Silje Andresen, Aina Fossum,
Jon Rogstad og Bjørn Smestad

På prøve

Evaluering av matematikkeksamen på 10. trinn våren 2017

Fafo

Silje Andresen, Aina Fossum,
Jon Rogstad og Bjørn Smestad

På prøve

Evaluering av matematikkeksamen på 10. trinn våren 2017

Fafo-rapport 2017:36

ISBN 978-82-324-0410-0 (papirutgave)
ISBN 978-82-324-0411-7 (nettutgave)
ISSN 0801-6143 (papirutgave)
ISSN 2387-6859 (nettutgave)

Omslagsbilde: Colourbox.com
Omslag: Fafos Informasjonsavdeling
Trykk: Allkopi AS

Innhold

Forord

Dette er første delrapport i evalueringen av matematikkeksamen på 10. trinn, som Fafo gjennomfører i samarbeid med Høgskolen i Oslo og Akershus. Hele prosjektet skal være ferdig høsten 2019. Prosjektet er finansiert av Utdanningsdirektoratet, og vi vil takke oppdragsgiver ved Grethe Hovland, Gregorios Brogstad, Jan Eivind Sodeland og Marianne Elgen Bøhler for gode innspill og kommentarer underveis. Vi har også mottatt verdifulle kommentarer fra ekstern forsker i Utdanningsdirektoratets forskerpanel. Kaja Reegård ved Fafo har kvalitetssikret rapporten, og vi vil takke for uvurderlige kommentarer. Størst takk går likevel til informantene som har stilt opp og vært generøse både med sin tid og sin kunnskap. En siste takksigelse er til publikasjonsavdelingen på Fafo.

Fafo, november 2017
Jon Rogstad, prosjektleder

Sammendrag

Denne rapporten handler om eksamenen i matematikk på 10. trinn, som ble gitt våren 2017. Rapporten er den første av tre rapporter. De to neste handler om matematikkeksamenene i 2018 og 2019. Temaet for evalueringen er hvorvidt eksamensoppgavene er utformet forståelig, og om oppgavene som ble gitt, var tilpasset undervisningen elevene hadde hatt. I årets rapport har vi i tillegg sett på sensorenes erfaringer når det gjaldt skolering i forkant og vurdering av besvarelsene. Ut fra et overordnet perspektiv spør vi altså om årets eksamen er og oppleves å være rettferdig. Det forutsetter at spørsmålene er tilgjengelige i form og innhold, slik at det er de matematiske ferdighetene elevene prøves i. Framstillingen er organisert for å besvare fem hovedspørsmål.

1. Er eksamensoppgaven utformet på en forståelig måte slik at det er elevens matematikkompetanse som vurderes?
2. Hvor god sammenheng er det mellom læreplanen, eksamenen og den opplæringen elevene har fått?
3. Inneholder eksamenen oppgaver av ulik vanskegrad, som kan måle alle kompetansenivå?
4. Hvordan vurderer elevene eksamenens arbeidsmengde i forhold til den tiden de hadde til rådighet?
5. I hvilken grad er det samsvar mellom ulike sensorers vurdering?

Datainnsamlingen for undersøkelsen har vært sammensatt. Vi har sendt ut et elektronisk spørreskjema til matematikklærere på 10. trinn som har hatt elever oppe til eksamen i 2017, gjennomført casestudier på fire forskjellige skoler, vært til stede rett etter eksamen på en skole, bedt om å få vurderingsskjemaer tilsendt fra sensorer og analysert eksamenen i lys av lærebøkene og teori om språk og layout i matematikkoppgaver. I tillegg har vi lagt til spørsmål i Utdanningsdirektoratets spørreundersøkelse til sensorene og deltatt på sensormøter.

Basert på dataene vi har samlet inn, framstår matematikkeksamenen på 10. trinn våren 2017 som god og rettferdig. Dette var en oppfatning som var gjennomgående blant elever, lærere og sensorer og utgjør følgelig vår hovedkonklusjon. Det er videre en rekke andre og mer detaljerte funn som først og fremst underbygger hovedkonklusjonen, men som også nyanserer bildet.

Et viktig tema i rapporten er hvorvidt elevene forstår spørsmålene i eksamensoppgaven. Dersom de ikke forstår oppgavene de gis, er det heller ikke overraskende at de svarer feil. Sagt på en annen måte, kan det være at de i større grad prøves i norsk enn i matematiske ferdigheter. Denne erkjennelsen var utgangspunktet for at vi satte språkets betydning for elevens prestasjoner til eksamen som et overordnet tema. Men forståelse, eller mangel på sådan, handler også om bruken av illustrasjoner. I en del tilfeller er de meget viktige og opplysende. I andre tilfeller kan de skape forvirring. Ved årets eksamen var illustrasjonene i hovedsak oppklarende, men ikke alltid. I en av oppgavene var det en feil i en illustrasjon, noe som kunne gjøre at elever ga opp å besvare den.

Et annet tema i rapporten er at årets eksamen ble oppfattet som god, fordi oppgavene ble ansett å berøre sentrale emner i kompetansemålene. De fleste lærerne ga også uttrykk for at de mente det var samsvar mellom kompetansekrav og hva som ble prøvd til eksamen. De kvalitative intervjuene viste at over tid er det ingen temaer som systematisk er blitt utelatt fra eksamensoppgavene. Ut fra informasjonen vi fikk, var elevene også svært godt fornøyde med undervisningen de har fått, målt opp mot de spørsmålene de fikk til eksamen. Det synes som om de var skolert, noe som innebærer at forskjeller i prestasjoner først og fremst skyldes at de har ulike matematiske kunnskaper på eksamenstidspunktet. En grunn til forskjellene i ferdigheter kunne være at elevene bruker forskjellige læreverk. Vi fant imidlertid ikke noen systematisk favorisering av elever som har hatt en bestemt lærebok. Dataene indikerer imidlertid at det er en utfordring knyttet til å gi elevene et likt utgangspunkt for å løse eksamensoppgaver som krever digitale hjelpemidler.

Et tredje tema er hvor vanskelig årets eksamen var, både når det gjelder det rent innholdsmessige og om kandidatene fikk nok tid til å vise alt de kan. Analysene viser at det er stor variasjon i vanskegrad, men at alle kandidatene fikk vist fram noe. Det ble imidlertid framhevet at de svakest presterende elevene nær på ikke mestret noen av oppgavene i del 2. Knapphet på tid kan være en viktig faktor for å forklare forskjellene i prestasjoner på del 2. Andelen ubesvarte oppgaver øker mot slutten av settet, noe som kan skyldes at kandidatene fikk dårlig tid. De fleste av lærerne mente imidlertid at arbeidsmengden var passe. Dette er til tross for at om lag en tredjedel av lærerne hadde elever som hadde sagt at de fikk for dårlig tid.

Rapportens fjerde tema er språk. Språket i årets eksamensoppgave var i hovedsak godt, men det var også enkelte ord som få forstod, som «jerrykanne» og «klaffebro», men begge deler var illustrert. I den sammenheng er det relevant å framheve at flere av matematikklærerne mente at oppgaver med mye tekst hindrer noen av elevene i å vise sine matematiske ferdigheter. Det gjelder særlig elever med minoritetsspråklig bakgrunn.

Det femte og siste temaet er sensorenes vurderinger. Sensorskoleringen synes å være svært vellykket. Alle 129 sensorer som hadde valgt å forklare hva de mente, var positive. Det skal understrekes at det er godt samsvar mellom sensorene i deres karakterforslag før fellessensurmøtet. Når det gjelder ulike vurderinger, synes det først og fremst å være knyttet til del 2 av eksamenen. Enkelte sensorer etterlyste bedre veiledning i sensur av enkelte oppgaver. Det gjaldt særlig å få klarere retningslinjer for sensurering av oppgaver som krever digitale hjelpemidler.

Summary

To the test. Evaluation of the tenth grade mathematics examination in the spring of 2017

This report describes the tenth grade mathematics examination given in the spring of 2017. The report is the first in a series of three; the two to follow will address the mathematics examinations in 2018 and 2019 respectively. The topic of this evaluation is whether the examination questions were comprehensibly formulated and whether they were congruent with the teaching that the students had received. In addition, this year's report investigates the experiences of the examiners with regard to the guidance provided and the assessment of the examination papers. In a general perspective we are thus asking whether this year's examination was fair and perceived to be fair. This presupposes that the questions are accessible in terms of their form and content, so that the students' mathematical skills are the ones that are tested. Our description is structured to answer five main questions:

1. Is the exam in mathematics designed in a comprehensible manner, so as to test the students' mathematics skills?
2. How good is the consistency between the curriculum, the examination and the teaching provided to the students?
3. Does the examination include questions with a varying degree of difficulty that can measure all skill levels?
4. How do the students assess the examination workload in relation to the time available?
5. To what degree is there concurrence between the assessments made by different examiners?

The data collection for this investigation was a complex one. We have sent electronic questionnaires to the tenth grade maths teachers whose students have sat the 2017 examination, undertaken case studies at four different schools, been present immediately after the examination in one school, requested assessment forms from examiners, and analysed the examinations in light of textbooks and theories on language and layout of mathematics exams. Moreover, we have added questions to the survey of examiners

by the Directorate of Education and Training, and we have participated in examiners' meetings.

Based on the data that we have collected, the tenth grade mathematics examination in the spring of 2017 appear to have been fair and proper. This opinion prevailed among students, teachers and examiners, and is accordingly our main conclusion. A number of other and more detailed findings serve to corroborate this conclusion, but they also add nuances to the main picture.

A key topic in the report concerns whether the students understand the questions in the question paper; if they fail to understand the questions, it should come as no surprise that they give the wrong answers. In other words, it might happen that their Norwegian language skills are tested more than their mathematical skills. This recognition was the reason why we made the importance of language for the students' examination performance into a general topic. However, understanding – or lack of such – is also associated with the use of illustrations. In some cases, they are highly important and illuminating, whereas in others they may cause confusion. In this year's exam the illustrations were mainly illuminating, but not without exception. The illustration for one of the questions contained an error that may have caused some students to give up answering it.

A second topic in this report describes how this year's examination was perceived as good, because the questions were considered to touch upon keys elements of the competence objectives. Most teachers also claimed that there was concurrence between the skills requirements and what was tested by the examination. The qualitative interviews showed that over time, no topics have been systematically excluded from the examination questions. According to information provided to us, the students were highly satisfied with the teaching that had been given to them, when seen in relation to the questions included in the examination. They appear to have been well trained, meaning that differences in performance were mainly caused by variations in their mathematical skill level at the time of the examination. Use of different sets of textbooks could be one reason for the differences in skills. However, we found no systematic bias in favour of students who had used a specific textbook. On the other hand, the data indicate that there are challenges associated with providing the students with equal opportunities when it comes to solving examination questions that require digital tools.

A third topic concerns the degree of difficulty of this year's examination, in terms of content as well as whether the allotted time was sufficient for the candidates to demonstrate their full range of skills. The analyses show that the degree of difficulty varied considerably, but also that all the candidates were able to demonstrate at least some of their skills. It was emphasised, however, that the lowest performing candidates were barely able to cope with any of the questions in the second part of the examination. Time constraints may explain the variability in performance in the second part. The proportion of unanswered questions increases towards the end of the set, which may indicate that the candidates ran out of time. Most teachers, on the other hand, believed that the workload was appropriate, despite the fact that approximately one-third of the teachers had students reporting to them that the allotted time was insufficient.

The fourth topic in this report is language. The language used in this year's exam was mainly appropriate, although there were some words that were not universally understood, such as 'jerrykanne' and 'klaffebro', although both of these were illustrated. In this context it is relevant to note that according to several maths teachers, questions with a lot of text in them prevent some of the students from demonstrating their mathematical skills. This applies to minority-language students in particular.

he fifth and final topic is the examiners' assessments. The guidance for external examiners appears to have been highly successful. All of the 129 examiners who had volunteered to explain their opinion took a positive view. It must be emphasised that there was a high degree of concurrence between the examiners in their grading proposals before the joint examiners' meeting. The differences in grading mainly appear to occur in the second part of the examination. Some examiners called for better guidance in the grading of some of the questions. This applied especially to clearer guidelines for grading of questions that require use of digital tools.

1 Innledning

Eksamen er en grunnleggende byggestein i utdanningssystemet. Sammen med standpunktkarakter utgjør eksamen en sluttvurdering av elevens kompetanse som får følger for opptak til videre utdanning. Eksamen bygger på læreplanen i faget og er regulert gjennom forskrift til opplæringsloven. Formålet med eksamen beskrives slik på Utdanningsdirektoratet sine nettsider:

> Formålet med eksamen er todelt. Kandidaten skal få anledning til å vise sin kompetanse i samsvar med læreplanen, og eksamenskarakteren skal gi informasjon om kandidatens individuelle kompetanse i faget, slik den ble uttrykt på eksamensdagen.[1]

I denne rapporten analyseres eksamenen i matematikk på 10. trinn våren 2017. Årets publisering er den første av totalt tre. De to neste vil handle om eksamenene i 2018 og 2019. Samlet vil dette gi oss et bilde av variasjon og eventuell endring. Årets rapport er derfor å regne som en delpublikasjon, men som kan leses som et eget bidrag til spørsmål om skoleutvikling og kvalitetsforbedringer.

Utgangspunktet for rapporten er en bekymring. Våren 2015 evaluerte Nasjonalt senter for matematikk et utvalg eksamensoppgaver og konkluderte med at eksamensoppgavene varierer fra år til år i blant annet anvendelse og vektlegging av kompetansemålene og vanskegraden på oppgavene, og at språket ikke er så presist som ønsket (Matematikksenteret 2015a). Utdanningsdirektoratet har fulgt opp Matematikksenterets anbefalinger med både kortsiktige og langsiktige tiltak. Regjeringens satsing på realfag er oppsummert i Kunnskapsdepartementets strategi «Tett på realfag – Nasjonal strategi for realfag i grunnopplæringen i barnehage og grunnskole (2015-2019)» (Kunnskapsdepartementet 2015). Det å evaluere matematikkeksamenen er ett av flere mål i realfagsstrategien.

Kvaliteten på en eksamen er betinget av at det er samsvar mellom opplæringen i løpet av skoleåret og det som gjøres til gjenstand for prøve. I tillegg må det ligge et likhetsprinsipp til grunn som sikrer likebehandling mellom kandidatene i vurderingen av eksamen. Videre må eksamen være, og oppleves å være, rettferdig fra år til år for å ha legitimitet som vurdering av deler av elevenes sluttkompetanse i faget og for å inngå som en del av vurderingen av kvaliteten i opplæringen. Problemstillingene i rapporten er som følger:

[1] https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-eksamen/2.hva-er-eksamen/

1) Utforming

a. Er eksamensoppgaven utformet slik at det er elevens matematikkompetanse som vurderes?
b. Hvordan er oppgavene er bygget opp (f.eks. rekkefølge og avhengighet mellom oppgaver), og hvordan kan dette påvirke elevenes prestasjoner på eksamen?
c. Er oppgavetekstene språklig sett gode og forståelige for elevene?
d. Hvordan kan oppgavenes design og layout ha betydning for elevenes prestasjoner?

2) Sammenheng med læreplanen og med opplæringen

a. Hvordan oppfatter lærere og sensorer sammenhengen mellom eksamenen og læreplanen i faget?
b. Er det samsvar mellom eksamenen og hva elevene har kjennskap til og erfaring med fra opplæringen?

3) Eksamenens vanskegrad

a. Inneholder eksamenen oppgaver av ulik vanskegrad, som kan måle alle kompetansenivåer (karakterene 1–6)?
b. Hvordan opplever elevene vanskegraden på eksamensoppgavene?

4) Eksamenens arbeidsmengde

a. Hvordan vurderer elevene eksamenens arbeidsmengde i forhold til den tiden de har fått til rådighet totalt og på del 1 og del 2 av prøven?

5) Vurdering av eksamenen

a. I hvilken grad er det samsvar mellom ulike sensorers vurdering?
b. Hvor god støtte opplever sensorene å ha fått til å vurdere eksamensbesvarelsene på en lik og rettferdig måte gjennom eksamensveileding, sensorveiledning og forhåndssensurrapport?
c. Hva er lærere og elevers oppfatning av en lik og rettferdig vurdering av eksamen i matematikk?

Dataene som analysene i rapporten bygger på, er flere. Gjennom dokumentanalyse, survey til sensorer og lærere, analyse av vurderingsskjemaer fra sensorene samt caseanalyse ved fire skoler hvor vi har intervjuet elever og lærere i matematikk, har vi mye informasjon om årets eksamen.

Bakgrunn

I utgangspunktet er det elevens matematikkompetanse som vurderes gjennom eksamen. Kompetanse er imidlertid et sammensatt begrep. I den danske rapporten «Kompetencer og matematiklæring» defineres åtte forskjellige kompetanser (Niss & Jensen 2002). Disse åtte kompetansene er også sentrale i de norske læreplanene i matematikk, og de skal følgelig testes gjennom den skriftlige eksamenen. Målet er at eksamenen skal gi et helhetlig bilde av en persons matematiske kompetanse, noe som fordrer at mange ferdigheter må prøves (ibid., s. 126). En skriftlig eksamen gir begrenset handlingsrom for matematisk aktivitet, og prosjektet vil belyse hvordan denne utfordringen møtes.

Alle elever i norsk skole har matematikk på timeplanen til og med 10. trinn og er således potensielt kandidater til en eksamen i matematikk etter endt grunnskole. Omtrent en tredjedel av årskullet på mellom 61 000 og 65 000 elever trekkes ut til skriftlig eksamen. Karakteren tar de med seg i konkurransen om plass på foretrukket skole og linje i videregående utdanning. I tillegg gir eksamensresultatene tilbakemelding til lærere, skoleledelse og skoleeiere om resultatet av opplæringen.

I den sammenheng er det viktig hvordan eksamenen oppfattes av elevene. Dersom elevene erfarer at en eksamen er uforutsigbar, kan det bidra til å øke stressnivået deres (Elwood, Hopfenbeck & Baird 2017). Økt stressnivå fører for en del elever til lavere prestasjon (Wang & Shah 2014). Dersom elever oppfatter eksamenen som uforutsigbar og/eller urettferdig, er det rimelig å anta at det kan påvirke deres motivasjon for å forberede seg til eksamen. Samtidig kan en eksamen som fra år til år tester de samme kompetansemålene på samme nivå, gi uønskede konsekvenser ved at undervisningen legger vekt på nettopp dette, og at kompetansemål og kompetanser som ikke er gjenstand for den samme testingen, blir utelatt fra undervisningen (Schoenfeld 2007).

I tillegg er det en pågående debatt i det offentlige ordskiftet om kvalitet i norsk skole generelt. Hægeland et al. (2010) finner at sammenhengen mellom fullføring av Vg1 og matematikkarakter fra grunnskolen er sterkere enn sammenhengen mellom grunnskolekarakterene i norsk og engelsk og fullføring av Vg1. Matematikk er også faget der flest elever ikke består, noe som bidrar direkte til at elever ikke fullfører videregående opplæring med bestått i alle fag. Derfor er mange opptatt av hvordan elevens prestasjoner i matematikk kan og bør forbedres. Et sentralt tema i denne debatten er hva den grunnleggende ferdigheten å lese i matematikk innebærer, og hvordan det aktualiserer etter- og videreutdanning av lærere og samtidig utfordrer elevene. Læreplanene sier at lesing er en grunnleggende ferdighet som også er en del av matematikkfaget. Forskningen på lesing i matematikk fører til en dypere forståelse av hvilke trekk ved matematikktekster som er vanskelige for (noen) elever, samtidig er det håp om at denne forståelsen kan bidra til at elevenes lesekompetanse på sikt vil øke og gjøre dem bedre i stand til å møte ulike fagtekster.

En gruppe svenske forskere har analysert eksamen i noen utvalgte matematikkfag og fant at kravene til kreativt resonnement var større i sentralgitte prøver enn i prøver gitt av lærerne (Palm, Boesen & Lithner 2011). I rapporten *Matematikk i norsk skole anno 2014* (Utdanningsdirektoratet 2014) pekes det på at dersom det samme er tilfelle i Norge, bidrar eksamen til å holde læringstrykket oppe (s. 87). Forskning tilknyttet TIMSS-studiene viser at økt læringstrykk kan bidra til å forklare de bedrede resultatene i matematikk og naturfag på 8. trinn (Nilsen, Grønmo & Hole 2013). Samlet aktualiserer denne forskningen at det er viktig å studere hva de faktiske og opplevde kravene i eksamensoppgavene er.

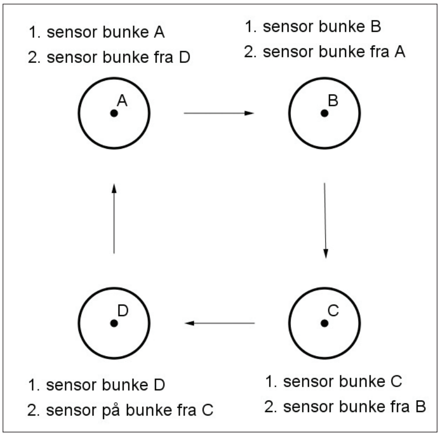
Hvordan skjer sensureringen i dag?

Våren 2017 ble eksamen i matematikk for 10. trinn gjennomført 16. mai. Spørsmålet er hva som skjer når kandidatene forlater eksamenslokalet og kjenner på vår, frihet og tilbakelagt eksamen. Hva er gangen i vurderingen før de får resultatet kunngjort?

Sensur av skriftlig eksamen i matematikk på 10. trinn:

Besvarelsenes gang:

Skolene sender besvarelsene til førstesensor som gjør sine vurderinger før bunken sendes videre til andresensor. Sensorene jobber i grupper på fire slik at de mottar bunke fra en annen sensor enn de sender til.



Sensorene vurderer besvarelsene uavhengig av hverandre og får ikke vite den andres vurderinger før de møtes på fellessensurmøtet og blir enige om en endelig karakter. Besvarelsene returneres så til skolen.

Hvilke føringer gis til sensorene?

Sensorene er bundet av retningslinjene som gis i eksamensveiledning (tilgjengelig i god tid før eksamen), sensorveiledning og vurderingsskjema der sensorene skriver inn sine poeng. De to siste offentliggjøres eksamensdagen.

Noen dager etter eksamen kommer det en bearbeidet versjon av sensorveiledningen: forhåndssensurrapporten. Den er utarbeidet av oppmennene i de ni sensorregionene og erstatter sensorveiledningen som bindende dokument for sensureringen.

Hvilken hjelp får sensorene?

Dokumentene som er nevnt over er også til hjelp i sensureringen. I tillegg arrangeres det sensorskoleringer i alle regioner. Her diskuterer sensorene både eksamensoppgavene og vurdering av eksempler på besvarelser som kan være utfordrende å vurdere.

Gangen i rapporten

Rapporten er organisert som følger. I neste kapittel presenteres datakildene og metodene som er benyttet. Analysene er basert på en nettundersøkelse til matematikklærere og sensorer, casestudier på fire forskjellige skoler, analyse av vurderingsskjemaer tilsendt fra sensorer, analyse av eksamenen i lys av lærebøkene og teori om språk og layout i matematikkoppgaver og deltakende observasjon på sensormøter. Alle kapitlene i rapporten inneholder analyser fra flere datakilder. I kapittel 3 analyseres eksamenens utforming, det vil si eksamenens oppbygning, svarformater og bruk av layout og illustrasjoner. I kapittel 4 ser vi på samsvaret mellom læreplanen og lærebøkene i matematikk og årets eksamen. I kapittel 5 og 6 retter vi søkelyset mot eksamenens vanskegrad og arbeidsmengde. Kapittel 7 inneholder en analyse av tekstoppgaver på årets eksamen samt læreres og elevers vurdering av bruk av tekstoppgaver i matematikk. Hvilke oppgaver sensorene finner det vanskeligst å vurdere, og sensorenes egen vurdering av årets sensurarbeid, er tema i kapittel 8. Kapittel 9 inneholder en avsluttende diskusjon og sammenfatning av funn.

Årets eksamen

Utdanningsdirektoratet

Eksamen

16.05.2017

MAT0010 Matematikk

Del 1

Skole:	Kandidatnr.:	Del 1 + __________ ark fra Del 2

Bokmål

Bokmål

Eksamensinformasjon	
Eksamenstid:	5 timer totalt. Del 1 og Del 2 skal deles ut *samtidig*. Del 1 skal du levere innen 2 timer. Del 2 skal du levere innen 5 timer.
Hjelpemidler på Del 1:	Ingen hjelpemidler er tillatt, bortsett fra vanlige skrivesaker, passer, linjal med centimetermål og vinkelmåler.
Framgangsmåte og forklaring:	Del 1 har 25 oppgaver. Skriv med penn når du krysser av eller fører inn svar i Del 1. I regneruter skal du vise hvordan du kommer fram til svaret. Du skal ikke kladde på oppgavearkene. Bruk egne kladdeark. På flervalgsoppgavene setter du bare ett kryss per spørsmål. **Eksempel:** Uttrykket $3\cdot(1+2\cdot 2)^2$ har verdien 35 ○ 50 ○ 62 ○ 75 ⊗
Veiledning om vurderingen:	Den høyeste poengsummen i Del 1 er 35, men den er bare veiledende i vurderingen. Karakteren blir fastsatt etter en samlet vurdering på grunnlag av Del 1 og Del 2. Sensor vurderer i hvilken grad du – viser regneferdigheter og matematisk forståelse – gjennomfører logiske resonnementer – ser sammenhenger i faget, er kreativ og kan anvende fagkunnskap i nye situasjoner – kan bruke hensiktsmessige hjelpemidler – forklarer framgangsmåter og begrunner svar – skriver oversiktlig og er nøyaktig med utregninger, benevninger, tabeller og grafiske framstillinger – vurderer om svar er rimelige
Andre opplysninger:	Kildeliste for bilder, tegninger mv.: • Forside Del 1, www.independent.ie (05.02.2017) • Mynter, www.norges-bank.no, www.roschberg.no (02.03.2017) • Hodetelefoner, www.apple.com (27.11.2016) • Måne, www.mathisenfoto.com (08.11.2016) • Panda og Tigergutt, www.br.no (30.10.2016) • Sko, www.norskbandysport.no (30.10.2016) • Andre bilder, tegninger og figurer: Utdanningsdirektoratet

Del 1 skal leveres innen 2 timer
Maks 35 poeng
Hjelpemidler: vanlige skrivesaker, passer, linjal med centimetermål og vinkelmåler

Oppgave 1 (2 poeng)

Regn ut

a) $657+468=$ ___________

b) $52\cdot 48=$ ___________

Oppgave 2 (2 poeng)

a) En kurv med jordbær veier 500 g.

12 kurver veier totalt _______ kg

b) 12 L saft skal helles over på flasker som hver rommer 4 dL.

Da trenger vi __________ flasker.

Oppgave 3 (1 poeng)

Hvilket uttrykk har den **laveste** verdien?

$(-2)^2\cdot 2^0$	$-2^2\cdot 2^1$	$-(2-2^2)$	$\dfrac{2\cdot(-2)}{2+2}$
O	O	O	O

Oppgave 4 (2 poeng)

Regn ut og skriv svaret så enkelt som mulig

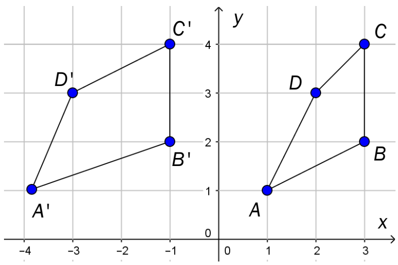
a) $\dfrac{1}{6}+\dfrac{1}{3}=$ _________

b) $\dfrac{0{,}2\cdot 0{,}4}{0{,}16}=$ _________

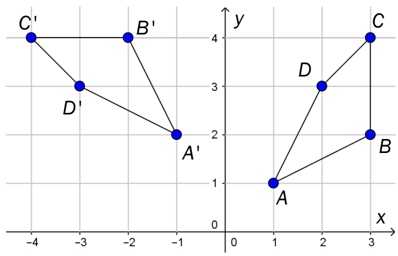
Oppgave 5 (1 poeng)

$\square ABCD$ skal speiles om y-aksen til $\square A'B'C'D'$.

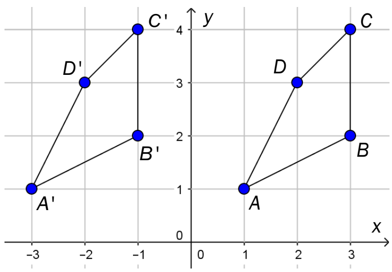
I hvilken figur nedenfor er dette gjort riktig?



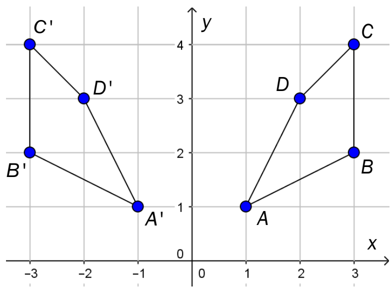
Figur 1



Figur 2



Figur 3



Figur 4

Figur 1	Figur 2	Figur 3	Figur 4
O	O	O	O

Oppgave 6 (1 poeng)

Vi kaster én terning.

Sannsynligheten for at terningen vil vise 3 eller 5 øyne, er

$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{5}{6}$
○	○	○	○

Oppgave 7 (1 poeng)

Vi har tre forskjellige mynter som hver viser mynt eller kron når vi kaster dem.

Vi kaster de tre myntene én gang.

Bestem sannsynligheten for at alle myntene viser mynt eller at alle myntene viser kron.

mynt mynt mynt

kron kron kron

Løs oppgave 7 her:

Oppgave 8 (1 poeng)

Christian kjøpte nye hodetelefoner da han var i England.
Han betalte £ 88,95 (engelske pund) for hodetelefonene.
Kurs: £ 1 = 10,21 norske kroner.

Omtrent hvor mye betalte Christian i norske kroner?

- ○ ca. 10 kroner
- ○ ca. 90 kroner
- ○ ca. 800 kroner
- ○ ca. 900 kroner

Oppgave 9 (1 poeng)

På hvor mange ulike måter kan åtte personer sette seg på åtte stoler?

- ○ 8
- ○ $8+8$
- ○ 8^8
- ○ $8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1$

Oppgave 10 (1 poeng)

Hvis $A=\dfrac{gh}{2}$, da er

- ○ $h=\dfrac{Ag}{2}$
- ○ $h=\dfrac{A}{2g}$
- ○ $h=\dfrac{2A}{g}$
- ○ $h=2A-g$

Oppgave 11 (2 poeng)

Skriv så enkelt som mulig

a) $\dfrac{a+a+a}{a}$

b) $\dfrac{a^2-b^2}{a-b}$

Løs oppgave 11 a) her:

Løs oppgave 11 b) her:

Oppgave 12 (2 poeng)

Løs likningene

a) $4x-4=11-x$

b) $\dfrac{x}{6}-\dfrac{2-x}{4}=2$

Løs oppgave 12 a) her:

Løs oppgave 12 b) her:

Oppgave 13 (1 poeng)

Avstanden fra jorda til månen er ca. 384 000 000 m.

Avstanden skrevet på standardform er

- O $3{,}84 \cdot 10^{-8}$ m
- O $3{,}84 \cdot 10^{6}$ m
- O $3{,}84 \cdot 10^{8}$ m
- O $38{,}4 \cdot 10^{8}$ m

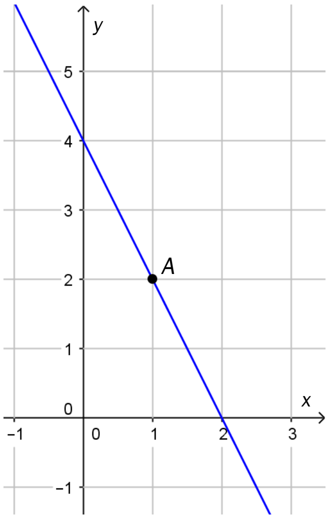
Oppgave 14 (1 poeng)

Et kart har målestokken 1 : 15 000. Avstanden mellom to steder er 3,0 cm på kartet.

Hvor langt er det mellom stedene i virkeligheten?

50 m	450 m	4 500 m	5 000 m
O	O	O	O

Oppgave 15 (2 poeng)



a) Skriv koordinatene til punktet *A*.

$A(\_\_\_\_, \_\_\_\_)$

b) Hvilken funksjon passer med grafen i koordinatsystemet?

- O $y = -2x - 4$
- O $y = -2x + 4$
- O $y = -4x + 4$
- O $y = 2x + 2$

Oppgave 16 (2 poeng)

 +  +  = 350 kroner

 +  +  +  = 500 kroner

a) Prisen for én  er ____________ kroner

b) Prisen for én  er ____________ kroner

Oppgave 17 (1 poeng)

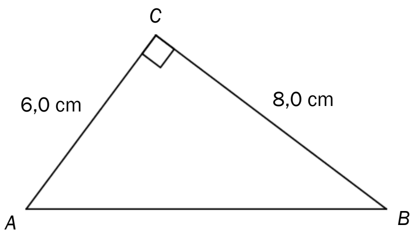
Et par sko kostet 990 kroner. Prisen ble satt ned med 20 %.

Omtrent hvor mange kroner ble prisen satt ned med?

ca. 20 kroner	ca. 50 kroner	ca. 150 kroner	ca. 200 kroner
O	O	O	O

Oppgave 18 (1 poeng)

Bruk Pytagoras-setningen til å regne ut lengden AB i $\triangle ABC$ nedenfor.

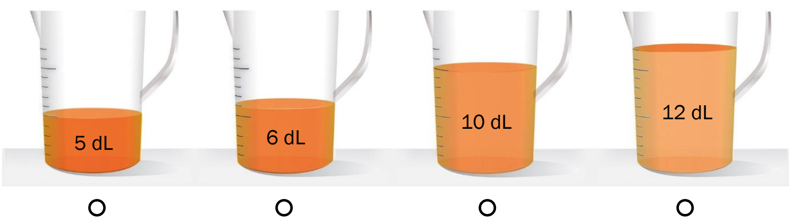


Løs oppgave 18 her:

Oppgave 19 (1 poeng)

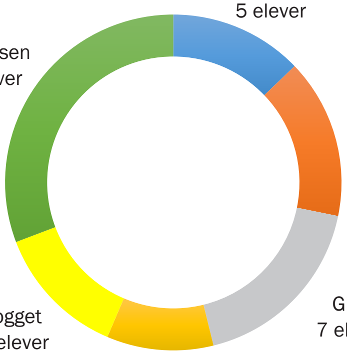
I hver mugge nedenfor har vi blandet 2 dL saft med vann. Volumet av blandingen er skrevet på hver mugge.

I hvilken mugge er blandingsforholdet mellom saft og vann 1 : 5?



Oppgave 20 (2 poeng)

Nedenfor ser du et hjuldiagram som viser hvordan 39 elever på en skole kom seg til skolen en dag.



a) Hvor stor del av elevene syklet eller gikk til skolen?

$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{13}$	$\frac{1}{39}$
O	O	O	O

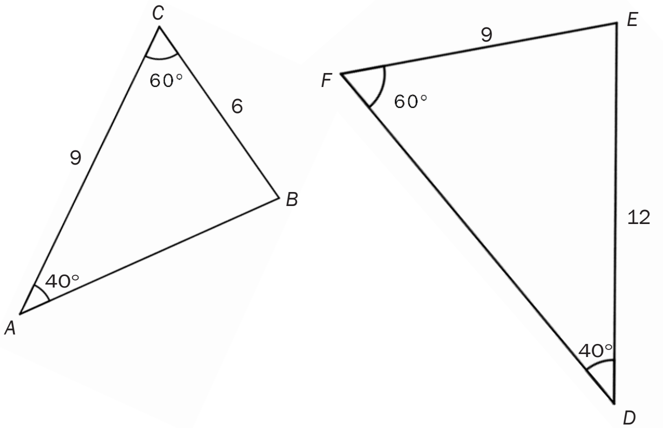
b) Hvor mange prosent av elevene tok bussen til skolen?

ca. 10 %	ca. 20 %	ca. 30 %	ca. 40 %
O	O	O	O

Oppgave 21 (1 poeng)

På figuren nedenfor er $\Delta ABC \sim \Delta DEF$ (formlike trekanter).

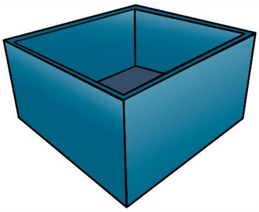
Regn ut lengden AB.



Løs oppgave 21 her:

Oppgave 22 (1 poeng)

Tegn perspektivlinjer for å finne forsvinningspunktene til figuren nedenfor.



Oppgave 25 (2 poeng)

Benedikte har laget disse tre figurene av fyrstikker:

Figur 1 Figur 2 Figur 3

Benedikte vil lage flere figurer etter samme mønster som figurene ovenfor.

a) Hvor mange fyrstikker trenger hun for å lage figur 5?

Figur 5: ____________ fyrstikker

b) Lag en formel som forteller hvor mange fyrstikker hun trenger for å lage figur *n*.

Løs oppgave 25 b) her:

Utdanningsdirektoratet

Eksamen

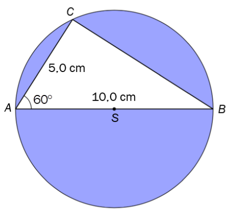
16.05.2017

MAT0010 Matematikk

Del 2

I trafikken





Geometri

Ada Lovelace

Bokmål

Bokmål

Eksamensinformasjon	
Eksamenstid:	5 timer totalt. Del 1 og Del 2 skal deles ut samtidig. Del 1 skal du levere innen 2 timer. Del 2 skal du levere innen 5 timer.
Hjelpemidler på Del 2:	Alle hjelpemidler er tillatt, med unntak av Internett og andre verktøy som tillater kommunikasjon, etter at Del 1 er levert inn. Før Del 1 er levert inn, er ingen hjelpemidler tillatt, bortsett fra vanlige skrivesaker, passer, linjal med centimetermål og vinkelmåler.
Framgangsmåte og forklaring:	Del 2 har 9 oppgaver. Der oppgaveteksten ikke sier noe annet, kan du fritt velge framgangsmåte. Vis hvordan du har kommet fram til svarene. Før inn nødvendige mellomregninger. Skriv med penn. I oppgaver der du bruker regneark, skal du vise hvilke formler du har brukt i regnearket. I oppgaver der du bruker digital graftegner, skal skala og navn på aksene være med på graftegningen.
Veiledning om vurderingen:	Den høyeste poengsummen i Del 2 er 49, men den er bare veiledende i vurderingen. Karakteren blir fastsatt etter en samlet vurdering på grunnlag av Del 1 og Del 2. Sensor vurderer i hvilken grad du – viser regneferdigheter og matematisk forståelse – gjennomfører logiske resonnementer – ser sammenhenger i faget, er kreativ og kan anvende fagkunnskap i nye situasjoner – kan bruke hensiktsmessige hjelpemidler – forklarer framgangsmåter og begrunner svar – skriver oversiktlig og er nøyaktig med utregninger, benevninger, tabeller og grafiske framstillinger – vurderer om svar er rimelige
Andre opplysninger:	Kildeliste for bilder, tegninger mv.: • Sykkeltur, www.perlebukta.no (30.10.2016) • Tanking av bensin, www.np.no (30.10.2016) • Sykkelutstyr, www.oslosportslager.no (07.02.2017) • Biler og CO2, www.tu.no (30.10.2016) • Bil og bremselengde, www.retten.no (05.11.2016) • Tower Bridge, www.voyage-londres.org (14.09.2016) • Ada Lovelace, www.google.com/doodles/ada-lovelaces-197th-birthday og www.amazon.com (30.10.2016) • Andre illustrasjoner og bilder: Utdanningsdirektoratet

Del 2 skal leveres innen 5 timer
Maks 49 poeng
Hjelpemidler: Se side 2

Oppgave 1 (6 poeng)

Noen elever i klasse 10 A registrerer hvor mange personer som sitter i hver personbil som kjører forbi skolen.

Elevene teller antall personer i de 30 første personbilene og får følgende resultat:

Antall personer	Frekvens
1	13
2	5
3	3
4	4
5	3
6	2
Sum	30

a) Framstill antall personer i personbilene med et passende diagram.

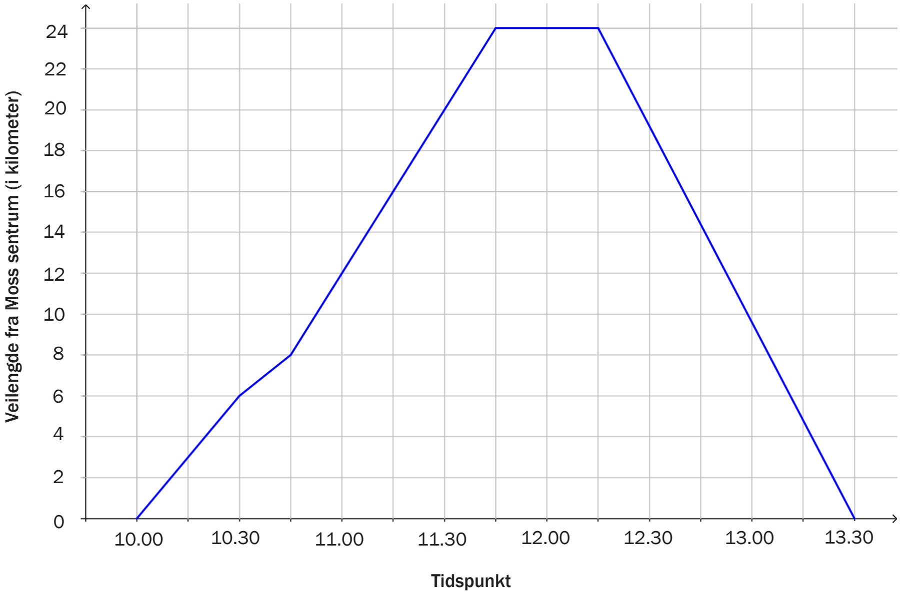
b) Bruk tabellen ovenfor til å bestemme typetall og median.

c) Regn ut gjennomsnittet for antall personer i personbilene.

Oppgave 2 (5 poeng)

Mari og David er på sykkelferie. En av turene deres går fra Moss sentrum til Nordre Jeløy og tilbake til Moss sentrum.

Diagrammet nedenfor viser sammenhengen mellom tid og veilengde på denne turen.



a) Hvor lang er veien fra Moss sentrum til Nordre Jeløy ifølge diagrammet?

b) Hvor stor del av sykkelturen har Mari og David pause?

c) Hvor stor gjennomsnittsfart har de fra Nordre Jeløy til Moss sentrum?

Oppgave 3 (6 poeng)

a) Georg stopper på en bensinstasjon når bilen har 16,5 L bensin igjen på tanken. Bensintanken rommer 60 L. Prisen for bensin er 14,30 kroner per liter.

Hvor mye må Georg betale for å fylle opp tanken?

b) Kathrine fyller bensin i en jerrykanne med en form som er tilnærmet lik et rett, firkantet prisme.

Bruk målene på figuren.

Hvor mange hele liter rommer jerrykannen?

c) Bensinstasjonen solgte en dag til sammen 28 000 L diesel og bensin. Forholdet mellom antall liter diesel og antall liter bensin var 3 : 5.

Hvor mange liter diesel, og hvor mange liter bensin solgte bensinstasjonen denne dagen?

Oppgave 4 (5 poeng) REGNEARK

Oversikten nedenfor viser:

- hvilke sykkelprodukter Bike Shop solgte en dag
- pris (i kroner) per stykk før rabatt
- hvor mange av hvert produkt som ble solgt
- hvor mange prosent rabatt som ble gitt til kundene for hvert produkt

Produkt	Sykkel Terreng	Sykkel Racer	Sykkelsko (par)	Sykkelhjelm
Pris per stykk før rabatt	12 000,00	15 500,00	1 500,00	750,00
Antall solgte produkter	3	2	4	5
Rabatt	23 %	32 %	50 %	23 %

Lag et regneark som vist nedenfor. Bruk formler til å regne ut det som mangler i de grå cellene i regnearket. Vis hvilke formler du har brukt.

	A	B	C	D	E
1	**Bike Shop**	**Sykkel Terreng**	**Sykkel Racer**	**Sykkelsko (par)**	**Sykkelhjelm**
2	Pris per stykk før rabatt	12000,00	15500,00	1500,00	750,00
3	Antall solgte produkter	3	2	4	5
4	Samlet pris før rabatt				
5	Rabatt	23 %	32 %	50 %	23 %
6	Rabatt (kroner)				
7	Salgsinntekt etter rabatt				
8	**Total salgsinntekt etter rabatt**				
9					

Oppgave 5 (6 poeng) GRAFTEGNER

Antall gram CO_2 (karbondioksid) som en bestemt bil slipper ut per kilometer, er gitt ved funksjonen

$$f(x) = 0{,}046x^2 - 6{,}7x + 386$$

der x er farten til bilen målt i kilometer per time.

a) Bruk graftegner til å tegne grafen til f for x - verdier fra og med 20 km/h til og med 110 km/h.

b) CO_2 - utslippet til bilen er 180 g/km. Hvor stor fart kan bilen ha?

c) Hvilken fart vil gi minst CO_2 - utslipp per kilometer?
Hvor mange gram CO_2 slipper bilen ut per kilometer da?

Oppgave 6 (5 poeng)

a) En bil har en gjennomsnittsfart på 60 km/h.

Hvor langt kan bilen kjøre på 1,5 h?

Bremselengden s målt i meter for en bil er gitt ved formelen

$$s = \frac{v^2}{19{,}62 \cdot f}$$

- v er farten til bilen målt i meter per sekund (m/s)
- f er et friksjonstall, dette avhenger av veidekket

b) En bil kjører på tørr asfalt. Da er $f = 0{,}9$. Farten er 21 m/s. Så bremser bilen.

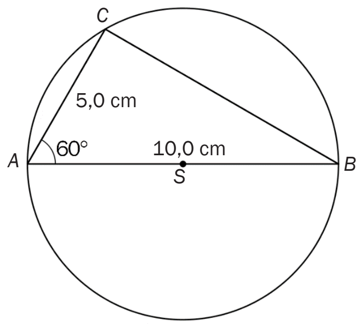
Bruk formelen ovenfor og vis at bremselengden til bilen er ca. 25 m.

c) En bil kjører på våt asfalt. Da er $f = 0{,}6$. Så bremser bilen, og bremselengden er 15 m.

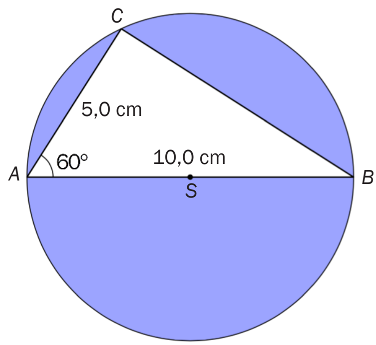
Bruk formelen ovenfor til å bestemme farten til bilen.

Oppgave 7 (6 poeng)

Figur 1 nedenfor viser en $\triangle ABC$ som er innskrevet i en sirkel med sentrum i S. AB er diameteren i sirkelen. Figur 2 viser figur 1 som delvis fargelagt.



Figur 1



Figur 2

a) Velg én av framgangsmåtene nedenfor til å konstruere ***eller*** tegne figur 1.

Med passer, linjal og blyant:

- Konstruer $\triangle ABC$ med $AB = 10{,}0\,\text{cm}$, $\angle A = 60^\circ$ og $AC = 5{,}0\,\text{cm}$.
- Konstruer midtnormalen til AB. Midtnormalen skjærer AB i punktet S.
- Slå en sirkel om S gjennom punktene A, B og C.

Med dynamisk geometriprogram:

- Tegn $\triangle ABC$ med $AB = 10{,}0\,\text{cm}$, $\angle A = 60^\circ$ og $AC = 5{,}0\,\text{cm}$.
- Tegn midtnormalen til AB. Midtnormalen skjærer AB i punktet S.
- Tegn en sirkel om S gjennom punktene A, B og C.

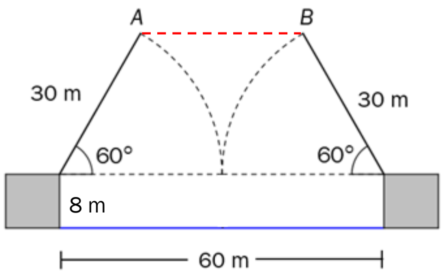
b) Forklar kort hvorfor $\angle ACB = 90^\circ$ uten å måle vinkelen.

c) Bestem arealet av den delen av sirkelflaten som er blå på figur 2.

Oppgave 8 (4 poeng)

Tower Bridge er en klaffebro over Themsen i London. Avstanden mellom tårnene er 60 m. Broen har to like store klaffer som kan heves for å la båter kjøre forbi.

På et tidspunkt er det 8 m fra lukket bro til vannoverflaten. Så blir begge klaffene hevet 60° for å la en seilbåt kjøre forbi. Se figuren nedenfor.



a) Forklar at bredden AB på åpningen mellom klaffene er 30 m.

b) Bestem ved regning avstanden fra A til vannoverflaten.

Oppgave 9 (6 poeng)

a) Løs likningssystemet

$$5x + 4y = 9$$
$$6x + 7y = 13$$

Ada Lovelace (1815 – 1852) regnes som verdens første dataprogrammerer. Hun skrev programmer for en stor, planlagt regnemaskin som skulle hete *Den analytiske maskinen*.

I 1979 fikk dataprogrammeringsspråket Ada navn etter Ada Lovelace.

Et likningssystem med to ukjente, x og y, kan skrives på formen

$$ax + by = c$$
$$dx + ey = f$$

der a, b, c, d, e og f er konstanter.

Ada Lovelace laget formler for løsning av et likningssystem med to ukjente:

$$x = \frac{ce - bf}{ae - bd} \quad \text{og} \quad y = \frac{af - cd}{ae - bd}$$

b) Bruk formlene til Ada Lovelace ovenfor til å løse likningssystemet

$$5x + 4y = 9$$
$$6x + 7y = 13$$

c) Løs likningssystemet

$$ax + by = c$$
$$dx + ey = f$$

og vis at du får formlene til Ada Lovelace som løsninger.

2 Metode

Datainnsamlingen for undersøkelsen har vært sammensatt. Vi har sendt ut et elektronisk spørreskjema til et utvalg matematikklærere på 10. trinn som har hatt elever oppe til eksamen i 2017, gjennomført casestudier på fire skoler, vært til stede rett etter eksamen på en skole, analysert vurderingsskjemaer tilsendt fra sensorer og analysert eksamenen i lys av lærebøkene og teori om språk og layout i matematikkoppgaver. I tillegg har vi lagt til spørsmål i Utdanningsdirektoratets spørreundersøkelse til sensorene og deltatt på sensormøter.

Spørreskjemaet

For å få kunnskap om sammenhengen mellom læreplan og opplæring, eksamenens vanskegrad og arbeidsmengde fant vi det hensiktsmessig å spørre matematikklærere om deres oppfatninger. Blant alle ungdomsskolene som ble trukket ut til eksamen, samlet vi inn e-postadresser til matematikklærere på 10. trinn på skoler hvor ti eller flere elever var trukket ut til eksamen. Det var til sammen 692 ungdomsskoler. Av disse skolene fikk vi tilsendt e-postadresser fra 510 skoler. Det vil si 74 prosent av skolene hvor ti eller flere elever var trukket ut til eksamen. Spørreundersøkelsen ble sendt ut til 1531 mulige respondenter. Av disse var det 1106 som svarte. Vi slettet respondenter som svarte at de ikke underviste i matematikk på 10. trinn eller ikke hadde hatt elever oppe til eksamen i matematikk våren 2017. Vi slettet også respondenter som kun hadde åpnet undersøkelsen og ikke svart på spørsmålene. Da gjenstod det 795 informanter. Fordi vi ikke vet hvor mange lærere som underviser i matematikk på 10. trinn og har hatt elever oppe til eksamen, kan vi ikke oppgi svarprosent. Spørreundersøkelsen ble sendt ut uken etter at eksamen var gjennomført, og lærerne ble bedt om å svare på spørsmålene basert på årets eksamen.

Vi la også til et knippe spørsmål i Utdanningsdirektoratet sitt spørreskjema til sensorene, som de sender umiddelbart etter at sensurarbeidet er ferdig. Spørreskjemaet ble sendt ut til 316 sensorer, og svarprosenten var på 46.

Case

I tillegg til spørreskjemaet gjennomførte vi en kvalitativ casestudie hvor vi intervjuet matematikklærere på 10. trinn, elever som hadde hatt matematikkeksamen, og ledelsen ved fire skoler. Formålet med casestudien var å gå i dybden på spørsmålene vi stilte matematikklærerne i spørreskjemaet, og i tillegg intervjue elever og skoleledelsen. Siden flere av elevene var under 16 år på intervjutidspunktet, ba vi om skriftlig tillatelse fra foresatte om å intervjue deres barn. Alle skolene og informantene fra casene er anonymisert i analysekapitlene, og vi fikk informert samtykke til å gjennomføre intervjuene av informantene.

Vi har innhentet data fra fire ungdomsskoler (se tabell 2.1). Skolene lå spredt rundt i landet og varierer med hensyn til sammensetning av både lærere og elever. En av skolene skilte seg ut ved å ha en høy andel minoritetsspråklige elever. Størrelsen på skolene varierte, hvor den minste hadde i overkant av 200 elever, og den største hadde over 500 elever. En av skolene var en kombinert barne- og ungdomsskole. Forskjellen mellom skolene etter disse variablene er viktig fordi det gjør det mulig å identifisere ulike mekanismer, selv om vi ikke kan generalisere funnene på bakgrunnen av et lite antall case. Til sammen intervjuet vi 56 informanter hvorav 15 var matematikklærere, 36 var elever og fem satt i skoleledelsen. Intervjuene med elevene ble gjort som fokusgruppeintervjuer. Intervjuene med skoleledelsen og matematikklærerne gjorde vi som enten enkeltintervjuer eller fokusgrupper avhengig av hva som passet best for skolene. På den ene skolen var vi også til stede rett etter eksamen for å fange opp «stemningen» blant elevene etter at eksamen var gjennomført. Her snakket vi med åtte elever, rektor, assisterende rektor og en lærer.

Tabell 2.1 Oversikt over caseskolene.

Skole 1	Skole 2	Skole 3	Skole 4
Ca. 300 elever	Ca. 500 elever	Ca. 400 elever	Ca. 200 elever
3 paralleller	6 paralleller	3 paralleller	3 paralleller
Intervjuet: rektor, 2 lærere og 9 elever	Intervjuet: rektor, IKT-ansvarlig, avdelingsleder, 5 lærere og 9 elever	Intervjuet: 4 lærere og 15 elever	Intervjuet: assisterende rektor, 3 lærere og 3 elever

Sensormøter

For å få best mulig inntrykk av sensorenes arbeid deltok vi på deler av forsensurmøte 30. og 31. mai der oppmennene i de ni sensorregionene møttes for å skrive forhåndssensurrapport og bli enige om poenggrenser for karakterene. Vi deltok også på en sensorskolering som fant sted omtrent midtveis mellom eksamen og fellessensurmøtet.

Vurderingsskjema

Et av verktøyene i sensureringsarbeidet er vurderingsskjemaene der sensorene legger inn sine poeng (se faktaboks i kapittel 1). For å undersøke om det er oppgaver der sensorene har ulik vurdering, ønsket vi å få tilsendt vurderingsskjemaene fra sensorer før fellessensurmøtet i 2017. Dette var for å få en mer detaljert oversikt over ulik vurdering også der sensorene har gitt samme karakter, og for å undersøke om det er enkelte oppgaver og/eller oppgavetyper som oftere gir ulik vurdering. Vi henvendte oss til sensorene i Oslo og Akershus fordi dette er den sensorregionen som har flest kandidater, de har sensorskoleringen etter forsensurmøtet, og det ville være mulig for oss å delta på sensorskoleringen. 59 av de 64 sensorene sendte alle eller noen av sine vurderingsskjemaer til oss. Vurderingsskjemaene inneholder kandidatnummer og sensorens poeng på de enkelte oppgavene, og i tillegg kan de inneholde kommentarer knyttet til den enkelte besvarelse. Vi lagret kun kandidatnummer og poengene. Sensorenes øvrige kommentarer, tilknytning til skole og sensor ble ikke lagret.

Noen vurderingsskjemaer ble delt med oss på en slik måte at sensorene kunne redigere dem etter at vi hadde fått tilgang til dem. Disse ble utelatt fra analysene siden de i teorien kunne være redigert etter at sensorene hadde møttes. Det ble også skjemaer hvor retningslinjene fra forhåndssensurrapporten ikke ble fulgt, ved at det ble gitt halve poeng. Kandidater der det var registrert flere poeng på en oppgave enn det ifølge forhåndssensurrapporten er mulig å oppnå, ble slettet. I tillegg ble skjemaer med feilføringer som t, 10, X og 12 i stedet for tillatte poeng på oppgavene slettet. Vi fant også andre feilføringer, og disse ble behandlet på følgende måte: 11, | og |1 ble rettet til 1, 22 ble rettet til 2, og 9 ble rettet til 0.

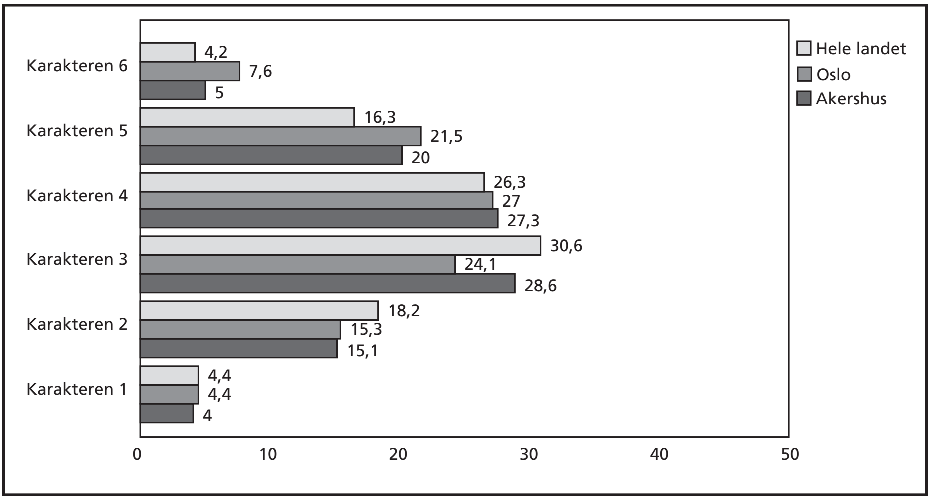
Et kandidatnummer består av tre tall og tre bokstaver etterfulgt av -V, og tre ulike kandidatnummer kan for eksempel se slik ut: 234ABC-V, 243ABC-V, 243BAC-V. Det viste seg å være en del feilføringer av kandidatnummer, og vi valgte å utelate alle vi var usikre på om representerte unike kandidater. Kvalitetssikringen ble gjort ved å sammenlikne med liste fra Utdanningsdirektoratet i forbindelse med at vi skulle koble kandidat mot kjønn. 175 kandidater med registrering fra bare den ene sensoren ble slettet fordi vi ikke fant samsvar mellom kandidatnummeret på skjemaet og listen fra

Utdanningsdirektoratet. Listen fra Utdanningsdirektoratet var ikke fullstendig for vår bruk, og vi valgte å beholde kandidatene der vi hadde registrering fra to sensorer på identiske kandidatnummer.

Det var potensielt mulig å få to registreringer for hver av de 4250 kandidatene i Oslo og Akershus. Våre analyser baserer seg på 2814 kandidater med vurdering fra to sensorer og 1154 kandidater med vurdering fra én sensor. I analyser som gjelder samsvar i vurdering mellom sensorene, og i analysene knyttet til blanke svar er kun materialet med to sensorer brukt. I de øvrige analysene har vi brukt både vurdering fra én sensor og et gjennomsnitt av poengene fra to sensorer.

Sensorene var bedt om å la ruten for poeng stå tom dersom oppgaven ikke var besvart. Noen sensorer har misforstått dette og latt både feil løsning som ikke gir uttelling, og ubesvart oppgave stå med tom rute, mens andre har gitt 0 poeng for både feil og ubesvart oppgave. Dette har ingen betydning for sensorenes vurdering, men kun for analyse av arbeidsmengde. Skjemaene som ikke skiller mellom feil og ubesvart er med i analysene som ser på ulik vurdering av samme besvarelse med 0 og blank som samme verdi, men er tatt ut i analysene som handler om arbeidsmengde.

Figur 2.1 Eksamenskarakterer, matematikk skriftlig eksamen 2016-2017, 10. trinn. Nasjonalt og for fylkene Oslo og Akershus.



Kilde: https://skoleporten.udir.no/

I tillegg til å se på oppgaver som gir opphav til ulik vurdering og arbeidsmengde, ble vurderingsskjemaene brukt til å gi informasjon om hvilke oppgaver som viste seg å være vanskelige, og poengfordeling på del 1 og del 2 for kandidater på ulike nivåer. En begrensning i materialet vårt er at det er innhentet i én sensorregion: Oslo og Akershus. Fylkesoversikten fra Utdanningsdirektoratet (figur 2.1) viser færre på de laveste og flere på de høyeste karakterene i disse to fylkene enn det er i hele landet. For analysene av hvilke oppgaver som er krevende å sensurere, ser vi det ikke som utslagsgivende at elevgrunnlaget i våre data avviker noe fra landet som helhet. I analysene av arbeidsmengde og vanskegrad har vi sammenholdt våre data med data fra hele landet samlet inn i forbindelse med IRT-analysene[2] som Institutt for lærerutdanning og skoleforskning (ILS) ved Universitetet i Oslo har utført for Utdanningsdirektoratet, for å se om disse viser samme mønster. Siden disse ble innhentet etter at sensorene hadde snakket sammen, kan de ikke brukes i analysen av hvilke oppgaver som var krevende å sensurere.

Språk, illustrasjoner og layout

For å få et grunnlag for vurdering av eksamenssettets språk, illustrasjoner og layout har vi foretatt en gjennomgang av nyere forskningslitteratur på feltet. Basert på denne gjennomgangen har vi valgt ut noen trekk som vi har analysert ut fra (se kap. 3 og 7 for detaljer om dette). Forskningen som er gjort, gir ikke grunnlag for å konkludere entydig om hvilke trekk som gjør eksamensoppgavene vanskeligere for ulike elever, trekkene må ses i sammenheng. Trekkene har vi dels behandlet kvantitativt (antall ord, setninger, lange ord, lavfrekvente ord, sammensatte ord, generelle akademiske ord), og dels kvalitativt (hvilke lavfrekvente ord, sammensatte ord, akademiske ord, hvordan brukes nominaliseringer, lange substantivgrupper, preposisjoner, ordstilling, paratekstlige elementer, illustrasjoner). Vi har i disse analysene valgt å sammenlikne med to tidligere eksamenssett (2009 og 2016). Vi har valgt ut eksamenen i 2016 fordi den er nær i tid, og eksamenen i 2009 fordi den kom dårlig ut i Matematikksenterets gjennomgang av eksamensoppgaver (Matematikksenteret 2015a).

Vi har tatt utgangspunkt i bokmålsversjonene i analysene, men i kap. 7 ser vi i tillegg på forskjeller i formuleringer mellom bokmåls- og nynorskversjonen av eksamenen i 2017.

Bruk av læremidler

Spørsmålet om det er samsvar mellom eksamenen og det elevene har kjennskap til fra opplæringen, kan blant annet besvares gjennom en analyse av eksamensoppgavene i lys av de mest brukte lærebøkene på 10. trinn.

[2] IRT står for Item Response Theory.

I analysen av eksamensoppgavene opp mot lærebøkene har vi tatt utgangspunkt i rammeverket som er utarbeidet av Johan Lithner og beskrevet i «A research framework for creative and imitative reasoning» (Lithner 2008), og gjort noen tilpasninger som er nærmere beskrevet i kapittel 4.

3 Utforming av eksamenen

I dette kapitlet skal vi se nærmere på hvordan eksamensoppgaven er utformet. For det første er vi interessert i ulike sider ved oppgavesettets *oppbygging*. Det vil si hvordan de ulike oppgavene er presentert og organisert. For det andre setter vi søkelyset på illustrasjoner og layout, altså oppgavenes formmessige uttrykk. Lærere og elevers subjektive opplevelse av eksamenens utforming er også en sentral del av kapitlet.

Eksamenens oppbygging

Vi vil drøfte eksamenens oppbygging ut fra tre perspektiver: hvordan vanskegraden varierer gjennom settet, om oppgaver bygger på hverandre, og hvilke svarformater (f.eks. åpen, flervalg osv.) som er valgt. Vi vil dessuten reflektere rundt oppdelingen i en del 1 og en del 2.

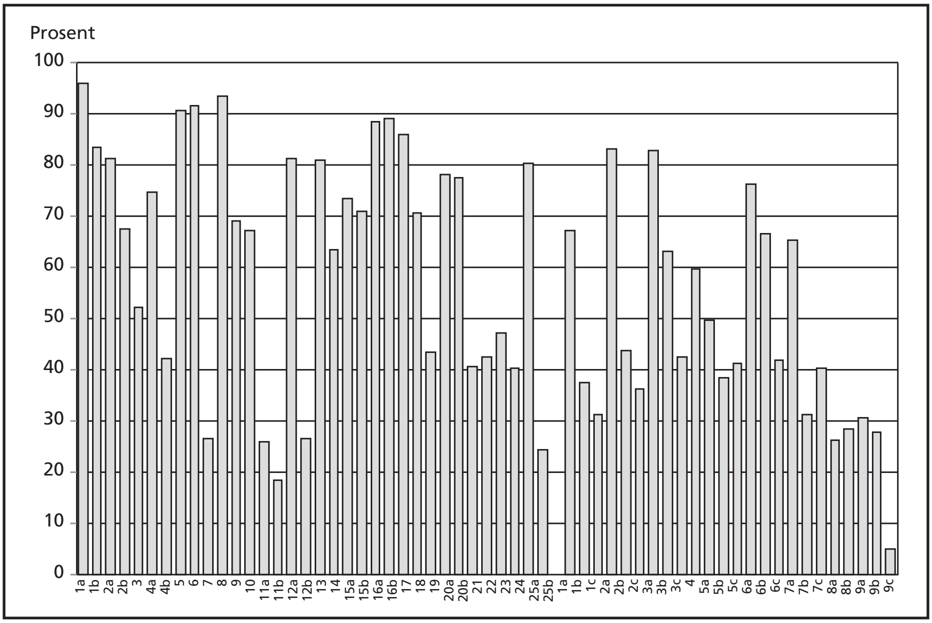
Stigende vanskegrad

Det er gjerne sett på som et ideal at oppgavesett skal ha stigende vanskegrad, slik at ikke elever mister motet før de kommer til de oppgavene som de kan få til. Dette gjelder både mellom deloppgaver i samme enkeltoppgave og for settet som helhet. Samtidig vil det være umulig å oppnå dette fullt ut.

I kapittel 5 har vi en inngående diskusjon av vanskegrad basert på flere ulike tilnærminger. Her i kapittel 3 vil vi bruke kandidatenes gjennomsnittlige uttelling på deloppgavene som mål på vanskegraden. Figur 3.1 viser uttelling på de enkelte deloppgavene basert på de vurderingsskjemaene vi har hatt tilgjengelig.[3]

[3] Vårt materiale er basert på vurderingsskjemaene fra sensorene i Oslo og Akershus. Utdanningsdirektoratet har samlet inn materiale fra noen sensorer i forbindelse med sin IRT-analyse. Et diagram basert på dette materialet gir samme bilde som vårt materiale, men med noe lavere uttelling på de fleste oppgavene.

Figur 3.1 Vanskegrad: elevenes gjennomsnittlige poengskår på oppgavene i settet.



Vi ser at uttellingen varierer gjennom eksamenssettet. Det mest påfallende i eksamenens del 1 er at et par av de vanskeligste oppgavene, oppgave 11a og 11b, kommer tidlig i settet. I del 2 kommer et par av de vanskeligste oppgavene allerede i oppgave 1, og heller ikke oppgave 1a er svært enkel. Ut fra prinsippet om stigende vanskegrad, og i etterpåklokskapens klare lys, kunne det vært mer gunstig å starte del 2 med oppgave 3.

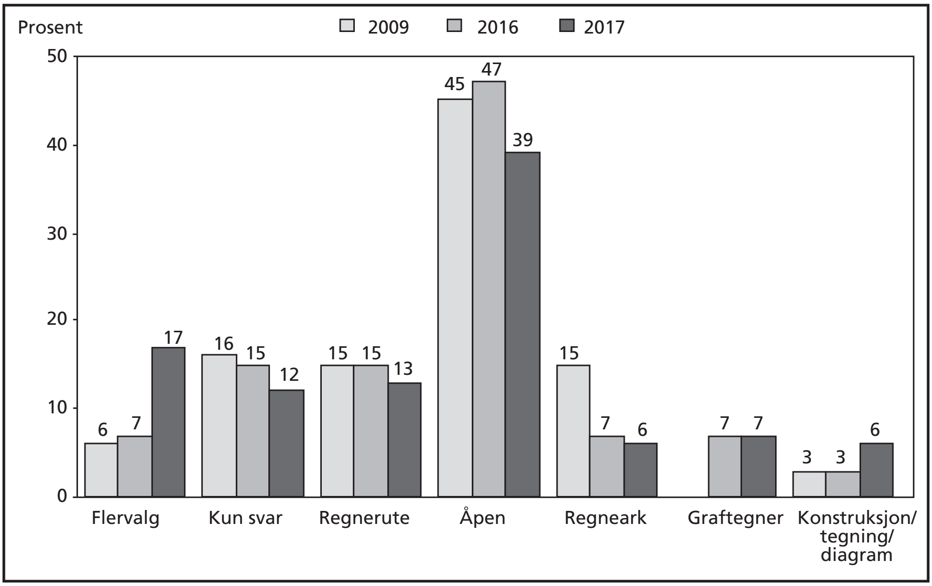
Oppgaver som bygger på hverandre

Oppgaver som bygger på tidligere deloppgaver, kan være problematiske, fordi elever som ikke har fått til den tidligere oppgaven, heller ikke prøver seg på den neste. I dette eksamenssettet er det bare oppgave 2c i del 2 som bygger på et tidligere svar, og det er oppgave 2a. En måte å unngå slike avhengigheter på er å omformulere den første av oppgavene slik at eleven skal vise/forklare et oppgitt svar istedenfor selv å finne svaret. I dette tilfellet ville det vært pussig å be elevene forklare at veien fra Moss sentrum til Nordre Jeløy er 24 km. I forhåndssensurrapporten er det også presisert at følgefeil fra 2a skal gi full uttelling dersom framgangsmåten er korrekt.

Svarformater

På eksamenens del 1 er det flervalgsoppgaver, oppgaver der eleven kun skal gi svar, og oppgaver med plass til utregning og begrunnelse for svaret. I denne delen fører eleven hele sin besvarelse i oppgaveheftet. Noen år skal elevene også konstruere for hånd i del 1. Del 2 føres på egne ark for hånd og på datamaskin. Her er det noen oppgaver som krever bruk av regneark og digital graftegner. I tillegg kan det være en konstruksjonsoppgave som åpner for å bruke dynamisk geometriprogram. Andelen av de ulike svarformatene varierer lite over tid, men i 2017 er det flere flervalgsoppgaver. Figur 3.2 viser hvor stor andel av poengene som kan oppnås på oppgaver med ulike svarformater. Det er velkjent at gutter ofte skårer bedre enn jenter på flervalgsoppgaver, mens jenter skårer bedre enn gutter på åpne svarformater – i naturfag (Eriksson 2005). I vårt materiale er det ikke signifikante forskjeller mellom jenters og gutters svar på flervalgs- og åpne oppgaver i del 1.

Figur 3.2 Andel av poengene på eksamenssettet fordelt på svarformater for avsluttende eksamen på 10. trinn 2009, 2016 og 2017.



Del 1 og del 2

Oppdelingen i del 1 (med begrensede hjelpemidler) og del 2 (med hjelpemidler) er grundig beskrevet i eksamensveiledningen (Utdanningsdirektoratet 2017), og oppgavesettet følger i hovedsak eksamensveiledningens beskrivelser.

Elevene får bruke maksimalt to timer på del 1, og den gir maksimalt 35 poeng. Til del 2 har elevene resten av eksamenstiden på fem timer til rådighet, og denne delen gir maksimalt 49 poeng. En analyse av vurderingsskjemaene vi har hatt tilgjengelig (fra Oslo og Akershus), viser at til tross for at del 2 er større enn del 1, får de faglig svakeste elevene svært lav uttelling på del 2. De elevene som i vårt materiale har skåret til karakteren 1, har fått 7,1 poeng i snitt på del 1 og 1,7 poeng i snitt på del 2. «Toerelever» har 13,4 på del 1 i gjennomsnitt, mot 6,4 poeng på del 2 (se tabell 3.1). Dette innebærer altså at de faglig svakeste elevene i liten grad får vist kompetanse etter at de to første timene er unnagjort. Det kan derfor være grunn til å spørre om ikke også del 2 i større grad burde inneholde oppgaver som gjorde det mulig for «toerkandidatene» å få vist mer av sin matematikkompetanse på de tre siste timene av eksamen.

Tabell 3.1 Karakterer til eksamen for 10. trinn 2017 og oppnådd poengskår på eksamenssettets del 1 og del 2. Poeng.

	Del 1	Del 2
Karakter 1	7,1	1,7
Karakter 2	13,4	6,4
Karakter 3	19,3	15,9
Karakter 4	25,2	27,8
Karakter 5	29,9	38,0
Karakter 6	33,5	45,3

Illustrasjoner og layout

Illustrasjoner

Oppgavetekster består gjerne av både tekst og andre elementer (figurer, bilder, diagrammer). Disse ulike elementene kan spille sammen på ulike måter – de kan si det samme eller ulike ting. Heuvel-Panhuizen (2005) oppgir seks forskjellige roller bilder kan ha i matematikkoppgaver: motivator, situasjonsbeskriver, informasjonsgiver, handlingsindikator, modelltilbyder, løsnings- og løsningsstrategikommunikator. Martiniello (2009) viser hvordan «schematic representations» (visuelle bilder som representerer enten romlige eller matematiske forhold mellom objekter eller symboler, slik som formler, diagrammer og tabeller) slår positivt ut for ELL-elever (English Language Learners). I vår analyse vil vi i utgangspunktet dele illustrasjonene inn i tre kategorier: de som er avgjørende for å kunne løse oppgaven, de som kan være til hjelp for å løse oppgaven, og de som er «til pynt» eller motivasjon.[4]

[4] I «writing guidelines» for TIMSS-oppgaver (Mullis & Martin 2011) er det presisert at «All grafikk inkludert i en oppgave bør være nødvendig for å løse oppgaven og bør være adekvat forklart og referert til direkte i oppgaven» (s. 10, vår oversettelse). Dette er altså ikke gjeldende policy for matematikkeksamen.

Forskning har vist at en del elever (på 4. og 5. trinn) velger å overse illustrasjonene, fordi de er vant til å finne det de trenger i teksten (pga. redundans) (Løvland i Maagerø & Tønnessen 2015, s. 137–140). Måten disse elementene bindes sammen på, kan derfor være vesentlig for forståelsen – hvis det bare er implisitt at elevene må trekke sammen informasjon fra flere elementer, viser det seg vanskelig for enkelte (Gürsoy et al. 2013, s. 11; Prediger et al. 2013, s. 54). Prediger et al. (2015, s. 93) viser på sin side et eksempel på at elever får til en oppgave med en språklig kronglete formulering, sannsynligvis fordi de klarer seg med kun å se på diagrammet, mens Schoultz, Säljö og Wyndhamn (2001) beskriver et eksempel hvor elever måtte ha muntlig hint om å se på figuren for å forstå hva et ukjent ord i teksten kunne bety. I eksamensoppgaver vil det derfor være viktig at det i teksten vises eksplisitt til illustrasjonene.

Det er i tillegg en selvfølge at teksten og illustrasjonen ikke skal motsi hverandre. Nyström (2008, s. 9) viser et eksempel hvor det er inkonsistens mellom tekst og illustrasjon, noe som potensielt kan være problematisk.[5]

Vi har foretatt en gjennomgang av illustrasjonene og gjort en skjønnsmessig kategorisering. Illustrasjonene i de to første kategoriene som det er vist til eksplisitt i teksten, er markert med fet skrift. Når det er flere illustrasjoner i samme oppgave, angis illustrasjonsnummeret i parentes.

Ut fra tabell 3.2 ser vi at de fleste illustrasjonene som er avgjørende for løsning av oppgaven, er vist til eksplisitt i teksten. I 2017-eksamenen er det bare to avgjørende illustrasjoner som ikke er vist til eksplisitt, og i begge tilfellene er det på grunn av layouten helt umulig å forestille seg at elevene skal kunne overse illustrasjonene. For illustrasjonene som er kategorisert som «til hjelp», er det sjeldnere at de vises til direkte i teksten – selv om de altså for eksempel kan bidra til at eleven forstår konteksten bedre. For eksempel mener vi at den første figuren i oppgave 8 i del 2 (som viser Tower Bridge og en seilbåt som passerer under – se eksamenssettet i kap. 1) kan være til hjelp for å forstå konteksten. Det kan da være uheldig for lesesvake elever at den ikke er vist til eksplisitt i teksten, hvis det fører til at de tror at de kan overse illustrasjonen.

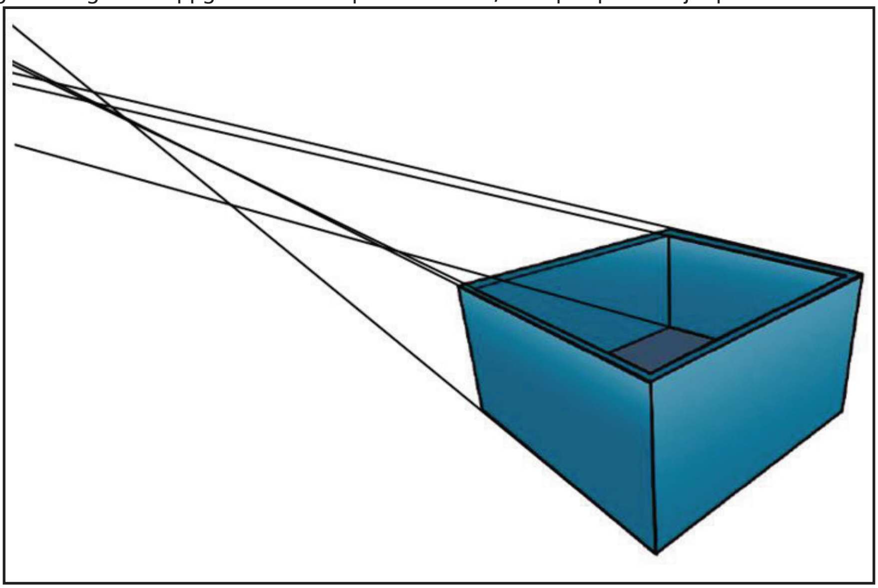
En god del illustrasjoner er kategorisert som «til pynt». I en del oppgaver er det både illustrasjoner som er «til pynt», og illustrasjoner som er avgjørende for å få til oppgaven, og det gjør det kanskje unødig vanskelig for elevene å trekke ut det de trenger av informasjon. Vi vil også understreke at en del av illustrasjonene som er markert som «til pynt», kan fungere som en påminnelse om enkelte ord – dette kan i 2017-eksamenen for eksempel være bildet av hodetelefoner i oppgave 8 i del 1.

[5] I oppgaven Nyström viser til, beskriver teksten en dynamisk situasjon, mens illustrasjonen etter Nyströms mening ikke støtter opp under dynamikken i hendelsen.

Tabell 3.2 Illustrasjoner i 2009-, 2016- og 2017-eksamenene, fordelt etter om de er avgjørende, til hjelp eller til pynt. Illustrasjoner i fet skrift er vist til eksplisitt i oppgaveteksten.

2009			2016			2017		
Avgjørende	Til hjelp	Til pynt	Avgjørende	Til hjelp	Til pynt	Avgjørende	Til hjelp	Til pynt
1:8a	**1:11**	1:3a	1:7	1:9	1:6	**1:5**	1:6	1:8
1:9	1:20	1:7	**1:14**	1:10	1:8	**1:15**	1:7	1:13
1:10	2:1 (1)	1:9	**1:16 (2)**	1:20	1:13	1:16	1:9	1:17
1:15	**2:1 (4)**	1:19	**1:18**	2:2 (1)	1:15	**1:18**	2:1 (1)	2:2 (1)
1:17	**2:5 (2)**	2:1 (3)	**1:19 (1)**	2:2 (2)	1:16 (1)	**1:19**	2:4 (1)	2:3 (1)
1:21		2:3 (1)	**1:19 (2)**	2:4 (2)	2:3	**1:20**	2:8 (1)	2:5
1:22		2:4 (1)	**1:19 (3)**	2:6 (2)	2:5 (1)	**1:21**	**2:8 (2)**	2:6
2:1 (2)		2:5 (1)	1:21	**2:7**	2:6 (1)	**1:22**		2:9 (1)
2:2		2:6 (2)	**2:1**	**2:9 (1)**	2:8	1:23		2:9 (2)
2:3 (2)			**2:2 (3)**			1:25		
2:3 (3)			**2:4 (1)**			2:1 (2)		
2:3 (4)			**2:5 (2)**			2:2 (2)		
2:3 (5)			**2:5 (3)**			2:3 (2)		
2:4 (2)			**2:9 (2)**			2:4 (2)		
2:6 (1)						2:4 (3)		
						2:7 (1)		
						2:7 (2)		

Figur 3.3 Figuren i oppgave 22 i del 1 på 2017-settet, med perspektivlinjer påført.



Illustrasjoner kan også være direkte misvisende. I 2017-settet er det en oppgave med perspektivtegning (del 1 oppgave 22) hvor figuren er feil. I figur 3.3 har vi tegnet inn seks perspektivlinjer som skulle ha møttes i ett punkt. At de ikke gjør det, kan forvirre elever og gjøre at de gir opp oppgaven – mens elever som tilfeldigvis velger to–tre passende linjer, kan få full skår.

I oppgave 19 i del 1 i 2017-eksamenen (se kap. 1) har man gjort seg flid med å få ulik fargenyanse på de ulike muggene, for å unngå at det skulle se ut som at konsentrasjonen var den samme i alle, noe som kunne ha vært forvirrende.

Paratekstlige elementer

Dagrun Skjelbred skriver om paratekstlige elementer – tekst som omgir brødteksten (f.eks. overskrifter) og formatering av brødteksten (f.eks. kursivering) (Maagerø & Tønnessen 2015, s. 40f.). Slike elementer er naturligvis ikke i seg selv faktorer som bidrar til å gjøre teksten vanskeligere å lese – snarere er de lagt inn for å gjøre teksten enklere å forstå. Samtidig kan overforbruk av slike elementer bidra til å gjøre oppgaven uoversiktlig.

Vi har sett på de paratekstlige elementene i eksamensoppgavene i 2009, 2016 og 2017 (Utdanningsdirektoratet 2009, 2016, 2017). Det er interessant å se disse over tid, fordi det kan forvirre elever om elementene brukes ulikt på deres eksamen enn på de eksamensoppgavene de har øvd seg på.

Figur 3.4 Eksempler på paratekstlige elementer i oppgaver fra flere år.

Disse typene paratekstlige elementer forekommer i eksamensoppgavene:

- En overskrift til hver oppgave med oppgavenummer og poengsum.
- Tekst i egne fargelagte bokser. I 2009 var dette ofte tematiske overskrifter (for eksempel «Stortinget»), i 2017 brukes disse isteden til informasjon om hvordan oppgaven skal løses (for eksempel «Regneark»). I 2016 var det en mellomting: Det var en overskrift om tema («Vi reiser til Italia») uten fargelagt boks, mens de fargelagte boksene ble brukt til løsningsinformasjon (for eksempel «Regneark»).
- I 2017-settet gjør man bruk av fet skrift for å markere spesielt viktige ord i enkelte oppgaver. («Hvilket uttrykk har den **laveste** verdien?», del 1, oppg. 3). Dette kan bidra til at færre elever taper poeng ved å svare på noe annet enn det oppgaven spør om. I 2016-settet ble det brukt fet skrift på viktig informasjon om hvordan oppgaven skulle føres (for eksempel «**Vis hvilke formler du har brukt**», del 2, oppg. 4).
- I 2016-settet ble elementer av tallfølgene markert med blå og rød skrift (del 2, oppg. 8). I oppgave b kan dette formodentlig være til hjelp for å se mønsteret. I 2017-settet er koeffisientene i likningene markert med blå skrift (del 2, oppgave 9). Vår skjønnsmessige vurdering er at det er tvilsomt om dette er tydelig nok til at alle elevene legger merke til det, og derfor er det usikkert om det er til hjelp for elevene.
- I oppgave 18 i del 1 av 2009-settet var det en overskrift i blå skrift: «Kjøp 3, betal for 2!» Dette var også gjengitt i teksten. Fargebruken gjør at opplysningen knyttes til figuren som står et par centimeter høyere opp på siden, hvor tre av åtte ruter er markert blått (oppgave 17), men det er vel tvilsomt om dette har forvirret elever.
- I del 1 av settene er det dessuten en del tekst av typen «Løs oppgave 1 her:».

Vi ser altså at det brukes en del paratekstlige elementer i oppgavesettene, både i form av overskrifter, farget skrift og fet skrift, og at disse dels er brukt til tematisk informasjon, dels for å framheve viktig informasjon eller mønstre i oppgavene, og dels til informasjon om hvordan oppgaven skal løses/føres. Ut fra vårt skjønn er det en fare for at det helhetlige bildet kan bli for rotete for enkelte elever, især når nye paratekstlige elementer kommer til uten at det er varslet før eksamen.

Layout

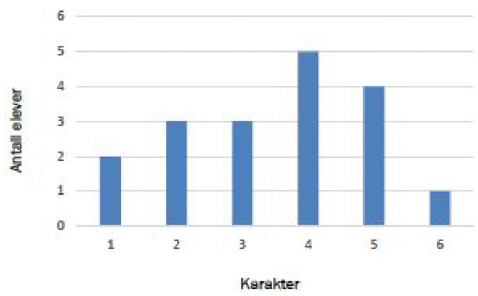
Vi har funnet lite litteratur som handler om de mer layoutmessige sidene ved matematikkprøver. Martiniello (2008) viser et eksempel på hvordan tekstbrytning kan gjøre at elever ikke ser sammenhenger som de normalt ville se – når uttrykket «even number» blir delt på to linjer, klarer ikke eleven å se dette som ett uttrykk. Vi har ikke sett eksempler på dette i eksamenssettene.

Noe av det aller viktigste ved et oppgavesett er at elevene ser oppgavene de er ment å svare på. I 2016-settet stod en deloppgave (del 1 oppgave 19c) nederst på en side under en stor tabell og et stort diagram og kunne være mulig å overse. I oppmannsrapporten for Sogn og Fjordane står å lese at «Fleire har hoppa over c, tyder på at sterke elevar kan ha oversett denne?» (Waag & Myklebust 2017, s. 3). Det er ikke tilsvarende eksempler i 2017-settet.

Figur 3.5 Oppgave 19 i del 1 av avgangseksamen for 10. trinn i 2016.

Oppgave 19 (2,5 poeng)

På en matematikkprøve fordeler karakterene seg slik i en klasse med 18 elever:

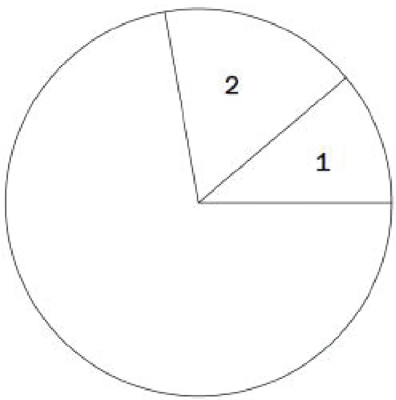


a) Typetallskarakteren er ________

b) Gjør beregninger, og fyll inn det som mangler i tabellen nedenfor.
Lag ferdig sektordiagrammet som viser karakterfordelingen i klassen.

Karakter	Grader for sektor
1	40°
2	60°
3	
4	
5	
6	

Sektordiagram



c) Gjennomsnittskarakteren er ________

Når oppgaver har regneruter, kan elever tenke at størrelsen på rutene signaliserer hvor mye regning som skal gjøres. Flere sensorer har kommentert at regneruten til oppgave 24 i del 1 (i 2017-settet) var for liten (se kapittel 1). Denne oppgaven henger sammen med oppgave 23, og illustrasjonen gjelder begge oppgavene. I figuren står sylinderen til venstre og prismet til høyre, men i svaralternativene har de byttet plass. Det samme skjer i oppgave 24, hvor sylinderen og prismet nevnes i to ulike rekkefølger. Dette kan være en unødvendig kilde til forvirring.

Elever og læreres oppfatning av illustrasjonene

Illustrasjoner i eksamensoppgaver har flere funksjoner og blir som nevnt brukt som pynt eller som hjelpetegninger for å løse oppgavene. Under intervjuene med elevene hadde vi med eksamensoppgaven, og vi lot elevene bla igjennom oppgaven slik at de ble minnet på illustrasjonene. Elevene var svært fornøyde med bruken av illustrasjoner på årets eksamen, både som pynt og som støtte til å løse oppgaven. En elev forklarte at det var fint «å ha noe å hvile blikket på», og at «det hadde vært tyngre og mer komplisert uten illustrasjoner».

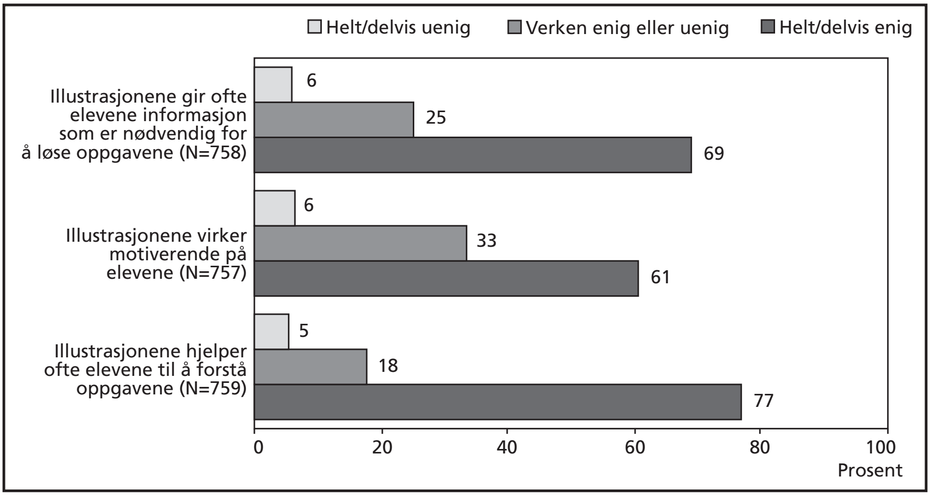
Pyntebilder kunne også ha en funksjon utover å være rent dekorative. Bildene kunne være med å illustrere hva tematikken i oppgaven var. En lærer mente at denne typen bilder kunne hjelpe elever med dårlig språkforståelse til å forstå hva oppgaven handlet om. Den samme læreren forklarte at en del av hans minoritetsspråklige elever ikke hadde de samme referanserammene som elever som var født og oppvokst i Norge av etnisk norske foreldre, og illustrasjoner som viste hva temaet i oppgaven var, kunne være til hjelp for denne gruppen elever. Et eksempel var oppgave 2, del 2, om et par på sykkeltur. Oppgaven har to bilder, et av et syklende par og et bilde av et diagram. Pyntebildet av paret som sykler, hjalp elevene til å skjønne temaet i oppgaven (se eksamensoppgaven). Vi intervjuet også en elev med dysleksi, hun fortalte at hun ofte måtte lese oppgavetekstene flere ganger for å forstå hva hun skulle gjøre, og at det var nyttig med pynteillustrasjoner som hjalp henne å forstå hva oppgaven handlet om.

Designet og layouten med bruk av illustrasjoner ble oppfattet som gode av både lærere og elever. Det var spesielt tre illustrasjoner som ble trukket fram av lærerne og elevene som nyttige når elevene skulle «sette seg inn i» og løse oppgavene. Den ene var illustrasjonen av jerrykannen i oppgave 3b, del 2. Illustrasjonen av jerrykannen hadde to funksjoner. Illustrasjonen viste målene som skulle regnes ut, og at en jerrykanne var en bensinkanne. Den andre illustrasjonen som ble diskutert, var illustrasjonen av Tower Bridge i oppgave 8 i del 2. Denne oppgaven hadde et reelt bilde av Tower Bridge og en hjelpetegning med målene som skulle regnes ut. Flere av elevene forklarte at de ikke ville klart å løse oppgaven hvis det ikke var for hjelpetegningen. Den siste illustrasjonen som ble trukket fram i intervjuene som ekstra nyttig, var bildet av saftkannene i oppgave 9 i del 1. I oppgaven fikk elevene spørsmål om forholdet mellom saft og vann og skulle krysse av for riktig svar. Hvert svaralternativ hadde en illustrasjon med et bilde av et desilitermål med en blanding av saft og vann hvor fargekonsentrasjonen varierte. Lærerne mente at den detaljen gjorde det lettere for elevene å krysse av på riktig svaralternativ.

Illustrasjoner som ikke er tydelige, eller hvor målet med illustrasjonen oppleves som uklart, fører kun til forvirring blant elevene. Det var et tema som opptok lærerne. Det var en illustrasjon, oppgave 22 i del 1, som ble oppfattet som slurvete av lærerne vi intervjuet. Flere av lærerne fortalte at illustrasjonen var direkte feil. På figur 3.3 i dette kapitlet ser vi at det stemmer. Lærerne fortalte at de hadde fått tilbakemeldinger av elever som hadde blitt forvirret av denne oppgaven, og at det er uheldig med slike feil, fordi elevene bruker unødig mye tid på å forsøke å løse en oppgave som ikke har en løsning.

I spørreundersøkelsen ble matematikklærerne bedt om å svare på i hvilken grad de var enige eller uenige i forskjellige påstander om bruken av illustrasjoner i eksamensoppgavene. Resultatene er presentert i figur 3.6.

Figur 3.6 Matematikklæreres svar på påstander om bruk av illustrasjoner på årets eksamen. Prosent.



En stor andel av lærerne (77 prosent) mente at illustrasjonene hjalp elevene til å forstå oppgavene. Over halvparten mente at illustrasjoner virket motiverende på elevene (61 prosent), og at og at illustrasjonene ofte ga elevene informasjon som er nødvendig for å løse oppgavene (69 prosent). Det er færre enn én av ti lærere som var delvis eller helt uenige i noen av påstandene ovenfor, og de resterende lærerne svarte at de verken var enige eller uenige. Svarene sammenfaller med det lærerne og elevene fortalte i de kvalitative intervjuene.

Oppsummering

- Eksamenen har en tendens til noe stigende vanskegrad, men to av de vanskeligste oppgavene i del 2 kommer allerede som oppgave 1b og 1c.
- De svakeste elevene får det meste av sine poenger på del 1 og nesten ingen uttelling på del 2.
- Paratekstlige elementer brukes på en konsekvent måte innad i hvert eksamenssett, men ikke over tid. Derfor bør det vurderes om bruken av disse for det enkelte år bør beskrives i eksamensveiledningen.
- Illustrasjoner som er avgjørende for å løse oppgaven, er i hovedsak tydelig vist til i teksten, og det er viktig å fortsette med det. I 2017-settet var det ett tilfelle hvor en feil i en illustrasjon i seg selv kunne gjøre at elever ga opp oppgaven (del 1, oppgave 22), og dette bør åpenbart unngås.
- Hovedinntrykket blant lærere og elever var at bruken av illustrasjoner på årets eksamen var god, og et flertall av lærerne mente at illustrasjonene hjalp elevene til å forstå oppgavene, og virket motiverende på elevene, og at illustrasjonene ofte ga elevene informasjon som er nødvendig for å løse oppgavene.

4 Samsvar mellom eksamenen og undervisning

Eksamenen er den samme for alle elevene som trekkes ut, men det vil være variasjon i innholdet i undervisningen. For å undersøke om elevene har fått undervisning i det de blir prøvd i til eksamen, har vi både sammenliknet de mest brukte lærebøkene med oppgavene gitt til eksamen, og vi har snakket med lærere og elever om deres opplevelse av samsvar mellom eksamenen og undervisningen.

Læremidler

Hva elever har erfaring med fra opplæringen, kan variere. «I skolen spiller læremidler og læringsressurser en vesentlig rolle for hvordan læreren velger ut, presenterer og strukturerer innholdet i faget» (Kunnskapsdepartementet 2015, s. 22). Blant læremidler og læringsressurser finner vi den trykte læreboken, og noen få bruker samme bok i digital versjon. I Gilje et al. (2016, s. 69) svarte lærerne på spørsmål om hvilke læremidler de hadde brukt siste time, at de i hovedsak hadde benyttet læreboken og elevenes skrive- og arbeidsbøker (s. 69).

På spørsmål til læreren om hvordan de betegner sin bruk av læremidler, svarer hele 82 prosent i vår undersøkelse at de i hovedsak bruker papirbaserte læremidler, men supplerer med noe bruk av digitale læremidler i sin undervisning. I tillegg kommer nærmere 3 prosent som kun bruker papirbaserte læremidler.

Det brukes mange ulike læreverk i matematikkundervisningen i norsk skole. 83 prosent av lærerne som har besvart vår spørreundersøkelse, rapporterer at de bruker et av disse læreverkene: *Faktor 10, Grunntall 10, Maximum 10, Nye Mega 10* og *Tetra 10*. De øvrige rapporterer bruk av andre lærebøker eller annet materiale. Disse fem mest brukte læreverkene er tatt med i analysen av eksamenen fra våren 2017. Forlagene tilbyr også prøver lærerne kan bruke som støtte for sin vurdering. I undersøkelsen lærerne besvarte etter eksamen, oppgir 85 prosent at de bruker forlagsprøver som en del av undervisningen, og alle bruker tidligere eksamensoppgaver. Dette er det tatt hensyn til i evalueringen av de enkelte oppgavene.

I læreplanen i matematikk er ikke kompetansemålene knyttet til hvert enkelt skoleår, men til perioder i opplæringen. Læreplanen beskriver kompetanse etter 2. trinn, 4. trinn, 7. trinn og 10. trinn i grunnskolen. Det betyr at det er opp til lærerne og skolen å fordele opplæringen innenfor disse intervallene, og det er gjort ulike prioriteringer av fagområder som behandles i lærebøkene. Av de fem lærebøkene vi har sett på, har to av bøkene repetisjonskapitler i hovedboken som også tar med stoff fra bøkene på 8. og 9. trinn. For alle forlagene gjelder at de har med områder fra tidligere trinn, men i varierende grad. I del 2 av eksamenen kan elevene bruke alle hjelpemidler bortsett fra de som tillater kommunikasjon. Her vil det være en fordel å ha bok som har med alle emnene som testes på eksamen.

Gjennomgang av lærebøkene

Alle læreverkene har oppgavesamlinger enten som en del av hovedboken eller i egen bok. Det er sannsynlig at elevene har regnet oppgaver fra oppgavesamlingen, men vi antar at variasjonen her er større enn når det gjelder innholdet i hoveddelen av læreboken. Ut fra svarene lærerne gir om bruk av andre ressurser, vil elevene også i ulik grad ha erfaring med oppgaver fra forskjellige nettsteder og egne ressurser utarbeidet på den enkelte skole og av den enkelte lærer. Det vil derfor være en større variasjon i hva elevene har erfaring med fra opplæringen, enn det som kommer fram i vår analyse som baserer seg på det vi må anta er et minimum av det elevene har erfaring med fra opplæringen. Lærerne rapporterer at de bruker tidligere eksamensoppgaver i sin undervisning.[6] Vi har derfor tatt med eksamenssettene fra 2015 og 2016 når vi har vurdert hva elevene har kjennskap til fra opplæringen.

Rammeverk for analysen

I analysen av eksamen opp mot lærebøkene har vi tatt utgangspunkt i rammeverket som er utarbeidet av Johan Lithner og beskrevet i «A research framework for creative and imitative reasoning» (Lithner 2008), og slik det er brukt i flere studier.[7] I dette rammeverket er imitativt og kreativt resonnement sentrale begreper. Resonnement kan her defineres som den tankerekken som brukes til å formulere antakelser og komme fram til en konklusjon i forbindelse med oppgaveløsning (Palm, Boesen & Lithner 2011, s. 224). I rammeverket til TIMSS 2015 (Mullis & Martin 2013) og i beskrivelsen av matematisk kompetanse i «Kompetencer og matematiklæring – ideer

[6] Det var kun én lærer som rapporterte at eksamensoppgaver ikke ble brukt.

[7] Se blant annet Boesen 2006 og Palm, Boesen & Lithner 2011.

og inspiration til udvikling af matematikundervisning i Danmark» (Niss & Jensen 2002) er resonnement brukt i en annen betydning. En av de åtte kompetansene Niss og Jensen definerer, er resonnementskompetanse som er karakterisert ved evnen til å følge og vurdere et matematisk resonnement og spesielt forstå hva et matematisk bevis er, kunne avdekke de bærende ideene i et matematisk bevis og kunne gjennomføre uformelle og formelle matematiske bevis (Niss & Jensen 2002, s. 54). I rammeverket til TIMSS er det tre kognitive områder. Her er «å resonnere» knyttet til evnen til å tenke logisk, analysere situasjoner og sammenhenger, generalisere situasjoner, kombinere informasjon, begrunne påstander og løse problemer som ikke er rutinepregede (Institutt for lærerutdanning og skoleforskning 2015). Resonnementskompetanse brukes også som begrep flere steder i eksamensveiledningen for skriftlig eksamen i matematikk på 10. trinn (Utdanningsdirektoratet 2017a) uten at begrepet er nærmere definert, men vi tolker det til å være i tråd med beskrivelsene hos Niss og Jensen og i rammeverket for TIMSS.

For å unngå at samme begrep brukes med ulik betydning, har vi valgt å bruke begrepene algoritmisk løsning og kreativ løsning i klassifiseringen av eksamensoppgavene. Løsning er i denne sammenhengen definert som hele prosessen med å tolke oppgaven, velge løsningsmetode og gjennomføre oppgaveløsningen. En løsning vil i de fleste tilfeller være et svar og en begrunnelse for at svaret er korrekt, der begrunnelsen gjerne kan være en utregning.

Kategorisering av oppgavene

Først er det gjort en analyse av selve oppgaven med identifikasjon av eventuelle algoritmer som kan brukes i løsningen, kjennetegn ved oppgaven som eksplisitte instruksjoner eller hint og svarformat. Oppgavene vurderes så opp mot hver av de fem lærebøkene vi har tatt med. Vi ser etter like og liknende oppgaver og eksempler som kan brukes i hele eller deler av løsningen. Dersom det for oppgavene i del 1 finnes en algoritme for løsningen, og boken har tre eller flere like og liknende eksempler og oppgaver, vil oppgaven karakteriseres som en oppgave som krever algoritmisk løsning. Tre eksempler eller oppgaver er valgt fordi det gir mulighet til noe trening, og når oppgaven gis så mye plass i hovedboken, er det sannsynlig at samme type oppgaver også finnes i oppgavesamlingene. For del 2, hvor alle hjelpemidler er tillatt, kan det være nok med ett eksempel eller én oppgave dersom likheten til eksamensoppgaven er så stor at hele eller deler av eksemplet/oppgaven kan brukes som guide i løsningen. I de tilfellene der liknende oppgaver og eksempler ikke finnes i lærebøkene og rutinepreget, algoritmisk løsning ikke er mulig, vil en kreativ matematisk løsning være nødvendig på deler av eller i hele oppgaven.

Etter en gjennomgang av lærebøkene var det noen oppgaver som klart falt i den ene eller andre kategorien for løsning. For en stor del av oppgavene var det tvil etter første gjennomgang fordi bøkene er forskjellige. Det betyr også at elevenes kjennskap til disse oppgavene vil variere med hvilken lærebok de har brukt.

Forlagsprøvene og tidligere eksamensoppgaver brukes til å justere klassifiseringen i disse tilfellene. Tabell 4.1 viser fordelingen av oppgaver etter gjennomgang av bøkene og etter korrigering mot forlagsprøver og eksamen.

Tabell 4.1 Eksamensoppgavenes fordeling mellom algoritmisk og kreativ løsning før og etter korrigering mot forlagsprøver og eksamen.

	Etter gjennomgang av bøkene		Korrigert mot forlagsprøver og eksamen	
	Del 1	Del 2	Del 1	Del 2
Algoritmisk løsning	40 %	16 %	74 %	37 %
Ulikhet mellom bøkene	49 %	41 %		
Kreativ løsning	11 %	43 %	26 %	63 %

Eksamensveiledningen fra Utdanningsdirektoratet (2017a), kapittel 1.3, beskriver innholdet i eksamensoppgavene. Både del 1 og del 2 skal dekke et bredt spekter av kompetansemålene i læreplanen og ha oppgaver av ulik vanskegrad. Mens et viktig element i del 1 er ferdighetsregning, vil del 2 inneholde oppgaver som prøver både bredden og dybden i elevenes matematiske kompetanse, og oppgavene tar utgangspunkt i dagligdagse situasjoner og matematikkfaglige temaer (ibid.). Det er derfor ikke overraskende at andelen oppgaver som krever kreativ løsning, er større i del 2 enn i del 1.

Algoritmisk og kreativ løsning i del 1

Oppgave 12a er løsning av en lineær likning. Liknende eksempler og oppgaver finnes i alle bøkene, på forlagsprøver og i tidligere eksamensoppgaver. Dette er derfor en oppgave som uten tvil kategoriseres under «algoritmisk løsning». Løsningen består i å kjenne igjen oppgavetypen og implementere en kjent metode. Hvis eleven husker algoritmen og har trening i å bruke den, er det bare en eventuell slurvefeil som kan hindre eleven i å løse oppgaven riktig.

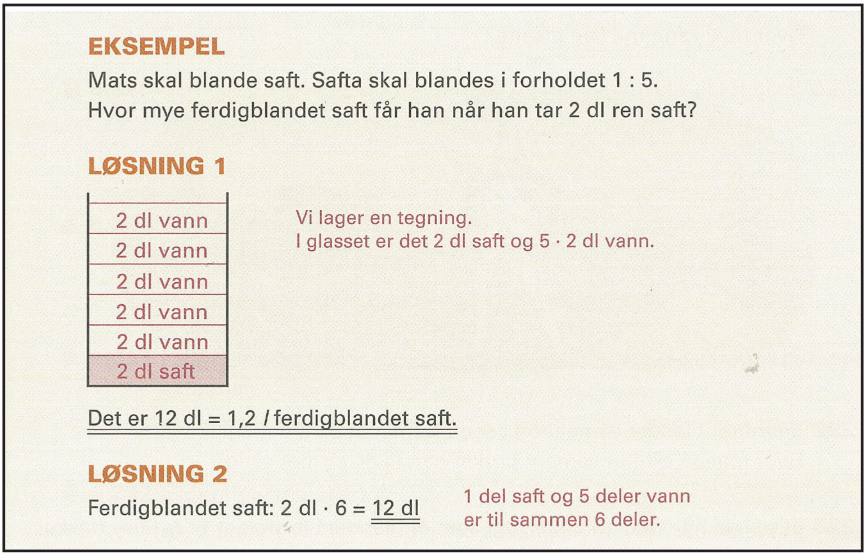
Oppgave 11b er en forenkling av algebrauttrykk som krever at eleven kjenner til og bruker konjugatsetningen i faktoriseringen av uttrykket $a^2 - b^2$. Selv om eleven skal kjenne kvadratsetningene, finner vi i liten grad eksempler på at den er brukt til forenkling av algebrauttrykk, selv om det er enkelte eksempler på forlagsprøver fra siste skoleår. Selv om oppgaven løses ved bruk av algoritmer, krever den kombinasjon av disse på et nivå vi vurderer til å være «kreativ løsning» på 10. trinn. Det forsterkes av at dette ikke er en flervalgs- eller kortsvaroppgave, men krever begrunnelse for svaret i en regnerute.

Oppgave 25b er av en type som tradisjonelt er mer forbundet med kreativt resonnement og løsning. Oppgaven går ut på å lage en formel for figur *n*, og eleven må lete etter et mønster, finne en generalisering og begrunne denne eller vise hvordan hun eller han har resonnert seg fram til løsningen. De fleste bøkene behandler ikke figurtall og tallmønstre, og liknende oppgave finnes ikke på de siste forlagsprøvene eller eksamensoppgavene. Innen hovedområdet «Tall og algebra» i læreplanen finner vi i kompetansemålene etter 7. trinn at eleven skal kunne utforske og beskrive strukturer og mønstre i geometriske mønstre og tallmønstre med figurer, ord og formler, og etter 10. trinn at eleven skal kunne bruke tall og variabler i utforskning, eksperimentering og praktisk og teoretisk problemløsning (Utdanningsdirektoratet 2013). Denne oppgaven ligger derfor innenfor læreplanens kompetansemål, men krever kreativ løsning.

For en stor del av oppgavene var det tvil om kategorisering basert på ulikhet i lærebøkene. Oppgave 13 i del 1 er en oppgave med omgjøring av et stort tall til standardform. To av bøkene behandler ikke standardform, men oppgavetypen finnes på forlagets terminprøver og i tidligere eksamensoppgaver. Oppgaven ble plassert i kategorien for algoritmisk løsning.

Det eneste eksemplet på at en oppgave er nesten identisk med et eksempel i læreboken, fant vi for oppgave 19. Selv om boken ikke er tilgjengelig i del 1 av eksamenen, vil oppgaven være enklere for de som har sett dette eksemplet og regnet flere tilsvarende oppgaver (se figur 4.1 og oppgave 19 i eksamenssettet).

Figur 4.1 Eksempel fra lærebok som er nesten identisk med oppgave 19 i del 1 av grunnskoleeksamenen i matematikk 2017.


EKSEMPEL

Mats skal blande saft. Safta skal blandes i forholdet 1 : 5.
Hvor mye ferdigblandet saft får han når han tar 2 dl ren saft?

LØSNING 1

2 dl vann
2 dl vann
2 dl vann
2 dl vann
2 dl vann
2 dl saft

Vi lager en tegning.
I glasset er det 2 dl saft og 5 · 2 dl vann.

Det er 12 dl = 1,2 *l* ferdigblandet saft.

LØSNING 2

Ferdigblandet saft: 2 dl · 6 = 12 dl

1 del saft og 5 deler vann
er til sammen 6 deler.

Kilde: Bakke & Bakke 2015, s. 271

De fleste bøkene behandler forhold kun i forbindelse med formlikhet i geometri, og det er kanskje ikke så rart siden forhold ikke står i kompetansemålene utover det som står om formlikhet under dette hovedområdet. Dette kompetansemålet finnes derimot under hovedområdet «Måling» etter 7. trinn: «Mål for opplæringa er at eleven skal kunne bruke forhold i praktiske samanhengar, rekne med fart og rekne om mellom valutaer.» (Utdanningsdirektoratet 2013). Eksamensveiledningen har følgende formulering: «Videre forutsettes det at elevene behersker grunnleggende formler og fremgangsmåter fra tidligere skolegang.» (Utdanningsdirektoratet 2017a, s. 25). Oppgaven er likevel kategorisert til å kreve kreativ løsning ut fra det vi finner i de fleste lærebøkene, forlagsprøver og eksamenen for 10. trinn.

Algoritmisk og kreativ løsning i del 2

I del 2 er alle hjelpemidler tillatt bortsett fra kommunikasjon. Dersom det i bøker og løsningsforslag finnes tilsvarende oppgaver med karakteristika som gjør likheten med eksamensoppgaven tydelig, kan oppgaven løses ved å bruke disse som guide, og oppgaven vil derfor ikke kreve kreativ løsning, men det kan ta tid å lete seg fram til eksempler i boken eller i egne løsninger av oppgaver. Med stor arbeidsmengde vil effekten av tilgjengelige hjelpemidler bli mindre. Det er også færre deloppgaver som krever utelukkende algoritmisk løsning, siden oppgavene ofte går på tvers av læreplanens hovedområder.

I del 2 av eksamenen er det krav om bruk av digital graftegner og regneark. Det er også mulig å bruke dynamisk geometriprogram, men dette er ikke obligatorisk. Kravet om bruk av digital graftegner har kommet etter at de fleste lærebøkene er skrevet, så her er elevene avhengige av at de får undervisning i dette utover det som står i læreboken. I klassifiseringen av disse oppgavene har vi derfor sett spesielt på forlagsprøvene og eksamensoppgavene. Kategoriseringen av oppgavene er gjort på samme måte som for del 1. Vi tar likevel med to eksempler der alle eller de fleste deloppgavene på en oppgave i utgangspunktet hadde ulik kategorisering basert på gjennomgang av bøkene.

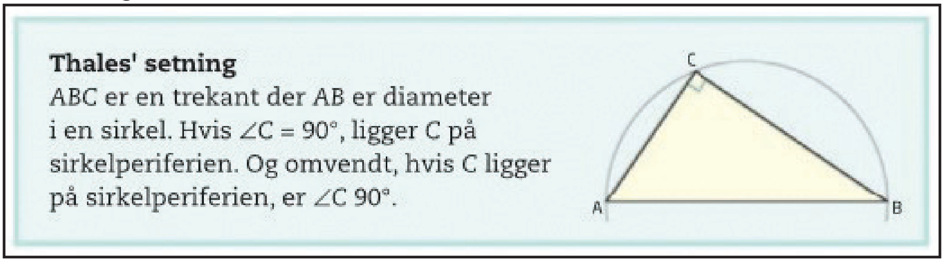
I oppgave 5 er det obligatorisk å bruke digital graftegner for å få full uttelling, og oppgaver med dette kravet finnes på alle forlagsprøvene vi har hatt tilgang til, og også på eksamenene for 2015 og 2016. Eksamensveiledningen har et kapittel som beskriver hva som kreves av en fullgod løsning på denne typen oppgaver (Utdanningsdirektoratet 2017a, kap 1.6.3.3), men det er ingen oversikt over hvilke kommandoer det forutsettes at elevene har kjennskap til. Eksemplet i eksamensveiledningen har med «skjæring mellom to objekter». For å besvare c-oppgaven er det nødvendig å kjenne kommandoen «ekstremalpunkt» i GeoGebra (eller tilsvarende for andre digitale graftegnere).

Det å tegne en begrenset graf med digital graftegner som a-oppgaven krever, vil være en oppgavetype elevene har møtt gjennom opplæringen. De strenge kravene til hva som skal til for å få full uttelling, kan det være en utfordring å imøtekomme, men det er snakk om å kopiere eksempler og løsningsforslag fra tidligere terminprøver og eksamener. Når det gjelder b-oppgaven, er det bare i én av bøkene vi finner andregradsfunksjoner med avlesing av to skjæringspunkter, så her må de fleste elevene bruke kunnskap om ett skjæringspunkt som er kjent fra tidligere oppgaver, kreativt i løsning av denne oppgaven. Det er ikke presisert at man er på jakt etter begge skjæringspunktene, og det ligger derfor ikke noe hint i oppgaveteksten som kunne plassert oppgaven i en annen kategori. Dersom formuleringen hadde vært «For hvilke hastigheter er CO_2-utslippene 180 g/km?», ville det ligget et hint i teksten om at det er flere svar. Hvis vi i oppgave c forutsetter at eleven kjenner nødvendig kommando og derfor har forutsetning for å løse oppgaven, er det i dette tilfellet ikke bare koordinatene til et bunnpunkt, men også å tolke betydningen av avlesingen eleven må gjøre. Dette finnes det i så liten grad eksempler på i bøker og forlagsprøver at vi vurderer det som at det krever kreativ løsning.

I oppgave 7 i del 2 skal eleven konstruere en trekant innskrevet i en sirkel og kan velge å gjøre dette enten for hånd med passer, linjal og blyant eller digitalt med dynamisk geometriprogram. Bruk av dynamisk geometriprogram er ikke obligatorisk på eksamen, og vi finner det ikke på forlagenes terminprøver, så det er usikkert i hvilken grad elevene kjenner til det gjennom opplæringen, selv om det er en del av et kompetansemål under hovedområdet geometri. Konstruksjon for hånd har vi funnet både på eksamen og på forlagsprøver. Det gjelder også forlagsprøvene knyttet til bøkene som ikke har med konstruksjon i boken for 10. trinn. Det varierer hvor mye konstruksjonene likner på den som skal gjøres på eksamen, men hovedtyngden ligger på algoritmisk løsning.

Begge de to siste deloppgavene på oppgave 7 er vurdert til å kreve kreativ løsning. Én av bøkene har med Thales setning og et bilde helt tilsvarende det vi finner i oppgaven (se figur 4.2 og oppgave 7 i del 2 av eksamenen), men i de andre bøkene er ikke Thales setning eller innholdet i den berørt.

Figur 4.2 Eksempel fra lærebok på figur som er nesten identisk med illustrasjonen til oppgave 7 i del 2 av grunnskoleeksamenen i matematikk 2017.


Thales' setning
ABC er en trekant der AB er diameter i en sirkel. Hvis $\angle C = 90°$, ligger C på sirkelperiferien. Og omvendt, hvis C ligger på sirkelperiferien, er $\angle C$ 90°.

Kilde: Tofteberg, G. N., m.fl., 2015, s. 57

Det er regning på trekanter og sirkler hver for seg i bøkene, men akkurat denne typen sammensatt figur som oppgave 7 har, finner vi ikke, og vi har vurdert det slik at elevene her må kombinere kunnskap på en for dem ny måte, og at oppgaven krever kreativ løsning.

Til slutt kan det være interessant å se på en oppgave som ikke har vakt stor oppmerksomhet i kommentarer fra sensorer, elever og lærere, men som inneholder en feil. Det er oppgave 6b og c som handler om fart og bremselengder der det er presisert at farten er målt i m/s, men friksjonstallet og konstanten 19,62 oppgis uten benevning. Innsetting i formel og utregning gir benevningen m^2/s^2 på bremselengden og fart med benevning $\sqrt{m}$. I noen bøker finner vi kun innsetting i formler uten benevning, og benevningen er først med når endelig svar oppgis. I andre bøker er benevning med fra start enkelte ganger, og andre ganger kommer den på i tekstsvaret til slutt. Alle elevene har derfor møtt formelregning uten benevning og blir neppe forvirret av manglende benevning på konstant og friksjonstall i dette tilfellet. Men elevene har i liten grad møtt eksempler og oppgaver med omgjøring av formler annet enn litt i forbindelse med geometri. For de som ikke benytter CAS[8], vil oppgave 6c kreve kreativ løsning.

Språk og layout i lærebøkene

Vi ønsker å se om det er samsvar mellom eksamensoppgavene og oppgaver elevene kjenner fra lærebøkene når det gjelder språk og layout. Å skulle analysere lærebøkene i sin helhet ville være for tidkrevende. Gjennom de andre analysene i rapporten har vi funnet fram oppgaver i lærebøkene som tester det samme som eksamensoppgavene. Vi har derfor valgt å gjøre en analyse av utvalgte konkrete oppgaver i lærebøkene og sammenlikne med de tilhørende eksamensoppgavene. For hver deloppgave i del 2 av eksamenssettet har vi valgt opptil tre oppgaver fra lærebøkene som tester det samme. Disse har vi valgt tilfeldig, dog slik at flest mulig læreverk er representert per deloppgave. I alt gir dette et materiale på 22 lærebokoppgaver fordelt på åtte eksamensdeloppgaver.

Analysene som vi har gjort på eksamensoppgavene (se kapittel 7), er så gjentatt for lærebokoppgavene. (For de rent kvantitative analysene har vi tatt gjennomsnittet av resultatene for hver enkelt lærebokoppgave som svarer til en eksamensoppgave.) Denne analysen antyder at lærebokoppgavene skiller seg en del fra eksamensoppgavene: Det er noe færre ord per setning (7,2 mot 7,5) og noe lavere andel lange ord (20,4 prosent mot 24,9 prosent) i lærebokoppgavene. Viktigere er det nok at hyppigheten av lavfrekvente, matematikkfaglige ord er vesentlig høyere (nær sju ganger så høy som i eksamensoppgavene), samtidig som andelen sammensatte ord bare er halvparten av den som er i eksamensoppgavene. Matematikkfaglige ord kan, selv når de er lavfrekvente, gjøre oppgaven enklere å løse, siden de signaliserer tydelig hva oppgaven går ut på. I lærebøker er dessuten disse ordene nettopp brukt i teksten før oppgaven. Materialet

8 CAS står for Computer Algebra System. For eksempel inneholder det mye brukte programmet GeoGebra en CAS-del.

er lite, men det kan indikere at matematikkoppgavene i lærebøkene slik sett kan være noe enklere å forholde seg til enn eksamensoppgavene.

I forberedelsene fram mot eksamen spiller også tidligere eksamensoppgaver og forlagsprøver en rolle. I analysene av 2017-eksamenssettet (se kapittel 7) har vi sett at det ikke skiller seg sterkt fra de tidligere eksamenssettene vi har sammenliknet med. Arbeid med tidligere eksamenssett kan derfor fungere godt som forberedelse til eksamen. Forlagsprøver brukes mye i skolen, men vi har ikke analysert disse her.

Opplever lærere og elever at det er samsvar mellom opplæring og eksamen?

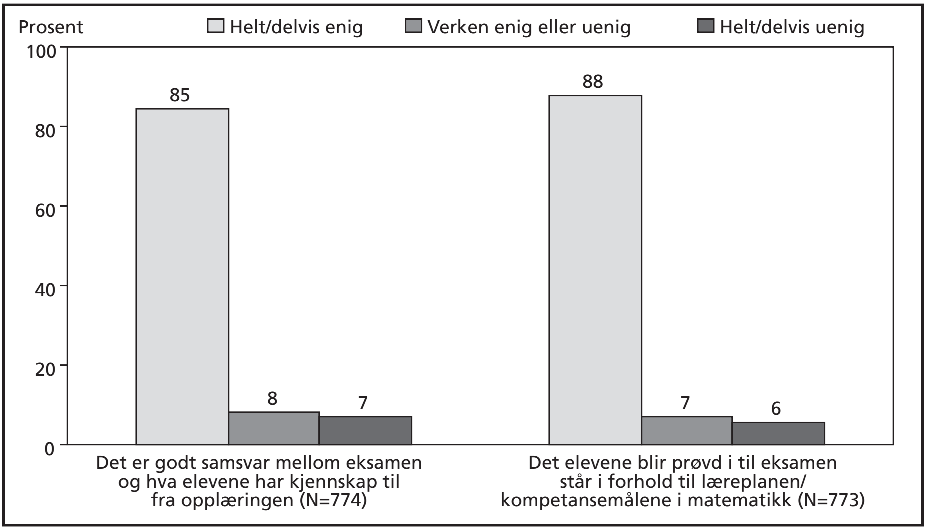
En viktig forutsetning for å kunne prestere godt under eksamen er at kandidatene har fått undervisning i de emnene som de prøves i. I den grad det er systematiske forskjeller i emner som er tatt opp i undervisningen, er det rimelig å anta at det vil resultere i ulike prestasjoner på eksamensdagen. Forskjellig undervisning er følgelig en potensiell kilde til intern urettferdighet mellom kandidater ved ulike skoler, men også mellom kandidater med forskjellige lærere ved samme skole. Følgelig er sammenhengen mellom læreplanen og opplæringen som faktisk er gitt, et tema både i den kvantitative og den kvalitative delen av denne studien. Mer konkret har vi spurt om det var spørsmål på årets eksamen som elever og/eller lærere mener at ikke har vært behandlet i timene.

Metodisk er ikke denne strategien perfekt. Trolig vil lærere unngå å si at de ikke har undervist i temaer som de skulle ut fra læreplanen. Viktigere er det at denne typen subjektive vurderinger ikke kan brukes for å vurdere eventuelle forskjeller i kvaliteten på undervisningen, og strengt talt forteller det heller ikke noe om eventuell ulik vekt av deler i undervisningen. Vi mener likevel det er svært relevant å identifisere i hvilken grad det er en oppfatning at årets eksamen består av spørsmål elevene har ulike forutsetninger for å besvare.

Basert på våre data, både de kvalitative og de kvantitative, er det rimelig å konkludere at det er en gjennomgående oppfatning om at årets eksamen var god. Og i den sammenheng er det viktig å understreke at en stor del av kvaliteten nettopp synes å være fraværet av oppgaver som lå på utsiden eller i periferien av kompetansemålene. Med ett unntak ble spørsmålene oppfattet som sentrale, og ikke minst var både elever og lærere enige i at undervisningen hadde tatt opp emnene som ble gjenstand for prøve.

I figur 4.3 har vi satt inn svarene på hvordan vurderingen er knyttet til to sentrale mål på samsvar. Det første dreier seg om opplæringen elevene har fått, mens det andre målet tematiserer samsvaret mellom eksamenen og kompetansemålene.

Figur 4.3 Grad av samsvar mellom kompetansekrav og hva som prøves til eksamen. Prosent.



Resultatene er påfallende enstemmige. Nær ni av ti lærere sier seg helt eller delvis enige i at årets eksamen både traff hva elevene har blitt undervist i, og de generelle kompetansemålene. Dette må ses som et svært godt resultat. At det er stor grad av samsvar mellom de to, er heller ikke overraskende. Det ville være påfallende om det var en gjennomgående oppfatning om at eksamenen var tett på kompetansemålene, men ikke på den undervisningen som ble gitt.

Nettopp av den grunn er det kanskje verdt å gjøre en liten randbemerkning om andelen som er helt eller delvis uenige. Når 7 prosent – nesten én av ti – er helt eller delvis uenige i at årets eksamen traff opplæringen som ble gitt, er det fristende – og viktig – å gå mer inn i denne gruppen senere. Er dette bestemte skoler eller lærere? Og hvordan er undervisningen lagt opp når resultatene blir så kontrære til majoriteten?

Det overordnede bildet vi har gitt, er også det som kom fram i det kvalitative materialet. To elever forteller:

> Jeg synes oppgavene var veldig gode.

En annen sa det slik:

> Vi hadde tatt opp alt i timene, og særlig med øvingene på tidligere eksamener … Den var faktisk mye lettere enn jeg trodde. Eneste problemet var siste spørsmål i del 2. Den skjønte jeg ikke noe av.

Den siste oppgaven i del 2 (9c) ble av både lærere og elever omtalt som eksemplet på en oppgave som man ikke kunne forvente at de skulle kunne. Denne var både lærere og elever enige om at lå utenfor det man hadde behandlet i timene.

Med denne oppgaven som utgangspunkt dreide samtalene seg gjerne inn på hva ulike informanter mente at skulle være nivået og ikke minst, forskjellene i nivå, på ulike oppgaver på eksamenen. På direkte spørsmål var også elevene enige i at det hadde vært en dårlig eksamen om alle hadde kunnet svare på alle oppgavene:

> Det må være variasjon for at også de beste skal kunne vise seg fram.

I den sammenheng kan vi legge til at det for mange ble en diskusjon hvor dette spørsmålet ble tatt opp, hvor de nærmest argumenterte for at de hadde klart alle andre spørsmål. I praksis var det neppe slik. Det er derfor mer nærliggende å skille mellom *vanskelige* oppgaver på den ene siden og oppgaver som ligger *utenfor* kompetansemålene, på den andre. Når det gjelder den siste oppgaven, synes den å være vanskelig for mange fordi den lå utenfor det de hadde lært. Samtidig var det lærere som i etterpåklokskapens lys sa at de nok kanskje burde ha tatt opp mer som hadde gjort elevene bedre i stand til å løse nettopp denne oppgaven.

I forlengelsen av dette punktet er det også interessant å trekke fram spørsmålet om det er emner som bortimot aldri blir prøvd til eksamen. I utgangspunktet er det nærliggende å konkludere med at dersom det er emner som det ikke spørres om, så bør det kanskje også vurderes om de bør stå på pensum. Resultatene indikerer at dette ikke er noe stort problem. Kun 13 prosent mener at det er stoff som ikke prøves på eksamen. Det store flertallet – ni av ti – sitter ikke med en slik opplevelse.

Oppsummering

- Det er ikke funnet systematisk favorisering av én lærebok. Metoden som er brukt til å kategorisere oppgavene, kunne også gitt oppgaver som falt helt utenfor. Det ville gjelde oppgaver der kunnskap utover det som er beskrevet i læreplanen, er nødvendig for å løse oppgaven, men det er ikke funnet slike oppgaver på eksamenen for 2017.
- Det er en utfordring å gi elevene et rimelig likt utgangspunkt for å løse eksamensoppgavene som krever digitale hjelpemidler, siden de fleste bøkene er utgitt før det ble et krav om bruk av digital graftegner på eksamen. Mer presise føringer i eksamensveiledningen kunne hjulpet på dette.
- Det er en gjennomgående oppfatning blant lærerne om at årets eksamen var god, mye på grunn av fraværet av oppgaver som lå på utsiden eller i periferien av kompetansemålene.
- De fleste lærerne (85 prosent) var også enige i at det er samsvar mellom kompetansekrav og hva som prøves til eksamen.

- Elever ga uttrykk for at de var fornøyde med årets eksamen, og at eksamenen dekket emner de hadde hatt om i undervisningen.

5 Eksamenens vanskegrad

I dette kapitlet ser vi på eksamenens vanskegrad ut fra tre ulike perspektiver. Først diskuterer vi vanskegraden ut fra om oppgavene er algoritmiske eller krever kreativ løsning. Deretter ser vi på vanskegraden ut fra hvor mange poeng elevene klarer på oppgavene. Til slutt ser vi på elevenes og lærernes syn på vanskegraden.

Hvor vanskelig en eksamen er, vil være subjektivt. Dersom det er godt samsvar mellom eksamenen og hva elevene har kjennskap til fra opplæringen, vil eksamenen oppfattes som enklere enn om dette ikke er tilfelle. Selv om samsvaret er godt, kan det være oppgaver som likevel er krevende og gir mulighet til å vise kompetanse på høyt nivå.

I eksamensveiledningen (Utdanningsdirektoratet 2017a) er det en matrise som beskriver kjennetegn på måloppnåelse og kommentarer til denne. Matrisen er delt i tre områder der begreper, forståelse og ferdigheter er ett område, problemløsning og kommunikasjon er de to andre. For hvert område beskrives hva som kjennetegner måloppnåelse tilsvarende karakteren 2, karakterene 3 og 4 og karakterene 5 og 6. Gjennomgående i denne matrisen er at kompetansen som beskrives for de laveste karakterene, kan knyttes til algoritmisk løsning, mens kompetansen som kreves for de høyere karakterene, kan knyttes til kreativ matematisk løsning. Gjennomgangen av eksamensoppgavene i kapittel 4 gir derfor også en indikasjon på vanskegraden på eksamenen.

Variasjon i vanskegrad

I formålet for fellesfaget matematikk heter det blant annet:

> «Problemløysing høyrer med til den matematiske kompetansen. Det er å analysere og omforme eit problem til matematisk form, løyse det og vurdere kor gyldig det er.» (Utdanningsdirektoratet 2013).

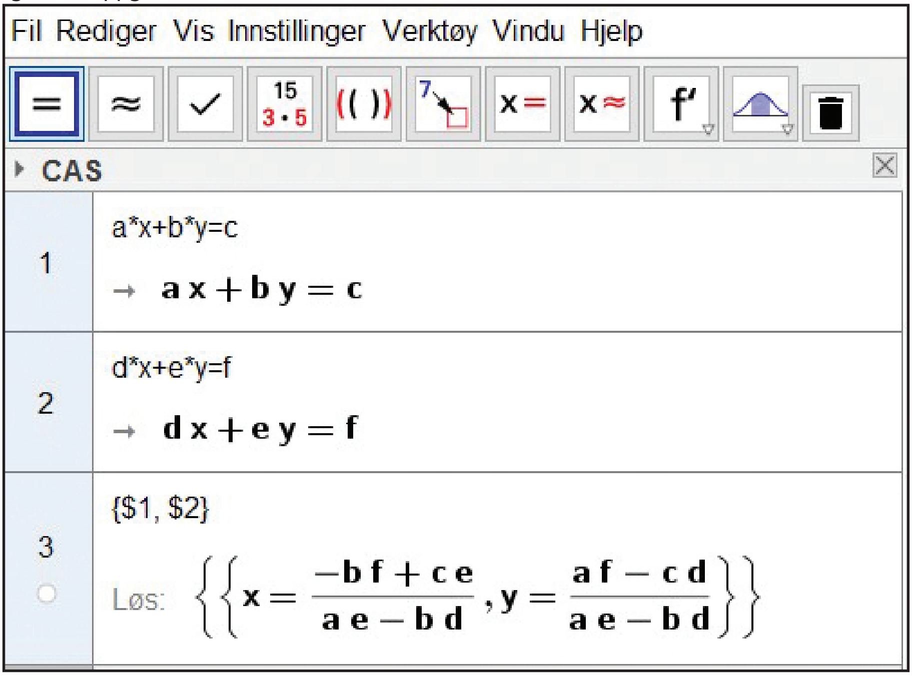
For å kunne vurdere om en matematikkoppgave krever problemløsning eller ikke, må det være definert hva det er som er et matematisk problem. Dersom elevene på eksamen møter oppgaver der de ikke har en kjent algoritme som kan brukes i løsningen, vil vi hevde at oppgaven krever problemløsning. Det vil gjelde alle oppgavene som er kategorisert som kreative i kapitlet om samsvar mellom eksamensoppgavene og det elevene har kjennskap til fra opplæringen, men både disse og oppgavene som krever algoritmisk løsning, varierer i vanskegrad.

Oppgavene med algoritmisk løsning har et stort spenn i vanskegrad. De innledende oppgavene i del 1 krever enkel bruk av de fire regneartene, mens oppgaver lenger ut i settet kan gi vanskeligere utregninger. Selv om forenkling av bokstavuttrykk finnes i de fleste lærebøkene og i tidligere eksamensoppgaver, vil oppgave 11a være vanskeligere enn om det var en oppgave med tall. I oppgave 24 i del 1 er det selve anvendelsen av algoritmen som kan være krevende og oppgaven derfor vanskelig. Dette er en fortsettelse på oppgave 23 uten at det er avhengighet mellom oppgavene. Tallene settes inn i formlene for overflate av sylinder og prisme. Mange elever vil starte utregningen fra venstre mot høyre og får vesentlig mer komplisert regning enn en elev som bruker regnelover og algebrakunnskaper aktivt.

Den siste oppgaven i eksamenssettet har likningssett som tema. Dette behandles i alle lærebøkene, og elevene skal derfor være kjent med algoritmer for å løse likningssett med tall i tillegg til to ukjente, men her er det veldig stor variasjon i hvor anvendelige eksempler elevene finner i læreboken de har tilgang til på eksamen. På denne oppgaven er det også mulig å se løsningen direkte ved betraktning av likningssettet. Dersom eleven ikke er kjent med metodefriheten som ligger i formuleringen «løs», blir en utregning for hånd krevende på dette nivået. Da testes både metode for løsning av likningssett og brøkregning, og det vil kreve at «Eleven kan kombinere begreper og kunnskap fra ulike områder og behandle forskjellige matematiske representasjoner og formler på en sikker måte» som er kjennetegn på høy måloppnåelse innen begreper, forståelse og ferdigheter (Utdanningsdirektoratet 2017a). Hvis eleven derimot behersker CAS, vil ikke noe av denne kunnskapen testes. «Eleven kan velge og bruke en rekke hjelpemidler med stor sikkerhet» er kjennetegn på høy måloppnåelse innen problemløsning og er kanskje det bruk av CAS i dette tilfellet representerer. Med CAS er det nok å skrive inn likningene, markere radene og trykke «Løs». Tilsvarende gjelder for oppgave 9c som er en kreativ oppgave fordi bokstaver i stedet for tall i likningssett ikke forekommer i bøkene. Denne er krevende å løse for hånd, men ikke mer komplisert enn løsningen av 9a dersom eleven bruker CAS. Den eneste eksplisitte ekstra kunnskapen som er viktig, er at det må være et tegn (mellomrom eller *) mellom bokstav og variabel ved bruk av CAS i GeoGebra.

Vi finner også variasjon i vanskegrad på oppgavene som krever kreativ løsning. Flervalgsoppgavene og oppgavene der det bare skal gis et svar, er enklere enn oppgavene der svaret må begrunnes i form av utregning eller annen forklaring. I del 2 føres alle oppgaver med utregning/forklaring, men også her er det variasjon i vanskegrad. For eksempel vil nok formelregningen i oppgave 6c være noe enklere enn forklaring, valg av framgangsmåte og utregning i oppgave 8.

Figur 5.1 Oppgave 9c i del 2 av eksamenen våren 2017 løst med CAS i GeoGebra.

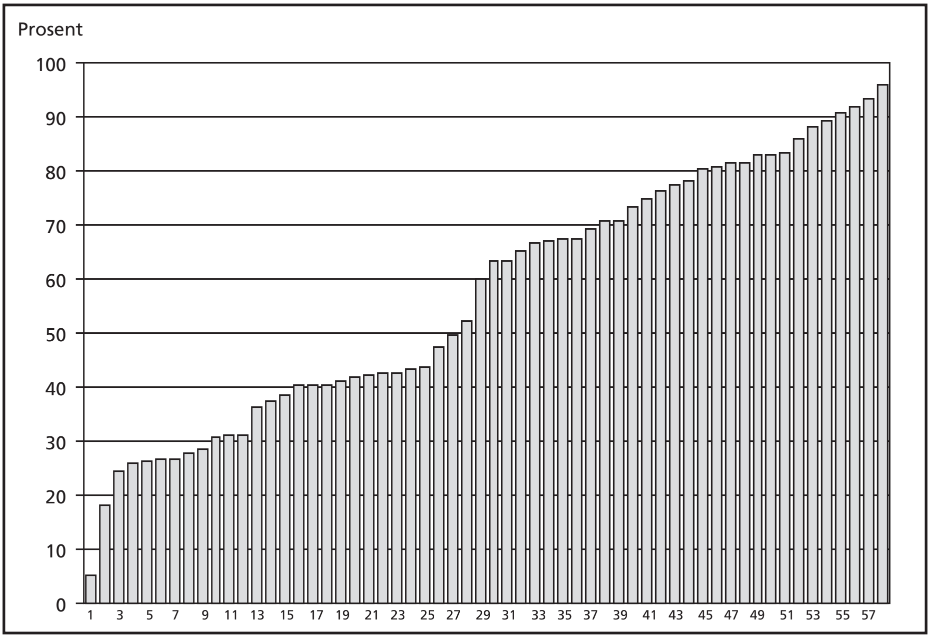


Vi har hatt tilgang til foreløpige datautskrifter fra IRT-analyser basert på innsamlede sensorskjemaer. Det er sterk sammenheng mellom oppgavenes vanskegrad ut fra IRT-analysene og vår kategorisering av oppgavene som algoritmiske eller kreative. For eksempel er alle de ti enkleste oppgavene i settet ut fra IRT-analysene enten rent algoritmiske eller delvis algoritmiske, mens kun tre av de ti vanskeligste oppgavene er enten rent algoritmiske eller delvis algoritmiske.

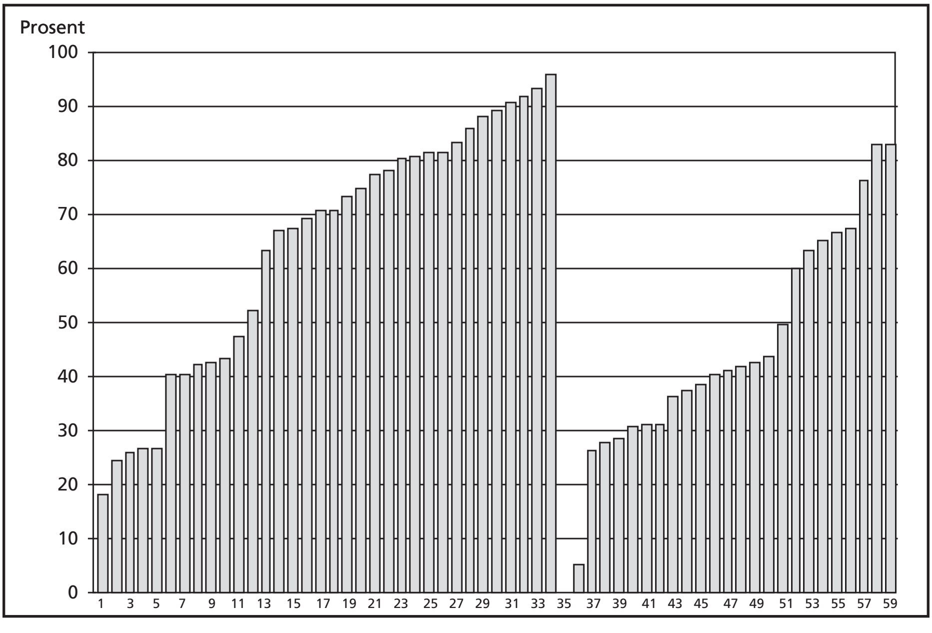
Vanskegrad ut fra vurderingsskjemaene

I tillegg til de kvalitative analysene kan vi støtte oss på analyser av sensorenes vurderingsskjemaer. Materialet er fra sensorene i Oslo og Akershus. På figur 5.2 er deloppgavene sortert ut fra hvor stor andel av mulige poeng som ble utdelt. Vi ser at det er en noenlunde jevn fordeling av oppgavene fra den nest laveste, hvor 24,4 prosent av mulige poeng ble utdelt, til den høyeste, hvor 95,9 prosent av mulige poeng ble utdelt. Det er imidlertid bare én oppgave som svært få klarer, nemlig oppgave 9c i del 2.

Figur 5.2 Andel av mulige poeng som ble gitt på deloppgavene ved avgangseksamen 2017, sortert i stigende rekkefølge.



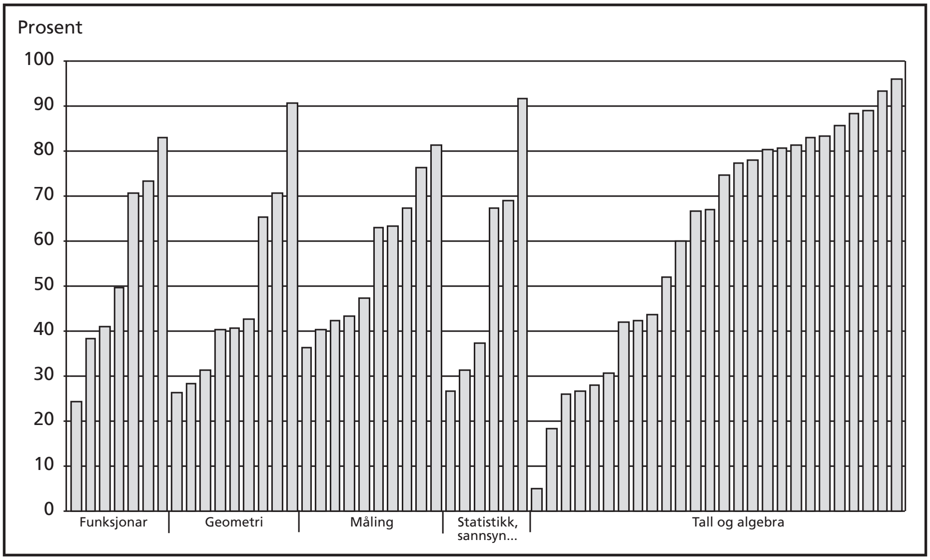
Figur 5.3 Andel poeng som ble gitt på deloppgavene ved avgangseksamen 2017, sortert i stigende rekkefølge, del 1 og del 2 hver for seg.



Ser vi på del 1 og del 2 hver for seg (figur 7.3), ser vi et liknende mønster, men vi ser at i del 2 er maksimal uttelling 83,1 prosent. Det er altså ingen oppgaver i del 2 som nesten alle elever får til.

Ser vi på hovedområdene i planen, blir bildet et litt annet (figur 5.4).[9] Det er ingen «vanskelige» oppgaver i måling – selv den vanskeligste får elevene 36,3 prosent uttelling på.[10] De vanskeligste oppgavene er konsentrert om hovedområdet tall og algebra. Alle hovedområdene har enkle oppgaver hvor uttellingen er på over 80 prosent, men geometri har bare én slik oppgave.

Figur 5.4 Andel poeng som ble gitt på deloppgavene ved avgangseksamen 2017, sortert i stigende rekkefølge innad i hvert hovedområde.



Ser man på alle oppgavene i et hovedområde under ett (se tabell 5.1), ser man at hovedområdene ikke har samme vanskegrad (målt i andelen av mulige poeng som deles ut). Mens tall og algebra er det «letteste» hovedområdet, er geometri det «vanskeligste».

Diagram 5.2, 5.3 og 5.4 er sammenliknet med tilsvarende diagrammer basert på innsamlede sensorskjemaer i forbindelse med IRT-analysene som er gjort for Utdanningsdirektoratet. De ser til forveksling like ut, noe som tyder på at det ikke påvirker konklusjonene her at vårt materiale kun er fra Oslo og Akershus.

9 Noen oppgaver kan falle innunder flere hovedområder, så det er brukt noe skjønn i fordelingen av oppgavene på hovedområder.

10 Oppgaven gir 32,2 prosent uttelling i materialet som er samlet inn fra hele landet i forbindelse med IRT-analysene.

Tabell 5.1 Poengene i eksamenssettet våren 2017 fordelt på hovedområdene og hvilken uttelling kandidatene i vårt materiale fikk på de ulike områdene.

	Antall poeng i settet	% uttelling
Tall og algebra	36,5	57,3 %
Geometri	14	47,2 %
Måling	14,5	54,5 %
Statistikk og sannsynlighet	9	51,1 %
Funksjoner	10	51,8 %

Forskjellen i uttelling mellom hovedområdene kommer tydeligst fram blant de svakeste elevene. Mens de elevene som lå an til å få karakteren 1, fikk 14,3 prosent uttelling på tall og algebra-oppgavene, fikk de kun 4,7 prosent uttelling på geometrioppgavene.

Elever og læreres oppfatning av vanskegraden på eksamenen

Som vi ser i avsnittene ovenfor, har eksamenen oppgaver med ulik vanskegrad. Allikevel finner vi variasjon mellom de ulike hovedområdene hvor geometri er området med færrest «enkle» oppgaver på årets eksamen. Et av målene med eksamen er at alle elever, uavhengig av kompetansenivå, skal gis en mulighet til å vise hva de kan. For at elevene skal ha mulighet til å vise sin kompetanse, må alle elever, på alle grader av karakterskalaen, ha oppgaver på eksamen som de klarer å løse. I de følgende avsnittene skal vi se hvordan matematikklærere og elever oppfattet vanskegraden på årets eksamen.

Blant de 36 elevene vi intervjuet, var det både elever med toppkarakterer som oppga matematikk som deres favorittfag, elever som skåret midt på treet, og elever som mislikte matematikk sterkt og strevde med å klare å få en ståkarakter. Felles for alle elevene var at de opplevde at de fikk «vist hva de kunne» på eksamen. Selv de «svakest» presterende elevene mente at de fikk til noen oppgaver. Det var gjennomgående i alle intervjuene, både med elever og lærere, at eksamenen var «god», og en lærer beskrev årets eksamen som en «eksamen for alle». Det at eksamenen var «god», handlet ikke om at alle elevene klarte alle oppgavene, men at de erfarte at gitt deres kompetansenivå var det oppgaver de klarte å løse. En lærer fortalte at de elevene han visste slet med matematikken, «hadde jobbefjes gjennom hele eksamen».

Det å mestre noen av oppgavene på eksamen ble beskrevet av lærere som avgjørende for at man ikke «ga opp». Måten denne eksamenen var strukturert på, hvor oppgavene startet enkelt på oppgave a, og vanskegraden økte fra b og utover, gjorde at elevene lettere holdt motivasjonen oppe gjennom hele eksamenen.

Et positivt element som ble trukket fram ved årets eksamen, var at hver oppgave hadde sitt eget tema. For eksempel handlet én oppgave om en sykkeltur og en annen om eksosutslipp og en tredje om broen London Bridge (se eksamensoppgaven i kapittel 1). Mestret ikke elevene en oppgave på årets eksamen, hadde de mulighet til å starte «på ny frisk» på neste oppgave. Lærerne fortalte at i tidligere eksamensoppgaver ble samme tema brukt over flere oppgaver. Da var det en fare for at elever som ikke forstod den første oppgaven, hoppet over alle oppgaver med samme tematikk, selv om oppgavene handlet om forskjellige matematiske utfordringer.

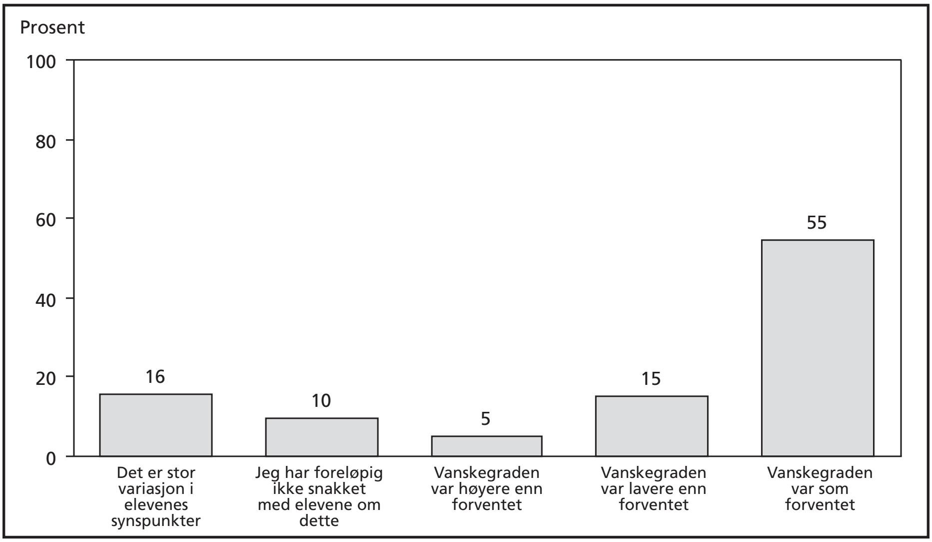
Årets eksamen hadde flere flervalgsoppgaver i del 1. Det ble oppfattet som et nytt grep av lærerne, og reaksjonene varierte. Som nevnt tidligere i kapitlet er flervalgsoppgaver enklere enn oppgavene der svaret må begrunnes i form av utregning eller annen forklaring. På den ene siden fortalte både elever og lærere at flervalgsoppgaver var motiverende, spesielt for elevene som strevde med matematikk. På den andre siden var ikke lærerne kun positive til flervalgsoppgavene, fordi de mente de «sterkeste» elevene mistet muligheten til å vise utregningene bak løsningsforslaget.

Enkelte av lærerne var også kritiske til poenggivningen på noen av oppgavene på eksamenen. I del 1 var det oppgaver lærerne mente burde gi større uttelling, og andre mindre. Et eksempel er en lærer som ikke kunne forstå at oppgave 1 i del 1 ga 2 poeng når oppgave 3 kun ga 1 poengs uttelling.

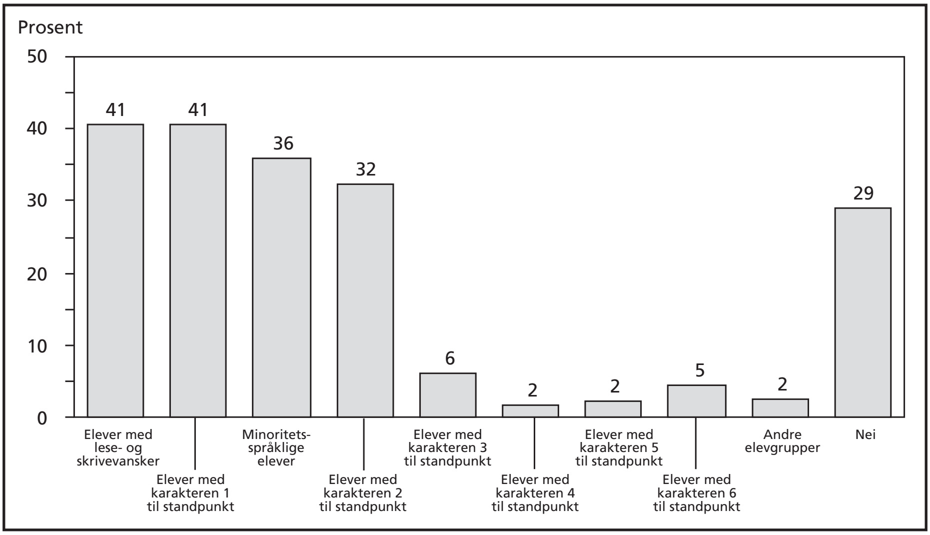
På alle de fire caseskolene ble oppgave 9 i del 2 raskt et tema i intervjuene med både lærerne og elevene. Spesielt oppgave 9c opplevde elevene som vanskelig og annerledes, og ingen av dem vi intervjuet, hadde brukt CAS for å løse oppgaven. Oppgaven ligger innenfor gjeldende kompetansemål i matematikk, men var ukjent for elevene siden de ikke hadde erfaring med den samme typen likningssett fra undervisningen. Flere lærere var enige med elevene i at det var en unødvendig oppgave, men det ble også poengtert at oppgaven var en oppgave for å kunne sile ut sekserelevene.

I spørreskjemaet til matematikklærerne spurte vi om hva elevene hadde sagt til lærerne om vanskegraden på årets eksamen. Det var 770 matematikklærere som svarte på dette spørsmålet. Over halvparten av lærerne svarte at elevene hadde sagt at vanskegraden på årets eksamen var som forventet, 15 prosent har fått høre at vanskegraden var lavere enn forventet, og 5 prosent at elevene har sagt at vanskegraden var høyere enn forventet. Resultatene er presentert i figur 5.5.

Figur 5.5 Matematikklæreres svar på påstander om vanskegraden på årets eksamen. Prosent.



Figur 5.6 Elevgrupper som ikke får vist fram sin kompetanse på eksamen. Prosent.



Vi stilte også et spørsmål om det var noen bestemte elevgrupper som ikke fikk vist sin kompetanse på eksamen. Svarene er presentert i figur 5.6. De elevgruppene lærerne i størst grad mener ikke får vist sin kompetanse på eksamen, er elever med lese- og skrivevansker og minoritetsspråklige elever. Det er følgelig de elevene som strever med tekstoppgaver og språk, som sliter mest med å få vist sine matematikkferdigheter på eksamen. Det er en tredjedel av lærerne som mente at de svakest presterende elevene i matematikk, de med 1 eller 2 som standpunktkarakter, ikke fikk vist sin kompetanse på eksamen. Svarene fra spørreskjemaet står i kontrast til svarene i de kvalitative intervjuene hvor lærerne mente at årets eksamen nettopp ga de svakest presterende elevene en mulighet til å få vist fram sin kompetanse.

Oppsummering

- Det er god variasjon i vanskegrad eksamenssettet sett under ett. Imidlertid er variasjonen i vanskegrad langt mindre innen enkelte hovedområder, som geometri.
- I del 2 er det ingen oppgaver som «nesten alle» elevene får til.
- Litt over halvparten av lærerne mente vanskegraden på eksamenen var som forventet.
- I de kvalitative intervjuene var tilbakemeldingen at vanskegraden på eksamenen var der den bør være. Lærerne var av den oppfatning at eksamenen var «en eksamen for alle».
- De elevene som lærerne i størst grad mente at ikke fikk vist sin kompetanse på eksamen, er minoritetsspråklige elever, elever med lese- og skrivevansker og elever med karakteren 1 eller 2 i standpunkt.

6 Eksamenens arbeidsmengde

I dette kapitlet er temaet eksamenens arbeidsmengde. Vi vurderer arbeidsmengden ved å sammenlikne med tidligere eksamener, ved å se på hvordan ubesvarte oppgaver fordeles i settet, og ved å innhente elevers og læreres synspunkter.

Vurderingen av arbeidsmengde på eksamen er krevende. Innledningsvis kan vi slå fast at eksamenen i 2017 er sammenliknbar med eksamenene de siste årene når det gjelder antall oppgaver, antall deloppgaver og mengden tekst elevene må sette seg inn i. Det er imidlertid noe flere flervalgsoppgaver, noe som kan være mindre arbeidskrevende enn oppgaver hvor man selv må finne svarene.

En annen tilnærming er å analysere hvilke oppgaver elevene har besvart, siden stor arbeidsmengde gjerne vil medføre at mange oppgaver blir stående ubesvart. For denne analysen har vi benyttet innsamlede vurderingsskjemaer fra sensorer, hvor sensorene skulle la poengrutene stå tomme hvis en deloppgave var ubesvart. Vi ser i materialet at ikke alle sensorer har vært konsekvente her, og begrenser derfor denne analysen til de 902 kandidatene hvor vi har to sensorer som har markert de samme oppgavene som ikke besvart.

Det er 266 elever (29,5 prosent) som har besvart alle deloppgavene i settet. På del 1 er det 604 elever (67,0 prosent) og på del 2 er det 274 elever (30,4 prosent) som har svart på alle deloppgavene. I figur 6.1 framgår andelen blanke besvarelser på de ulike deloppgavene i settet.[11]

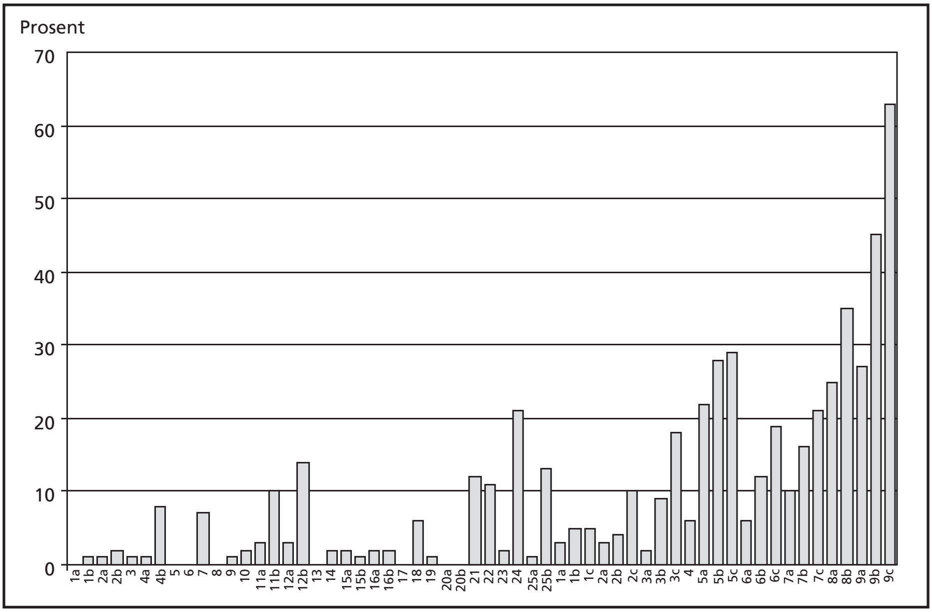
Vi ser at andelen blanke besvarelser er lav for del 1 og høyere for del 2. Vi ser også at andelen blanke besvarelser øker mot slutten av settet. Imidlertid ser vi også at de aller fleste elever (90 prosent) prøver seg på for eksempel oppgave 7a i del 2, noe som tilsier at de har hatt tid til å se på oppgavene og vurdere hvilke de ville prioritere å gjøre. Tilsvarende ser vi i del 1 at 99 prosent av elevene har prøvd seg på oppgave 25a.

Det er vanskelig å skille mellom vanskegrad og arbeidsmengde: At en elev har latt være å svare på en oppgave, kan skyldes at oppgaven oppfattes som vanskelig, eller at eleven ikke har hatt tid til å se på den. En analyse av besvarelsene kan ikke fortelle oss om det er vanskegrad eller arbeidsmengde som er årsaken. Vi vil forvente at stor arbeidsmengde spesielt vil gi seg utslag på de siste oppgavene i settet, men det er samtidig ofte

[11] Utdanningsdirektoratet har samlet inn materiale fra noen sensorer i forbindelse med sin IRT-analyse. I dette materialet er andelen blanke noe høyere enn det vi ser basert på vurderingsskjema fra Oslo og Akershus, men bildet av hvilke oppgaver som i størst grad besvares, og hvilke som står ubesvart, er det samme.

der vi finner de vanskeligste oppgavene – og også elevene forventer at oppgavene sist i settet skal være vanskelige. Når 63 prosent av elevene i vårt materiale lar oppgave 9c i del 2 stå ubesvart, skyldes det formodentlig hovedsakelig at den er vanskelig, samtidig som det kan være elever som ville ha fått den til hvis de hadde hatt mer tid til rådighet. Vi finner ikke eksempler på enkle oppgaver med uforklarlig lav svarprosent.

Figur 6.1 Andelen blanke besvarelser på deloppgavene ved avgangseksamen for 10. trinn 2017. I eksamenssettets rekkefølge.



Det er for øvrig en signifikant (på 99-prosentnivå) forskjell på jenter og gutter når det gjelder det å la oppgaver være blanke: Jenter har i snitt 4,4 blanke svar, mens gutter i snitt har 6,6 blanke svar. Det er signifikante (på 99-prosentnivå) forskjeller mellom gutter og jenter på deloppgavene 7, 12b, 18, 21, 24 og 25b i del 1 og deloppgavene 6b, 6c, 7b, 7c, 8b, 9a, 9b og 9c i del 2. I alle disse deloppgavene er det guttene som svarer blankt i større grad. Men siden vi vet at jenter gjorde det bedre enn gutter på denne eksamenen (også i Oslo og Akershus), er det vanskelig å komme til noen klar konklusjon ut fra dette om hva forskjellen mellom gutter og jenter skyldes.

Hva synes elevene og lærerne?

En siste tilnærming til å se på arbeidsmengden på eksamenen er å se hva elevene selv synes om arbeidsmengden, og hva lærerne oppgir at de har snakket med elevene om. Lærerne på de fire skolene fortalte at forberedelsen til eksamen bestod av å gjøre tidligere eksamensoppgaver. Et viktig aspekt ved å benytte seg av tidligere eksamensoppgaver i undervisningen var å lære elevene å disponere tiden. En av elevene fortalte at i oppkjøringen til eksamen hadde hun brukt stoppeklokke når hun regnet gjennom tidligere eksamensoppgaver, for å øve på hvor lang tid hun kunne bruke på ulike typer oppgaver. Elevene vi intervjuet, fortalte at de hadde blitt opplært til å gå videre til neste oppgave hvis det var noe de ikke forstod eller opplevde som vanskelig. Hadde de tid til overs på slutten av eksamen, kunne de gå tilbake til de vanskelige oppgavene og prøve å løse dem da. Elevene framhevet at det å ha tilstrekkelig tid var avgjørende for å ha mulighet til å bruke den siste timen på slutten av eksamen til å se over svarene og gruble på de vanskeligste oppgavene.

En av caseskolene hadde et høyt antall minoritetsspråklige elever, og mange strevde med å kunne lese og forstå norsk godt nok. De brukte derfor tid på å forstå tekstoppgaver i matematikk. Allikevel mente lærerne på denne skolen at mer tid på eksamen ikke ville hjelpe elever med dårlig norskforståelse, fordi hovedproblemet var at de ikke forstod begrepene i oppgaven. Mer tid ville ikke gi større språkforståelse, men som vi viser i kapitlet om tekstoppgaver, var det å ha et enkelt, kortfattet språk på eksamen avgjørende for at minoritetsspråklige, og elever med lese- og skrivevansker, skulle klare å vise sine matematiske ferdigheter.

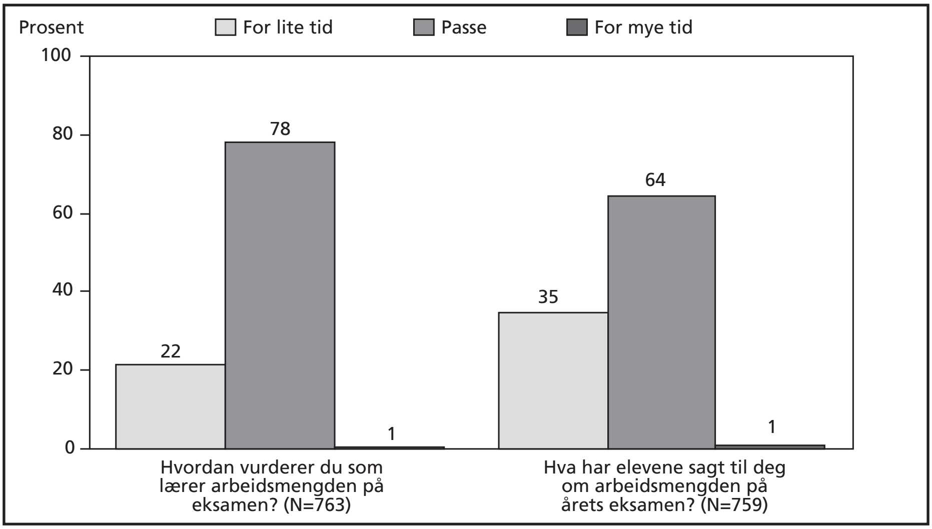
Elevene og lærerne som ble intervjuet på de fire skolene, var alle godt fornøyde med forholdet mellom arbeidsmengden og tid på eksamenen. Lærerne var av den oppfatning at del 1 på årets eksamen hadde flere oppgaver enn tidligere eksamensoppgaver, men siden mange av oppgavene var flervalgsoppgaver, gjorde det at hver enkelt oppgave tok kortere tid å løse. I spørreundersøkelsen som ble sendt ut til matematikklærere over hele landet, ble dette inntrykket problematisert. Et av spørsmålene i spørreskjemaet var: «Hvis elever kommer til deg og klager på at eksamen er urettferdig, hvordan begrunner de det?» Spørsmålet hadde en åpen svarkategori. Flere lærere rapporterte at de hadde fått klager fra elever om at de hadde for dårlig tid til å komme igjennom alle eksamensoppgavene.

I spørreundersøkelsen ble lærerne spurt om hvordan de som lærere vurderte arbeidsmengden på eksamen. Vi spurte også om hva elevene hadde sagt til dem om arbeidsmengden på årets eksamen. Resultatene er presentert i figur 6.2.

78 prosent av lærerne mente at arbeidsmengden på årets eksamen var passe der hvor 22 prosent svarte at det var for lite tid. På spørsmålet om hva elevene hadde sagt til dem om arbeidsmengden, svarte over halvparten av lærerne, 64 prosent, at elever

hadde sagt at arbeidsmengden var passe, mens 34 prosent av lærerne oppga at elever mente at det var for lite tid på årets eksamen.

Figur 6.2 Læreres svar på spørsmål om arbeidsmengden på eksamenen. Prosent.



Det er som nevnt tidligere utfordrende å skille mellom vanskegrad og arbeidsmengde siden utfordrende oppgaver matematisk sett vil ta lengre tid å løse. Imidlertid er hovedinntrykket, både fra de kvalitative intervjuene og spørreskjemaet, at arbeidsmengden på årets eksamen var som forventet, og at forholdet mellom tid og arbeidsmengde var godt.

Oppsummering

- Andelen ubesvarte oppgaver øker mot slutten av settet. Dette kan skyldes at elevene fikk dårlig tid, eller at oppgavene mot slutten er vanskeligere.
- Lærerne fortalte at en stor del av elevenes forberedelse til eksamen bestod av å gjøre tidligere eksamensoppgaver. Et viktig aspekt ved å benytte seg av tidligere eksamensoppgaver i undervisningen var å lære elevene å disponere tiden.
- Nesten åtte av ti lærere mente at arbeidsmengden på eksamenen var passe, og et flertall av lærerne oppga at elevene hadde sagt at det var passe arbeidsmengde på eksamenen. På samme tid opplyste 35 prosent av lærerne at elever hadde sagt det var for dårlig tid til å gjennomføre eksamenen.

7 Bruk av tekstoppgaver i matematikkeksamen

I dette kapitlet vil vi diskutere bruken av tekstoppgaver i årets matematikkeksamen. Først vil vi presentere en gjennomgang av tidligere forskning på språk, for så å undersøke eksamensoppgavene i lys av tidligere forskning. Til slutt skal vi se hvordan den språklige utformingen av tekstoppgavene oppfattes av lærere og elever.

Lesing er en grunnleggende ferdighet i matematikkfaget, og i læreplanen (Utdanningsdirektoratet 2013) står det:

> «Å kunne lese i matematikk inneber å forstå og bruke symbolspråk og uttrykksformer for å skape meining i tekstar frå daglegliv og yrkesliv så vel som matematikkfaglege tekstar. Matematikkfaget er prega av samansette tekstar som inneheld matematiske uttrykk, grafar, diagram, tabellar, symbol, formlar og logiske resonnement. Lesing i matematikk inneber å sortere informasjon, analysere og vurdere form og innhald og samanfatte informasjon frå ulike element i tekstar. Utvikling i å lese i matematikk går frå å finne og bruke informasjon i tekstar med enkelt symbolspråk til å finne meining og reflektere over komplekse fagtekstar med avansert symbolspråk og omgrepsbruk.»

Det er derfor naturlig og riktig at en eksamen tester denne grunnleggende ferdigheten. Samtidig er det en balansegang, siden også elever med svake språklige ferdigheter bør få vist sin øvrige matematiske kompetanse. Denne balansegangen er en sentral bakgrunn for vurderingene i dette kapitlet.

Litteraturgjennomgang: forståelse og språk

Bakgrunn

I matematikk er det en klar sammenheng mellom elevers språklige ferdigheter på prøvespråket og deres resultat på prøven (Ercikan et al. 2015; Prediger, Wilhelm, Büchter, Gürsoy & Benholz 2015; Townsend, Filippini, Collins & Biancarosa 2012). Dette kan delvis skyldes at elever med svake språklige ferdigheter har hatt dårligere utbytte av undervisningen i stoffet som prøven tester (Kieffer, Lesaux, Rivera & Francis 2009).

Men det finnes mye forskning som viser at språklig kompliserte oppgaver kan være uforholdsmessig vanskelige for elever med svake språkferdigheter på prøvespråket, og i særdeleshet for elever som tar prøver på et annet språk enn sitt morsmål (Abedi et al. 2005; Haag, Heppt, Stanat, Kuhl & Pant 2013; Wolf & Leon 2009). For eksempel viser en analyse av 4.-trinnselevers svar på TIMSS- og PIRLS-oppgaver at svakere lesere skårer dårligere på oppgaver som er tyngre å lese – dette gjelder også for elever i Norge (Martin & Mullis 2011). «Mens de svakeste leserne gjennomgående presterte på et lavere nivå i matematikk enn de beste leserne, var de ytterligere negativt påvirket på matematikkoppgavene som krevde mer lesing.» (Martin & Mullis 2011, s. 106, vår oversettelse) Dette kan bety at disse elevene i mindre grad får vist sin matematikkferdighet på eksamen enn andre elever. Å undersøke i hvilken grad eksamensoppgaver inneholder unødvendige språklige hindringer, er derfor en naturlig del av en evaluering av eksamensoppgavene.

Da Matematikksenteret i 2015 evaluerte matematikkeksamen for grunnskolen for 2009–2014 og to eksamener i videregående skole, så de også på om språket var «tilpasset elevgruppa på en slik måte at oppgavene oppfattes som entydig definert» (Matematikksenteret 2015a, s. 4). Hver oppgave ble gitt en skår, og for grunnskolens eksamen var konklusjonen at «Resultatene våre viser at med unntak av 2010 er presisjonen i språket lavere enn ønsket» (s. 17). I et eget vedlegg (Matematikksenteret 2015b) presiserte de vurderingskriteriene på dette punktet: 1) virkelighetsnær kontekst, 2) kontekst tilpasset målgruppen, 3) språk tilpasset målgruppen, 4) konsekvent begrepsbruk, 5) korrekt bruk av *tid*, 6) all informasjon i oppgaveteksten, 7) presise formuleringer, 8) spørsmålet, 9) unødvendig støy i teksten, 10) forkortelser for enheter, 11) illustrasjoner, 12) annet. Til tross for dette vedlegget framstår det ikke som transparent hvordan konklusjonene er framkommet, og det gjør heller ikke det faglige grunnlaget for å velge disse kriteriene.

I denne rapporten velger vi derfor å lage en egen kunnskapsoppsummering som grunnlag for våre analyser. Formålet med kunnskapsoppsummeringen er å få fram bredden av språktrekk som forskningen mener er relevante for forståelsen av matematikkoppgaver, og få grunnlag for å velge hvilke trekk ved eksamensoppgavene vi vil se på. Det finnes lite forskning på språket i norske eksamensoppgaver i matematikk.[12] I litteraturgjennomgangen har vi derfor også sett på forskning som gjelder andre språk, andre situasjoner enn eksamen og/eller andre fag enn matematikk. I sum vil dette gi en kunnskapsbase som er relevant for dette prosjektet.

[12] Det finnes generelt lite forskning på språk og matematikkprøver, selv om situasjonen har bedret seg noe siden 2008, da Nyström skrev (sitert i Persson 2016a, s. 40): «Systematisk og forskningsbasert kunnskap om teksttrekk som påvirker lesbarhet i matematikk- og naturfagoppgaver er sjelden i forskningslitteraturen. Til tross for det aksepterte faktum at språkferdigheter, og spesielt leseferdighet, er et viktig aspekt av alle vurderingssituasjoner, har overraskende få studier handlet om fagspesifikk diskurs i skriftlige tester.» (Nyström 2008, s. 3, vår oversettelse)

For litteraturgjennomgangen er det brukt klassisk snøballtaktikk, hvor litteratursøk i databaser på sentrale emneord på ulike språk har gitt et utgangsmateriale av tekster, hvoretter utvalgte referanser i disse og til disse tekstene er blitt brukt for å utvide materialet. På grunn av begrensede språkferdigheter er litteratursøket begrenset til skandinaviske språk samt engelsk og tysk. Et viktig utgangspunkt er den norske forskningen på lesing av matematikktekster fra Høgskolen i Vestfold (Maagerø & Skjelbred 2010; Maagerø & Tønnessen 2015; Skjelbred & Aamotsbakken 2010). Nedenfor er resultatene av litteraturgjennomgangen strukturert ut fra de ulike språktrekkene som trekkes fram i litteraturen, og ikke for eksempel ut fra hva vi vet om ulike elevgrupper. Språktrekk inndeles gjerne i trekk på ordnivå, trekk på setningsnivå og trekk på tekstnivå. Trekk på tekstnivå er mindre relevante i eksamensoppgaver (som ofte er på bare én–tre setninger) enn i lengre tekster, men det kan i visse tilfeller likevel være relevant og behandles kort her.

Overordnet oppsummering av litteraturgjennomgangen

Litteraturgjennomgangen som følger, viser en stor bredde i språklige trekk som kan anses relevante, og også en stor bredde i måter disse språklige trekkene konkret undersøkes. Imidlertid er det få klare konklusjoner om hvilke elever som «rammes» av de ulike språklige trekkene – det varierer fra undersøkelse til undersøkelse hvilke språklige trekk som gir størst effekt. For eksempel skriver Martiniello (2008, s. 336) at «Selv om mange av disse studiene, som forventet, fant en kobling mellom språklig kompleksitet og ELLs prestasjoner på matematikktekstoppgaver, varierer effekten av de spesifikke språklige trekkene fra test til test og fra klassetrinn til klassetrinn» (vår oversettelse). En mulig tolkning av dette er at bildet er komplisert, av minst fire grunner: For det første må vi anta at ulike språktrekk spiller sammen. For det andre må vi anta at ulike språktrekk har ulik effekt for ulike (grupper av) barn. For det tredje vil de ha ulik effekt innen ulike fag/temaer. Og for det fjerde er kontekstene som undersøkes, ulike – noen undersøkelser ser på oppgaver som har vært gjennom en nitid pretestingsprosess (f.eks. Persson 2016a; Shaftel, Belton-Kocher, Glasnapp & Poggio 2006), mens andre ser på oppgaver som ikke har vært pretestet.

Om en matematikktekst er enkel eller vanskelig å lese, avhenger i stor grad av eleven – for eksempel barnets leseferdighet og barnets fritidsinteresser (Maagerø & Skjelbred 2010, s. 131). Mye av forskningen som behandles nedenfor, ser spesielt på elever som ikke har prøvespråket som førstespråk. Analysen av eksamensoppgavenes språk skal se på om oppgavene er forståelige for alle elevene, og da er dette en viktig gruppe å ha i tankene. Samtidig er det viktig å være klar over at gruppen elever som ikke har prøvespråket som førstespråk, i seg selv er en svært uensartet gruppe, fra elever som har lite tidligere skolegang, til elever som behersker norsk flytende. Siden vi i dette prosjektet skal vurdere matematikkoppgavers vanskegrad uten å ha direkte tilgang til mer enn noen få av de mange tusen elevene som skal ha eksamen, må vi i stor grad rette søkelyset mot forhold ved teksten. Samtidig må vi se dette opp mot det vi vet om elevmassen som helhet, og det vi måtte vite om enkelte elevgruppers spesielle utfordringer. For de enkelte språktrekk vil det variere hvilke elevgrupper forskningen er knyttet til, og dette er forsøkt tydeliggjort i teksten.

Generelt om akademisk språk i matematikk

For å lykkes i skolen er det viktig å beherske akademisk språk:

> Måten akademiske tekster er strukturert, og de former for meninger de skaper, er forskjellige fra de former for språkstrukturering og meningsskaping som allerede er kjent for de fleste barn i deres hverdagsliv. Derfor påvirker evnen til å håndtere akademiske tekster direkte deres læringsopplevelse, bruk av skole, tilgang til utdanning på høyere nivåer og hvilke muligheter de har utenfor skolen. (Fang, Schleppegrell & Cox 2006, s. 250, også sitert i Persson 2016a, s. 34, vår oversettelse)

Akademisk språk er forskjellig fra fag til fag. Matematisk akademisk språk har sine egne særegenheter, og mye fagdidaktisk forskning handler for tiden om hvordan noen elever kommer inn i fagdiskursen mens andre ikke gjør det (Shanahan & Shanahan 2012). Mye av forskningen vi finner, handler om elever som ikke har prøvespråket som sitt morsmål, men også for andre grupper er språket et hinder for forståelse, for eksempel elever med lav sosioøkonomisk status (f.eks. Heppt, Haag, Böhme & Stanat 2015).

Ordnivået / leksikalsk nivå

Maagerø og Skjelbred (2010, s. 133–134) peker på en del aspekter på ordnivå som kan gjøre tekster vanskelige å lese, her supplert med andre aspekter fra litteraturgjennomgangen:

Lange ord, for eksempel definert som ord med mer enn seks bokstaver, kan vanskeliggjøre lesing, og antall lange ord er derfor en faktor i mange undersøkelser (f.eks. Persson 2016a). Lepik (1990) påpeker imidlertid at i matematikkoppgaver er lange ord ofte matematikkord som «fungerte som verbale hint til løsningen og reduserte på den måten løsningstiden» (s. 89, vår oversettelse). Selv om det kan være problematisk med mange lange ord, kan faktisk enkelte lange ord bidra til å gjøre oppgavene enklere.

Fagterminologi kan gjøre matematikkfaglige tekster vanskelige å lese (Gürsoy et al. 2013, s. 11; Maagerø & Skjelbred 2010, s. 75–82). Imidlertid tyder undersøkelser på at fagterminologi som elevene har møtt i opplæringen, ikke bidrar vesentlig til redusert lesbarhet (Haag et al. 2013). Fagterminologien kan variere mellom ulike deler av matematikkfaget, og det kan av den grunn være enkelte språktrekk som er spesielt viktige i enkelte deler av faget. For eksempel har Niederhaus et al. (2015) undersøkt språkbruk knyttet til prosentbegrepet, og de viser at det finnes en lang rekke ord som brukes for å uttrykke sammenhenger som angår prosent (for eksempel «prosentsats», «helheten», «den opprinnelige prisen», «fettinnhold», «øker med»). Tilsvarende viser Prediger (2013, s. 7–8) at det finnes svært mange måter å formulere funksjonssammenhenger på. For elever som tar eksamen, vil det da være en utfordring å kjenne igjen de ulike måtene å formulere seg på, især dersom disse ikke er kjent fra opplæringen.

I tillegg til fagterminologi finnes det en del «ord og uttrykk som er typiske for fagtekster, men som ikke tilhører ett bestemt fag» (Maagerø & Skjelbred 2010, s. 11). Maagerø og Skjelbred (2010) kaller disse «de ikke-faglige abstrakte ordene», men vi vil følge internasjonal litteratur og kalle dem «**generelle akademiske ord**». Slike ord (f.eks. «i sammenheng med») viser seg å kunne være utfordrende for elever å forstå (Gürsoy et al. 2013, s. 11; Maagerø & Skjelbred 2010, s. 84–86). Dette henger sammen med at slike ord sjeldnere blir forklart i matematikkundervisningen, i motsetning til rent matematisk vokabular (Haag et al. 2013, s. 25). Townsend et al. (2012) fant signifikant sammenheng mellom elevers forståelse av generelle akademiske ord og deres akademiske prestasjoner. Wolf og Leon (2009) fant signifikant sammenheng mellom antall akademiske ord i en oppgave og hvor svakt ELL (English Language Learners) skåret i forhold til andre elever. Spesielt var dette tydelig i oppgaver som var relativt enkle faglig sett, men som inneholdt generelt akademisk språk. Der skåret ELL-elever uforholdsmessig dårlig.

Andre **ikke-matematiske ord som er ukjente** for eleven, kan også skape problemer (Gürsoy et al. 2013, s. 6). Flere forskere bruker ulike korpus for å få en indikasjon på hvilke ord som kan være ukjente (f.eks. Persson 2016a), mens andre bruker intervjuer med elever (f.eks. Prediger 2013). Laufer (1997) skiller mellom to grupper ukjente ord: ord elevene vet de ikke kan, og ord elevene tror de kan. I denne siste kategorien kommer ord som elevene tror de forstår ut fra hva de er satt sammen av («shortcomings» tolket som «short» + «comings»), faste uttrykk («sit on the fence» tolket bokstavelig), falske venner (tolke ut fra et ord som likner på førstespråket, som «rar» på norsk og dansk), ord med flere betydninger og ord som likner på ord elevene kan. Siden leseforståelse ofte bygger på å gjette på betydningen av ukjente ord ut fra de ordene du forstår, kan ord du tror du forstår, gi ringvirkninger til andre ord (Laufer 1997, s. 27). I en analyse av eksamensoppgavene vil det være mulig å påpeke lavfrekvente ord, men umulig å gi noen fullstendig analyse av hvilke ord som kan være ukjente for ulike grupper av elever.

Sammensatte ord kan være krevende å tolke, for ordene er ofte lange, og delene kan ha ulike roller i ulike ord (Gürsoy et al. 2013, s. 11; Maagerø & Skjelbred 2010, s. 82–84). En isselger selger is, men en telefonselger bedriver salg via telefon. Prediger (2013, s. 2) bruker ordet «Zuschauer|schnitt» (tilskuersnitt) som eksempel som er vanskelig for elever å forstå – og hvor det lengre «gjennomsnitt av tilskuertallene» kunne vært enklere å gjennomskue.

Et særtrekk ved akademisk språkbruk er at verb ofte gjøres om til substantiver – handlinger gjøres om til ting. Dette kalles **nominaliseringer**. Bruk av nominaliseringer innebærer at vi i neste omgang kan diskutere «rotasjonen», ikke bare snakke om «å rotere» (Maagerø & Skjelbred 2010, s. 78–92). Dernest kan man snakke om rotasjoner og speilinger som matematiske objekter som sammenliknes. Persson (2016a, s. 36) viser til Fang et al. (2006) og skriver at «Nominaliseringer har dog ulempen at de ikke bare medfører abstraksjoner, men også medfører en viss tvil og usikker i teksten, hvilket kan senke forståeligheten av teksten veldig og da også interessen og engasjementet hos leseren.» (vår oversettelse). Nominalisering bidrar også til at det kan bli uklart hvem som er aktøren (se nedenfor).[13]

I tillegg pekes det i litteraturen på at **metaforer**, «**packing**» og meningsbærende **preposisjoner** kan være krevende for noen elever. Askeland i Maagerø og Tønnessen (2015, s. 115–117) peker på at særlig sjeldne og svake metaforer kan være vanskelige å forstå. Persson, af Geijerstam og Liberg (2016, s. 182) bruker antall substantiver (bortsett fra «proper names» og «descriptive names») og antall lange ord[14] – begge deler dividert med antall ord i deloppgaven – som mål på «packing», som de viser at kan gjøre teksten vanskeligere å forstå for enkelte elever (Persson et al. 2016, s. 183). Gürsoy et al. (2013) peker på at det i matematikk kan være mye «bruk av preposisjoner som strukturord med særlig relevans for løsning av oppgaven.» (s. 7, vår oversettelse). Dermed blir oppgaven vanskelig å forstå hvis man ikke forstår alle preposisjonene. Eksempel: «Med hvor mange prosent ligger forbruket ved 180 km/t over forbruket ved 100 km/t?» (s. 8, vår oversettelse)[15]. De minner samtidig om at preposisjoner ofte overses av elever, og siterer Jorgensen på at «Preposisjoner er de små ordene som ofte ignoreres av lesere, men som har en betydelig verdi i matematikk» (Jorgensen 2011, sitert i Niederhaus et al. 2015, s. 140, vår oversettelse).

Setningsnivået / grammatikalsk nivå

Vi har nå gått gjennom mange språklige trekk på ordnivå som alle kan påvirke elevers forståelse negativt. Problemene som oppstår på setningsnivå, kan imidlertid ofte være vel så store. Persson (2016a, s. 35) viser til Halliday (1993) som mener at det snarere er grammatikken enn vokabularet som er problemet (i naturvitenskap, riktignok).

[13] Persson (2016a, s. 36) har følgende morsomme selvrefererende eksempel: «Skrivingen av denne setningen er et forsøk på eksemplifisering av naturvitenskapsspråkets unnvikelse av tekniske termer i form av verb.» – en setning som til tross for at både skriving og eksemplifisering foregår, bare inneholder ett eneste finitt verb: det kursiverte er.

[14] Definert som ord med mer enn seks bokstaver.

[15] Se også Prediger, Renk, Büchter, Gürsoy og Benholz (2013, s. 53).

Problemene som oppstår, er ikke ordene selv, men «de komplekse relasjonene de har med hverandre» (vår oversettelse).

Maagerø og Skjelbred (2010, s. 134–144) peker på en del aspekter på setningsnivå som kan gjøre tekster vanskelige å lese. Blant disse er uvant ordstilling og lange, komplekse ytringer (med mange komplekse ledd og underordnede leddsetninger). Martiniello (2008, s. 354) nevner som eksempel hvordan en elev har problem med setningen «Eksakt ¾ av klinkekulene i bagen er blå» (vår oversettelse), og kobler ordet «blå» til «bagen», ikke til «klinkekuler».

Pakking av informasjon i substantivgrupper handler om at ord kan kobles til substantiver både i forkant og etterkant for å gi mer informasjon (Maagerø & Skjelbred 2010, s. 93–96). Fang et al. (2006, s. 253) gir eksemplene «the alligator in the swamp» og «those two very ferocious Florida alligators» som to substantivgrupper med utgangspunkt i substantivet «alligator». Slike substantivgrupper fungerer i setningen som et substantiv, og setninger kan godt bestå av to slike substantivgrupper med et verb som binder dem sammen. Dette kan da gi en veldig stor informasjonstetthet som kan være vanskelig å avkode.

Maagerø og Skjelbred (2010, s. 144–147) peker videre på at realfaglige tekster ofte framstår som **uten aktører.** Dette skjer blant annet ved bruk av nominalisering (der man skriver «gassblanding», i stedet for «vi blander gassen») og bruk av passivformer. Når teksten framstår som uten aktører, kan teksten virke abstrakt og fremmedgjørende på elevene. Abedi og Lord (2001, s. 221) peker også på at bruk av betingede setninger kan komplisere språket unødvendig. «Hvis to av batteriene var tomme for strøm [...]» er mer komplisert enn «to av batteriene var tomme for strøm».

Tekstnivået

Maagerø og Skjelbred (2010, s 140–144) legger også vekt på at det kan vanskeliggjøre forståelsen når det er mangel på sammenheng mellom setningene. Slik sammenheng skapes ofte av ord som «fordi», «altså» osv. Det er også vanlig å bruke pronomener (for eksempel «den»), og det er da viktig at disse viser entydig til noe som er beskrevet i tidligere setninger.

I tillegg er det viktig at det er sammenheng mellom tekst og illustrasjoner og paratekstlige elementer, se kapittel 3. De ulike elementene kan komplementere hverandre, gjenta den samme informasjonen på ulike måter (redundans) eller motsi hverandre. For svake lesere kan det være en fordel at den samme informasjonen finnes på flere måter.

En opplagt faktor å se på er antall ord en oppgave består av. Det vil ofte være slik at jo flere ord en oppgave har, desto mer er det å forstå. Derfor er det vanlig å ha med antall ord i oppgaven i analyser av lesbarhet (f.eks. Martin & Mullis 2011). I en del mål på lesbarhet inngår isteden lange setninger som et element. Gürsoy, Benholz, Renk, Prediger og Büchter (2013, s. 9) peker imidlertid på at ingen av forståelsesproblemene i deres intervjuer med elever kunne knyttes til setningslengde alene.[16]

Samspill mellom faktorer

Vi har vist til en mengde forhold ved teksten som kan gjøre matematikkoppgaver vanskelige å forstå for enkelte elever. Det ideelle ville vært om vi også kunne gi en oversikt over hvor utslagsgivende de enkelte språktrekkene er. Dette viser seg å være vanskelig. For eksempel viser Haag et al. (2013) at selv om tekstlengde, generelt akademisk vokabular, antall substantivfraser og antall preposisjonsuttrykk er forklaringsvariabler for at noen elever skårer dårligere på oppgaver enn andre, så kan en betydelig del av variasjonen mellom elevene forklares av en kombinasjon av slike språklige faktorer. De drøfter også en del annen forskning på feltet og konkluderer slik: «Det er derfor uklart hvilke trekk ved akademisk språk eller kombinasjoner av trekk som mest sannsynlig fører til 'differential validity' i tekstoppgaver for andrespråkselever» (ibid., s. 26, vår oversettelse). Tilsvarende sier Martiniello (2009, s. 164, vår oversettelse) at «forskning som bruker ulike nivåer av sammensetning av språkkompleksitetsindikatorer, tyder på at deres effekt er signifikant når de ses i sammenheng i en samlet vurdering av helhetlig språkkompleksitet.» Dette leser vi slik at vi ikke kan se oss blinde på enkeltfaktorer, men være spesielt på vakt for en opphoping av slike faktorer i enkeltoppgaver eller i oppgavesettene som helhet.

Læreplanene sier at lesing er en grunnleggende ferdighet som også er en del av matematikkfaget. Det er derfor ikke noe grunnleggende galt med at matematikkeksamen også tester elevers evne til å lese en matematikktekst for å kunne svare på den. Men det blir et problem hvis elever ikke får vist sin øvrige matematikkompetanse fordi det språklige kommer i veien. Dette kan delvis skje ved at elevene ikke får til den enkelte oppgave med vanskelig språk, delvis ved at de bruker lengre tid på å forstå den enkelte oppgaven og dermed går tom for tid på eksamenen som helhet. Derfor vil enhver analyse av tekst måtte ende opp med en faglig vurdering av om de språklige «snublesteinene» er formålstjenlige eller ikke.

En måte å komme nærmere en forståelse av hvordan språket påvirker prestasjonene, er å lage ulike versjoner av oppgavene. Abedi og Lord (2001) er et eksempel på en studie hvor man har identifisert språklig problematiske oppgaver og laget modifiserte versjoner av disse. I studien ble de modifiserte versjonene prøvd ut på 1174 elever (på 8. trinn). Undersøkelsen viste signifikant bedre resultat på de modifiserte versjonene enn på de opprinnelige. Bedringen er større for elever som er i ferd med å lære seg språket, og for elever med lav sosioøkonomisk status. Andre studier viser vekslende resultater (Kieffer et al. 2009; Li & Suen 2012), for eksempel viser Haag, Heppt, Roppelt og

[16] Se Persson (2016b, s. 5–6) for en diskusjon om ulike lesbarhetsindekser.

Stanat (2015) ingen signifikant effekt av språklig modifisering i en studie som omfatter nesten 18 000 elever på 4. trinn. «For å oppsummere er forskning på effektiviteten av språklig forenkling som en hjelp på tester fortsatt uten konklusjon.» (Haag et al. 2015, s. 148, vår oversettelse). Det ville vært interessant å gjøre tilsvarende forskning på norsk, men innenfor rammen av dette evalueringsoppdraget har vi ikke kapasitet til å gjennomføre undersøkelser med et stort nok antall deltakere til at det er håp om signifikante resultater.

Liberg (2001, s. 12) oppsummerer forskning på andrespråksleseres lesing slik:

> «Som faktorer som har framstått som svært viktige, kan man regne tekstinnholdet, leserens forkunnskaper innen fagområdet og kunnskaper om den kulturelle basen for dette, hans/hennes måte å forholde seg til faget på og den kulturelle basen for dette og de sosiale levevilkår som han/hun lever under. Tekstens oppbygning og leserens lingvistiske kompetanse har framstått som noe mindre viktige.» (vår oversettelse)

Luykx et al. (2007) betrakter lingvistikk og kultur som sammenfiltret.

Vurdering av kontekstuelle og kulturelle aspekter som påvirker forståelsen av matematikkoppgaver, ligger ikke inne i vårt oppdrag – utover hvordan ukjente ord kan påvirke forståelsen. Det er imidlertid viktig å være klar over at dette også er et aspekt ved forståelse av tekster.[17]

At oppgaver settes inn i en kontekst, bidrar ikke nødvendigvis til å gjøre dem lettere for elevene.[18] Samtidig må det nevnes at det er en myte at tekstoppgaver *nødvendigvis* er vanskeligere enn oppstilte oppgaver. Det finnes rikelig med eksempler på at det kan være motsatt: at flere elever får til oppgaver hvor det er gitt en kontekst (f.eks. Heuvel-Panhuizen 2005, s. 7–8).

[17] For eksempel viser Solano-Flores og Trumbull (2003) hvordan elever fra lavinntektsfamilier i stor grad misforstod en setning i en matematikkoppgave, «His mother has only $1.00 bills», ut fra egen families økonomiske situasjon (hvor det ikke var urimelig at mor hadde bare én dollarseddel), mens hvite elever fra høyinntektsfamilier tolket den som intendert: at mor bare hadde éndollarsedler. Oppgaven var satt i en kontekst hvor man skulle kjøpe skolelunsj, mens mange fattige barn i USA har gratis skolelunsj, noe som gjorde oppgaven enda mer irrelevant for dem.

[18] Persson (2016b) skriver om naturfagoppgaver som prøver å knytte faget til hverdagslivet og konkrete personer, og viser hvordan dette i hans undersøkelse fører til svakere resultater: «Høyere nivåer av personifisering reduserer sannsynligheten for at oppgaver besvares korrekt.» (s. 17, vår oversettelse). Imidlertid så det i den studien ut til at faglig høytpresterende elever også var bedre til å håndtere den hverdagslige konteksten uten å bli distrahert. «Dette kan tolkes som at høytpresterende elever klarer å 'akseptere' plasseringen av naturfaglige oppgaver i en hverdagslig kontekst, uten å bli distrahert eller irritert av bruken av personlige pronomener eller navn på fiktive steder eller elever, som indikert i tidligere forskning [...]» (s. 17, vår oversettelse). Serder og Jakobsson (2015) beskriver elever som blir irritert på de merkelige ungdommene i naturfagoppgavene som har så sterk interesse for å undersøke naturfagfenomener.

Relasjon til læreverkene

Hvorvidt språket i oppgavene oppfattes som krevende for elevene, vil avhenge av hva elevene har møtt gjennom opplæringen. Hvordan eksamenen samsvarer med innholdet i opplæringen, er da også et av temaene vi er bedt om å evaluere. Zevenbergen (2000) mener at det er ulike kommunikative strategier som ligger til grunn for «texts and tests» og «classroom talk».

Mange lærere forenkler fagspråket i sin undervisning i så stor grad at elevene får liten tilgang til det presise fagspråket (Ernst-Slavit & Mason 2011; Hipkiss 2014).[19] Dette betyr at lærebøkene kan bli hovedkilden til det mer akademiske språket for mange elever – i den grad elevene leser dem. (Det er naturligvis utenfor rammene av dette prosjektet å undersøke om det også er vanlig at norske lærere forenkler språket i en slik grad.)

Vi har derfor som utgangspunkt at hvorvidt bestemte språkfenomener oppfattes som krevende for elevene, også avhenger av i hvilken grad språkfenomenene er kjent gjennom læreverkene elevene har arbeidet med.[20] Å sammenlikne språket i sentralgitte prøver med lærebøkenes språk er derfor relevant – for eksempel viser Niederhaus et al. (2015) hvordan enkelte språkfenomener som forekommer i prøvene, er vanlige i enkelte lærebøker og ikke i andre. Liknende analyser er interessant å gjøre i en norsk kontekst.

Å vurdere eksamenen i lys av lærebøkene innebærer vel og merke ikke en påstand om at «læreboken har rett» – dersom en evaluering av eksamenen viser at språkbruken i eksamenen ikke samsvarer med språkbruken i lærebøkene, kan det naturligvis like gjerne argumenteres for at det er lærebøkene som bør endres. Vi er kjent med forskning som er kritisk til matematikklærebøker på ulike områder (Kongelf 2015; Smestad 2002). Dersom det skulle vise seg at skolen i for liten grad bringer elevene inn i matematikkdiskursen så de blir i stand til å løse eksamensoppgavene, kan en mulig konklusjon være at det er skolen som må bli bedre, ikke at eksamenen må bli enklere.

Framgangsmåte

På bakgrunn av det ovenstående har vi utviklet følgende framgangsmåte for analyse av språklige og layoutmessige aspekter ved eksamenen (tabell 7.1 er inspirert av Nie-

[19] For eksempel viser Ernst-Slavit og Masons studie at lærere i stor grad bruker hverdagsspråk og sier «that» og «this one» kombinert med peking når de viser til for eksempel teller, nevner osv. «[…] lærere som forsøker å hjelpe elever gjennom å holde seg innenfor det hverdagslige domenet, med dets hverdagslige termer og språkbruk, forringer snarere verdien av utdanning i dens kulturelle kontekst og etterlater elevene i en skoleversjon av hverdagskunnskap.» (Persson 2016a, s. 28, vår oversettelse). Gee (2005, s. 33) konkluderer rett og slett med at «et ansikt-til-ansikt-samtale-rammeverk er problematisk for tilegnelsen av vitenskapelig akademisk språk» (vår oversettelse).

[20] «Når eksamensoppgavene benytter mer komplekse eller varierte språklige hjelpemidler enn elevene har møtt i læringssituasjonene tidligere, så kan de med enkelte skolebøker ikke forberede seg på en adkvat måte på de språklige utfordringene i eksamenen» (Niederhaus et al. 2015, s. 153, vår oversettelse).

derhaus et al. (2015) og Haag et al. (2013, s. 32)). En del av faktorene analyseres fullt ut (for eksempel telle alle ord i alle oppgaver), for andre faktorer er vi bare på utkikk etter spesielt interessante eksempler.

Tabell 7.1 Oversikt over språktrekk som, på bakgrunn av litteraturgjennomgangen, er utgangspunkt for analysene av eksamenssettene, med kort operasjonalisering og eksempler.

Språktrekk	Beskrivelse*	Eksempler
Deskriptive trekk		
Antall ord	Antall ord i hver deloppgave. Teller ikke med tall som ord. (Tekst som står før en deloppgave, telles med til den deloppgaven.)	
1.2 Antall setninger	Antall setninger i hver deloppgave	
1.3 Antall lange ord	Ord med mer enn 6 bokstaver	seilbåt
Leksikalske trekk		
2.1 Lavfrekvente ord	(se under)	
2.1a Matematikkfaglig terminologi		formlike
2.1b Generell terminologi		fyrstikker
2.2 Sammensatte ord		hodetelefoner
2.3 Generelle akademiske ord	(se under)	sammenhengen
2.4 Nominaliseringer	Substantiver som er omformet fra verb	framgangsmåtene
2.5 Pakking av informasjon i substantivgrupper	Trekker fram eksempler	hjuldiagram som viser hvordan 39 elever på en skole kom seg til skolen en dag
Grammatiske trekk		
3.1 Uvant ordstilling	Trekker fram eksempler	Vi har tre forskjellige mynter som hver viser mynt eller kron når vi kaster dem
3.2 Lange, komplekse ytringer	Mer enn én frase Trekker fram eksempler	Antall gram CO_2 (karbondioksid) som en bestemt bil slipper ut per kilometer, er gitt ved funksjonen $f(x) = 0{,}046x^2 - 6{,}7x + 386$ der x er farten til bilen målt i kilometer i timen
3.3 Mangel på sammenheng(ord) mellom setningene	Trekker fram eksempler	Tower Bridge er en klaffebro over Themsen i London. Avstanden mellom tårnene er 60 m.
3.4 Betydningsfulle preposisjoner	Trekker fram eksempler	□ABCD skal speiles om y-aksen til □A′B′C′D′.
Andre trekk		
4.1 Uklar aktør	Trekker fram eksempler	Prisen ble satt ned med 20 %.
4.2 Paratekstlige elementer	Formatering av brødteksten og tekst som omgir brødteksten	(Overskrifter, fet skrift, farget skrift osv.)
4.3 Illustrasjonenes rolle (og hvordan det sammenbindes)	Vurderer om illustrasjonene er avgjørende for å løse oppgaven, til hjelp eller «til pynt». Også om de vises eksplisitt til i teksten.	

* Se punktet «Operasjonalisering» for nærmere beskrivelse nedenfor.

I noen forskningsprosjekter ser man på slike aspekter og oppsummerer dem så i en slags «indeks», se for eksempel Persson (2016a, s. 59). Å lage en slik indeks vil innebære å foreta en vekting av de ulike aspektene, og slik kunnskapsstatus er i øyeblikket, er det ikke grunnlag for å gjøre det for vårt formål. Det er dessuten grunn til å tro at ulike aspekter virker ulikt på ulike elevgrupper. Persson (2016a, s. 87) skriver for eksempel om sine resultater:

> «Det vises i avhandlingen at ulike språklige trekk har signifikante korrelasjoner med ulike elevgruppers resultat på oppgavene. Språklige trekk som har signifikant betydning for lavtpresterende elevers resultat, viser seg samtidig ofte å mangle slike signifikante korrelasjoner for høytpresterende elever. Også det omvendte gjelder, at språklige trekk som har signifikant betydning for høytpresterende elever, i visse tilfeller ikke har betydning for de lavtpresterende. På den måten viser avhandlingens resultat ikke bare at språkbruken har betydning for elevers meningsskaping, men også at den varierer mellom ulike elevgrupper. Det innebærer at den språklige drakten stoffet har, får ulike konsekvenser for ulike elever.» (vår oversettelse).

Persson finner dessuten ulikheter mellom ulike deler av et fag – i hans tilfelle naturfag – og vi kan tenke oss tilsvarende ulikheter mellom deler av matematikkfaget (Persson et al. 2016). Vi vil derfor isteden gjøre enkle sammenlikninger mellom de eksamenene vi evaluerer, og andre tekster – i første omgang tidligere eksamensoppgaver.

Operasjonalisering

For denne studien må det gjøres noen valg som både bygger på kunnskapen vi har, og er praktisk mulig innenfor rammene til prosjektet. For eksempel har vi valgt å ikke gjøre en fullstendig grammatisk analyse av alle oppgavene, selv om det kunne ha gitt grunnlag for interessante kvantitative analyser. Vi har landet på følgende operasjonaliseringer:

I den kvantitative analysen ser vi ikke på forsiden av eksamenssettet eller siden med eksamensinformasjon. Vi tar heller ikke med eksamensinfo om form eller hjelpemidler som står i egne rammer gjennom oppgavesettet, ei heller tekst om hvordan oppgaven skal føres. Tekst om løsningsmåte som layoutmessig framstår som del av oppgaven, tas med. Oppgavenummer og poengsum (overskrift til hver oppgave) tas ikke med. Informasjonen om kilder til bilder tas ikke med. Tekst i fotografier tas ikke med, mens tekst i tabeller, figurer og diagrammer tas med. Overskrifter tas med, også om de er i egen boks.

Antall ord: Vi teller ikke tallsymboler som ord. Heller ikke «A» (som navn på punkt) eller formler. Måleenheter regnes som ord. Forkortelser telles som om de var skrevet ut («dvs.» er tre ord, «f.Kr.» er to ord, «g/cm^3» er tre ord, «KrF» er to ord). Valutasymboler («€») telles som ord. Tekst som ikke står inne i en deloppgave, teller med til den deloppgaven som kommer rett etter. Matematikktegn («=», «%») teller ikke som ord.

Antall setninger: Fullstendige setninger er enkle å telle, men mye tekst i oppgavene er ikke gitt som fullstendige setninger. Hovedregelen er da at hver linje som inneholder ord, teller som en setning. Men enkeltopplysninger i figurer («5 cm») teller ikke som egne setninger (men ordene telles som ord).

Lavfrekvente ord: Inspirert av databasen «Ordforrådet» (Norsk tidsskrift for logopedi 1/13, s. 24[21]) definerer vi som «lavfrekvente» ord den fjerdedelen av ordene som forekommer sjeldnest i korpuset NoWaC[22] – det vil si ord som forekommer færre enn 1800 ganger i korpuset. Vi inkluderer alle former av ordet i tellingen (eks: «kofferter» – vi teller da alle forekomster av «koffert», «kofferten» osv.). Sammensatte ord vil ofte være lavfrekvente, men de registrerer vi som sammensatte ord. (Men senere i analysen ser vi på delene av sammensatte ord for å se om de er lavfrekvente.)

Denne metoden fanger vel og merke ikke opp ord som har flere betydninger, og som er lavfrekvente i den bestemte betydningen som brukes i oppgavene.

Lavfrekvente ord sjekkes mot de vanligste lærebøkene for 10. trinn.

Generelle akademiske ord: Vi definerer som generelle akademiske ord de ordene som forekommer i Akademisk ordliste[23] (Hagen, Johannessen & Saidi 2016). Denne ordlisten skiller mellom «høyfrekvente», «middels frekvente», «mindre frekvente» og «lite frekvente» akademiske ord.

Tekstanalyse

Ved evalueringen av de språklige trekkene ved eksamenen i matematikk for 10. trinn våren 2017 har vi på basis av en gjennomgang av forskningen på området lagt vekt på å se på en del språktrekk som kan gjøre oppgaveteksten vanskeligere for enkelte elevgrupper. Forskningen kan ikke si noe om hvor mye som er *for mye* av de enkelte trekkene. Vi vil derfor blant annet evaluere 2017-eksamenen ved å sammenlikne med to tidligere sett. Vi har valgt ut eksamenen fra 2016, fordi den er nær i tid, og eksamenen fra 2009, fordi den kom dårlig ut av Matematikksenterets gjennomgang av eksamensoppgaver (Matematikksenteret 2015).

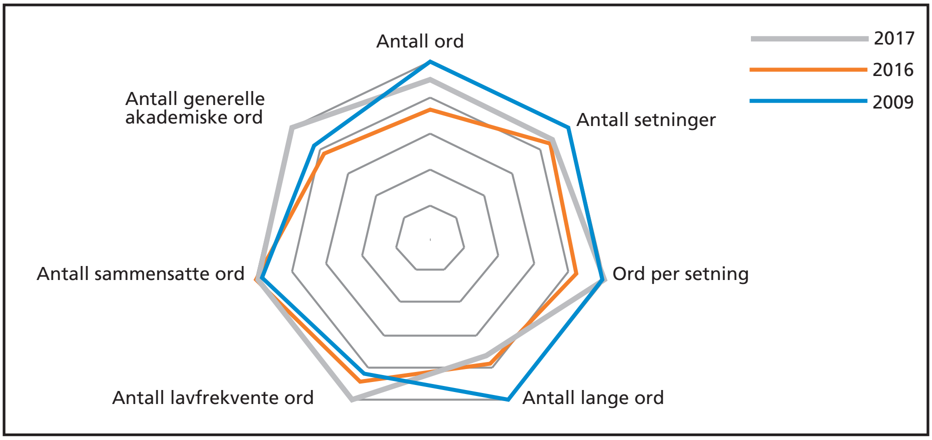
Vi vil først se på eksamenssettene samlet for de ulike aspektene, mens vi til slutt vil gå inn på enkelte oppgaver som utmerker seg.

[21] http://norsklogopedlag.no/uploads/docs/logopeden/1_2013/Logopeden_1-13_ordforradet.pdf

[22] http://www.hf.uio.no/iln/om/organisasjon/tekstlab/prosjekter/nowac/

[23] http://www.tekstlab.uio.no:4000/academic

Figur 7.1 Sammenlikning av noen kvantitative resultater av analysene av 2009-, 2016- og 2017-eksamenene. I prosent av høyeste verdi i den enkelte kategori.



I radardiagrammet (figur 7.1) sammenlikner vi eksamenene basert på sju kvantitative kategorier (hvor alle tall er omgjort til prosent av maksimumsverdien for de tre årene i den angjeldende kategorien). Diagrammet er altså kun egnet til å se hvordan 2017-eksamenen plasserer seg i forhold til 2016 og 2009.

Overordnet ser vi fra radardiagrammet at 2017-eksamenen plasserer seg mellom de to referanseeksamenene på flere av indikatorene. Eksamenen i 2017 har flere lavfrekvente ord enn 2016- og 2009-eksamenene, men som vi skal se nedenfor, har den færre ulike lavfrekvente ord enn 2016-eksamenen. Den har også noe flere generelle akademiske ord enn 2009- og 2016-eksamenene.

Etter opptelling av oppgaveord og setninger ut fra våre kriterier hadde 2009-eksamenen 1384 ord fordelt på 178 setninger. I 2016 var dette redusert til 1013 ord, men fordelt på 154 setninger. I 2017 var setningene blitt noe lengre igjen: 1240 ord fordelt på 158 setninger. Tekstmengde og gjennomsnittlig setningslengde var altså i 2017 mellom de to referanseårene og utmerker seg dermed ikke.

Leksikalske trekk

Tabell 7.2 viser hvordan de lavfrekvente (og ikke-sammensatte) ordene fordelte seg på de tre årene (i parentes er antall treff i NoWaC-korpuset).

De matematikkfaglige lavfrekvente ordene i denne listen er ord tett knyttet til den matematikken som er rimelig å teste på en eksamen for 10. trinn. Dette er ord som stort sett er å finne i lærebøkene også; unntakene er kube (som benevnes som «terning» i to av læreverkene), parabel og tangere (som det er to læreverk som ikke nevner) samt median (som mangler i ett læreverk, men det er fordi de ikke har statistikk i boken for 10. trinn).

Tabell 7.2 Matematikkfaglig og generell terminologi i avgangseksamenen for 10. trinn 2009, 2016 og 2017.

	2009	2016	2017
Matematikkfaglig terminologi	*kube* (1294)	*likningene* (266)	*y-aksen* (152)
	rektanglet (355)	*kongruent* (28)	*prisme* (x2) (203)
	desimal (841)	*parabel* (250)	*prismet* (203)
		tangerer (650)	*median* (1222)
Generell terminologi	*Arkimedes* (x9) (167)	*Firenze* (x3) (1297)	*kurv* (1200)
	Syrakus (x2) (1)	*arrabiata* (x2) (13)	*kurver* (1200)
	Hieron (1)	*bolognese* (x2) (124)	*mugge* (1175)
	skraver (159)	*stracotto* (x2) (0)	*muggen* (1175)
	Sandal (1061)	*vitruvisk* (x2) (3)	*Benedikte* (x2) (815)
	Starheim (252)	*Galileo* (x2) (1199)	*fyrstikker* (x5) (247)
	Heggø (64)	*Galilei* (476)	*Jeløy* (x4) (814)
	Gunvald (264)	*Fibonacci* (x2) (41)	*Themsen* (423)
	Sicilia (1384)	*Palazzo* (186)	*Lovelace* (x5) (58)
	arbelos (0)	*Vendramin-Calergi* (1)	

De andre lavfrekvente ordene («generell terminologi») er ord av flere typer: Dels er det navn på personer, steder eller matretter som det ikke er nødvendig å ha møtt før for å kunne håndtere. Mer interessant er en annen kategori ord: «kurv», «mugge» og «fyrstikker» er ord som vi lett kan tenke at alle kan, men som – fordi de er lavfrekvente – kan være ukjente for en del elever. De er dermed ord som kan gjøre oppgavene mindre forståelige. Ordene «mugge» og «fyrstikker» er illustrert i oppgavene, mens ordet «kurv» ikke er det. Både ordet «fyrstikk» og «kurv» er dessuten å finne i de fleste av læreverkene, mens ordet «mugge» ikke finnes i noen av dem.

I 2009-eksamenen forekommer det lavfrekvente ordet «skraver», og det er også avgjørende for å forstå hva man skal gjøre i oppgaven. Det kan derfor være kritisk at dette er et lavfrekvent ord. I 2017-eksamenen forekommer det ikke lavfrekvente ord som er kritiske for å forstå hva man skal gjøre i oppgaven.

Sammensatte ord forekommer oftere enn lavfrekvente ikke-sammensatte ord – på alle de tre eksamenene er det mer enn dobbelt så mange sammensatte ord som lavfrekvente ikke-sammensatte ord. Sammensatte ord er ofte lavfrekvente i seg selv, men kan likevel være mulige å tolke fordi de er sammensatt av frekvente ord. Vi velger derfor å se på hvilke av de sammensatte ordene som har deler som er lavfrekvente. Tabell 7.3 viser de sammensatte ordene som forekommer i oppgavene i de tre eksamenssettene (ord i kursiv er de som inneholder deler som selv er lavfrekvente ord).

Tabell 7.3 Sammensatte ord i avgangseksamenen for 10. trinn 2009, 2016, 2017. Ord som inneholder deler som selv er lavfrekvente ord, er i kursiv.

2009	gjennomsnitt (x2), middelverdi, typetall (x2), gullsmed, massetettheten (x2), gullsmeden, gullkrone, stortingsbygningen, *symmetrilinjene*, flytoget, standardform, normalform, trekant, funksjonsuttrykk, koordinatsystemet, folketallet (x2), funksjonsuttrykk (x3), skoletur, trekanten, klesbutikk, fotballkampen, målestokken, speedbåter, nasjonalforsamling, stortingsvalg, stortingsplassene (x3), Stortingssalen, Grunnloven (x2), *desimaltall*, regneark (x2), stolpediagram, stortingsplasser (x6), stortingsvalg, valgdistrikt, oddetallene, stortingsplass, stortingsvalget, stortingsbygningen, halvsirkel, halvsirkler, vitenskapsmann, skomakerkniv, halvsirkel (x3), skomakerkniven (x3)
2016	standardform, funksjonsuttrykket, prisavslag, hundekjeks (x3), hjelpefigur (x2), badevakt, matematikkprøve, typetallskarakteren, sektordiagrammet, karakterfordelingen, gjennomsnittskarakteren, luftlinje, *hermetikkboks, hermetikkboksen, vekslingsgebyr*, eurosedlene, valutakalkulator, valutakalkulatoren, kodelås (x2), håndbagasje, gjennomsnitt, bensinprisen, varmretter, varmrettene, salgsinntekten, salgsinntektene, salgsinntekt, regneark, regnearket, vitenskapsmannen, perspektivlinje, forsvinningspunktet, kunstverket, langfingertupp, armspennet, langfingertuppen, blykuler, fallhøyden, luftmotstanden, kanonkuler, kanonkule, havoverflaten, graftegner (x2), kanonkulen, *Fibonacci*-tallene (x4), tallfølge, tallfølgen (x2), halvsirkler, halvsirklene, halvsirkelen, halvsirklene, linjestykket
2017	jordbær, hodetelefoner, hodetelefonene, standardform, blandingsforholdet, hjuldiagram, perspektivlinjer, forsvinningspunktene, romfigurene (x2), personbil, personbilene (x3), typetall, gjennomsnittet, sykkelferie, veilengde, sykkelturen, gjennomsnittsfart, bensinstasjon, bensintanken, jerrykanne, jerrykannen, bensinstasjonen (x2), sykkelprodukter, sykkelsko (x2), sykkelhjelm (x2), regneark, regnearket, salgsinntekt, *karbondioksid*, graftegner, x-verdier, *CO_2-utslippet, CO_2-utslipp*, gjennomsnittsfart, bremselengden (x2), friksjonstall, veidekket, innskrevet, fargelagt, framgangsmåtene, midtnormalen (x4), geometriprogram, sirkelflaten, klaffebro, tidspunkt, vannoverflaten, seilbåt, likningssystemet, likningssystem, *dataprogrammerer*, regnemaskin, dataprogrammeringsspråket

Vi ser at det er et betydelig antall sammensatte ord i hver eksamen. Vi ser også at en del av dem har deler som er lavfrekvente. Mens disse hørte til den matematikkfaglige terminologien i 2009 (symmetrilinjene og desimaltall), ser vi at det i 2016 og 2017 er flere sammensatte ord fra den generelle terminologien som inneholder lavfrekvente deler. Det er rimelig å forvente at når lavfrekvente ord som hermetikk, veksling, CO_2 og programmerer brukes i sammensatte ord, vil dette gjøre tolkningen av noen oppgaver mer krevende for en del elever.

I 2017-eksamenssettet var det 29 ord som vi karakteriserer som «generelle akademiske ord» (dvs. som finnes i *Akademisk ordliste*). Dette er flere enn i 2009 (24) og 2016 (22). Alle er i kategoriene høyfrekvente eller middels frekvente akademiske ord, det finnes ikke eksempler på mindre frekvente eller lite frekvente akademiske ord i de tre settene som er undersøkt.

De generelle akademiske ordene i 2017 var: bestemt, forholdet, forklar (2), form, formen, funksjon, funksjonen, konstruere, mellom (7), mønster, oppgave, ovenfor (5), sammenhengen, ulike, uttrykk, verdien, verdier, virkeligheten.

Bruk av generelle akademiske ord kan gjøre oppgaver vanskeligere å løse for elever med svake språkkunnskaper på prøvespråket. Det er derfor positivt at eksamensoppgavene i liten grad har slike ord, og at man unngår lavfrekvente generelle akademiske ord i oppgavene.

Et særtrekk ved akademisk språk er antallet nominaliseringer – altså at handlinger (verb) gjøres om til ting (substantiver). I eksamensoppgavene er det forholdsvis få av disse. Eksempler fra 2017: blandingen, CO_2-utslipp, framgangsmåtene, løsninger. Det er ingen ting som tyder på noen opphoping av nominaliseringer som kan gjøre oppgavene vanskelige å tolke.

Preposisjoner har ofte en viktig rolle i matematikkoppgaver, og den samme preposisjonen kan bety svært forskjellige ting i ulike sammenhenger. Derfor kan meningsbærende preposisjoner være en utfordring for de elevene med svakest norskkunnskaper. For eksempel opptrer preposisjonen «til» i en rekke situasjoner i 2017-eksamenssettet – her er noen eksempler:

- «□ABCD skal speiles om y-aksen til □A'B'C'D'.» (del 1, oppg. 5)
- «Avstanden fra jorda til månen» (del 1, oppg. 13)
- «Skriv koordinatene til punktet A» (del 1, oppg. 15a)
- «forsvinningspunktene til figuren nedenfor» (del 1, oppg. 22)
- «CO_2-utslippet til bilen» (del 2, oppg. 5)
- «bremselengden til bilen» (del 2, oppg. 6)
- «bruk formlene til Ada Lovelace ovenfor til å løse likningssystemet» (del 2, oppg. 9)

«Til» kan altså bety å tilhøre noen – delvis i ganske abstrakt betydning (koordinatene til et punkt, forsvinningspunkt til en figur), eller det kan bety en bevegelse mot noe (hvor langt det er til månen). «CO_2-utslippet til bilen» kan sikkert misforstås som å være noe som slippes i retning bilen (på samme måte som vi sier «utslipp til luft»), og bremselengden til bilen kunne noen elever forstå som lengden vi bremser før vi er framme ved bilen. I dette tilfellet vil «bilens bremselengde» være mindre tvetydig. I oppgave 5 i del 1 vil nok noen i første omgang tenke at y-aksen hører til □A'B'C'D' – det er ikke så lett gjennomskuelig at setningen prøver å si at □A'B'C'D' er resultatet av speilingen om y-aksen.

I tillegg til disse språklige utfordringene kommer mer faglige problemstillinger: Når oppgave 2 i del 2 i 2016 oppga at «Koden består av fire sifre fra 0 til 9», er da 0 og 9 med eller ikke? Matematisk sett er de ikke det, men illustrasjonen i margen tydet på at i hvert fall 0 var med – og elevenes hverdagserfaringer tyder nok på at både 0 og 9 skal være med.

Generelt bør man være oppmerksom på preposisjonene og unngå oppgavetekster hvor forståelse avhenger av å manøvrere mellom ulike betydninger av preposisjonene.

Grammatiske trekk

I akademisk språkbruk kan uvant ordstilling, lange, komplekse ytringer og mangel på sammenheng mellom setningene bidra til å redusere lesbarheten. I analysene har vi sett etter eksempler på dette, samtidig som vi har vært forberedt på at vi sannsynligvis ikke vil finne mange eksempler, siden eksamensoppgavetekster er korte og dessuten nøye gjennomarbeidede. I 2017-eksamenene har vi funnet noen eksempler:

- «Vi har tre forskjellige mynter som hver viser mynt eller kron når vi kaster dem» (del 1, oppg. 7)
- «Nedenfor ser du et hjuldiagram som viser hvordan 39 elever på en skole kom seg til skolen en dag» (del 1, oppg. 20a)
- «Antall gram CO_2 (karbondioksid) som en bestemt bil slipper ut per kilometer, er gitt ved funksjonen $f(x) = 0{,}046x^2 - 6{,}7x + 386$ der x er farten til bilen målt i kilometer i timen» (del 2, oppg. 5a)

Ut fra forskningen kan man anta at hver av disse setningene kan skape problemer for svake lesere.

Beslektet med nominaliseringer er pakking av informasjon i substantivgrupper (nominalgrupper). Dette dreier seg om at et substantiv bygges ut foran og/eller bak. Resultatet kan bli kompakte setninger hvor mye innhold er presset inn på lite plass. Eksempler fra 2017-eksamenen:

- «tre forskjellige mynter som hver viser mynt eller kron når vi kaster dem» (del 1, oppg. 7)
- «hjuldiagram som viser hvordan 39 elever på en skole kom seg til skolen en dag» (del 1, oppg. 20)
- «arealet av den delen av sirkelflaten som er blå på figur 2» (del 2, oppg. 7)
- «en stor, planlagt regnemaskin som skulle hete Den analytiske maskinen» (del 2, oppg. 9)

Hver av disse substantivgruppene kunne ha vært omformulert ved å bruke flere setninger – for eksempel kunne den første vært splittet opp slik: «Vi har tre forskjellige mynter. Hver av dem viser mynt eller kron når vi kaster dem.» Ut fra faglitteraturen er det grunn til å tro at utstrakt bruk av lange substantivgrupper kan bidra til å gjøre tekst vanskeligere å lese. Det ser ut til at det er færre eksempler på lange substantivgrupper i 2017-eksamenen enn i 2016- og 2009-eksamenene.

Et særtrekk med akademisk språkbruk er også at man ved hjelp av nominaliseringer eller passivformer gjør det utydelig hvem som er aktøren, noe som gjør situasjonen mer abstrakt for leserne. Også her er det noen eksempler i 2017-settet:

- «12 L saft skal helles over på flasker som hver rommer 4 dL. Da trenger vi ____ flasker.» (del 1, oppg. 2b)
- «Prisen ble satt ned med 20 %. Omtrent hvor mange kroner ble prisen satt ned med?» (del 1, oppg. 17)
- «Volumet av blandingen er skrevet på hver mugge» (del 1, oppg. 19)
- «Så blir begge klaffene hevet 60° for å la en seilbåt kjøre forbi» (del 2, oppg. 8a)
- «Et likningssystem [...] kan skrives på formen [...]» (del 2, oppg. 9b)

Det er ikke tilrådelig å unngå passivform for enhver pris – det kan bli kunstige formuleringer av det. Å introdusere en aktør (David Wilson, technical officer ved Tower Bridge) i det fjerde eksemplet vil neppe gjøre oppgaven mer leselig. Men i den første oppgaven, hvor «vi» likevel introduseres i den andre setningen, ville det sannsynligvis vært en fordel å skrive oppgaven som «Vi skal helle 12 L saft [...]»

Nynorskversjonen

Det har ved tidligere eksamener forekommet forskjeller på bokmåls- og nynorskversjonene. I 2016 var oppgave 2b i del 2 ulikt formulert: I bokmålsversjonen var oppgaven «Skriv opp de ulike kombinasjonene», mens i nynorskversjonen var den «Skriv opp dei seks ulike kombinasjonane».

Analysene i denne rapporten er gjort med utgangspunkt i bokmålsversjonene. For 2017-eksamenen har vi også gjort en sammenlikning mellom nynorsk- og bokmålsversjonen for å se om det er vesentlige ulikheter. De viktigste forskjellene vi har funnet, er gjengitt i tabell 7.4.

I prinsippet kan enhver omformulering gjøre en oppgave vanskeligere eller enklere. Siden formuleringer med «hvor mange» / «kor mange» er fullt gangbart både på bokmål og nynorsk, er det vanskelig å se faglige grunner til å bruke formuleringer med «antall» på bokmål og «kor mange» på nynorsk. Tilsvarende er det vanskelig å se at det bør hete «I hver mugge» på bokmål hvis det skal hete «I kvar av muggene» på nynorsk.

Tabell 7.4 Forskjeller i formulering mellom nynorsk- og bokmålsversjonene av avgangseksamenen for 10. trinn 2017.

Oppgave	Bokmål	Nynorsk
Del 1, oppgave 7	«Bestem sannsynligheten for at alle myntene viser mynt eller at alle myntene viser kron.»	«Bestem sannsynet for at alle myntane viser mynt, eller at alle myntane viser kron.»
Del 1, oppgave 19	«I hver mugge nedenfor […]»	«I kvar av muggene nedanfor […]»
Del 2, oppgave 1	«Elevene teller antall personer i de 30 første personbilene og får følgende resultat:»	«Elevane tel kor mange personar som sit i dei 30 første personbilane, og får dette resultatet:»
Del 2, oppgave 3	«Forholdet mellom antall liter diesel og antall liter bensin var 3:5.»	«Forholdet mellom liter med diesel og liter med bensin var 3:5.»
Del 2, oppgave 5	«Antall gram CO_2 (karbondioksid) […]»	«Kor mange gram CO_2 (karbondioksid) […]»
Del 2, oppgave 9	«Ada Lovelace (1815-1852) regnes som […]»	«Ada Lovelace (1815-1852) blir rekna for å vere […]»

Det er ikke grunnlag for å si at det er språklige ulikheter i de to versjonene av 2017-eksamenen som påvirker elevenes resultater, men det finnes ulikheter i formuleringer som framstår som unødvendige. Unntaket er oppgave 7 i del 1, hvor plasseringen av komma påvirker betydningen av innholdet. Oppgaven i bokmålsutgaven av eksamenen har formuleringen «*Bestem sannsynligheten for at alle myntene viser mynt eller at alle myntene viser kron.*» og på nynorsk har det kommet inn et komma og oppgaven lyder «*Bestem sannsynet for at alle myntane viser mynt, eller at alle myntane viser kron.*» Formuleringen i bokmålsutgaven vil tolkes slik at det spørres etter summen av sannsynlighetene for disse to utfallene, mens formuleringen i nynorskversjonen vil tolkes slik at det spørres etter sannsynligheten for ett av utfallene.

Interessante enkeltoppgaver i 2017-settet

Vi har valgt å se nærmere på oppgave 8 i del 2 i 2017-settet (se kapittel 1), fordi denne oppgaven har mange av de språklige trekkene som det er interessant å se nærmere på. Oppgave 8 er middels lang (75 ord), og mange av ordene er lange (17 ord). Den har kun ett lavfrekvent ord, «Themsen», som vel neppe forvirrer noen. Derimot har den en del sammensatte ord (klaffebro, tidspunkt, vannoverflaten, seilbåt), og ikke minst ordet «klaffebro» kan være ukjent for en del. Det er to illustrasjoner til oppgaven, den nederste er kategorisert som «til hjelp» – for selv om alle opplysningene i den er å finne i teksten også, er den nok grei å ha for å systematisere opplysningene. Det vises til den i teksten. Den øverste illustrasjonen kategoriserer vi som «til pynt», men for elever som lurer på hva en klaffebro er, eller hvordan en seilbåt ser ut, kan den være nyttig.

Oppgaven har mange preposisjoner som man må holde orden på: over, mellom, for, fra, til, på, mellom, ved. En setning som «Forklar at bredden AB på åpningen mellom klaffene er 30 m.» kan være tung hvis man sliter med preposisjoner, men AB er i tillegg markert med rødt på figuren, og det kan være til hjelp.

Selv om oppgaven er litt innfløkt og noen av formuleringene kan være vanskelige, spiller altså illustrasjoner og tekst sammen og gjør at oppgaven neppe utmerker seg som språklig vanskelig.

Vi vil også se på oppgave 6a i del 2 (se kapittel 1). Som nevnt tidligere i rapporten har de svakeste elevene svært lav uttelling på del 2 av eksamenen – i gjennomsnitt får «enerelevene» 1,7 poeng på hele del 2. For eksempel får «enerelevene» i gjennomsnitt 0,10 poeng på oppgave 6a, men gjennomsnittet for alle elevene er 0,76 (oppgaven gir maksimalt 1 poeng). Det er i utgangspunktet en enkel problemstilling: Hvis en bil kjører i 60 kilometer per time, hvor langt kjører den da på en og en halv time? Elevene har mange hjelpemidler, for eksempel kalkulator, til rådighet. Det ville vært interessant å finne ut mer om hvorfor så få av disse faglig svake elevene får til denne oppgaven. I denne omgang kan vi bare spekulere: Har de på dette tidspunktet gitt opp å finne noe de får til i del 2? Er de ukomfortable med desimaltallet (1,5)? Er de usikre på hva h står for (i hverdagsspråk bruker vi time, ikke h, for å beskrive denne måleenheten)? Eller får ord som «gjennomsnittsfart» og spørsmålet om hvor langt den *kan* kjøre – ikke hvor langt den faktisk kjører – elevene til å tro at oppgaven er mer komplisert enn den er? Eller synes de at oppgaven ser så lett ut at de tror at de må ha misforstått noe? Å finne ut mer om hvordan de svakeste elevene tilnærmer seg oppgaver som denne, ville si oss mer om hvilken kompetanse eksamenen faktisk måler – og ikke måler – hos de svakeste elevene.

Eksamensspråket

Eksamensveiledningen definerer bestemte ord på måter som legger føringer for hvordan eksamenen skal løses, og det forutsettes at elevene har satt seg inn i denne. Figur 7.2 viser en tekst fra side 8 i eksamensveiledningen for 2017.

Figur 7.2 Formulering om «Finn […]», «Løs […]» og «Bestem […]» i eksamensveiledningen (Utdanningsdirektoratet 2017a).

> Ved formuleringer som **"Finn …"**, **"Løs …"** og **"Bestem …"** legges det ikke opp til bestemte framgangsmåter eller spesielle hjelpemidler. Eleven kan velge å løse oppgaven grafisk, ved regning (algebraisk) eller ved å benytte ulike kommandoer i et digitalt verktøy. Her har eleven *full* metodefrihet.

Et eksempel på en oppgave med formuleringen «Løs» er oppgave 9 i del 2. Det kinkige med dette likningssystemet er at mange elever nok ser løsningen umiddelbart, siden 5 + 4 = 9 og 6 + 7 = 13. Siden det er «*full* metodefrihet», bør det da holde å skrive at løsningen $x = 1$ og $y = 1$ er åpenbar ut fra visuell betraktning av likningssystemet. Det er likevel uklart om elevene kan stole på å få full poengsum ved en slik løsning.

Elever og læreres oppfatning av eksamenens tekstoppgaver

Det er et mål at tekstoppgavene på en matematikkeksamen skal være språklig gode og forståelige for elevene. I denne delen av kapitlet settes det søkelys på hvordan elever og lærere oppfattet språket og tekstoppgavene på årets eksamen.

Språkbruk og tekstoppgaver i matematikk var noe alle matematikklærerne vi intervjuet, var opptatt av. Lærerne brukte mye tid på å lære elevene å hente ut det matematiske problemet i tekstoppgaver både i undervisningen og i forberedelsen til eksamen.

Inntrykket fra årets eksamen var at eksamensoppgaven hadde et presist språk og ikke for store tekstbolker, og at setningene ikke var for lange sammenliknet med tidligere eksamensoppgaver. Lærerne var gjennomgående positive når det gjaldt språkets vanskegrad. Til tross for at årets eksamen ble oppfattet som vellykket med et tydelig, presist språk som gjorde det enklere for elevene å hente ut den matematiske informasjonen i tekstoppgavene, fortalte lærerne at det var elevgrupper som systematisk hadde problemer med tekstoppgaver. Det er elever med svak leseforståelse og minoritetsspråklige elever med svake norskkunnskaper. I spørreundersøkelsen til matematikklærerne stilte vi spørsmål om bruk av tekstoppgaver på eksamen. Lærerne ble bedt om å svare på hvor enige eller uenige de var i påstander om tekstoppgaver. Resultatene er presentert i figur 7.3.

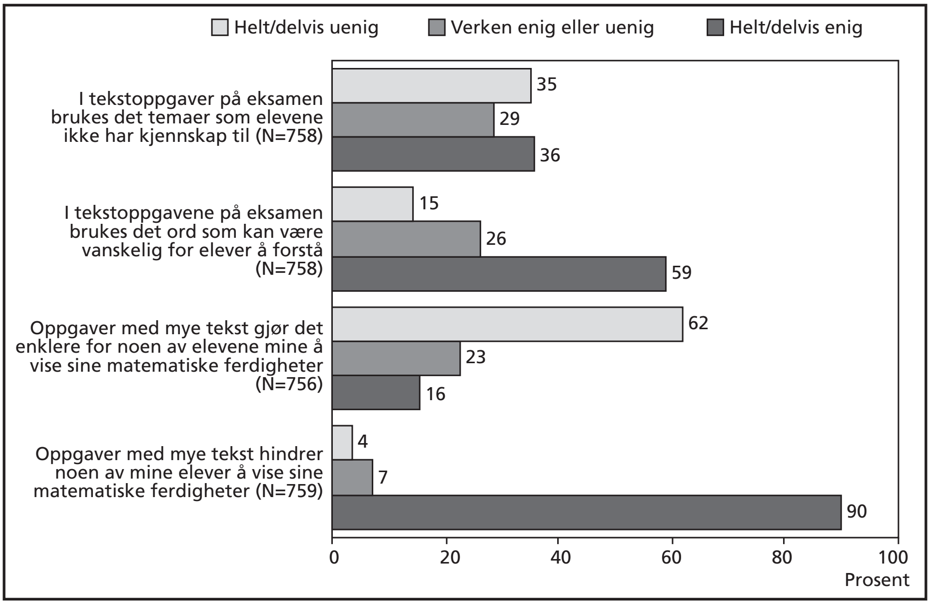
Resultatene fra spørreundersøkelsen viser at matematikklærere mener at bruk av oppgaver med mye tekst hindret noen elever i å vise sine matematiske ferdigheter. Ni av ti lærere svarte at de var helt eller delvis enige i den påstanden. Over halvparten av lærerne er også helt eller delvis enige i at det brukes ord i tekstoppgavene på eksamenen som kan være vanskelige for elever å forstå. En matematikklærer på en skole med en høy andel minoritetsspråklige elever mente at det for noen av hans elever var svake norskkunnskaper og ikke matematikkforståelsen som gjorde at de presterte svakt i matematikkfaget. Han fortalte:

> Matematisk sett kunne de fått toppkarakterer, det handler om språk.

En annen lærer ved en annen skole sa det slik om de elevene hun hadde med svak norskforståelse:

> Det er språket, ikke matematikken, som er problemet.

Figur 7.3 Matematikklæreres svar på påstander om tekstoppgaver. Prosent.



Det vil ikke si at alle elevene med svake norskkunnskaper ville fått bedre karakterer i matematikk gitt bedre språkkunnskaper, men for lærere kan det være krevende å skille mellom hva som handler om svak norskforståelse, og hva som er manglende kompetanse i matematikk.

For elever med dårlig leseforståelse og svake norskferdigheter er skriftlig eksamen krevende. I matematikktimene kan lærerne hjelpe elevene med å finne fram til det matematiske problemet, men i en prøvesituasjon må de finne ut av det selv. Lærerne på skolen med en høy andel minoritetsspråklige elever bruker lang tid i undervisningssituasjonene på å forklare hverdagsbegreper elevene ikke skjønner, før de kan begynne på matematikkundervisningen. En lærer fortalte at for mye tekst og bruk av vanskelige ord i oppgaveteksten kunne føre til at elever som forstod det matematiske problemet, ikke klarte å sortere ut den nødvendige informasjonen for å løse oppgaven. Det var spesielt bruk av ord som var ukjente for elevene, som kunne skape problemer. Eksempler på ord elever med svak leseforståelse og svake norskkunnskaper kunne ha problemer med å forstå, var sum, kjøregodtgjørelse og fortjeneste.

I spørreskjemaet til matematikklærerne stilte vi spørsmålet «Mener du det er noen bestemte elevgrupper som gjennomgående ikke får vist sin kompetanse på eksamen?» (se figur 5.6 i kapittel 5). De kunne krysse av for flere alternativer. 41 prosent av lærerne krysset av for elever med lese- og skrivevansker og 41 prosent for minoritetsspråklige elever. Svarene fra spørreskjemaet sammenfaller med svarene i de kvalitative intervjuene.

Det overordnede inntrykket var at årets eksamen var preget av et presist språk og ikke for store tekstbolker, og at setningene ikke var for lange. Det var imidlertid oppgaver på årets eksamen hvor både elever og lærere mente oppgavene var for utydelige. Et eksempel som ble trukket fram, var oppgave 2b i del 2 (se også drøfting i kapittel 8). I oppgaven ble elevene spurt om sammenhengen mellom tid og veilengde på en sykkeltur. Veilengden og tiden var presentert i et diagram. Spørsmålet i oppgave 2b var «Hvor stor del av sykkelturen har Mari og David pause?» Formuleringen «hvor stor del» ble oppfattet som uklar, og lærerne mente mange elever ville svare i minutter. En lærer sa det slik:

> Hvis de ønsket at elevene skulle svare i brøk eller prosent, er det bedre å spørre om det.

En annen oppgave hvor teksten var forvirrende, var i del 2, oppgave 9. Informasjonen i en tekstboks om Ada Lovelace sitt liv mente en lærer var unødvendig. Elevene ble forvirret og lurte på om det var informasjon de trengte for å løse oppgaven. Det ble framhevet av både lærere og elever at i en stressende situasjon, som en eksamen, leter elevene etter den informasjonen de trenger for å løse oppgaven, slik at de kan gå videre til neste oppgave. Ekstra informasjon virker mot sin hensikt og oppleves forvirrende.

En siste tekstoppgave som ble nevnt i intervjuene med både lærere og elever, var oppgave 3b i del 2. Her ble elevene bedt om å regne ut hvor mange liter bensin som fikk plass i en jerrykanne. Det var svært få elever som visste hva en jerrykanne var for noe. Problemet løste seg fordi oppgaven hadde en illustrasjon ved siden av med et bilde av en jerrykanne som gjorde at elevene forstod oppgaven.

Oppsummering

- Språket i eksamensoppgavene er i hovedsak hensiktsmessig. Det brukes beskjedent med lavfrekvente ord, og de som brukes, er ofte illustrert. Andelen sammensatte ord, spesielt de som har deler som er lavfrekvente, kan med fordel reduseres.
- Når nynorskversjonen er laget, bør bokmålsversjonen gjennomgås for å få størst mulig språklig samsvar mellom de to versjonene, siden enhver ulikhet i formulering i prinsippet kan gjøre at oppgavene får ulik vanskelighetsgrad.
- En oppfatning blant matematikklærerne var at oppgaver med mye tekst hindrer noen av elevene i å vise sine matematiske ferdigheter.
- Elever med minoritetsspråklig bakgrunn er den gruppen som sliter mest med tekstoppgaver i matematikk.
- Inntrykket fra årets eksamen var at mengden tekst og språket i tekstoppgavene var tilstrekkelig og forståelig for elevene.

8 Utfordringer i sensorenes arbeid

Tradisjonelt har retting av matematikkprøver vært sett på som veldig enkelt: riktig eller galt svar. For flervalgsoppgaver og oppgaver der det kun er et svar som skal oppgis, er det riktig, men for alle andre måter å besvare oppgavene på er det rom for sensorenes skjønn. En lik og rettferdig vurdering av eksamen krever at sensorene er samstemt. Vi har undersøkt hvilke oppgaver det kan være vanskelig å sensurere, og om sensorene opplever å få den veiledningen de trenger i sensurarbeidet.

Analyse av sensorers vurderingsskjemaer

Det er i stor grad samsvar mellom sensorenes vurdering i form av forslag til endelig karakter på eksamenen. Likevel kan det være ulik vurdering av de enkelte oppgavene, og det er interessant å se hvilke oppgaver som får ulik vurdering fra sensorene.

Våre data er fra sensorene i Oslo og Akershus (se kapittel 2). I vårt materiale var det 2655 kandidater der samlet avvik mellom sensorene var mindre enn eller lik 10 poeng[24]. Denne avgrensningen ble gjort for å fjerne ekstreme tilfeller og feilføringer. Vi kommenterer videre de oppgavene der mer enn 10 prosent av disse kandidatene har fått ulik vurdering fra de to sensorene. På del 1 måtte sensorene gi 0 eller 1 poeng på hver deloppgave (bortsett fra én), mens på del 2 var det mulig å få flere poeng på de fleste deloppgavene. På oppgave 4 kunne sensorene for eksempel gi 0, 1, 2, 3, 4 eller 5 poeng.

Fra del 1 av eksamenen er det bare én oppgave av totalt 34 deloppgaver som utmerker seg: oppgave 7 der 14,7 prosent fikk ulik vurdering. Dette er en sannsynlighetsoppgave som tester samme kompetanse som oppgaven foran, men på høyere nivå. Oppgaven hadde også språklige utfordringer som er kommentert i kapittel 7. I forhåndssensurrapporten (Utdanningsdirektoratet 2017d) ble føringen til sensorene endret fra at korrekt framgangsmåte og svar gav uttelling, til at også feil svar med korrekt framgangsmåte skulle gi full uttelling. Det ble også tatt med i diskusjonen at et komma i nynorskversjonen av oppgavesettet gir ordet «eller» ulik betydning i de to versjonene.

[24] Vi summerte absoluttverdiene av forskjellene på hver deloppgave.

Tabell 8.1 Oversikt over hvilke deloppgaver på eksamenen der mer enn 10 prosent av sensorene har ulik vurdering.

Oppgave		Andel med ulik vurdering
Del 1	7	14,7 %
Del 2	1a	29,5 %
	2b	30,6 %
	2c	10,6 %
	3b	14,4 %
	4	39,9 %
	5a	19 %
	5b, c	20,8 %
	7a	27 %
	7b	14,8 %
	7c	14,1 %
	8a	20,5 %
	9a	11 %

Svarformatene i del 1 av eksamenen der en stor del av oppgavene er enten flervalg (14 av 34) eller oppgaver der det bare skulle oppgis svar uten begrunnelse (10 av 34), gir ikke så stort rom for skjønn i vurderingen. I del 2 skal alle svar kommuniseres på vanlige innføringsark eller som utskrift fra graftegner eller regneark. Denne delen har også oppgaver som krever løsning i flere trinn, og har større krav til kommunikasjon av framgangsmåte og begrunnelse for svarene. Det fører til at det er flere oppgaver der det er mer utfordrende å vurdere besvarelsene, og tabell 8.1 viser at det bare er oppgave 6 som ikke har minst én deloppgave som ser ut til å være utfordrende å sensurere.

Vurdering av oppgaver med krav om digitale hjelpemidler

Én av oppgavene i del 2, oppgave 4, krever bruk av regneark, og denne oppgaven kan gi inntil 5 poeng ved full uttelling. På denne oppgaven får hele 39,9 prosent av kandidatene ulik vurdering fra de to sensorene, men i de fleste tilfellene er forskjellen kun på 1 poeng. Regneark i del 2 av eksamenen er behandlet i et eget delkapittel i eksamensveiledningen (Utdanningsdirektoratet 2017a, kap. 1.6.3.4). På sensorskoleringen vi deltok på, forutså sensorene at vi kom til å se stor variasjon i vurderingen av denne oppgaven, og etterlyste enda tydeligere veiledning. Mange sensorer tar også opp dette i sin tilbakemelding til Utdanningsdirektoratet. Eksamensveiledningen sier for eksempel på side 15 at «Selv om det primært er det faglige innholdet som skal vurderes, vil også presentasjonen av løsningen bli vurdert.» Spørsmålene som ble reist på sensorskoleringen, var blant annet: Hvilken uttelling skal gis for formelutskrifter over tre–fire ark eller i mikroskrift som knapt er leselig? Dersom kandidatene leverte Excel-filer i PAS, hvordan skulle kravet om formelutskrift vurderes? Hva gjør sensorene der de ser at det er gjort systematisk feil fra skolene? Det har siden våren 2015 vært krav om bruk av graftegner i del 2 av eksamenen, og kapittel 1.6.3.3 i eksamensveiledningen (Utdanningsdirektoratet 2017a) beskriver kravene og har med et eksempel i tillegg til at det eksplisitt står at det vil komme en oppgave med krav om bruk av graftegner på eksamenen. Det er oppgave 5 i del 2 i dette settet. Oppgave a er å tegne grafen til en andregradsfunksjon innenfor et begrenset område, og det skal gis 2 poeng for korrekt tegnet graf innenfor oppgitte x-verdier med skala og navn på aksene. En korrekt tegnet graf uten begrensning og uten skala og navn på aksene skal gi 1 poeng, og det er ifølge forhåndssensurrapporten (Utdanningsdirektoratet 2017d) ikke mulig å gi ekstra uttelling for å ha fått til begrensning eller skala og navn på aksene hvis kandidaten ikke har klart begge deler. 19 prosent av kandidatene fikk ulik vurdering. Oppgave 5b og c har vi sett under ett siden det er presisert at det kun skal trekkes én gang for ikke å oppgi kommandoer. I oppgave b ble det på sensorskoleringen kommentert at det ikke var åpenbart at begge skjæringspunktene skulle oppgis som svar. Til sammen kan disse to deloppgavene gi 4 poeng, og 20,8 prosent av kandidatene fikk ulik vurdering – for de fleste, 16,6 prosent, var forskjellen på 1 poeng.

I løsningen av oppgave 7a kan kandidaten velge enten å løse oppgaven for hånd med passer, blyant og linjal eller å løse den med dynamisk geometriprogram. I eksamensveiledningen står det at en slik oppgave vil komme på eksamenen våren 2017 (Utdanningsdirektoratet 2017a, s. 11), og det er vist et eksempel på digital løsning av en oppgave fra et tidligere sett. I veiledningen, under punktet om konstruksjon med passer, linjal og blyant (s. 10), står det at konstruksjoner i del 2 skal gjøres på blankt papir. Dette gjentas ikke i eksamensteksten, og inntrykket fra sensorskoleringen er at en del skoler ikke har fulgt opp dette. 27 prosent av kandidatene har fått ulik vurdering på denne oppgaven, som maksimalt kan gi 3 poeng i uttelling, og for de fleste er det en forskjell på 1 poeng. Ut fra vårt materiale kan vi ikke si om den ene eller andre måten å løse oppgaven på har vært mer utfordrende å sensurere.

Vurdering av oppgaver med valgfri metode

På hver av de to siste deloppgavene på oppgave 7 er det også ulik vurdering, men for færre kandidater enn på oppgave a (se tabell 8.1). Oppgave b gir enten 0 eller 1 poeng for å «forklare kort» at en vinkel er 90°, uten rom for å gi noe uttelling til en kandidat som har skrevet «Thales' setning» uten å gjengi denne. Forhåndssensurrapporten presiserer at måling av vinkler ikke godtas, men hva med de som måler lengden på den tredje siden og bruker Pytagoras' setning til å vise at trekanten er rettvinklet? I siste deloppgave er det tre utregninger – areal av trekant og sirkel og differansen mellom disse – og 2 poeng å dele ut. Ulik vektlegging kan gi opphav til ulik uttelling.

Oppgave 1a i del 2 inneholder en tabell som skal brukes som utgangspunkt for å tegne et «passende diagram». Her er 29,5 prosent av sensorene uenige om hvilken uttelling besvarelsen skal få. I eksamensveiledningen (Utdanningsdirektoratet 2017a) står det ikke noe spesielt om krav til diagrammer, men på forsensurmøtet ble det lagt til krav om at aksenavn og tittel må være informativ. Dette kan tolkes ulikt av sensorene.

I eksamensveiledningen (Utdanningsdirektoratet 2017, s. 7) står det at oppgavene i del 2 tar utgangspunkt i dagligdagse situasjoner. Sykkelturen i oppgave 2 kan ses på som en dagligdags situasjon, og det dagligdagse svaret på spørsmål b) «Hvor stor del av sykkelturen har Mari og David pause?» ville være 30 minutter av 3,5 timer. I sensorveiledningen slik den forelå på eksamensdagen, må svaret være eller ca. 14,3 prosent i tillegg til korrekt framgangsmåte for å få full uttelling. Dette ble endret på forsensurmøtet slik at 30 minutter ga 1 poeng og 30 minutter av 3,5 timer ga full uttelling. Siden disse svarene kan leses direkte av grafen i oppgaven, kan det være ulik vurdering av «korrekt framgangsmåte» hos sensorene. Hele 30,6 prosent har ulik vurdering av denne oppgaven, mens 10,6 prosent har ulik vurdering av c-oppgaven. Her skal elevene beregne gjennomsnittsfart på hjemturen. Siden det er en oppgave som krever løsning i flere trinn, kan sensorene vektlegge de ulike trinnene ulikt. Vi så også på hele oppgave 2 under ett i tilfelle sensorene hadde valgt å fordele uttellingen forskjellig på deloppgavene, men fant da ulik vurdering for 37,8 prosent av kandidatene.

I oppgave 3b i del 2 er spørsmålet «Hvor mange hele liter rommer jerrykannen?». Et bilde der målene er oppgitt, er utgangspunkt for utregningen. På forsensurmøtet ble det bestemt at 20,16 L skulle godtas som korrekt svar, mens på sensorskoleringen var det flere som var uenige i dette og mente hverdagsforståelsen ble borte når det ikke ble trukket for å oppgi svaret med desimaler. Både den opprinnelige formuleringen i sensorveiledningen der kravet til full uttelling var hele liter, og uenigheten hos noen sensorer kan være utslagsgivende for at 14,4 prosent av kandidatene ble vurdert ulikt av de to sensorene.

Oppgave 8a er den andre av to oppgaver som krever forklaring. På sensorskoleringen ble det kommentert at det var mange fine forklaringer, men også mange lange. 20,5 prosent av kandidatene fikk ulik vurdering av sine forklaringer. De lange forklaringene kan gjøre vurderingen vanskelig. Lange forklaringer kan inneholde en riktig forklaring, men i tillegg mye som ikke er riktig, og vektleggingen kan være utfordrende. Forhåndssensurrapporten (Utdanningsdirektoratet 2017d) gir ingen føringer utover at det er mulig å gi 0, 1 eller 2 poeng, og at flere framgangsmåter må godtas.

I sensorveiledningen som ble publisert om ettermiddagen eksamensdagen (Utdanningsdirektoratet 2017c), var kravet løsning av likningssettet i oppgave 9a ved regning for hånd eller med CAS. Etter forsensurmøtet ble dette endret til full metodefrihet på grunn av oppgavetekstens formulering «løs». Siden løsningen er så åpenbar ved ren betraktning av likningssettet, er den vanskelig å vurdere. I eksamensveiledningen står det at «prøve og feile»-metode / verifisering ved innsetting kan gi noe uttelling dersom det argumenteres for strategien og vises en systematisk tilnærming. Hva dette innebærer, kan det kanskje være vanskelig både for eksamenskandidater og sensorer å avgjøre.

Eksamenens siste deloppgave (9c) er det få som har forsøkt seg på. Av de 217 kandidatene som har fått uttelling fra den ene eller begge sensorene på denne oppgaven, er det 94 som har samsvarende vurdering, og 123 som har fått ulik vurdering. Dette fordeler seg med 23 kandidater som har fått full uttelling fra én sensor og ingen uttelling fra den andre, 70 har fått 1 poeng fra den ene og 0 poeng fra den andre sensoren, mens de siste 30 har fått 1 poeng fra den ene og 2 poeng fra den andre sensoren. Dette har åpenbart vært en vanskelig oppgave å vurdere.

Det er i stor grad samsvar mellom hvilke oppgaver som krever kreativ løsning, og oppgavene der det er ulik vurdering fra de to sensorene. Det er kun to av oppgavene som vi har vurdert at krever algoritmisk løsning, oppgave 3b og 9a i del 2, som ofte får ulik vurdering av sensorene. Begge disse er oppgaver det var diskusjon rundt på sensorskoleringen vi deltok på. De øvrige oppgavene der sensorene har ulik vurdering, er kategorisert som oppgaver som helt eller delvis krever kreativ løsning. På oppgaver der kreativ løsning er nødvendig, vil det også være større variasjon i elevenes besvarelser, og det kan være vanskelig å ta høyde for alle varianter i forhåndssensurrapporten. Dette åpner derfor i større grad for bruk av skjønn hos sensorene.

Oppgaver med bruk av digitale hjelpemidler, oppgaver hvor elevene selv velger en hensiktsmessig metode, og oppgaver som stiller høyere krav til kommunikasjon og begrunnelse, tester viktige deler av den matematiske kompetansen. Dette er deler av den matematiske kompetansen som andre oppgaver ikke tester. Slike oppgaver tester også ofte flere kompetanser på en gang, noe som gjør det nødvendig å gi delvis uttelling. Det er derfor ikke et faglig holdbart alternativ å slutte med slike oppgaver som et tiltak for å få større samsvar mellom sensorene. Ytterligere forbedring av skoleringen av sensorene kan imidlertid være et tiltak for å øke samsvaret.

Hva sier sensorene?

Intern rettferdighet, altså det at besvarelser vurderes ut fra identiske kriterier, som er uavhengige av bosted, er avgjørende for rettssikkerheten. Det avgjørende er følgelig at det ikke skal ha betydning hvilken sensor kandidatene tilfeldigvis får tildelt. Dette universalistiske idealet vil alle være enige i og slutte opp om. Spørsmålet er hvordan utdanningsmyndighetene kan bidra til å sikre at det også skjer i praksis. Det viktigste grepet er ulike tiltak som kalibrerer sensorene, deriblant sensorskolering, som arrangeres på ulike steder rundt om i landet. Her møtes sensorene, og møtet ledes av oppmennene i sensorregionen. Hovedinntrykket fra den tilbakemeldingen vi har fått fra sensorene, på spørreskjemaet vi sendte dem, er at de fleste deltok på sensorskoleringen, og de som deltok, opplevde møtet som viktig og meningsfullt.

På direkte spørsmål til sensorene om deres deltakelse, er det hele ni av ti (91 prosent) som rapporterte at de har vært med på sensorskolering. Når det gjelder den lille andelen som ikke deltok, var det flere forklaringer som ble gitt. Det vet vi fordi vi ba dem utdype. 17 av sensorene benyttet seg av muligheten. Svarene kan kategoriseres i tre ganske ulike grupper. Den første dreier seg som *prioriteringer*: «Andre ved skolen ble prioritert denne gangen», «Det var mange andre fra kollegiet som deltok». En annen forklaring handler om andre og *utenforliggende forhold:* «Kollisjon», «Jeg var påmeldt et annet kurs», «Ferie». Den siste grunnen som oppgis, er at enkelte sensorer tilsynelatende *mener at de kan sensurere*, fordi de har vært på møtet før: «Tenkte jeg hadde nok erfaring», «Jeg deltok i fjor».

At det er noe fravær, er ikke overraskende. Det er derfor grunn til å understreke at fraværet var svært lite, noe som er betryggende. Samtidig forteller ikke dette i seg selv noe om hvorvidt sensorskoleringen gjør en forskjell. Det er derfor viktig å innhente informasjon om hvorvidt sensorene opplevde skoleringen som meningsfull, og i tilfelle hva de legger vekt på som viktig.

For å få svar på sensorenes opplevelse har vi både sendt ut spørsmål som omhandler dette direkte, og bedt dem om å utfylle mer hvordan de opplevde møtet. Totalt var det 129 som benyttet seg av muligheten til å fylle ut åpne svar. I motsetning til spørreskjemaet, hvor vi benytter oss av faste svaralternativer, innebærer åpne svaralternativer at de må bruke egne ord for å beskrive sine erfaringer. Dette kan gi andre innblikk, som er viktige, fordi det ikke er predefinerte kategorier av vår forutgående kunnskap til tematikken.

Den gjennomgående forklaringen fra sensorene var at de opplevde sensormøtet som svært nyttig. Faktisk var det ingen av deltakerne som ga uttrykk for at de var negative til sensormøtet, eller at de ikke så poenget, tvert om. Mange skrev at «det var meget bra gjennomført», eller en annen: «skoleringen er konstruktiv, lærerik og innholdsmessig bra».

Det synes også som om sensormøtet i en del tilfeller hadde direkte implikasjoner på karaktersettingen:

> «Sensorskoleringen og føringene i sensorveiledningen var veldig nyttig. Pga dette hadde jeg og mine medsensorer veldig få besvarelser med ulik karakter. Av 118 besvarelser var det bare 11 ulike karakterer.»

Eller som en annen skrev:

> «Det var veldig nyttig med sensorskolering før vi begynte sensureringen. Det var god veiledning i forhåndssensur-rapporten. Det var noen forandringer som gjorde at en måtte justere en del på den første gruppen en sensurerte.»

I lys av den betydningen flere tilla sensormøtet, er det også interessant å trekke fram noen av forslagene til forbedringer. To innspill gikk igjen. Det som forekom flest ganger, var at flere mente at møtet burde vært avholdt litt tidligere. Det burde vært «midtveis», var det flere som mente, samtidig som de også ga uttrykk for at det kunne være noen utfordringer med å flytte møtet fram i tid.

Det andre innspillet som gikk igjen, var at man ønsket noe mer hjelp til vippekarakterene. Det vil si hvordan man skulle vurdere besvarelser som lå mellom to karakterer. Særlig synes det som det kan være vanskelig å skille mellom 3 og 4.

Oppsummering

- Det er godt samsvar mellom sensorene i deres karakterforslag før fellessensurmøtet, men det er mange oppgaver i del 2 av eksamenen som gir opphav til ulik vurdering av besvarelsene hos de to sensorene.
- Mange av oppgavene som gir ulik vurdering, er oppgaver som krever bruk av digitale hjelpemidler, oppgaver hvor elevene selv velger en hensiktsmessig metode, og oppgaver som stiller høyere krav til kommunikasjon og begrunnelse. Dette er sentrale områder av matematisk kompetanse.
- Det er ikke grunn til å tro annet enn at sensorene etterstreber en mest mulig rettferdig og lik sensur, og flere etterlyser i sin tilbakemelding til Utdanningsdirektoratet bedre veiledning i sensur av enkelte oppgaver. Det gjelder særlig å få klarere retningslinjer for sensurering av oppgaver som krever digitale hjelpemidler.
- Hvis sensorene, som i stor grad også underviser i ungdomsskolen, er usikre på vurderingen, er det sannsynlig at lærer og elever generelt også ville ha nytte av klarere retningslinjer for å vite hva som kreves til eksamen.
- Sensorene ga uttrykk for at de var svært fornøyde med sensorskoleringen. Av 129 sensorer som hadde valgt å forklare hva de mente, var alle positive.

- Flere sensorer ga uttrykk for at sensorskoleringen hadde ført til at de justerte karakterene.
- Et gjennomgående ønske var å ha sensorskoleringen litt tidligere.

9 Avslutning

Det gjennomgående temaet i denne rapporten er hvorvidt eksamenen i matematikk på 10. trinn 2017 har vært rettferdig eller ikke. Nå er ikke rettferdighet et helt enkelt begrep, men i denne sammenhengen har vi valgt å analysere eksamenen med utgangspunkt i spørsmål om *samsvar*. Mer konkret har vi sett på om det er samsvar mellom oppgavene som er gitt til eksamen, og de kompetansemålene elevene har i matematikk. Vi har også vært opptatt av om elevene faktisk har hatt undervisning i alle emnene, eller om det er temaer som de er mindre rustet for å kunne besvare. Samsvar dreier seg imidlertid ikke bare om spørsmålenes innhold, men også deres utforming. Er det rimelig å anta at alle elever forstår hva det spørres om, og fungerer illustrasjonene klargjørende, eller skaper de mer forvirring i unge sinn på prøve?

En rettferdig eksamen

Ut fra de dataene vi har samlet inn, var et hovedfunn i vår analyse at årets eksamen var svært god. Analysene viser i hovedsak godt samsvar med opplæringen. Flere av informantene mente at denne eksamenen burde sette en standard for hvordan oppgavene burde se ut i årene som kommer. Noe som er særlig interessant, er at denne vurderingen er delt av alle de involverte gruppene, fra elever til lærere og sensorer.

Det var flere trekk ved årets eksamen som ble framhevet som avgjørende. For det første var oppgavene av en art hvor alle elevene mestret noen spørsmål. Samtidig var oppgavene satt sammen slik at besvarelsene skilte mellom elever med høyt og lavt mestringsnivå.

For det andre var det ikke for mange spørsmål. De fleste elevene klarte derfor å bli ferdige med oppgavene. Fra de kvalitative dataene gikk det fram at elevene som var intervjuet, opplevde at de hadde fått vist fram det de kunne. Nettopp det var en avgjørende grunn til at flere skrøt av eksamenen. De gikk ikke hjem med en følelse av at de «brant inne» med mye kunnskap de ikke var blitt spurt om.

Et tredje avgjørende forhold var at informantene opplevde at spørsmålene berørte kjernen i kompetansemålene. Det var spørsmål om sentrale emner, og de var stort sett formulert på måter som var forståelige for de fleste av elevene.

Språk er en utfordring for noen elevgrupper

Til tross for at årets eksamen er vurdert som god, er det rom for forbedring. I rapporten har vi lagt særlig vekt på utfordringer knyttet til språk. Mer konkret refererer de språklige utfordringene her til at det i noen oppgaver ble brukt ord som ikke var intuitivt forståelige for en del av eksamenskandidatene, sammen med en del andre språklige trekk som gjør tekster vanskelige. Blant eksemplene på ord fra årets eksamen var «jerrykanne» og «klaffebro». Samtidig skal det trekkes fram her at problemet i stor grad ble løst ved at oppgavene både inneholdt tekst og bilder eller illustrasjoner, som gjorde det mulig å forstå oppgaven uten at man hadde kjennskap til ordene som ble benyttet.

Når vi likevel trekker språk fram, er det fordi informantene var opptatt av tematikken. En del elever har problemer med å forstå hva det spørres om, til forskjell fra å forstå det matematiske spørsmålet. I intervjuene ble det derfor flere ganger tematisert om det er riktig og rimelig at krav til norsk språk skal være en del av kompetansekravene i matematikk. Spørsmål om dette var særlig aktualisert på skoler med en høy andel elever med innvandrerbakgrunn, som i noen tilfeller kunne svare på spørsmålene om de bare skjønte hva de ble spurt om. Generelt er det grunn til å diskutere alle emner hvor bestemte grupper i samfunnet systematisk har mindre mulighet til å lykkes enn andre grupper.

Vellykket sensorskolering

I årets evaluering har vi inkludert spørsmål til sensorene. Svarene vi har fått, viser at sensorskoleringen blir vurdert som svært vellykket. Av 129 sensorer som hadde valgt å forklare hva de mente, var alle positive. Men det skal understrekes at det er godt samsvar mellom sensorene i deres karakterforslag før fellessensurmøtet. Når det gjelder ulike vurderinger, synes det først og fremst å være knyttet til del 2 av eksamenen.

Det var imidlertid flere som etterlyste noe bedre veiledning og fortolkning av enkelte oppgaver. Det gjaldt særlig oppgaver som krever digitale hjelpemidler.

Et forslag – pilotering

Avslutningsvis er det grunn til igjen å understreke at årets eksamen er vurdert som god. Samtidig er det enkelte utfordringer knyttet til enkelte ord. Ut fra erfaringene vi har gjort oss, er det nærliggende å ta til orde for nødvendigheten av pilotering. Nå leses selvsagt oppgavene av en rekke mennesker slik systemet er i dag. Vi mener derfor at det er behov for pilotering av likesinnede. Det vil si at man må hente inn unge folk som skal lese oppgavene og vurdere hvorvidt spørsmålene er forståelige. Nå er det åpenbart noen praktiske og sikkerhetsmessige forhold knyttet til en slik pilotering, men det vil det alltid være. Når vi skriver at unge må trekkes inn i piloteringsarbeidet, impliserer heller ikke det at de må være jevngamle med eksamenskandidatene. Trolig vil det være tilstrekkelig om man rekrutterer folk som er noen få år eldre.

Referanser

Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219–234.

Abedi, J., Bailey, A., Butler, F., Castellon-Wellington, M., Leon, S. & Mirocha, J. (2005). *The validity of administering large-scale content assessments to english language learners: An investigation from three perspectives.* Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Bakke, B. & Bakke, I. N. (2015). *Grunntall 10*. Drammen: Elektronisk Undervisningsforlag AS Tofteberg.

Boesen, J. (2006). *Assessing mathematical creativity: Comparing national and teacher-made tests, explaining differences an examining impact.* Umeå, Sweden: Department of Mathematics and Mathematical Statistics, Umeå University.

Elwood, J., Hopfenbeck, T. & Baird, J.-A. (2017). Predictability in high-stakes examinations: students' perspectives on a perennial assessment dilemma. *Research Papers in Education, 32*(1), 1–17. doi:10.1080/02671522.2015.1086015

Ercikan, K., Chen, M. Y., Lyons-Thomas, J., Goodrich, S., Sandilands, D., Roth, W.-M., et al. (2015). Reading proficiency and comparability of mathematics and science scores for students from english and non-english backgrounds: An international perspective. *International Journal of Testing, 15*(2), 153–175.

Eriksson, N. (2005). *Prestationsskillnader mellan flickor och pojkar i no: En studie av uppgiftsformatets betydelse i timss 2003.* Umeå: Institutionen för beteendevetenskapliga mätningar, Umeå universitet.

Ernst-Slavit, G. & Mason, M. R. (2011). "Words that hold us up:" teacher talk and academic language in five upper elementary classrooms. *Linguistics and Education, 22*(4), 430–440.

Fang, Z., Schleppegrell, M. J. & Cox, B. E. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research, 38*(3), 247–273.

Gee, J. P. (2005). Language in the science classroom: Academic social languages as the heart of school-based literacy. I: R. K. Yerrick & W.-M. Roth (red.), *Establishing scientific classroom discourse communities. Multiple voices of teaching and learning research* (s. 19–37). New Jersey: Lawrence Erlbaum.

Gilje, Ø., Ingulfsen, L., Dolonen, J. A., Furberg, A., Rasmussen, I., Kluge, A., . . . Skarpaas, K. G. (2016). *Med ARK&APP – Bruk av læremidler og ressurser for læring på tvers av arbeidsformer*. Oslo: Universitetet i Oslo, IPED.

Gürsoy, E., Benholz, C., Renk, N., Prediger, S. & Büchter, A. (2013). Erlös = erlösung? Sprachliche und konzeptuelle hürden in prüfungsaufgaben zur mathematik. *Deutsch als Zweitsprache*(1), 14–24.

Haag, N., Heppt, B., Roppelt, A. & Stanat, P. (2015). Linguistic simplification of mathematics items: Effects for language minority students in germany. *European Journal of Psychology of Education, 30*(2), 145–167.

Haag, N., Heppt, B., Stanat, P., Kuhl, P. & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction, 28*, 24–34.

Hagen, K., Johannessen, J. M. & Saidi, A. (2016). Constructing a Norwegian academic wordlist. I: N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk & S. Piperidis (red.), Tenth international conference on language resources and evaluation (lrec 2016). Portorož, Slovenia: European Language Resources Association (ELRA). fra http://www.lrec-conf.org/proceedings/lrec2016/pdf/532_Paper.pdf.

Halliday, M. A. K. (1993). On the language of physical science. I: M. A. K. Halliday & J. R. Martin (red.), *Writing science: Literacy and discursive power* (s. 54–68). Pittsburgh, NJ: University of Pittsburgh Press.

Heppt, B., Haag, N., Böhme, K. & Stanat, P. (2015). The role of academic-language features for reading comprehension of language-minority students and students from low-ses families. *Reading Research Quarterly, 50*(1), 61–82.

Heuvel-Panhuizen, M. v. d. (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics, 25*(2), 2–9.

Hipkiss, A. M. (2014). *Klassrummets semiotiska resurser. En språkdidaktisk studie av skolämnena hem- och konsumentkunskap, kemi och biologi*. Umeå Universitet, Umeå.

Hægeland, T., Kirkebøen, L. J. & Raaum,O. (2010). *Skolebidragsindikatorer for videregående skoler i Oslo.* Rapport 2010/36. Statistisk sentralbyrå.

Institutt for lærerutdanning og skoleforskning ved Universitet i Oslo (2015). Rammeverket for TIMSS – kortversjon. Hentet fra https://www.uv.uio.no/ils/forskning/prosjekt-sider/timss-norge/TIMSS/2015/rammeverk2015.html

Jorgensen , R. (2011). Language, culture and learning mathematics: A bourdieuian analysis of indigenous learning. I: C. Wyatt-Smith, J. Elkins & S. Gunn (red.), *Multiple perspectives on difficulties in learning literacy and numeracy* (s. 315–329). Dordrecht: Springer Netherlands. fra http://dx.doi.org/10.1007/978-1-4020-8864-3_15.

Kieffer, M. J., Lesaux, N. K., Rivera, M. & Francis, D. J. (2009). Accommodations for english language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168–1201.

Kongelf, T. R. (2015). Introduksjon av algebra i matematikkbøker for ungdomstrinnet i norge. *Nordisk Matematikkdidaktikk, 20*(3–4), 83–109.

Kunnskapsdepartementet. (2015). Tett på realfag. Nasjonal strategi for realfag i barnehagen og grunnopplæringen. https://www.regjeringen.no/no/dokumenter/tett-pa-realfag/id2435042/

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. I: J. Coady & T. Huckin (red.), *Second language vocabulary acquisition: A rationale for pedagogy* (s. 20–34). Cambridge: Cambridge University Press.

Lepik, M. (1990). Algebraic word problems: Role of linguistic and structural variables. *Educational Studies in Mathematics, 21*(1), 83–90.

Li, H. & Suen, H. K. (2012). The effects of test accommodations for english language learners: A meta-analysis. *Applied Measurement in Education, 25*(4), 327–346.

Liberg, C. (2001). Svenska läromedeltexter i ett andraspråksperspektiv – möjligheter och begränsningar. I: K. Nauclér (red.), *Symposium 2000 – ett andraspråksperspektiv på lärande* (s. 108–128). Stockholm: Sigma Förlag.

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

Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J. & Deaktor, R. (2007). Cultural and home language influences on children's responses to science assessments. *Teachers College Record, 109*(4), 897–926.

Maagerø, E. & Skjelbred, D. (2010). *De mangfoldige realfagstekstene: Om lesing og skriving i matematikk og naturfag*. Bergen: Fagbokforlaget.

Maagerø, E. & Tønnessen, E. S. (2015). *Å lese i alle fag (2. utg.)*. Oslo: Universitetsforlaget.

Martin, M. O. & Mullis, I. V. S. (2011). *Timss and pirls 2011: Relationships among reading, mathematics, and science achievement—implications for early learning*. TIMSS and PIRLS, 1–11.

Martiniello, M. (2008). Language and the performance of english-language learners in math word problems. *Harvard Educational Review, 78*(2), 333–368.

Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for english language learners in math tests. *Educational Assessment, 14*(3–4), 160–179.

Matematikksenteret. (2015a). *Vurdering av eksamen i matematikk*. Hentet fra http://www.matematikksenteret.no/content/5769/Vurdering-av-eksamen-i-matematikk

Matematikksenteret. (2015b). *Vedlegg til rapport «vurdering av eksamen i matematikk, matematikksenteret 2015».*

Mullis, I. V. S. & Martin, M. O. (2011). *Timss 2011 item writing guidelines.* Boston: Boston College.

Mullis, I. V. S. & Martin, M. O. (red.) (2013). *TIMSS 2015 Assessment Frameworks.* Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/frameworks.html (01.08.2017).

Niederhaus, C., Pöhler, B., Prediger, S., Tschirner, E., Bärenfänger, O. & Möhring, J. (2015). *Relevante sprachmittel für mathematische textaufgaben–korpuslinguistische annäherung am beispiel prozentrechnung. Kompetenzprofile Deutsch als fremde Bildungssprache*. Stauffenberg: Tübingen.

Nilsen, T., Grønmo, L. S. & Hole, A. (2013). Læringstrykk og prestasjoner i matematikk og naturfag. I L. S. Grønmo & T. Onstad (red.), *Opptur og Nedtur. Analyser av TIMSS data fra Norge og Sverige* (ss. 19–51). Oslo: Akademika forlag.

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

Nyström, P. (2008). Identification and analysis of text-structure and wording in timss-items. Paper presentert på 3rd IEA International Research Conference.

Palm, T.,Boesen, J. & Lithner, J. (2011). Mathematical Reasoning Requirements in Swedish Upper Secondary Level Assessments, *Mathematical Thinking and Learning*, *13*(3), 221–246. doi: 10.1080/10986065.2011.564994

Persson, T. (2016a). *De naturvetenskapliga ämnesspråken de naturvetenskapliga uppgifterna i och elevers resultat från timss 2011 år 8*. Uppsala: Uppsala Universitet.

Persson, T. (2016b). The language of science and readability. Correlations between lingustic features in timss science items and the performance of different groups of swedish 8th grade students. *Nordic Journal of Literacy Research, 2*(1).

Persson, T., af Geijerstam, Å. & Liberg, C. (2016). Features and functions of scientific language(s) in timss 2011. *NorDiNa, 12*(2), 176–196.

Prediger, S. (2013). Sprachmittel für mathematische verstehensprozesse – einblicke in probleme, vorgehensweisen und ergebnisse von entwicklungsforschungsstudien. I: A. Pallack (red.), *Impulse für eine zeitgemäße mathematiklehrer-ausbildung. Mnu-dokumentation der 16. Fachleitertagung mathematik* (s. 26–36). Neuss: Seeberger.

Prediger, S., Renk, N., Büchter, A., Gürsoy, E. & Benholz, C. (2013). Family background or language disadvantages? Factors for underachievement in high stakes tests. I: A. Lindmeier (red.), *Mathematics learning across the life span : Proceedings of the 37th conference of the international group for the psychology of mathematics education ; july 28 – august 02, 2013. 4* (s. 49–56). Kiel.

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E. & Benholz, C. (2015). Sprachkompetenz und mathematikleistung – empirische untersuchung sprachlich bedingter hürden in den zentralen prüfungen 10. *Journal für Mathematik-Didaktik, 36*(1), 77–104.

Schoenfeld, A. (2007). *Assessing mathematical proficiency*. New York: Cambridge University Press.

Schoultz, J., Säljö, R. & Wyndhamn, J. (2001). Conceptual knowledge in talk and text: What does it take to understand a science question? *Instructional Science, 29*(3), 213–236.

Serder, M. & Jakobsson, A. (2015). "Why bother so incredibly much?": Student perspectives on pisa science assignments. *Cultural Studies of Science Education, 10*(3), 833–853.

Shaftel, J., Belton-Kocher, E., Glasnapp, D. & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of english language learners and students with disabilities. *Educational Assessment, 11*(2), 105–126.

Shanahan, T. & Shanahan, C. (2012). What is disciplinary literacy and why does it matter? *Topics in language disorders, 32*(1), 7–18.

Skjelbred, D. & Aamotsbakken, B. (2010). *Lesing av fagtekster som grunnleggende ferdighet*. Oslo: Novus.

Smestad, B. (2002). *Matematikkhistorie i grunnskolens lærebøker: En kritisk vurdering*. [Alta]: Høgskolen i Finnmark Avdeling for nærings- og sosialfag. (Hif-rapport 2002:1).

Solano-Flores, G. & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of english-language learners. *Educational Researcher, 32*(2), 3–13.

Tofteberg, G. N., Tangen, J., Stedøy-Johansen, I. M., Alseth, B. (2015). *Maximum 10*. Oslo: Gyldendal Undervisning.

Townsend, D., Filippini, A., Collins, P. & Biancarosa, G. (2012). Evidence for the importance of academic word knowledge for the academic achievement of diverse middle school students. *The Elementary School Journal, 112*(3), 497–518.

Utdanningsdirektoratet (2009). Eksamen MAT0010 Matematikk, 13.05.2009.

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

Utdanningsdirektoratet. (2014). Matematikk i norsk skole anno 2014. http://www.udir.no/tall-og-forskning/finn-forskning/rapporter/Matematikk-i-norsk-skole-anno-2014/

Utdanningsdirektoratet (2015). Eksamen MAT0010 Matematikk, 20.05.2015. https://dok.udir.no/EksamensOppgaver.aspx?proveType=EG&provePeriode=Alle

Utdanningsdirektoratet (2016). Eksamen MAT0010 Matematikk, 20.05.2016.

Utdanningsdirektoratet (2017a). Eksamensveiledning – om vurdering av eksamensbesvarelser 2017, MAT0010 Matematikk, Sentralt gitt skriftlig eksamen. https://sokeresultat.udir.no/eksamensoppgaver.html?k=MAT0010&r1=%C7%82%C7%824772756e6e736b6f6c65&r1val=Grunnskole&r3=%C7%82%C7%82323031372d31&r3val=V%C3%A5r%202017

Utdanningsdirektoratet (2017b). Eksamen MAT0010 Matematikk, 16.05.2017. https://sokeresultat.udir.no/eksamensoppgaver.html?k=MAT0010&r1=%C7%82%C7%824772756e6e736b6f6c65&r1val=Grunnskole&r3=%C7%82%C7%82323031372d31&r3val=V%C3%A5r%202017

Utdanningsdirektoratet (2017c). Sensorveiledning MAT0010 Matematikk 16.05.2017. https://sokeresultat.udir.no/eksamensoppgaver.html?k=MAT0010&r1=%C7%82%C7%824772756e6e736b6f6c65&r1val=Grunnskole&r3=%C7%82%C7%82323031372d31&r3val=V%C3%A5r%202017

Utdanningsdirektoratet (2017d). Forhåndssensurrapport MAT 0010 Matematikk, 31.05.2017. https://sokeresultat.udir.no/eksamensoppgaver.html?k=MAT0010&r1=%C7%82%C7%824772756e6e736b6f6c65&r1val=Grunnskole&r3=%C7%82%C7%82323031372d31&r3val=V%C3%A5r%202017

Waag, T. & Myklebust, F. (2017). *Oppmannsrapport etter fellessensur i Region 7, Sogn og Fjordane/Møre og Romsdal, 13.06.2017.* https://www.fylkesmannen.no/PageFiles/857044/Oppmannsrapport%202017%20-%20Matematikk.pdf

Wang, Z. & Shah, P. (2014). The effect of pressure on high- and low-working-memory students: An elaboration of the choking under pressure hypothesis. *British Journal of Educational Psychology, 84*(2), 226-238. doi:10.1111/bjep.12027

Wolf, M. K. & Leon, S. (2009). An investigation of the language demands in content assessments for english language learners. *Educational Assessment, 14*(3-4), 139-159.

Zevenbergen, R. (2000). 'Cracking the code' of mathematics: School success as a function of linguistic, social and cultural background. I: J. Boaler (red.), *Multiple perspectives on mathematics teaching and learning* (s. 201-221). New York: JAI/Ablex.

På prøve

I denne rapporten evaluerer vi matematikkeksamen gitt på 10. trinn våren 2017. Overordnet spør vi om årets eksamen er og oppleves å være rettferdig. Mer konkret undersøker vi om det er samsvar mellom oppgavene som er gitt til eksamen, elevenes kompetansemål og den opplæringen de har fått. I tillegg ser vi på hvordan eksamen er utformet, både med tanke på illustrasjoner, layout og bruken av tekstoppgaver, og vi vurderer arbeidsmengde og vanskegrad. Til slutt tar vi for oss sensorenes erfaringer og utfordringer med hensyn til skolering i forkant og deres vurdering av besvarelsene. I to følgende rapporter skal vi se på eksamen i 2018 og i 2019.

HØGSKOLEN I OSLO
OG AKERSHUS

Fafo
Borggata 2B/Postboks 2947 Tøyen
N-0608 Oslo
www.fafo.no

Fafo-rapport 2017:36
ISBN 978-82-324-0410-0
ISSN 0801-6143
Bestillingsnr. 20644