

Hvordan kan vi forholde oss til dette?

Hvis man skal gi et absolutt og entydig svar på spørsmålet i tittelen for dette kapittelet, kan man vanskelig svare noe annet enn «Nei». Ovenfor har jeg omtalt noen direkte og lett identifiserbare årsaker til at data fra 1995-undersøkelsen ikke kan brukes til sammenlikninger med resultater fra den siste undersøkelsen i 2023. For det første sammenlikner vi elever som representerer to veldig forskjellige populasjoner. De norske elevene i den siste undersøkelsen har i gjennomsnitt gått 2,5 år lenger i skolen, og de er i gjennomsnitt 1,5 år eldre enn de som deltok i den første TIMSS-undersøkelsen. For det andre har det vært tre relativt store endringer i prøvedesignet siden den tid. Dette har det blitt justert for i egne tilleggsstudier, men slike justeringer vil alltid medføre usikkerheter i estimatene – som ikke er inkludert i usikkerhetsestimatene som i dag rapporteres for elevenes matematikk og naturfagprestasjoner. Dessuten har spørreskjemaene blitt så mye endret over tid at man til slutt står igjen med et svært begrenset datasett som kan anvendes for å si noe om endringer over tid. I

tillegg har jeg ovenfor drøftet, hovedsakelig ved å bruke noen eksempler, at tidsserier med denne typen datainnsamlinger helt uunngåelige vil være utsatt for små, snikende og til dels usynlige endringer.

Så hva kan vi egentlig gjøre? Og for å reformulere spørsmålet i tittelen til et som er mer meningsfullt: Er tidsseriedata av denne typen ubrukelige? Selvfølgelig ikke, er mitt svar. For Norge, og mange andre land, er resultater fra de internasjonale undersøkelsene de beste, og ofte de eneste, informasjonskildene som kan brukes for å vurdere kvalitetsutviklingen i skolen. Disse undersøkelsene har en metodisk rigiditet og kvalitet som på mange måter representerer en gullstandard for andre som ønsker å samle data som er representative for og kan generaliseres til en veldefinert populasjon. I tillegg rapporterer og dokumenterer organisasjonene bak studiene detaljer om gjennomføringen og gjør data tilgjengelig for andre, hovedsakelig for å stimulere til at dataene blir maksimalt utnyttet til forskning, men også for å gjøre det mulig for andre å ettergå og undersøke kvaliteten i dataene.

Som jeg har vist til, er konsekvensene av disse små og store, synlige og usynlige, trege og raske endringene sannsynligvis spesielt store for sammenlikninger hvor man bruker data som spenner et svært stort tidsrom. Sammenlikninger over korte tidsrom er mindre problematiske. Men man bør erkjenne at feilmarginene som beregnes for endringene med stor sannsynlighet egentlig er litt større, og de er større jo lengre tid som er gått mellom målingene som sammenliknes. Hvor mye større har jeg ikke grunnlag for å si, men en enkel kjøreregel i tolkningen av endringer kan være å ikke legge særlig vekt på knapt signifikante resultater, men heller vektlegge de endringene som framstår som betydelige og substansielt meningsfulle. Og i tråd med vanlige standarder for forskning, kan nok en kort omtale av denne utfordringen med sammenlikninger over tid nevnes som en begrensning i publikasjoner av analyser som benytter data for en lang tidsperiode.

De som er ansvarlige for de internasjonale undersøkelsene, bør etter min mening prioritere enda sterkere å bevare større deler av spørreskjemaene uendret over lengre tid. Det er forståelig at det er et press fra deltakende land om å inkludere målinger av nye fenomener som er av stor betydning for politikkutformingen her og nå, men det er viktig å knesette noen prinsipper. Et prinsipp kan for eksempel være at alle spørsmål som inngår i samlevariabler (konstrukter) skal være inkludert i minst tre eller fire sykluser. Utsiftingen av spørsmål i disse skjemaene må derfor skje

over lang tid og gradvis. Tilsvarende prinsipper finnes for øvrig allerede for utskifting av de faglige oppgavene over tid.

I tillegg bør TIMSS ta det andre tidsdesignet i studien på mer alvor: Gjennom denne undersøkelsen samler man faktisk data som er genuint longitudinelle for noen bestemte alderskohorter. Det som er populasjon 1 i ett år (for eksempel i 2019), er så å si identisk med det som fire år senere utgjør populasjon 2 (som i dette eksempelet er i 2023). I dag er det dessverre store begrensninger i hvordan dette designet kan anvendes i analyser. Den aller største begrensningen for å kunne gjøre analyser av hvordan utfall for en bestemt populasjon har endret seg fire år senere er at de to faglige prøvene er helt separate – de er ikke lenket sammen gjennom et sett med felles oppgaver (såkalt vertikal skalering). I Norge har vi sett at slike design definitivt er mulig å etablere for de nasjonale prøvene i matematikk (Ræder et al., 2022). I tillegg er spørreskjemaer til de to populasjonene så ulike at det er vanskelig å bruke data fra disse for å modellere endringer over tid. Jeg er ikke kjent med i hvilken grad en bedre samordning mellom instrumentene i populasjon 1 og 2 har blitt drøftet og vurdert av de som beslutter og utformer undersøkelsene, men personlig ser jeg at TIMSS her har et stort uutløst potensial for å gi deltakerlandene enda bedre informasjon for å evaluere sine utdanningssystemer.

Det finnes også statistiske grep som allerede er i bruk når man gjør større endringer i design eller metodikk. Jeg har tidligere nevnt et eksempel på en såkalt brostudie for å justere for større endringer i prøvedesignet i TIMSS. Et alternativ til å justere ved hjelp av en slik studie er i stedet å rekalkibrere historiske data når man endrer design eller gjør endringer i hvordan man beregner verdier for konstrukter. Dette er eksempelvis gjort noen ganger i PISA. Definisjonen av og metodikken for å konstruere samlevariabelen for SØS har i denne studien egentlig hele tiden vært i endring (Avvisati, 2020). I forbindelse med gjennomføringen av PISA i 2018 ble det introdusert en ny modell for å beregne denne samleskåren, og sammen med publiseringen av rapportene og dataene fra PISA 2022 inkluderte man reskalerte skårer tilbake til PISA-studien fra 2012² (OECD, 2024). Tilsvarende publiserte OECD reskalerte landsgjennomsnitt bakover i tid etter at de i 2015 endret den psykometriske modellen som benyttes for å beregne elevenes skårer for de faglige prestasjonene (OECD, 2016). Ulempen med dette er selvsagt at

2 <https://www.oecd.org/en/data/datasets/pisa-2022-database.html>

det kan skape forvirring og usikkerhet når et gjennomsnitt for et land som er allerede rapportert for mange år siden, viser seg å være litt annerledes dersom den hadde blitt beregnet på en litt annen måte.

Denne siste løsningen ligger tett opp til hvordan Statistisk sentralbyrå (SSB), og tilsvarende organisasjoner i andre land, beregner konsumprisindekser (KPI). Dette er jo veldig lange tidsserier, og de skal enkelt sagt fange inn hvor dyrt det er for oss å leve. Men forbruksmønsteret vårt er i endring over tid – noen store og synlige endringer, og noen små og snikende endringer. På et eller annet tidspunkt innser man at forbruksmønsteret er såpass mye endret at den gjeldende modellen ikke lenger er gyldig. Konsekvensen er at man da setter et nytt basisår, endrer modellen som brukes, og så forsøker man i den grad det er mulig å rekalkulere bakover. Sist gang dette skjedde i Norge var i 2015. SSB oppdaterer også løpende beskrivelser av de endringene som gjøres i KPI (SSB, 2025). I prinsippet bør det være mulig å etablere en tilsvarende praksis også for de mest sentrale indeksene som de internasjonale undersøkelsene har inkludert i alle studiene, det vil i praksis si samleskårer for de faglige områdene og for SØS.

Kapitlene i denne boka viser at forfatterne er ydmyke og forsiktige med å tolke små endringer over tid. Dette er en fornuftig praksis, gitt de utfordringene som er definert ovenfor. Personlig tror jeg at TIMSS, og noen av de andre internasjonale undersøkelsene, vil fortsette i tretti år til, og utfordringene jeg har nevnt ovenfor vil ikke bli mindre framover. Jeg ser også at de organisasjonene som er ansvarlige for å utvikle og lede den internasjonale gjennomføringen av studiene har økende oppmerksomhet om å bevare trendseriene som de rapporterer. Dette er sannsynligvis noen av de viktigste produktene de har å tilby forskere og politikkutformere som bruker data og resultater fra undersøkelsene. Jeg vil derfor ikke bli overrasket om en eller flere av disse undersøkelsene i nær framtid tar skrittet fullt ut og setter et nytt basisår og foretar en større rekalkulering av historiske data – nettopp for å bevare tilliten til at undersøkelsene gir viktige data for å undersøke endringer over tid.

Forfatterbiografi

Rolf Vegar Olsen er instituttleder ved Institutt for spesialpedagogikk ved Universitetet i Oslo. Han har en bred portefølje av forskning knyttet til nasjonale og internasjonale prøver. I tillegg har han vært involvert i en rekke komiteer som blant annet har gitt anbefalinger om nasjonale prøver i Norge og Danmark

Referanser

- Avvisati, F. (2020). The measure of socio-economic status in PISA: A review and some suggested improvements. *Large-scale Assessments in Education*, 8, Artikkel 8. <https://doi.org/10.1186/s40536-020-00086-x>
- Bergem, O. K., Haarstein, H. & Nilsen, T. (2016). TIMSS 2015. I O. K. Bergem, H. Kaarstein & T. Nilsen (Red.), *Vi kan lykkes i realfag* (s. 11–21). Universitetsforlaget.
- Blömeke, S., Nilsen, T., Olsen, R. V. & Gustafsson, J.-E. (2022). Conceptual and methodological accomplishments of ILSAs, remaining criticism and limitations. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education* (s. 603–655). Springer. https://doi.org/10.1007/978-3-030-88178-8_22
- Brekke, G., Kobberstad, T., Lie, S. & Turmo, A. (1998). *Hva i all verden kan elevene i matematikk? Oppgaver med resultater og kommentarer*. Universitetsforlaget.
- Engzell, P. (2019). What do books in the home proxy for? A cautionary tale. *Sociological Methods & Research*, 50(4), 1487–1514. <https://doi.org/10.1177/0049124119826143>
- Greger, D., Straková, J. & Martinková, P. (2022). Extending the ILSA study design to a longitudinal design. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education* (s. 955–977). Springer. https://doi.org/10.1007/978-3-030-88178-8_31
- Hastedt, D. (2020). History and current state of international student assessment. I H. Harju-Luukkainen, N. McElvany & J. Stang (Red.), *Monitoring student achievement in the 21st century* (s. 21–37). Springer. https://doi.org/10.1007/978-3-030-38969-7_3
- Hastedt, D. & Sibberns, H. (2022). Future directions, recommendations, and potential developments of ILSA. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education* (s. 579–601). Springer. https://doi.org/10.1007/978-3-030-88178-8_24
- Holley, C. E. (1916). *The relationship between persistence in school and home conditions*. University of Chicago Press.
- Hooper, M. (2022). Dilemmas in developing context questionnaires for international large-scale assessments. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.), *International handbook of comparative large-scale studies in education*. Springer. https://doi.org/10.1007/978-3-030-38298-8_29-1
- Kind, P. M., Kjærnsli, M., Lie, S. & Turmo, A. (1999). *Hva i all verden gjør elevene i realfag? Praktiske oppgaver i matematikk og naturfag*. Institutt for lærerutdanning og skoleutvikling, UiO.
- Kjærnsli, M., Lie, S., Stokke, K. H. & Turmo, A. (1999). *Hva i all verden kan elevene i naturfag? Oppgaver med resultater og kommentarer*. Universitetsforlaget.
- Lie, S., Kjærnsli, M. & Brekke, G. (1997). *Hva i all verden skjer i realfagene? Internasjonalt lys på trettenåringers kunnskaper, holdninger og undervisning i norsk skole*. Institutt for lærerutdanning og skoleutvikling, UiO.
- Lockheed, M. (2015). *Why do countries participate in international large-scale assessments? The case of PISA* (Policy research working paper; 7447). World Bank Group. <http://documents.worldbank.org/curated/en/692561468000587471>
- Martin, M. O. & Kelly, D. (1997). *Third International Mathematics and Science Study technical report: Vol. II. Implementation and analysis – primary and middle school years*. Center for the Study of Testing, Evaluation, and Educational Policy, Boston College. <https://timssandpirls.bc.edu/timss1995i/TechVol3.html>
- Martin, M. O., von Davier, M. & Mullis, I. V. S. (Red.). (2020). *Methods and procedures: TIMSS 2019 technical report*. TIMSS & PIRLS International Study Center, Boston College. <https://timss.bc.edu/timss2019/methods/index.html>
- Mullis, I. V. S. & Martin, M. O. (2022). IEA's TIMSS and PIRLS: Measuring long-term trends in student achievement. I T. Nilsen, A. Stancel-Piątak & J.-E. Gustafsson (Red.),

- International handbook of comparative large-scale studies in education*. Springer. https://doi.org/10.1007/978-3-030-38298-8_15-1
- OECD. (2016). *PISA 2015 results: Vol. I. Excellence and equity in education*. OECD Publishing. <http://dx.doi.org/10.1787/9789264266490-en>.
- OECD. (2024). *PISA 2022 technical report*. OECD Publishing. <https://doi.org/10.1787/01820d6d-en>.
- Olsen, R. V. & Björnsson, J. K. (2018). Fødselsmåned og skoleprestasjoner. I J. K. Björnsson & R. V. Olsen (Red.), *Tjue år med TIMSS og PISA i Norge: Trender og nye analyser*. Universitetsforlaget. <http://dx.doi.org/10.18261/9788215030067-2018-05>.
- Robitaille, D. F., Schmidt, W. H., Raizen, S., Mc Knight, C., Britton, E. & Nicol, C. (1993). *Curriculum frameworks for mathematics and science*. Pacific Educational Press.
- Ræder, H. G., Andersson, B. & Olsen, R. V. (2022). Numeracy across grades – vertically scaling the Norwegian national numeracy tests. *Assessment in Education: Principles, Policy & Practice*, 29(6), 653–673. <https://doi.org/10.1080/0969594X.2022.2147483>
- Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T. & Wiley, D. E. (1997). *Many visions, many aims: Vol. I. A cross-national investigation of curricular intentions in school mathematics*. Kluwer Academic Publishers.
- Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J. & Wolfe, R. G. (1997). *Many visions, many aims: Vol. II. A cross-national investigation of curricular intentions in school science*. Kluwer.
- Statistisk sentralbyrå. (2025, 28. april). *Endringer i konsumprisindeksen (KPI) – oppdateres fortløpende*. <https://www.ssb.no/priser-og-prisindekser/konsumpriser/statistikk/konsumprisindeksen/artikler/endringer-i-konsumprisindeksen-kpi>
- Wiberg, M. & Rolfman, E. (2021). Students' self-reported background SES measures in TIMSS in relation to register SES measures when analysing students' achievements in Sweden. *Scandinavian Journal of Educational Research*, 67(1), 69–82. <https://doi.org/10.1080/00313831.2021.1983863>

Denne antologien oppsummerer utviklingstrekk i norske elevers prestasjoner og læringsforutsetninger i matematikk og naturfag, basert på data fra TIMSS-undersøkelsen. Analysene viser at ulikhetene i skolen har økt, særlig knyttet til elevenes sosioøkonomiske bakgrunn, skolemiljø og motivasjon. Selv om gjennomsnittsprestasjonene har vært relativt stabile over tid, har forskjellene mellom elevene økt. Dette stiller økte krav til skolens evne til å tilby en opplæring som er likeverdig for alle.

Skoleledelse fremstår som en nøkkelfaktor. Rektorer med relevant utdanning og erfaring er positivt assosiert med høyere elevprestasjoner. I tillegg har undervisningskvalitet, høy selvtilit, motivasjon og læreplanimplementering stor betydning. TIMSS-data viser også at innsats i prøvesituasjonen henger sammen med prestasjon.

Studiene i antologien peker også på behovet for å styrke elevenes begrepsforståelse og kritiske tenkning, særlig innen algebra og tolkning av grafiske framstillinger.

Antologien gir et forskningsbasert grunnlag for videre skoleutvikling og politikkutforming, og understreker at skolen har et handlingsrom til å fremme sosial utjevning og faglig utvikling – dersom den får riktige rammer og ressurser.

ISBN ISBN 978-82-02-89476-4



www.cda.no