

Rapport fra piloteringer og gjennomføring av den nasjonale prøven i lesing på 5. trinn 2018

**Benedikte Homme Fevolden og
Jostein Andresen Ryen**



Enhet for kvantitative utdanningsanalyser
Institutt for lærerutdanning og skoleforskning

Universitetet i Oslo

1.03.19

Innhold

1. Innledning	1
Måling av leseforståelse i de nasjonale prøvene i lesing	1
Oppgaveformater	2
Leseprosesser og eksempeltekst	2
Generaliserbarhet	4
Prøveutvikling i perioden 2017–2018	5
Den nasjonale prøven i lesing for 5. trinn 2018	5
Analyser i rapporten	5
2. Første pilotering	7
Innholdet i prøve A og B	7
Klassisk item-analyse av prøve A og B	10
Utvalg av tekster og oppgaver til andre pilotering	12
3. Andre pilotering	15
Innholdet i prøve C	15
Endringer fra andre pilotering til endelig prøve	17
4. Forslag til endelig prøve for 2018	20
Interessante leseutfordringer i forslaget til endelig prøve	21
Oppgave til teksten «Eplepai»	21
Oppgave til teksten «Ensomme George»	22
Oppgave til teksten «Jeg hater lukten av røyk!»	23
Oppgave til teksten «Skyggen»	24
Oppgave til teksten «Hva er penger?»	25
IRT-analyse	26
Oppgaver med lav diskriminering	27
Oppgaver med DIF	28
Oppgavene fordelt på mestringsnivåer	28
Testinformasjon og måleusikkerhet	29
Prøvens TRF (Test response function)	30
5. Hovedgjennomføring	32
Resultater for prøven som helhet	32
Kjønnsforskjeller	34
Klassisk item-analyse av prøven 2018	35
DIF-analyser av oppgavene i prøven 2018	36
Oppgavene fordelt på mestringsnivåer	37
Vedlegg 1: Beskrivelser av lesekompetanse på de fem ulike nivåene	38
Referanser	39

1. Innledning

Denne rapporten beskriver og dokumenterer de lesefaglige og psykometriske vurderingene som ligger til grunn for utvikling av den nasjonale prøven i lesing på 5. trinn 2018, fra første pilotering til endelig prøve.

I innledningen gir vi først en oversikt over prøvespesifikasjonen og viser en eksempeltekst med oppgaver. Deretter gir vi en vurdering av generaliserbarhet fra årets leseprøve og presenterer kort hvilke analyser som er gjort – deriblant validitet, reliabilitet, utvalg til piloteringer, ankertekster og IRT-analyser.

I kapittel 1 – 4 dokumenterer vi utviklings- og kvalitetssikringsprosessen, som er grunnlag for dialog mellom prøveutviklere, eksterne kvalitetssikrere og Utdanningsdirektoratet. I kapittel 5 beskriver vi resultatene fra hovedgjennomføringen av prøven i oktober-november 2018.

Måling av leseforståelse i de nasjonale prøvene i lesing

De nasjonale prøvene i lesing måler i hvilken grad elevenes leseforståelse er i samsvar med beskrivelsen av lesing som grunnleggende ferdighet i læreplanene og i rammeverket for grunnleggende ferdigheter. Siden disse beskrivelsene er relativt generelle, uttrykker tekstutvalget og oppgavene i prøven prøveutviklernes fortolkning av dem. Kunnskap om den aktuelle aldersgruppas lesekompetanse ligger også til grunn for tekstutvalget og utformingen av prøven. Prøvene skal også gjenspeile de kravene som stilles til den leseforståelsen som elevene trenger for å klare seg, både på skolen og i samfunnet ellers.

Lesing som grunnleggende ferdighet i alle fag er beskrevet i «Læreplanverket for Kunnskapsløftet» (Utdanningsdirektoratet 2013) og spesifiseres nærmere i «Rammeverk for grunnleggende ferdigheter» (Utdanningsdirektoratet 2017a). Lesefaglige og psykometriske krav til de nasjonale leseprøvene defineres nærmere av rammeverket for nasjonale prøver (Utdanningsdirektoratet 2017b), både generell del (del 1), om prøvenes innhold (del 2.4) og de tekniske kravene (del 3), samt i metodegrunnlaget for nasjonale prøver (Utdanningsdirektoratet 2018). I rammeverket defineres lesing slik:

Å lese handler om å kunne forstå, bruke, reflektere over og engasjere seg i innholdet i tekster. Tekster inkluderer alt som kan leses i ulike medier, ikke bare ord, men også illustrasjoner, symboler eller andre uttrykksmåter. Kunnskap om hva som kjennetegner ulike typer tekster og deres funksjon, er en viktig del av lesing.

Prøvene skal måle lesing i alle fag i tråd med læreplanen, og dette skal gjenspeiles i et bredt utvalg av tekster. I beskrivelsen av lesing som grunnleggende ferdighet i ulike fag, legges det vekt på elevenes evne til å finne fram til informasjon, forstå og tolke tekster, samt reflektere over form og innhold i et bredt spekter av tekster fra ulike fagområder. Disse utdragene fra læreplanen i naturfag, norsk, matematikk og samfunnsfag er eksempler på hvordan lesing som grunnleggende ferdigheter er nedfelt i fagplanene (Utdanningsdirektoratet, 2013):

- Å kunne lese i naturfag innebærer å forstå og bruke begreper, identifisere, tolke og bruke informasjon og å kunne skille mellom antakelser, påstander og konklusjoner.

- Å kunne lese i matematikk innebærer å kunne forstå symbolspråk, sortere informasjon og sammenfatte informasjon fra sammensatte tekster med flere uttrykkselementer.
- Å kunne lese i norsk innebærer å kunne skape mening i et bredt utvalg sjangre og å kunne tilpasse lesingen til formålet og med ulike tekster.
- Å kunne lese i samfunnsfag innebærer å kunne sammenligne informasjon fra ulike kilder og vurdere relevans og troverdighet.

I eksemplene ovenfor går det fram at den kritiske og analytiske lesekompetansen, samt evne til å integrere informasjon og trekke slutninger på bakgrunn av ulike tekster og tekstformater er sentral.

Oppgaveformater

I de nasjonale leseprøvene for 5. trinn er det kun to aktuelle oppgaveformater: flervalgsoppgaver og sammensatte flervalgsoppgaver. En flervalgsoppgave består av et spørsmål og fire svaralternativer der det bare er ett svar som er riktig, mens en sammensatt flervalgsoppgave består av påstander som elevene skal ta stilling til ved å krysse av for ett av to svaralternativ, for eksempel «riktig» eller «galt». Sammensatte flervalgsoppgaver benyttes til å behandle spørsmål der flere elementer er likestilt i teksten, for eksempel å koble utsagn til riktig stemme i teksten.

Begge de to oppgavetyperne blir skåret automatisk og er velegnet for leseprøver, både fordi elevens skriveferdigheter ikke påvirker resultatet, og fordi oppgavene vurderes uten lærerens skjønn. Det er et mål at elevene skal få svaralternativer og påstander presentert i rekkefølger som *enten* gjenspeiler den logiske orden i teksten (etterfølgende årstall, forekomst i teksten eller andre logiske sammenhenger) *eller* i en tilfeldig rekkefølge som ikke fremmer gjetting (tilfeldig plassering av riktig svar A–D). I det endelige prøveforslaget er logiske rekkefølger kontrollert, og for øvrig er plassering av riktige svar tilfeldig. Vi gjør analyser av svarmønstre og dyktighet for alle oppgaver.

Leseprosesser og eksempeltekst

Alle oppgavene i prøven er knyttet til en lesemåte som sier noe om hvilken leseoperasjon som kreves for å løse oppgaven. En oppgave skal enten måle elevens evne til å finne informasjon, tolke og sammenholde informasjon eller å reflektere over og vurdere teksters form og innhold. Disse tre leseprosessene er omtalt som *aspekter* i Rammeverk for nasjonale prøver, og er beskrevet slik:

Å finne informasjon i tekster innebærer å finne fram til informasjon som er eksplisitt eller implisitt uttrykt. Selv om informasjon i tekster er eksplisitt, kan det være vanskelig å lokalisere informasjonselementer i tekster som inneholder konkurrerende og til dels motstridende informasjon.

Å tolke og sammenholde informasjon innebærer å trekke slutninger på bakgrunn av én eller flere tekster. Dette leseaspektet spenner fra å trekke enkle slutninger mellom innholdet i to setninger som følger etter hverandre, til de mest kompliserte slutningene vi vanligvis forbinder med å «lese mellom linjene».

Å reflektere over og vurdere teksters form og innhold innebærer å forholde seg selvstendig til tekster. Dette leseaspektet spenner fra å kommentere teksters innhold med utgangspunkt i egne meninger til å forholde seg kritisk til tekster som helhet og begrunne egne synspunkter, analyser eller vurderinger av tekster.

De tre leseaspektene finnes ikke som atskilte ferdigheter hos leserne. Det er derfor ikke et mål å vurdere disse som avgrensede delkompetanser. Alle oppgavene i de nasjonale prøvene i lesing skal likevel være kategorisert etter hvilket leseaspekt som er den mest fremtredende i oppgavene. Alle de tre leseaspektene skal være representert i en prøve, men ikke nødvendigvis i samme omfang.

(Utdanningsdirektoratet, 2017b)

Tre eksempeloppgaver til teksten «Plastkuler på tannbørsten?» som var en del av leseprøven for 5. trinn i 2017, illustrerer de tre ulike lesemåtene. De riktige svarene er markert med en sirkel.

Denne teksten er skrevet av Jenny Jordahl og Ole Mathismoen og stod på trykk i Aftenposten Junior i mai 2016. Bruk teksten når du svarer på oppgavene.

Plastkuler på tannbørsten?



Du eier garantert mange ting av plast! På kjøkkenet hjemme er det drossevis av plastdingser. Pennalet, kosedyret og skoene dine inneholder også plast. Vi er avhengig av plast.

Det tar mange, mange år før plast går i oppløsning hvis det havner i naturen. Plast som havner i havet, løses opp i tusenvis av nesten usynlige mikrobiter etter lang tid.

Spytter ut plast

De fleste vet at det er teit å kaste plast i naturen eller i havet. Men kanskje spytter du ut plast når du pusser tennene? Uten at du aner det. Det er nemlig plastkuler i noen tannkremer.

Og det er stor sjanse for at det er plast i sjampoene eller balsamen du bruker. Eller i pappas barberskum eller mammas maskara eller skrubbkrem. Og hva med familiens solkrem?

Umulig å se

I mange kosmetikk- og pleieprodukter er det nemlig tilsatt mikroskopiske plastbiter. De hjelper blant annet til å skrubbe tenner og hud.

De er veldig små, oftest mindre enn én millimeter. Det er nesten umulig å se dem. Du må ha mikroskop for å få øye på dem.

Ender opp i havet

Alt vann vi har brukt – i vasken, i dusjen og i do – renner i rør til et renseanlegg før det slippes ut i havet. Problemet er at renseanleggene ikke klarer å fjerne disse små plastkulene.

Derfor ender de opp i havet. Der suger de til seg miljøgifter. Ofte klumper de seg sammen og ser så spennende ut at fisk spiser dem. Miljøgiftene på kulene tas opp i fiskekjøttet. Og kanskje fisker vi fisken og spiser den.

Altså: Vi pusser tennene med tannkrem med plastkuler. Vi spytter ut og forurensrer havet og fisken. Da får vi fisk med gift i på middagstallerkenen. Ikke bra!

Hvordan finne riktig tannkrem?

I USA er det forbudt å bruke plastkuler i kosmetikk og pleieprodukter. I Norge er det fortsatt lov.

Men hvordan kan du vite om det er plastkuler i tannkremen eller solkremen? Den enkleste måten er å bruke en app på mobilen. Be voksne rundt deg installere appen «Beat the Microbead». Med den kan du skanne strekkoden bak på produktet og sjekke.

Vi skannet to barnetannkremer med appen. Én hadde plastkuler, én hadde det ikke.

1 Hva anbefaler forfatterne deg å gjøre for å finne ut om det er plastkuler i tannkremen?

(Finne informasjon)

- A å lese informasjonen som står på tannkremtuben
- ☒ B å skanne strekkoden på tannkremtuben med en app
- C å spørre en ansatt i en dagligvarebutikk
- D å studere tannkremen med et forstørrelsesglass

2 Hva vil den ene dykkeren vise fram gjennom forstørrelsesglasset sitt?

(Tolke og sammenholde informasjon)

- ☒ A en klump som inneholder plast og miljøgifter
- ☐ B tannkremen som jenta spyttet ut i vasken
- ☐ C små dyr i havet som fiskene spiser
- ☐ D småfisk som spiser miljøgifter

3 I teksten står det at det er forbudt å bruke plastkuler i tannkrem i USA. Hvorfor nevner forfatterne dette?

(Reflektere over form og innhold)

- ☒ A Forfatterne syns plastkuler burde vært forbudt i Norge også.
- ☐ B Teksten var opprinnelig skrevet for amerikanske lesere.
- ☐ C Forfatterne hadde vært i USA like før de skrev teksten.
- ☐ D Politikere i USA har blitt påvirket av det norske forbudet mot plastkuler.

Generaliserbarhet

Vurderinger av leseprøvenes validitet er undersøkelser av om det som måles, korresponderer med fenomener i virkeligheten og om det kan trekkes gyldige slutninger om elevenes leseforståelse i andre kontekster (Silverman 2011: 367). Tekst- og oppgaveutvalget i nasjonal leseprøve for 5. trinn 2018 er i tråd med beskrivelser av lesing som grunnleggende ferdighet og kompetansemål i fagene etter 4.trinn. Prøven inneholder både lette og vanskelige oppgaver, og måler dermed både svake og sterke elevers leseferdigheter. Prøvens innhold favoriserer verken gutter eller jenter i særlig grad, men måler balansert. Denne rapporten dokumenterer alle detaljer i de undersøkelsene av reliabilitet og validitet som er foretatt i løpet av prøveutviklingen.

Reliabilitet er forbundet med målesikkerhet, altså hvor sannsynlig er det at de samme funnene ville ha framkommet om den samme målingen ble gjentatt mange ganger. Vurderings-reliabilitet spiller ikke inn i en prøve med automatrettede flervalgsoppgaver. Det endelige prøvens reliabilitetskoeffisient er 0,860 (Cronbachs alpha), altså høy.

Trusler mot prøvens reliabilitet og validitet er i første rekke knyttet til prøvegjennomførings-plattformen. Den valgte presentasjonsformen setter rammer for hvordan autentiske tekster kan presenteres, og få oppgaveformater begrenser muligheten til å stille ellers aktuelle spørsmål som utfordrer elevenes leseferdighet.

Prøveutvikling i perioden 2017–2018

For å få et så representativt utvalg som mulig, inviterer vi pilotskoler med ulikt elevantall fra hele landet. Skolene er tilfeldig utvalgt fra en liste gitt av Utdanningsdirektoratet. Prøven for 2018 er pilotert i to omganger: første gang i september 2017 og andre gang i mars 2018. I tillegg ble det inkludert en tekst som ble pilotert for de nasjonale prøvene i lesing 2019. Denne teksten ble også pilotert i mars 2018. I alle tilfeller deltok elever på 5. trinn, henholdsvis like gamle og et halvt år eldre som målgruppen. Vi vet at elevene blir mer modne og utvikler bedre leseferdigheter i løpet av skoleåret, og det er derfor viktig å være klar over denne aldersforskjellen når man tolker resultatene fra de to piloteringene.

Ved siden av disse to større utprøvingene kan øyebevegelsesstudier, intervjuer med elever og lærere være verdifulle ressurser i prøveutviklingsarbeidet. Til denne prøven har vi foreløpig ikke tatt i bruk disse utprøvningsformene på grunn av omlegging til nytt årshjul. Prøvene er kvalitetssikret av eksterne lesefaglige eksperter ved Uppsala universitet, av en referansegruppe bestående av fire lærere på mellomtrinnet og av ekstern psykometrisk ekspert ved NTNU.

Ved gjennomføring inngår også såkalte ankertekster i leseprøven – altså tekster og oppgaver som er like fra år til år – og som behandles i en egen underrapport av hensyn til konfidensialitet.

Den nasjonale prøven i lesing for 5. trinn 2018

Den nasjonale prøven i lesing for 5. trinn 2018 består av 5 tekster med til sammen 29 oppgaver som måler elevenes leseforståelse, i tråd med læreplanen for 2.–4. trinn. Det har vært et mål at prøven skal gjenspeile det tekstmangfoldet som elevene møter i ulike fag, og de utvalgte tekstene kan knyttes til fagene norsk, samfunnsfag, naturfag og mat og helse. Det er som vanlig én skjønnlitterær tekst og fire sakprosattekster med i utvalget. Flertallet av sakprosattekstene er satt sammen av verbaltekst og andre tekstressurser som figurer, faktabokser og grafiske framstillinger. Prøven har en tradisjonell fordeling av oppgavene knyttet til de tre aspektene med 10 finne-, 13 tolke- og 6 refleksjonsoppgaver. Ved siden av at alle oppgaver skal tilfredsstille lesefaglige og psykometriske krav, legges det stor vekt på at prøven som helhet representerer en bredde i fagtilhørighet, sjangre, tema og leseutfordringer. Tekstene i årets prøve, og hvordan prøven måler lesing som grunnleggende ferdighet i flere fag, blir presentert i kapitlet «Utvalg av tekster og oppgaver til andre pilotering»

Analyser i rapporten

I tråd med rammeverket skal prøvene utformes etter etablerte standarder for prøvekonstruksjon, med bruk av egnede psykometriske metoder. Oppgaveutviklingen skal skje med klassiske mål for reliabilitet og oppgaveegenskaper, i tillegg til IRT-analyser av enkeltoppgaver og prøvene i sin helhet (Utdanningsdirektoratet, 2017b). Resultatene blir rapportert i form av p-verdier og mål på diskriminering (biserial correlation), samt mål for

reliabilitet/indre konsistens (Cronbachs alpha). De klassiske analysene er utført med programmet SPSS 25, IRT-analysene med programmet XCalibre 4,2.

P-verdi refererer til prosentandel elever som har riktig svar på en oppgave. For polytome oppgaver, som skåres med 0, 1 eller 2 poeng, beregnes p-verdien ved å legge sammen andelen elever som fikk full skår (2 poeng) og 50 prosent av elevene som fikk delvis riktig svar (1 poeng).

Diskriminering viser i hvor stor grad en oppgave måler det samme som helheten, altså hvor godt oppgaven skiller mellom sterke og svake lesere (totalt på hele prøven). Diskrimineringsmål viser korrelasjon mellom oppgaveskåre og totalskåre. For dikotome oppgaver, det vil si oppgaver som skåres med enten 0 eller 1 poeng, kalles en slik korrelasjonskoeffisient «point-biserial». Oppgaver med høy diskriminering gir høy reliabilitet for prøven. En oppgave med veldig lav diskriminering ($< 0,250$) kan være «bortkastet» ved ikke å skille godt mellom ulike lesere, men kan likevel være pedagogisk eller lesefaglig forsvarlig. Dette kan for eksempel gjelde svært lette oppgaver som nesten alle lesere mestrer, men som gir en god og oppklarende introduksjon til teksten og oppgavesettet. For høy diskriminering ($> 0,7$) kan være indikasjon på lav validitet for hele prøven ved at oppgavene kan være for like og derfor dekke konstruert for dårlig. Polytome oppgaver, som skåres med 0, 1, 2 eller flere poeng, gir som regel høyere diskriminering, fordi de gir mer informasjon.

Et **reliabilitetsmål** er en indikasjon på hvor konsistent prøven måler den latente variabelen, leseforståelse, som et uttrykk for sann varians. En korrelasjonskoeffisient som Chronbachs alpha beregnes fra gjennomsnittet av alle mulige korrelasjonskoeffisienter mellom to «halvdeler» av prøven. En tilfredsstillende alpha for leseprøven bør ligge mellom 0,8 og 0,9.

Ved hjelp av IRT kan man estimere oppgavens vanskegrader uavhengig av elevene som deltok, og man vil kunne beregne elevers dyktigheter uavhengig av nøyaktig hvilke oppgaver de besvarte. I IRT-analysene av oppgavene i prøven brukes en to-parameter-modell (2PL). Modellen viser oppgavens vanskegrad (b), oppgavens diskriminering (a), det vil si i hvilken grad oppgaven diskriminerer mellom svake og sterke elever (Utdanningsdirektoratet, 2018, s. 5). En IRT-analyse gjør det mulig å utforske om oppgavene og dermed prøven som helhet favoriserer en bestemt gruppe framfor en annen, for eksempel jenter framfor gutter, gitt at representanter for begge grupper har samme underliggende ferdighet (Utdanningsdirektoratet, 2018, s. 9).

IRT-analysen leverer resultater i form av såkalte theta-verdier (θ) som er omtrent fra -3 til +3 med et gjennomsnitt på 0 for estimering av elevprestasjoner. Oppgavens vanskegrader er også uttrykt på samme skala. Etter en enkel omregningsnøkkel regnes disse verdiene om til positive verdier, såkalte skalapoeng (Utdanningsdirektoratet 2018).

IRT-analysen leverer en sum av informasjonsverdier og målefeil for hele ferdighetsfordelingen i en prøve. Men den leverer også en såkalt TRF, en Test Respons Funksjon, som er en god beskrivelse av sammenhengen mellom de skalerte skårene i form av skalapoeng og antall riktig besvarte oppgaver på prøven (Utdanningsdirektoratet 2018).

2. Første pilotering

I den første piloteringen ble det prøvd ut 10 tekster og 73 oppgaver fordelt på prøve A og prøve B. Piloteringen ble gjennomført i *PAS – Prøver*, som er Utdanningsdirektoratets prøvegjennomføringssystem, i september 2017. Prøvetiden var 90 minutter, som ved ordinær gjennomføring. 560 elever fra 12 skoler i 9 fylker deltok i den første piloteringen. Skolene meldte selv opp elevene i PAS, som fordelte elevene jevnt på de to prøvene A og B.

Noen av tekstelementene i tekstene *Panda* og *Farger* ble satt i et såkalt svg-format i editor. Flere skoler meldte tilbake at de ikke klarte å laste inn akkurat disse tekstene, og vi antar at det skyldes svg-formatet. Tallene knyttet til disse tekstene er derfor usikre, ettersom elevene ikke klarte å besvare oppgavene. Usikkerheten rundt tallene, spesielt knyttet til teksten *Farger*, har gjort det vanskelig å plukke ut tekster til den siste piloteringen. Som vi kommer tilbake til senere i rapporten, foreslår vi å bytte ut denne teksten med en annen tekst som ble prøvd ut til den nasjonale prøven for 5. trinn i 2019.

Tabell 1: Oversikt over utvalget av elever (N=560) pilot prøve A og B.

	Alle
Antall inviterte elever	1049 (fra 22 skoler)
Antall oppmeldte elever	667 (fra 14 skoler)
Antall elever som gjennomførte (N)	Totalt 560 Prøve A: 289 (139 gutter, 150 jenter) Prøve B: 271 (140 gutter, 131 jenter)
Antall skoler	12

Innholdet i prøve A og B

Tekstene fordeler seg på fag og teksttyper som vist i tabell 2 og 3. Det er både oppgitt navn for tekstene som brukes i rapporten og original tittel. For hver tekst er det her oppgitt en p-verdi og kjønnsforskjell, basert på den gjennomsnittlige p-verdien til hver enkelt oppgave som er knyttet til teksten. Oppgavenes vanskegrad avhenger både av kvaliteter ved teksten og oppgaveformuleringen, og det vil som regel være både lette og vanskelige oppgaver knyttet til én og samme tekst. Den gjennomsnittlige p-verdien gir derfor ikke et presist bilde av hvor lett eller vanskelig en tekst er for elevene, men i vurderingen av hvilke tekster som passer sammen i en prøve, kan det være nyttig å se på gjennomsnittlig p-verdi for hver tekst og slik beregne vanskegraden totalt sett.

Tabell 2: Oversikt over tekstene i prøve A.

Tekstnavn i rapport	Original tittel	Fagtilhørighet	Teksttype	Antall oppgaver: 36 **	Gj. snittlig p-verdi	Kjønnsforskj. (j-g)*
<i>Humle</i>	«Hjelp humlene – lag et bol»	Naturfag	Fagtekst	9	0,65	0,02
<i>Penger</i>	«Hva er penger?»	Samfunnsfag	Fagtekst, sammensatt	6	0,53	-0,01
<i>Det regnet</i>	[Ingen tittel]	Norsk	Novelle	7	0,40	-0,01
<i>Panda</i>	«Forskere kler seg ut i drakter med panda-tiss for å redde pandaer»	Samfunnsfag, naturfag	Avisartikkel fra Aftenposten Junior	6	0,30	-0,06
<i>Skyggen</i>	«Skyggen»	Norsk	Eventyr	8	0,64	-0,03

* Positive verdier er verdier i jentenes favør, negative verdier i guttenes.

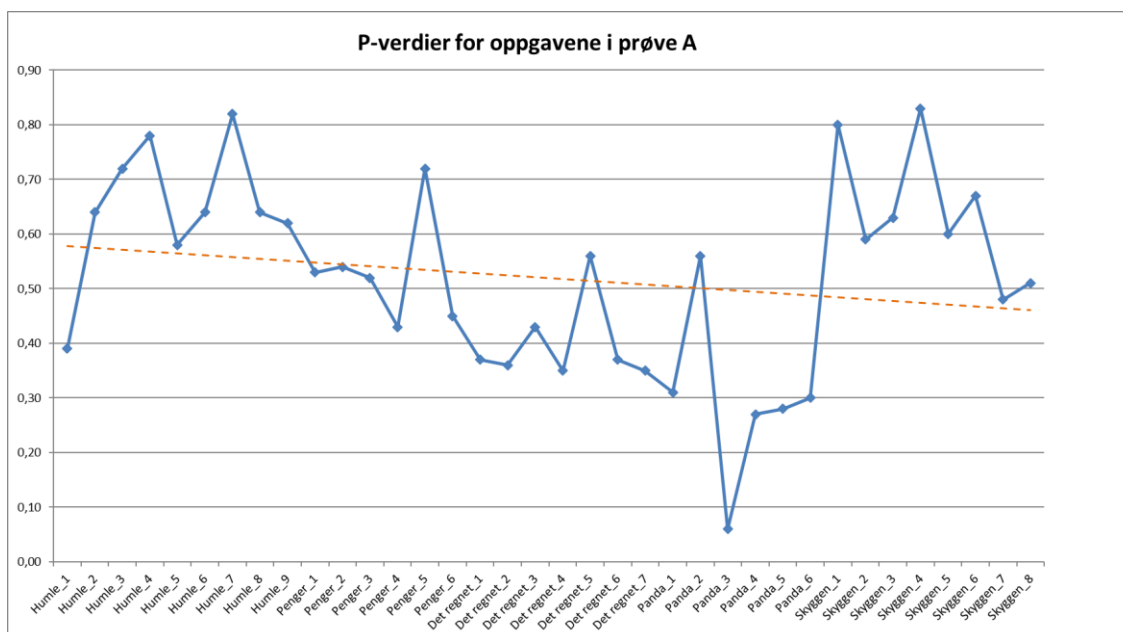
** 35 av oppgavene i prøven er flervalgsoppgaver, 1 oppgave er en sammensatt flervalgsoppgave

Tabell 3: Oversikt over tekstene i prøve B.

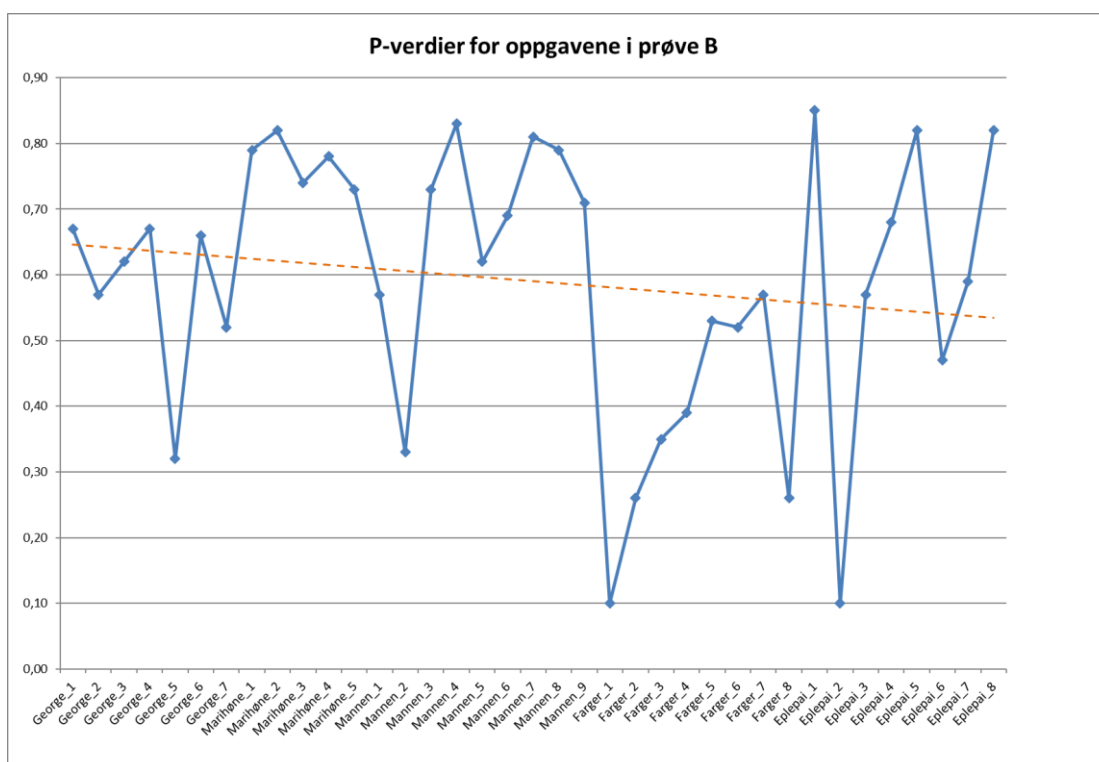
Tekstnavn i rapport	Original tittel	Fagtilhørighet	Teksttype	Antall oppgaver: 37	Gj. snittlig p-verdi	Kjønnsforskj. (j-g)*
<i>George</i>	«Ensomme George»	Naturfag, samfunnsfag	Utdrag fra fagbok for barn	7	0,58	-0,02
<i>Marihøne</i>	«Marihøne»	Naturfag	Utdrag fra fagbok for barn	5	0,77	0,04
<i>Mannen</i>	«Mannen og de kramlesyke sønnene»	Norsk	Eventyr	9	0,68	0,03
<i>Farger</i>	«Jente- og guttefarger»	Samfunnsfag	Avisartikkel fra Aftenposten Junior	8	0,37	0,03
<i>Eplepai</i>	Eplepai	Mat og helse, samfunnsfag	Oppskrift	8	0,61	0,02

* Positive verdier er verdier i jentenes favør, negative verdier i guttenes.

Figur 1 og 2 viser p-verdier for hver enkelt av oppgavene i henholdsvis prøve A og prøve B. Oppgavene presentert i samme rekkefølge som i prøvene.



Figur 1: Andelen elever med riktig svar (p-verdi) på oppgavene i prøve A, organisert etter rekkefølgen i prøven.



Figur 2: Andelen elever med riktig svar (p-verdi) på oppgavene i prøve B, organisert etter rekkefølgen i prøven.

Både figur 1 og 2 viser at det er både lette og vanskeligere oppgaver knyttet til alle tekstene i prøven. Trendlinjen viser en tendens til at p-verdien blir noe lavere mot slutten av begge prøvene. Det kan være flere årsaker til dette, for eksempel at elevene går lei eller mister motivasjonen utover i prøven – dette er normalt, og en slik effekt kan også ofte være større ved pilotering enn ved ordinær gjennomføring. Trenden kan i dette tilfellet også skyldes at flere oppgaver til teksten *Panda* og *Farger* ikke var gode nok og slo ut som svært vanskelige.

Klassisk item-analyse av prøve A og B

Tabell 4 og 5 viser resultater fra klassisk item-analyse av henholdsvis prøve A og B. Oppgavene er her gitt et løpenummer, og hver oppgave er tildelt en item-ID og et navn. Navnet viser til hvilken tekst oppgaven tilhører og den nummerplasseringen oppgaven har innenfor tekstenheten under utprøving. Item-ID er et unikt nummer som følger oppgaven uavhengig av endringer og omplussinger som eventuelt blir gjort mellom de ulike piloteringene. (Ved større endringer, blir den tildelt et nytt item-ID-nummer). Resultatene i tabellen viser prosentvis fordeling på de ulike svaralternativene for hver oppgave og gjennomsnittlig antall poeng på prøven for elevene som har valgt de ulike svaralternativene. Kolonnen «diskriminering» viser en biserial korrelasjon. De siste kolonnene viser de ulike oppgavens p-verdi og kjønnsforskjell, der positive verdier er i jentenes favør og negative verdier i guttenes favør. Tabell 4 viser resultatene fra prøve A.

Tabell 4: Resultater fra klassisk item-analyse, prøve A.

N: 289 (139 gutter, 150 jenter). Reliabilitet (Cronbachs Alpha): 0,868.

Gjennomsnittlig p-verdi: 0,52 (jenter: 0,51 / gutter: 0,53). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Oppg.	Item-ID	Prøve A	Svarfordeling i % (frekvens):						Gj.snittlig poengsum på prøven for de som har valgt svaralt.						Diskrim. (biserial)	P-verdi	Kjønnsforskj. (j-g)	Merk
		Flervalgsoppgave	A	B	C	D	E	Blank	A	B	C	D	E	Blank				
1	L042S01	Humle_1	8	1	45	39	6	1	16	10	20	24	19	3	0,198	0,39	0,08	
2	L042S02	Humle_2	1	29	64	5		1	9	17	24	14		3	0,412	0,64	-0,08	
3	L042S03	Humle_3	72	23	3	1		1	22	18	14	10		3	0,236	0,72	0,02	
4	L042S04	Humle_4	78	4	4	13		2	23	12	14	17		5	0,343	0,78	0,03	
5	L042S05	Humle_5	31	4	58	6		1	17	19	24	14		3	0,437	0,58	-0,05	
6	L042S06	Humle_6	64	6	10	20		1	24	18	17	16		0	0,365	0,64	0,07	
7	L042S07	Humle_7	12	3	82	2		1	15	14	22	15		0	0,333	0,82	0,07	
8	L042S08	Humle_8	15	64	12	8		1	19	24	15	13		0	0,396	0,64	0,03	
9	L042S09	Humle_9	5	62	28	5		1	15	23	19	14		0	0,217	0,62	-0,03	
10	L043S01	Penger_1	12	53	19	14		3	16	24	21	16		5	0,342	0,53	0,08	
11	L043S02	Penger_2	54	22	11	11		3	25	18	18	16		5	0,398	0,54	-0,12	
12	L043S03	Penger_3	19	16	11	52		2	17	18	18	25		3	0,407	0,52	0,06	
13	L043S04	Penger_4	10	28	43	16		3	15	20	23	23		3	0,180	0,43	-0,08	*
14	L043S05	Penger_5	72	17	5	2		4	24	15	14	13		5	0,493	0,72	0,01	
15	L043S06	Penger_6	27	9	45	15		4	20	16	24	18		7	0,316	0,45	0,01	
16	L044S01	Det regnet_1	30	15	13	37		5	21	18	21	24		8	0,220	0,37	0,08	
17	L044S02	Det regnet_2	11	38	9	36		6	17	20	17	26		8	0,385	0,36	0,02	
18	L044S03	Det regnet_3	15	8	43	29		5	15	18	25	21		7	0,351	0,43	-0,06	
19	L044S04	Det regnet_4	19	35	15	25		6	21	24	19	20		8	0,243	0,35	0,04	
20	L044S05	Det regnet_5	8	19	56	11		7	16	18	25	17		9	0,491	0,56	-0,06	
21	L044S06	Det regnet_6	37	23	16	17		7	25	21	21	18		9	0,291	0,37	0,00	
22	L044S07	Det regnet_7	19	13	35	25		8	19	16	27	21		9	0,464	0,35	-0,09	
23	L045S01	Panda_1	19	4	17	31		30	18	20	21	27		17	0,407	0,31	-0,10	
24	L045S02	Panda_2	56	3	4	8		29	25	12	17	15		17	0,429	0,56	-0,07	
25	L045S03	Panda_3	11	8	6	45		30	21	19	24	24		17	0,062	0,06	0,03	*
26	L045S04	Panda_4	14	8	27	17	5	30	19	21	28	21	17	16	0,424	0,27	0,00	
27	L045S05	Panda_5	28	3	11	28		29	22	16	17	26		17	0,327	0,28	-0,12	
28	L045S06	Panda_6	17	30	8	14		31	20	26	19	23		17	0,331	0,30	-0,07	
29	L046S01	Skyggen_1	4	3	80	4		9	12	14	23	15		8	0,522	0,80	-0,09	
30	L046S02	Skyggen_2	9	12	59	10		9	18	17	25	19		9	0,482	0,59	-0,08	
31	L046S03	Skyggen_3	11	7	10	63		9	15	14	19	25		9	0,565	0,63	-0,06	
32	L046S04	Skyggen_4	8	83				10	12	23				9	0,646	0,83	-0,03	**
33	L046S05	Skyggen_5	60	12	10	7		11	25	16	17	16		10	0,601	0,60	0,03	
34	L046S06	Skyggen_6	67	9	7	7		10	25	13	14	13		9	0,675	0,67	-0,02	
35	L046S07	Skyggen_7	15	48	7	18		12	17	27	15	18		10	0,622	0,48	-0,03	
36	L046S08	Skyggen_8	21	8	9	51		11	19	17	21	25		9	0,441	0,51	0,02	

* Lav diskriminering (under 0,250). ** Sammensatt flervalgsoppgave

Tabell 5 viser resultatene fra prøve B.

Tabell 5: Resultater fra klassisk item-analyse, prøve B.

N: 271 (140 gutter, 131 jenter). Reliabilitet (Cronbachs Alpha): 0,866.

Gjennomsnittlig p-verdi: 0,59 (jenter: 0,60 / gutter: 0,58). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Oppg.	Item-ID	Prøve B	Svarfordeling i % (frekvens):						Gj.snittlig poengsum på prøven for de som har valgt svaralt.						Diskrim. (biserial)	P- verdi	Kjønns- forskjell (j-g)	Merk
		Flervalgsoppgave	A	B	C	D	E	Blank	A	B	C	D	E	Blank				
1	L047S01	George_1	25	4	3	67		0	19	13	16	23		0	0,338	0,67	-0,06	
2	L047S02	George_2	57	15	13	6	9	0	24	18	18	17	24	0	0,393	0,57	0,03	
3	L047S03	George_3	11	10	62	17		0	18	18	23	18		13	0,312	0,62	-0,01	
4	L047S04	George_4	13	67	10	10		0	17	23	19	17		0	0,302	0,67	0,02	
5	L047S05	George_5	36	27	32	6		1	21	19	24	16		22	0,256	0,32	-0,05	
6	L047S06	George_6	10	66	10	14		0	18	23	19	19		13	0,237	0,66	-0,06	
7	L047S07	George_7	14	16	17	52		2	18	21	21	23		12	0,139	0,52	0,02	*
8	L048S01	Mariehøne_1	79	4	2	15		1	22	17	20	18		12	0,201	0,79	0,06	
9	L048S02	Mariehøne_2	5	10	82	3		0	14	16	23	10		1	0,403	0,82	0,02	
10	L048S03	Mariehøne_3	16	74	4	6		0	15	23	15	17		1	0,455	0,74	0,03	
11	L048S04	Mariehøne_4	7	78	4	11		0	14	23	15	19		1	0,319	0,78	0,05	
12	L048S05	Mariehøne_5	10	8	73	7		1	15	19	23	15		3	0,439	0,73	0,05	
13	L049S01	Mannen_1	11	57	24	6		2	17	24	20	17		4	0,358	0,57	0,06	
14	L049S02	Mannen_2	54	7	4	33		3	21	17	17	24		8	0,176	0,33	-0,04	*
15	L049S03	Mannen_3	7	17	2	73		2	14	17	15	23		5	0,462	0,73	0,11	
16	L049S04	Mannen_4	7	4	4	83		2	15	17	18	22		5	0,360	0,83	-0,02	
17	L049S05	Mannen_5	62	8	10	19		2	24	15	17	18		5	0,493	0,62	0,02	
18	L049S06	Mannen_6	2	4	17	69	7	2	15	15	21	23	15	5	0,292	0,69	-0,01	
19	L049S07	Mannen_7	11	81	2	4		2	18	23	11	13		5	0,434	0,81	0,02	
20	L049S08	Mannen_8	6	7	6	79		2	14	15	17	23		5	0,471	0,79	0,04	
21	L049S09	Mannen_9	1	17	71	9		3	9	19	23	15		7	0,417	0,71	0,05	
22	L050S01	Farger_1	21	17	23	10		27	24	22	22	24		17	0,125	0,10	0,01	*
23	L050S02	Farger_2	4	26	32	10		28	14	25	24	18		17	0,239	0,26	-0,03	
24	L050S03	Farger_3	10	16	11	35		28	19	21	21	25		17	0,378	0,35	0,02	
25	L050S04	Farger_4	39	10	12	11		28	25	21	20	21		17	0,365	0,39	0,14	
26	L050S05	Farger_5	53	15	1	3		28	25	17	19	18		17	0,546	0,53	0,00	
27	L050S06	Farger_6	13	5	52	2		28	19	17	24	20		17	0,466	0,52	0,02	
28	L050S07	Farger_7	5	6	57	4		28	17	16	25	16		17	0,529	0,57	0,02	
29	L050S08	Farger_8	4	34	26	7		29	17	23	26	18		17	0,358	0,26	0,03	
30	L051S01	Eplepai_1	2	4	85	3		7	10	15	23	14		12	0,499	0,85	0,01	
31	L051S02	Eplepai_2	38	39	6	10		7	23	21	16	23		12	0,061	0,10	-0,05	*
32	L051S03	Eplepai_3	13	57	8	14		8	20	23	16	22		12	0,277	0,57	-0,01	
33	L051S04	Eplepai_4	3	68	18	4		7	17	24	18	14		12	0,502	0,68	0,02	
34	L051S05	Eplepai_5	0	9	82	2		7	0	18	23	9		12	0,418	0,82	0,08	
35	L051S06	Eplepai_6	47	33	4	9		8	23	21	17	19		13	0,211	0,47	0,07	
36	L051S07	Eplepai_7	12	6	15	59		7	18	18	17	24		12	0,522	0,59	0,05	
37	L051S08	Eplepai_8	5	3	3	82		8	17	12	12	23		12	0,507	0,82	0,01	

*Lav diskriminering (under 0,250).

Utvalg av tekster og oppgaver til andre pilotering

I prosessen med å velge ut tekster til andre pilotering ligger resultatene og analysene fra første pilotering til grunn. I tillegg er det tatt hensyn til de tilbakemeldingene vi har fått fra fagpersonene ved Uppsala universitet, og innspill fra referansegruppen bestående av fire lærere på mellomtrinnet.

De fem tekstene som gikk videre til andre pilotering, blir presentert nærmere nedenfor, men først gir vi en kort redegjørelse for vurderingen som ligger til grunn for at fem tekster ble valgt bort:

- *Hjelp humlene – lag et bol:* Lærerne i referansegruppa savnet mer informasjon om humlenes betydning, men ellers mente både kvalitetssikrerne og lærerne at temaet var spennende og viktig. Kvalitetssikrerne påpekte at mange av distraktorene til oppgavene var hentet utenfor teksten. Bilder og bildetekst hang dårlig sammen med resten av teksten. Vi ble enige om å bearbeide teksten og bruke den ved en senere anledning.
- *Det regnet:* Både lærerne og kvalitetssikrerne likte teksten godt, men mente at det var vanskelig for en femteklassing å forstå det sentrale poenget i teksten. Teksten ble flyttet til 8. trinn.
- *Forskere kler seg ut i drakter med panda-tiss for å redde pandaer:* Teksten ble forkastet. Nærmere analyser av teksten og oppgavene avslørte både mangler på utdypende og relevant informasjon og en usammenhengende tekststruktur.
- *Marihøne.* Denne teksten har flere kvaliteter. Den er kort, den er multippel og den har et faglig innhold som er relevant og passende for aldersgruppa. Den ble likevel valgt vekk fordi vi ikke klarte å utnytte det multiple potensialet på en god måte. Oppgavene er gjennomgående litt for lette. Teksten ble flyttet til en annen prøve.
- *Mannen og de kranglesyke sønnene:* Både lærerne og kvalitetssikrerne likte denne teksten. Da valget stod mellom denne teksten og «Skyggen», ble sistnevnte valgt, da kvaliteten på spørsmålene ble vurdert som bedre i denne teksten. Den var også kortere.

De resterende tekstene ble alle vurdert som egnede for å måle elevenes leseferdigheter av både de eksterne kvalitetssikrerne og referanselærerne. De utvalgte tekstene kan knyttes til fagene norsk, samfunnsfag, naturfag og mat og helse. Det er som vanlig én skjønnlitterær tekst og fire sakprosa-tekster i utvalget. Nedenfor følger en mer utfyllende redegjørelse for de vurderingene som ligger til grunn for at disse fem tekstene ble valgt:

- *Ensomme George* er hentet fra fagboka *Dyrene som forsvant* av Line Renslebråten. Teksten forteller om de utryddete Pinta-kjempeskilpaddene. Den tar utgangspunkt i fortellingen om det siste gjenlevende eksemplaret av arten, populært kalt «Ensomme George», som døde i 2012. I det gule informasjonsfeltet står det at teksten er en del av en bok om utryddede dyr. Denne informasjonen, i kombinasjon med den første setningen i teksten, kan være til hjelp for å få en overordnet forståelse av hva teksten handler om. I fagtekster for barn er det et vanlig fortellergrep å formidle større og komplekse prosesser gjennom ett enkelt individs skjebne. Historien om «Ensomme George» er et eksempel på dette og viser hvordan mennesket har ansvaret for at mange

arter har dødd ut eller er i ferd med å dø ut. Flere av oppgavene til teksten krever at elevene forstår hvorfor dyr blir utryddet og menneskets rolle i denne utryddelsen.

Analysene avdekket at oppgaven «L047S07» diskriminerte dårlig. Denne oppgaven ble tatt bort og erstattet av en ny oppgave.

- *Jente- og guttefarger* er en reportasje hentet fra avisa Aftenposten Junior. Teksten handler om to jenter som har gjort en undersøkelse av hva slags farger gutter og jenter liker, og hva slags farger det er på gutte- og jenteklær i butikkene. De hevder at det ikke er samsvar mellom hva slags farger gutter og jenter liker, og hva slags farger det er på klærne de kjøper i butikkene. En ekspert er også intervjuet. En grafisk framstilling viser hvordan guttenes og jentenes fargepreferanser fordeler seg. En god forståelse av tekstens tema krever at leserne klarer å sette sammen informasjon fra ulike tekstelementer og ulike aktører i teksten.

Vi fjernet en tekstboks og forenklet grafikken (fra grafisk framstilling av prosent til mer tradisjonelt kakediagram). Oppgave «L050S02» diskriminerte dårlig, og ble fjernet. Det ble laget en ny oppgave.

- *Skyggen* er hentet fra en eventyrsamling av Erna Osland. Teksten handler om en hvithåret vismann med et stort problem: han klarer ikke å sove. Vismannen søker råd hos en rekke ulike mennesker, men ingen av rådene er til hjelp. En kveld smyger imidlertid en liten svart skygge – en katt – seg inn på soverommet og legger seg ved siden av den utmattede vismannen, som omsider får sove. Eventyret kan leses som en moralsk fortelling om at mye kunnskap ikke nødvendigvis er egnet til verken å løse praktiske problemer eller å redde folk fra ensomhet. Oppgavene til teksten måler i hvor stor grad leserne klarer å leve seg inn i og forstå vismannens tanker, følelser og handlinger ut i fra den informasjonen som blir gitt av fortelleren i teksten. Oppgavene krever også at elevene klarer å lese et mer overordnet budskap ut av den enkle handlingen i eventyret, noe som er vanlig i lesing av skjønnlitterære tekster.

De statistiske analysene viser at alle oppgavene fungerer tilfredsstillende. Oppgave «L046S06» ble tatt ut.

- *Hva er penger?* er et utdrag fra en fagtekst i boka *Hva er økonomi? Fakta og funderinger*. Teksten presenterer betalingsmidler fra ulike steder og til ulike tider og løfter fram morsomme opplysninger om et fenomen de fleste forholder seg til, men som ikke nødvendigvis er enkelt å forstå. Flere av oppgavene knytter seg til det sentrale poenget i teksten, nemlig at penger kan være omtrent hva som helst, så lenge folk har tillit til betalingsmiddelet. For å forstå teksten som helhet må elevene forstå hvordan de to tekstboksene og bildet i utdraget bidrar til å støtte opp om dette sentrale poenget. I denne prøven representerer «Penger» tekster med historiske temaer. Slike tekster krever at leserne evner å plassere hendelsene som beskrives, kronologisk, slik noen av oppgavene etterspør. I tillegg krever teksten at leserne har forståelse for

hvordan mennesker forholdt seg til penger før i tiden, og at de kan bruke denne innsikten til å forstå vår tids syn på det samme fenomenet.

Analysene avdekket at oppgave «L043S04» diskriminerte dårlig. Ordlyden i denne oppgaven ble endret fordi den var upresis og kunne feiltolkes.

- *Eplepai* er satt sammen av ulike tekster om pai: Brødteksten og oppskriften er hentet fra *Pomonas eplebok*, og tekstboksen om paikasting er hentet fra Wikipedia. Både i form og innhold likner denne sammensatte teksten andre kokebokoppskrifter. I slike tekster er oppskrifter ofte sammenstilt med utvalgte faktaopplysninger om opphavet til ulike retter, og ingrediensene i oppskriften er som regel ordnet på en bestemt måte. Leserne må dermed selv forstå hva som bestemmer rekkefølgen av ingrediensene, samt hva som er hensikten med tilleggsopplysningene som presenteres ved siden av oppskriftene. Flere av oppgavene til «Eplepai» krever at elevene leser nøyaktig og henter ut detaljinformasjon fra teksten.

Analysene avdekket at oppgave «L051S02» diskriminerte dårlig. Den ene distraktoren ble endret fordi den var for riktig. Spørsmålsformuleringen ble også endret noe. Oppgave «L051S06» ble tatt ut da denne kunne besvares ut i fra forkunnskaper.

Før andre pilotering ble det også gjort andre endringer enn de som er kommentert ovenfor. Etter råd fra kvalitetssikrerne og lærerne i referansegruppa ble noen oppgaveformuleringer endret, og noen svaralternativer ble justert og endret. Oppgaver som ikke fungerte godt, ble fjernet, og det ble lagt til nye oppgaver til tre av tekstene. Alle disse endringene er dokumentert ved hjelp av funksjonen «spor endringer» i Word, og tekstfilene er lagret i prøveutviklernes arkiver.

3. Andre pilotering

De fem utvalgte tekstene ble satt sammen til en ny prøve: prøve C. Denne prøven inneholdt 34 oppgaver, og piloteringen ble gjennomført digitalt i mars 2018. Tekstene og oppgavene ble prøvd ut i den rekkefølgen som var planlagt for den ordinære gjennomføringen. 1129 elever fra 25 skoler i 14 fylker deltok. Prøvetiden var 90 minutter.

Tabell 6: Oversikt over utvalget av elever (N=1129) pilot prøve C.

	Alle
Antall oppmeldte elever	1252 (fra 25 skoler)
Antall elever som gjennomførte (N)	Totalt 1129 (562 gutter, 567 jenter)
Antall skoler	25

Innholdet i prøve C

Tekstene fordeler seg på fag og teksttyper som vist i tabell 9.

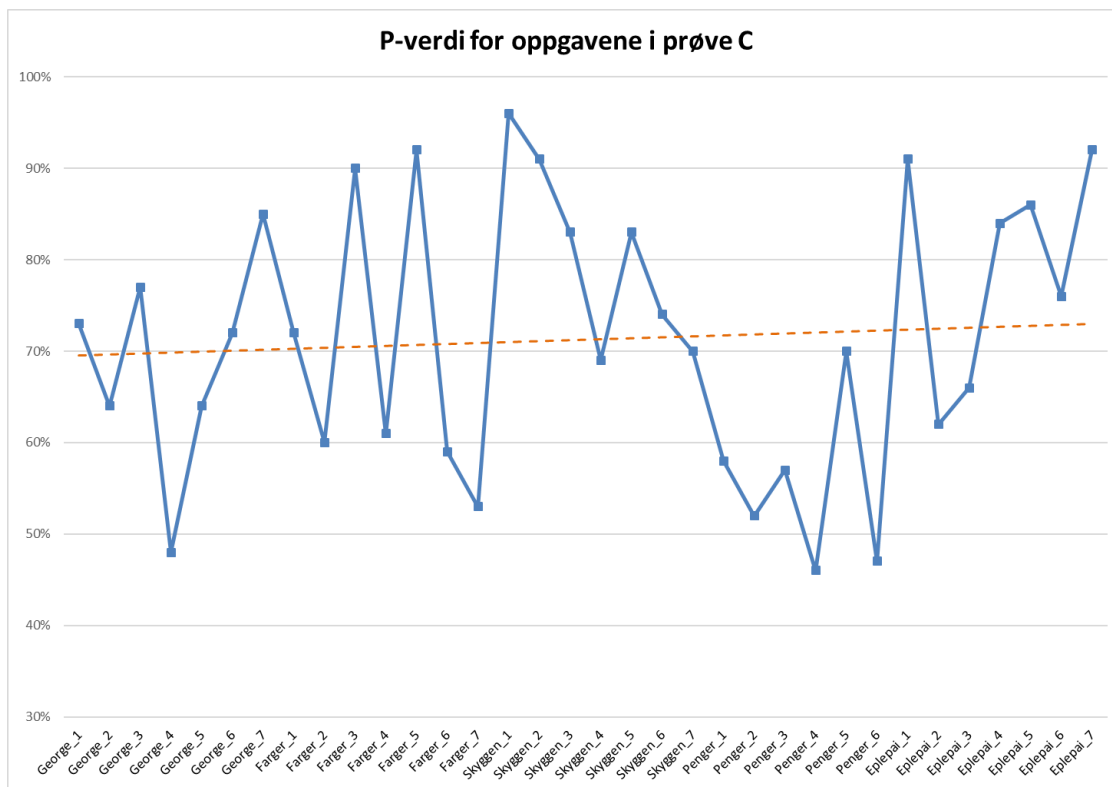
Tabell 7: Oversikt over tekstene i prøve C.

Tekstnavn i rapport	Original tittel	Fagtilhørighet	Teksttype	Antall oppgaver: 34 **	Gj.snittlig p-verdi	Kjønnsforskj. (j-g)*
<i>George</i>	«Ensomme George»	Naturfag, samfunnsfag	Utdrag fra fagbok for barn	7	0,69	0,03
<i>Farger</i>	«Jente- og guttefarger»	Samfunnsfag	Avisartikkel fra Aftenposten Junior	7	0,70	0,07
<i>Skyggen</i>	«Skyggen»	Norsk	Eventyr	7	0,81	0,04
<i>Penger</i>	«Hva er penger?»	Samfunnsfag	Fagtekst, sammensatt	6	0,55	0,01
<i>Eplepai</i>	«Eplepai»	Mat og helse, samfunnsfag	Fagboktekst, oppskrift	7	0,80	0,03

* Positive verdier i jentenes favør, negative i guttenes favør.

** 33 av oppgavene i prøven er flervalgsoppgaver, 1 oppgave er en sammensatt flervalgsoppgave

Figur 3 viser p-verdien for hver av oppgavene i prøve C. Oppgavene er presentert i samme rekkefølge som i prøven.



Figur 3: Andelen elever med riktig svar (p-verdi) på oppgavene i prøve C, organisert etter rekkefølgen i prøven.

Figur 3 viser at det også her er både lette og vanskeligere oppgaver knyttet til alle tekstene i prøven, men figuren viser også at oppgavene har en til dels betydelig høyere p-verdi enn de hadde i den første piloteringen. Trendlinjen viser at p-verdiene faktisk blir noe høyere mot slutten av prøven. Dette betyr at i tillegg til at prøven som helhet har blitt lettere, er det ingen ting som tyder på at elevene har hatt problemer med å fullføre prøven.

Klassisk item-analyse av prøve C

Tabell 8 viser resultater fra klassisk item-analyse av prøve C.

Tabell 8: Resultater fra klassisk item-analyse, prøve C.

Antall elever (N): 1129 (562 gutter, 567 jenter). Reliabilitet (Cronbachs Alpha): 0,861.

Gjennomsnittlig p-verdi: 0,71 (gutter: 0,69 / jenter: 0,73). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Oppg.	Item-ID	Prøve C	Svarfordeling i % (frekvens):					Gj.snittlig poengsum på prøven for de som har valgt svaralt.					Diskrim. (biseriell)	P- verdi	Kjønn- forskj. j-g	Merk
			A	B	C	D	Blank	A	B	C	D	Blank				
		Flervalgsoppgaver														
		Paringsoppgave	0	1				0	1							
1	L047S01	George_1	21	3	3	73	0	21	17	16	26	8	0,340	0,73	0,05	
2	L047S02	George_2	64	14	15	8	0	26	21	21	18	5	0,389	0,64	0,04	
3	L047S04	George_3	6	77	9	8	0	18	26	20	20	4	0,379	0,77	0	
4	L047S05	George_4	14	35	48	4	0	20	23	27	19	9	0,373	0,48	-0,01	
5	L047S07	George_5	18	2	15	64	1	21	19	22	26	9	0,301	0,64	0,06	
6	L047S06	George_6	4	72	9	15	1	18	26	20	20	11	0,403	0,72	0,04	
7	L047S09	George_7	10	3	1	85	1	20	17	13	25	8	0,367	0,85	0,03	
8	L050S02	Farger_1	5	72	6	18	0	18	25	20	23	1	0,225	0,72	0,11	
9	L050S03	Farger_2	3	21	16	60	1	16	23	23	26	5	0,197	0,60	0,17	
10	L050S06	Farger_3	8	1	90	0	0	19	19	25	17	2	0,259	0,90	0	
11	L050S04	Farger_4	61	9	16	14	1	26	20	22	21	7	0,360	0,61	0,11	
12	L050S07	Farger_5	3	3	92	2	1	15	16	25	14	4	0,429	0,92	0,04	
13	L050S08	Farger_6	3	31	59	6	1	15	22	26	19	5	0,359	0,59	0,06	
14	L050S0?	Farger_7	15	17	14	53	1	23	23	23	26	12	0,169	0,53	0,03	
15	L046S01	Skyggen_1	1	96	1	1	1	12	25	14	14	6	0,375	0,96	0	
16	L046S02	Skyggen_2	91	4	3	2	1	25	18	17	17	7	0,378	0,91	0,03	
17	L046S03	Skyggen_3	5	3	8	83	1	18	19	20	25	8	0,394	0,83	-0,02	
18	L046S04a-d	Skyggen_4	5	95			1	15	24			9	0,434	0,69	0,07	**
19	L046S05	Skyggen_5	83	5	6	5	2	26	16	19	17	8	0,528	0,83	0,05	
20	L046S07	Skyggen_6	12	74	4	9	2	20	26	17	20	8	0,491	0,74	0,04	
21	L046S08	Skyggen_7	16	6	6	70	2	21	19	20	26	10	0,428	0,70	0,11	
22	L043S01	Penger_1	10	58	16	14	2	19	27	22	21	9	0,443	0,58	0,01	
23	L043S02	Penger_2	52	27	8	11	2	27	22	22	21	9	0,354	0,52	0,01	
24	L043S03	Penger_3	17	13	12	57	2	21	22	21	27	9	0,415	0,57	0,03	
25	L043S04	Penger_4	11	26	46	15	2	21	23	27	23	9	0,312	0,46	-0,04	
26	L043S05	Penger_5	70	11	6	10	2	26	19	20	21	9	0,485	0,70	0,06	
27	L043S06	Penger_6	32	47	9	8	3	22	27	21	22	12	0,408	0,47	-0,02	
29	L051S01	Eplepai_1	1	3	91	2	2	14	21	25	16	10	0,374	0,91	0,01	
30	L051S02	Eplepai_2	6	23	6	62	3	17	23	18	27	11	0,413	0,62	0,06	
31	L051S03	Eplepai_3	66	14	5	12	3	26	23	19	24	10	0,232	0,66	-0,01	
32	L051S04	Eplepai_4	2	84	10	1	3	21	25	19	16	10	0,423	0,84	0,09	
33	L051S05	Eplepai_5	1	9	86	1	3	15	21	25	13	10	0,382	0,86	0,04	
34	L051S07	Eplepai_6	76	4	9	8	3	26	19	19	19	11	0,524	0,76	0,03	
35	L051S08	Eplepai_7	3	2	1	92	3	20	17	16	25	11	0,380	0,92	0,01	

** Sammensatt flervalgsoppgave. Siden piloteringen ble gjennomført digitalt, har vi ikke tilgang på svarfordelingen (frekvens) innad i oppgaven.

Resultatene i tabell 8 ligger til grunn for utvalget av oppgaver til endelig prøve. Oppgavene er også analysert ved hjelp av IRT, men parametrene fra denne analysen blir først vist i avsnittet som omhandler det endelige prøvforslaget.

Endringer fra andre pilotering til endelig prøve

I forslaget til ny leseprøve har vi foretatt noen endringer etter den andre piloteringen, på bakgrunn av resultatene fra denne. Endringene er gjort for å sørge for god spredning i

oppgavenes vanskegrad, at alle spørsmål og svaralternativ i prøven er så godt formulerte som mulig, samt at ingen av disse inneholder ord eller formuleringer som kan virke kompliserende for elevene. Ettersom ingen av oppgavene i prøve C hadde lavere p-verdi enn 0,46, og den samlede p-verdien for denne prøven ble svært høy (0,71), ble teksten *Farger* (markert med mørk farge tabell 8) fra prøve C tatt ut og erstattet med teksten *Jeg hater lukten av røyk!* (*Røyking*), en tekst med flere oppgaver med en lavere p-verdi. Teksten ble prøvd ut i førstepiloteringen for nasjonale prøver i lesing for 5. trinn for 2019, og er dermed pilotert av 433 elever på femte trinn.

Jeg hater lukten av røyk! er en reportasje hentet fra avisa Aftenposten Junior. Teksten handler om ti år gamle Luke som henger opp plakater på Haugen stasjon for å få folk til å slutte å røyke på stasjonen. I tillegg til intervjuet med Luke blir Kristin Paus, som jobber i Bane NOR, intervjuet om røyking på togstasjoner i Norge. En tekstboks presenterer røyking i et historisk perspektiv. En god forståelse av tekstens tema – røyking på offentlige steder – krever at leserne klarer å sette sammen informasjon fra ulike tekstelementer og ulike aktører i teksten. Flerstemmigheten i «Jeg hater lukten av røyk» er svært typisk for nyhetsartikler og fagtekster som elevene møter utenfor skolen, men er sjelden til stede i læreboktekster, som ofte har én styrende fortellerstemme. Flere av oppgavene til teksten krever at elevene forstår hvilke aktører som er representert i teksten, hvilke interesser disse personene har, samt hvor og hvordan de kommer til orde i teksten.

Tabell 9 viser hvordan hver enkelt oppgave til teksten «Jeg hater lukten av røyk» fungerte som prøvetekst for 5. trinn under gjennomføringen i mars 2018.

Tabell 9: Resultater fra klassisk item-analyse, første pilotering nasjonal prøve for femte trinn 2019

N: 432 (219 gutter, 213 jenter). Reliabilitet (Cronbachs Alpha): 0,835.

Gjennomsnittlig p-verdi: 0,60 (jenter: 0,61 / gutter: 0,59). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Oppg.	Item-ID	Tekst	Svarfordeling i % (frekvens):						Gj.snittlig poengsum på prøven for de som har valgt svaralt.						Diskrim. (biserial)	Gj.snittlig p-verdi	Kjønns-forskj. (j-g)	Merk
		Flervalgs-oppgaver	A	B	C	D	E	Blank	A	B	C	D	E	Blank				
1	L053S01	Røyking_1	45	45	1	2	6	1	20	25	12	18	21	27	0,297	0,45	0,13	
2	L053S02	Røyking_2	5	3	4	89		0	16	16	16	23		25	0,288	0,89	0,03	*
3	L053S03	Røyking_3	21	18	32	26		1	26	21	22	20		21	0,241	0,21	0,03	
4	L053S04	Røyking_4	7	10	78	5		1	19	18	23	15		16	0,318	0,78	0,01	
5	L053S05	Røyking_5	9	1	79	9		2	18	18	24	17		15	0,365	0,79	0,05	
6	L053S06	Røyking_6	25	15	6	53		1	25	20	24	21		20	0,217	0,25	0,06	
7	L053S07	Røyking_7	52	40	5	2		1	22	23	18	18		21	0,012	0,40	-0,09	**
8	L053S08	Røyking_8	17	2	18	62		1	19	17	18	24		22	0,383	0,62	0,09	

* Denne oppgaven var svært lett og er derfor ikke med i forslaget til endelig prøve.

** Denne oppgaven diskriminerte svært dårlig og er derfor ikke tatt med i forslaget til endelig prøve.

Avsnittene nedenfor beskriver kort endringene for hver enkelt av oppgavene som er endret etter den andre piloteringen. Alle oppgavene er flervalgsoppgaver.

- I «Ensomme George_5» (L047S07) har vi endret ordlyden i B-distraktoren noe for å gjøre denne mer plausibel.
- I «Ensomme George_7» (L047S09) har vi endret ordlyden i spørsmålet for å gjøre det litt tydeligere hva vi spør om.

- «Røyking_2» (L053S02) hadde i piloteringen en p-verdi på 0,89, og er i forslaget til endelig prøve tatt bort for å redusere antallet oppgaver som er veldig enkle.
- «Røyking_7» (L053S07) diskriminerte dårlig og er derfor ikke tatt med i forslaget til endelig prøve.
- «Skyggen_1» (L046S01) hadde i den andre piloteringen en p-verdi på 0,96, og er i forslaget til endelig prøve tatt bort for å redusere antallet oppgaver som er veldig enkle.
- «Skyggen_2» (L046S02) hadde i den andre piloteringen en p-verdi på 0,91, og er i forslaget til endelig prøve tatt bort for å redusere antallet oppgaver som er veldig enkle.
- I «Skyggen_3» (L046S03) har vi endret ordlyden, samt lagt til mer informasjon i spørsmålet. I tillegg har vi endret ordlyden i både i A- og B-distraktorene for å gjøre dem mer plausible.
- I «Eplepai_2» (L051S02) har vi lagt til mer informasjon i spørsmålet, for å gjøre det tydeligere hva vi spør om. I tillegg har vi endret ordlyden i både i A- og B-distraktorene for å gjøre dem mer plausible.
- «Eplepai_4» (L051S04) har vi endret tallet i A- og D-distraktorene for å gjøre dem mer plausible.
- «Eplepai_5» (L051S05) har vi tatt ut fordi den måler de samme leseferdighetene som «Eplepai_4».
- «Eplepai_7» (L051S08) hadde i den andre piloteringen en p-verdi på 0,92, og er i forslaget til endelig prøve tatt bort i hensikt av å redusere antallet oppgaver som er veldig enkle.

I tillegg til endringene som er beskrevet over, har vi byttet om rekkefølgen på svaralternativene på enkelte oppgaver, for å sørge for en god fordeling i hvor det riktige svaret er plassert, samt enkelte mindre endringer i ordlyder i oppgavene.

4. Forslag til endelig prøve for 2018

Leseprøven for 2018 består av 29 oppgaver til 5 tekster som til sammen måler elevenes leseforståelse, i tråd med læreplanen for 2.-4.trinn. Blant de 29 oppgavene er det 28 flervalgsoppgaver og 1 sammensatt flervalgsoppgave. De utvalgte tekstene kan knyttes til fagene norsk, samfunnsfag, naturfag og mat og helse. Det er som vanlig én skjønnlitterær tekst og fire sakprosa-tekster med i utvalget. Prøven har en tradisjonell fordeling av oppgavene knyttet til de tre leseaspektene med 10 finne-oppgaver, 13 tolke- og 6 refleksjonsoppgaver. Tabell 10 gir en oversikt over oppgavene i forslaget til endelig prøve.

Tabell 10: Oppgavene i forslaget til endelig prøve, riktig svar/oppgaveformat og leseaspekt.

Nr.	Item-ID	Tekst	Riktig svar	Aspekt
1	L047S01	George_1	D	Finne
2	L047S02	George_2	A	Tolke
3	L047S04	George_3	B	Finne
4	L047S05	George_4	C	Tolke
5	L047S07	George_5	D	Reflektere
6	L047S06	George_6	B	Finne
7	L047S09	George_7	D	Tolke
8	L053S01	Røyking_1	C	Tolke
9	L053S03	Røyking_2	C	Tolke
10	L053S04	Røyking_3	C	Finne
11	L053S05	Røyking_4	C	Reflektere
12	L053S06	Røyking_5	A	Tolke
13	L053S08	Røyking_6	D	Reflektere
14	L046S03	Skyggen_1	D	Tolke
15	L046S04a-d	Skyggen_2	Sammensatt flervalg: Imam – klok kone – stjernetyder – søsteren	Finne
16	L046S05	Skyggen_3	A	Tolke
17	L046S07	Skyggen_4	B	Reflektere
18	L046S08	Skyggen_5	D	Tolke
19	L043S01	Penger_1	B	Tolke
20	L043S02	Penger_2	A	Finne
21	L043S03	Penger_3	D	Tolke
22	L043S04	Penger_4	C	Finne
23	L043S05	Penger_5	A	Tolke
24	L043S06	Penger_6	B	Reflektere
25	L051S01	Eplepai_1	C	Finne
26	L051S02	Eplepai_2	D	Tolke
27	L051S03	Eplepai_3	A	Reflektere
28	L051S04	Eplepai_4	B	Finne
29	L051S07	Eplepai_5	A	Finne

Interessante leseutfordringer i forslaget til endelig prøve

I det følgende har vi valgt ut oppgaver fra hver av tekstene i prøven som illustrerer de ulike leseprosessene og som gjenspeiler relevante leseutfordringer for elevene i denne alderen.

Oppgave til teksten «Eplepai»

Denne teksten er hentet fra boka «Pomonas eplebok» av Görel Kristina Näsund og nettleksikonet Wikipedia. Bruk teksten når du svarer på oppgavene.

Eplepai

Kunsten å bake pai kommer fra England. Det engelske ordet er *pie* (uttales pai), som er blitt til vår pai.

Mange mener at paien er blitt oppkalt etter skjæra. Skjæra heter *magpie* (uttales mægpai) på engelsk. En pai kan inneholde mye rart og minner kanskje litt om alle de småtingene som skjæra samler i redet sitt. Derfor mener de at paien har fått navn etter skjæra.

I en kokebok fra 1845 står det oppskrifter på eplepai, både med og uten bunn. Den eldste svenske oppskriften på eplepai, fra 1647, har ikke bunn, bare lokk. Her kommer en oppskrift som ikke er fullt så gammel.



Smuldrepai 4 porsjoner

Ingredienser:

Fyll:

6 epler

1 ts kanel

Smuldredeig:

1 dl sukker

1 dl hvetemel

1 dl sammalt hvete

75 g kaldt smør

Framgangsmåte:

1 Sett ovnen på 175 °C.

2 Skrell og skjær eplene i båter. Skjær bort kjernehuset fra hver båt. Legg eplebåtene i en rund form (ca. 20 cm tvers over). Du trenger ikke smøre formen. Strø kanel over eplene.

3 Ha sukker, mel og smør i en bolle.

4 Smuldre smøret med fingertuppene. Det skal bli en kornete masse av smør og mel og sukker. Strø massen over eplene.

5 Stek paien på andre rille i ca. 25 minutter, eller til smuldredeigen har fått pen farge og eplene er blitt myke.

6 Server paien med krem, vaniljesaus eller vaniljeis.

Visste du at ...?

I filmen *Mr. Flip* fra 1909 fikk den amerikanske komikeren Ben Turpin en pai presset opp i ansiktet. Siden den gangen har kasting av kremfylte paier vært et vanlig innslag i mange filmkomedier.



I filmen *The Great Race* fra 1965 ble det kastet rekordmange paier. I løpet av en scene ble det brukt 4000 paier! Det tok hele fem dager å spille inn denne scenen.

Paikasting regnes for å være en enkel narrestrek, men blir også brukt som en form for protest. Noen kaster pai i ansiktet på politikere eller andre offentlige personer for å vise at man er uenige med dem, og for å gjøre narr av dem.

Oppgave 27

Står ingrediensene i oppskriften i noen bestemt rekkefølge?

- A Ja, de står i den rekkefølgen de skal brukes.
- B Ja, det som skal stekes, står først.
- C Ja, de viktigste ingrediensene står til slutt.
- D Nei, ingrediensene står i tilfeldig rekkefølge.

Aspekt: Reflektere over tekstens form og innhold

Nasjonal løsningsprosent: 65 %

Fag: Mat og helse

For å svare riktig på denne oppgaven må elevene reflektere over om og hvorfor ingrediensene i en oppskrift er plassert i en bestemt rekkefølge.

Det er en fordel å kjenne sjangeren oppskrift fra før, men det er ikke nødvendig for å svare riktig. Elevene kan også komme fram til rett svar ved å kombinere informasjonen under «Framgangsmåte» med ingredienslisten. Da vil de for eksempel se at man må begynne med eplene før man lager deigen og på den måten komme fram til at ingrediensene i denne oppskriften er plassert i den rekkefølgen de skal brukes (svaralternativ A) – noe som er vanlig i oppskrifter.

Det er som regel en bevisst plan bak utformingen av tekster. Kunnskap om sjangertrekk er derfor en viktig del av lesekompetansen, og det kan være nyttig å la elevene reflektere over hvorfor teksten er bygget opp på en spesiell måte.

Oppgave til teksten «Ensomme George»

Denne teksten er hentet fra fagboka «Dyrene som forsvant» av Line Renslebråten.
Bruk teksten når du svarer på oppgavene.

Ensomme George

Pinta-kjempeskilpadde



Den aller siste Pinta-kjempeskilpadden bodde på Pintaøya, en av Galápagosøyene i Stillehavet. Skilpadden, som var en hann, døde i 2012, og så lenge han levde var han ifølge Guinness' rekordbok verdens mest sjeldne dyr!

Det var fiskere som oppdaget Pinta-kjempeskilpaddene på Pintaøya på 1800-tallet. Flere av dem forsøkte å få med seg et eksemplar over til fastlandet, men ettersom kjøttet smakte så godt, kom de aldri så langt – hver gang ble skilpaddene spist opp på turen!

Sjøfolk fra land i Sør-Amerika som skulle ut på lange seilaser, gikk ofte i land på øya Pinta for å fylle opp matlagene før de dro videre over havet. For å sikre tilgang på fersk geitemelk og geitekjøtt fikk noen den strålende ideen å sette ut geiter på øya. Geitene trivdes over all forventning, og snart var tre geiter blitt til 40 000! Men alle disse geitene skulle ha mat, og de gikk løs på alt gresset og plantene som også kjempeskilpaddene trengte for å leve. Slik gikk det til at skilpaddene forsvant én etter én – på grunn av matmangel.

På 1970-tallet trodde forskerne at Pinta-kjempeskilpadden var totalt utryddet, helt til en ungarsk biolog oppdaget én enslig skilpadde som stabbet rundt på øya. Funnet av kjempeskilpadden skapte store overskrifter, og i USA begynte journalistene å omtale ham som *Lonesome George*, etter TV-komikeren George Gobel, som var populær på den tiden.

Ensomme George ble en slags maskot i arbeidet med bevaringen av Galápagosøyene. Og ettersom det var geitene som hadde drevet skilpaddene vekk, ble det igangsatt et prosjekt for å fjerne hver eneste drøvtygger fra Pinta.

I 2003 ble øya erklært «geitefri», og kanskje kan kjempeskilpaddene en dag komme tilbake. I 2012 ble det nemlig oppdaget skilpadder med DNA som ligner Pinta-kjempeskilpaddene på en annen av Galápagosøyene!

Visste du at:

- Det finnes tre andre typer kjempeskilpadder på flere av Galápagosøyene.
- Den eldste kjempeskilpadden man vet om, ble 152 år.
- Kjempeskilpaddene vokser helt til de blir ca. 50 år og kan veie over 200 kg.
- Kjempeskilpaddene kan gå i nesten et helt år uten å spise eller drikke fordi kroppen deres er så god til å lagre næring.
- Det latinske navnet er *Chelonoidis abingdoni*.

Utrydningstruede dyr

- En utrydningstruet dyreart er en art hvor det er så få eksemplarer igjen at det er en fare for at arten blir helt borte.
- Dersom én til fem arter forsvinner i året, er det et naturlig fenomen, men forskere har funnet ut at vi nå kanskje mister så mye som 50 000 arter hvert år på grunn av menneskelige handlinger!
- En art blir erklært utryddet når ingen har sett dyret på mange år og grundige undersøkelser er blitt gjort. Tidligere var regelen at en art ble erklært utryddet om ingen hadde sett den på 50 år.

Oppgave 4

Teksten forteller om et prosjekt på Pinta som handlet om ...

- A å utrydde alle kjempeskilpaddene på øya.
- B å flytte alle kjempeskilpaddene til et trygt sted.
- C å fjerne alle geitene fra øya.
- D å fjerne plantene som geitene spiste.

Aspekt: Tolke og sammenholde informasjon i teksten.

Nasjonal løsningsprosent: 48 %

Fag: Naturfag

For å svare riktig på denne oppgaven må elevene hente ut informasjon som er vanskelig å finne, og tolke innholdet riktig.

Det første elevene må gjøre for å løse oppgaven, er å lokalisere hvor i teksten det står om et prosjekt på øya Pinta. Teksten inneholder mange opplysninger om Pinta-kjempeskilpadden og hvorfor den ble utryddet. Selv om mennesker spiste skilpaddene, var det først og fremst innføringen av geiter på øya Pinta som førte til at skilpaddene døde ut. Prosjektet som nevnes i teksten, handler om å fjerne alle geitene fra øya (svaralternativ C), som er et tiltak for å bedre levevilkårene for eventuelle framtidige kjempeskilpadder. For å svare riktig på oppgaven må elevene finne opplysningene som er plassert i det nest siste avsnittet i teksten, og videre tolke og forstå at «drøvtyggerne» som er nevnt i den siste delen av setningen, henviser til geitene.

Løsningsprosenten indikerer at denne oppgaven har vært vanskelig. Mange elever valgte svaralternativ B, «å flytte alle kjempeskilpaddene til et trygt sted». Elever som valgte dette svaralternativet, har kanskje mistolket hvem «drøvtyggerne» er, og trodd at det var kjempeskilpaddene som skulle flyttes fra øya. Dersom man ikke har forstått at Pinta-kjempeskilpaddene faktisk er utryddet, gir det også mening å flytte dem til et annet sted der de har bedre tilgang på mat.

Oppgaven illustrerer hvor viktig det er å ha tekstens hovedpoeng, nemlig at Pinta-kjempeskilpadden er utryddet, klart for seg når man skal forstå informasjon som er mindre sentralt plassert. Elevenes feilsvar til denne oppgaven indikerer at mange ikke har klart dette.

Oppgave til teksten «Jeg hater lukten av røyk!»

Denne teksten stod på trykk i Aftenposten Junior i februar 2018. Bruk teksten når du svarer på oppgavene.

– Jeg hater lukten av røyk!

– Folk burde smile istedenfor å røyke, sier Luke. Han har tatt opp kampen mot røyking.

Tekst: Mari Lund Wictorsen
Foto: Jon T. Espedal

– Røyk gjør meg svimmel i hodet. Jeg hater lukten av røyk! sier Luke. Aftenposten Junior er med da han henger opp plakater han har skrevet, på Haugen stasjon. Mens vi er der, oppdager Luke at en av plakatene han hengte opp tidligere, er borte.

– Å, nei! Har noen kastet den? sier Luke og kikker i søppelkassen. Han ser ikke plakaten, men henger opp flere nye.

Løper forbi røykerne

Luke tar toget fra denne stasjonen til skolen. Han pleier å se minst fem personer som røyker mens han venter på toget. Også når han kommer med toget fra skolen, kjenner han lukten av røyk på stasjonen.

– Da løper jeg fort forbi røykerne, så slipper jeg å puste inn røyken. Røyking irriterer meg skikkelig! Jeg håper de som røyker skjønner at jeg løper på grunn av dem, sier Luke.

Færre røyker

Mange færre røyker i Norge hver dag enn for ti år siden, ifølge Statistisk sentralbyrå. For 2004 (14 år siden) var det lov å røyke inne på blant annet restauranter og kafeer. Men så kom det en lov som gjorde slutt på det. For fire år siden ble det også forbudt å røyke på togstasjoner.

På Haugen stasjon viser skilt at røyking er forbudt på perrongen. Men ifølge Luke er det ikke nok.

– I flere måneder irriterte jeg meg over røyken. En dag tenkte jeg «Dette gidder jeg ikke mer! Jeg må gjøre noe», sier han.

Han bestemte seg for å skrive plakater og henge dem opp, i håp om at folk skulle stoppe og lese hvorfor de ikke må røyke der.

– Jeg skriver at jeg er 10 år med stor skrift. Jeg tror folk vil høre mer på et barn enn på et vanlig skilt.

Har du sagt fra til noen av røykerne at det ikke er lov?

– Nei, det tør jeg ikke. Pappa har sagt at det er litt feilt.

Hvorfor tror du de røyker der når det ikke er lov?

– Jeg tror ikke de tenker over at det er ulovlig. Og hvis de er avhengig av røyk, klarer de ikke å slutte.

Ber dem gå

Bane NOR, som har ansvaret for de fleste togstasjonene i Norge, opplever ikke at det er noe stort problem med folk som røyker på stasjonene.

– På venterommet inne er det heldigvis nesten aldri noen som røyker, men vi ser litt oftere at noen røyker mens de står ute på plattformen og venter på toget, skriver Kristin Paus i Bane NOR en e-post til Aftenposten Junior.

Hva gjør dere for å hindre at folk røyker på stasjonene?

– Vi har satt opp skilt med «Røyking forbudt» både ute og inne på stasjonene. Ser vi at noen røyker, ber vi dem pent om å gå et annet sted.

Plakater flere steder

Akkurat nå har Luke nok med å prøve å bli kvitt røyking på stasjonen der han bor. Etter hvert kan det hende at han vil henge opp plakater andre steder også.

– Folk burde smile istedenfor å røyke. Det gjør andre lykkeligere.

Vil ha slutt på røyking



Hø? Hva betyr det?

Statistisk sentralbyrå teller hvor mye det finnes av forskjellige ting i Norge, som mennesker, biler, navn og mye annet.

Visste du dette om røyking?

- For rundt 50 år siden satt programledere og gjester og røykte på TV under debatter.
- Rundt samme tid fant forskere for første gang bevis for at sigaretter var svært skadelige.
- Før 2004 var det lov å røyke på fly, tog og restauranter i Norge.
- Stadig flere land forbyr røyking på offentlige steder.

Vi minner om at dette er en røykfri stasjon!

Fra Luke (10 år)

Hver gang jeg går forbi noen som røyker, må jeg holde pusten for at jeg ikke skal få dårlig luft i lungene. Jeg blir skikkelig sliten av å gjøre det. Jeg går forbi minst 5 folk som røyker på Haugen stasjon hver eneste dag. Det synes jeg ikke jeg skal trengte å gjøre. Egentlig er Haugen stasjon en røykfri stasjon. Men jeg tror ikke det er noen forskjell mellom en røykfri stasjon og en vanlig stasjon lenger.

De som røyker kan få lungeskade, de som puster inn røyken kan også få det. Jeg vil ikke ha det. Jeg som bare er 10 år, må skrive et slikt brev til de som røyker. Jeg er på stasjonen hver dag fra ca. 7:40-7:50, og når jeg ser noen som røyker, vil jeg bare kaste sigarettene på bakken.

Hvis du er en som røyker: Tenk før du røyker, for det er ikke bare du som puster inn luften. Du kan gjøre en stor forskjell!

Lukes plakat

Den som røyker på et sted der det ikke er lov til å røyke, kan bli anmeldt til politiet og få bot.



Oppgave 9

Hvilken setning stemmer med informasjonen i den blå tekstboksen?

- A For 50 år siden var det mange som røykte selv om det ikke var lov.
- B Mange røykere respekterer ikke røykeforbud på offentlige steder.
- C Synet på røyking har endret seg mye de siste tiårene.
- D Mange har lite kunnskap om hvor farlig røyking er.

Aspekt: Tolke og sammenholde informasjon i teksten.

Nasjonal løsningsprosent: 27 %

Fag: Samfunnsfag, norsk

For å svare riktig på denne oppgaven må elevene få et overblikk over innholdet i tekstboksen.

Tekstelementet som spørsmålet dreier seg om, er lett å finne, men vi vet av erfaring at mange elever synes det er vanskelig å sammenfatte innhold i en punktliste. I dette tilfellet må elevene velge et svaralternativ som oppsummerer fire kulepunkter med ulike faktaopplysninger om røyking. Det riktige svaret («Synet på røyking har endret seg mye de siste tiårene.») finnes ikke som en konkret opplysning i teksten. Leserene må selv tenke seg en plausibel underoverskrift som oppsummerer det kulepunktene i tekstboksen har til felles.

Det som gjør denne oppgaven spesielt krevende, er at flere av svaralternativene inneholder opplysninger som er riktige. Det som gjør dem feil, er at de ikke stemmer med informasjonen i tekstboksen. For å velge rett svar må elevene med andre ord kunne sortere vekk opplysninger som ikke er relevante for det aktuelle spørsmålet. Sterke lesere klarer å orientere seg i et mangfold av opplysninger og velge ut relevant informasjon for å skape mening, mens svake lesere i større grad lar seg forvirre og leser bare for å bli ferdig.¹

¹ Roe, Astrid: Lesedidaktikk – etter den første leseopplæringen. Oslo 2014, s. 89.

Oppgave til teksten «Skyggen»

Denne teksten er hentet fra boka «Gullfalken. Ni nye eventyr» av Erna Osland.
Bruk teksten når du svarer på oppgavene.

Skyggen

I landet med snø og lange netter bodde det en hvithåret vismann som kunne alle språk og alle slags bokstaver. Han hadde lest de fleste bøkene i verden. Hele dagen leste han, og hele natta, hver eneste natt.

For han fikk ikke sove!

Vismannen fortalte ingen hva han drev med om natta, han var så redd for at folk ville le av han. Tenk at han som hadde lest så mye og visste nesten alt, ikke visste hvordan han kunne sove!

Å sove var jo ingen kunst. Til og med de som ikke kunne en eneste bokstav og aldri hadde lest en bok, klarte det. Så vismannen holdt det hemmelig at han ikke fikk sove. Og for at ingen skulle se at han var våken hver natt, trakk han for gardinene så snart mørket kom. Helt alene satt han lys våken og leste til det lysnet neste morgen. Da gjorde han som andre folk og gikk på arbeid. Og der på kontoret sitt fortsatte han å lese.

Etter mange år med evig lesing, ble vismannen så trøtt at han ikke klarte å være oppe om natta. Han begynte å legge seg om kvelden i håp om å få sove.

Det hjalp ikke.

Time etter time lå han i mørket med øynene lukket. Ikke ett sekund sov han. Han ble så trøtt at det bare var én ting å gjøre: Søke råd hos andre.

Først gikk han til søsteren sin. Hun sa at han måtte drikke varm melk, og så gjorde han det. Han drakk og drakk, men han var like våken.

En klok kone sa at han ikke måtte drikke, bare tisse før han la seg. Å tisse var viktig, gjerne flere ganger før han gikk til sengs. Men samme hvor mye vismannen tisset, sovnet han ikke.

Da gikk vismannen til en tryllekunstner som rådet han til å si hokuspokus og stå på én fot før han la seg.

Sovnet han av det? Nei.

Vismannen dro fra den ene til den andre og spurte om råd. En imam sa at han måtte synge sju sanger i alle himmelens retninger. En stjernetyder sa at han måtte spise fjorten ulike frukter ved fullmåne.

Den stakkars vismannen fulgte alle slags råd, men ingen ting hjalp. Til slutt ble han så utmattet at han knapt visste hvem han var og hvor han bodde. Han spiste ikke, han vasket seg ikke, og han stelte ikke huset sitt. Han var for sliten til å bry seg om noe som helst.

De siste kreftene sine brukte han på å lese. Men nå leste han bare gode råd om å få sove. Og det gikk sakte, sakte. Han måtte stave hvert eneste ord, for han var så trøtt.

Så snart mørket kom, slepte han seg til sengs, og der lå han våken, utslitt og redd. Overalt i huset var det uleste bøker og skitne kopper og kar. Klærne hans lå strødd utover, og oppgangen var full av snø, han hadde ikke krefter til å måke. Han orket ingen ting. Ikke engang døra si orket han å lukke! Han lå gjemt under dyna og ventet på at det skulle bli dag, så han på nytt kunne håpe på å få sove om natta.

Og derfor merket han ikke at det en natt kom noen smygende inn i huset!

En svart liten skygge hoppet gjennom snøen, ristet seg og smatt inn den åpne døra. Stille smatt den svarte skyggen inn på soverommet til vismannen. Stille krøp han innunder den varme dyna. Og like stille la han seg ned. Han ble liggende inntil vismannen, tett inn mot det varme kinnet hans og sove.

Vismannen selv var så forkavet, varm og utmattet at han ikke klarte å protestere.

Og etter en stund sovnet også han.

Om morgenen, på den tiden han og andre brukte å stå opp, lå vismannen fremdeles og sov.

– Hva er det med vismannen vår? spurte naboene da de gikk forbi huset hans.

– Han er vel ikke død, han som aldri får sove?

De oppdaget sporene i snøen og gikk inn.

Og hva fikk de se?

Vismannen og en katt. Som begge sov!

Katten lå på armen til vismannen med snuten innunder kinnet hans.

Stille hvisket naboene god natt – *god natt, god natt* – før de listet seg ut i morgenlyset.

Slik ble det hver natt heretter. Den hvithårede vismannen og den vesle svarte katten sov i samme seng.

Natt etter natt fikk vismannen sove. Så sant han hadde katten der!

Vismannen ble uthvilt og lettet. Og så lykkelig at han gjorde alt han kunne for den vesle katten: Ga den fløte og klappet den. Ja, han prøvde til og med å lære den å lese. Men det ville ikke katten. Den ville bare ligge i fanget til vismannen og varme seg. Mens vismannen selv leste i rasende fart, for det var skrevet en hel stabel med nye bøker i den tiden han hadde vært for trøtt til å lese.

Oppgave 18

Hvilken setning oppsummerer best budskapet i eventyret?

- A Stol på deg selv, ikke be om råd fra andre.
- B Svaret på alle verdens problemer kan du finne i bøker.
- C Jo smartere du er, desto vanskeligere blir det å sove.
- D All verdens kunnskap løser ikke alltid et enkelt problem.

Aspekt: Tolke og sammenholde informasjon i teksten.

Nasjonal løsningsprosent: 61 %

Fag: Norsk

For å svare riktig på denne oppgaven må elevene få overblikk over hele teksten og velge en setning som representerer en gyldig tolkning av eventyret.

Mange fortolkninger av dette eventyret er mulige, og en flervalgsoppgave med korte svaralternativer kan ikke gi en dekkende oppsummering av budskapet. Det er likevel mulig å presentere elevene for noen tolkninger som står i sterk motsetning til det som er riktig, slik både svaralternativ A og B gjør. Disse svaralternativene viser til hva vismannen gjør for å løse problemet sitt i starten av eventyret. Han vil jo helst klare seg selv, og han er en mann som søker kunnskap i bøker, men verken svaralternativ A, B eller C representerer løsninger på vismannens søvnproblemer.

Budskapet i teksten er imidlertid knyttet til løsningen på vismannens problem. Svaralternativ D, «All verdens kunnskap løser ikke alltid et enkelt problem», presenterer en gyldig tolkning av eventyret. Dette svaret viser til motsetningen mellom vismannens store kunnskaper og det «lille» problemet han må løse. Oppgaven kan egne seg som utgangspunkt for å diskutere videre hvilket kunnskapssyn som formidles i teksten.

Oppgave til teksten «Hva er penger?»

Denne teksten er hentet fra boka «Hva er økonomi? Fakta og funderinger» av Gunnhild J. Ecklund og Bjørk Bjarkadóttir. Bruk teksten når du svarer på oppgavene.

Penger

Da du var liten, tenkte du sikkert at penger var det samme som sedler og mynter. Du sparte kanskje penger ved å putte dem på en sparebøsse? I dag er det mindre og mindre vanlig å ha mynter og sedler. I stedet handler vi ofte med bankkort når vi er i butikken. Vi kan også bruke mobiltelefon, nettbrett og datamaskiner til å kjøpe varer på Internett. Da er plutselig penger bare tall på en skjerm. Dermed kan det bli vanskeligere å holde oversikt over hvor mye penger du bruker. Det blir også vanskelig å forstå hva penger egentlig er.

Hva er penger?

Penger kan være nesten hva som helst. Så lenge folk flest i et område er enige om at noe er penger, så kan det brukes som penger. Derfor kalles penger med fine ord et *allment akseptert betalingsmiddel*. Det betyr altså noe som de fleste godtar at vi kan betale med. Opp gjennom tidene har folk blitt enige om å bruke mye rart som penger.

Pengenes historie

Vi vet ikke helt når de første pengene ble tatt i bruk. For 5000 år siden skrev folk i Mesopotamia det som nå er Irak, på leirtavler og brukte dem når de handlet eller lånte ut penger. På stillehavets øya Yap valgte innbyggerne en mer tungvint løsning. Der brukte de penger laget av svære, runde steiner. Ellers har folk til forskjellige tider og på forskjellige steder brukt både skjell, perler og pels som penger, i tillegg til edle metaller som gull, sølv, kobber og bronse. I Norge har vi tidligere brukt smør og tørrfisk som betalingsmiddel. De første myntene kom i bruk for rundt 2600 år siden og er funnet i det som i dag er Tyrkia. De første pengesedlene dukket opp i Kina for rundt 1000 år siden.



De store steinpengene på øya Yap i Stillehavet kalles *rai*. De største er over 3,5 meter høye og veier flere tonn. I dag er steinpengene mest til pynt.

Mynter og sedler i Norge

Den eldste norske mynten som er funnet, er en sølvpenning fra år 995. Den fikk Olav Tryggvason laget like etter at han ble konge i Norge. 700 år senere kom den første seddelen. Det skjedde i 1695. Da ga den dansk-norske kongen en kjøpmann fra Bergen tillatelse til å utstede sedler. Dette ble et lite vellykket eksperiment, blant annet fordi folk ikke stolte på at sedlene hadde noen verdi. Først i 1816, da Stortinget opprettet Norges bank, fikk Norge et nasjonalt system for å trykke og bruke sedler.

Oppgave 22

I hvilket år fikk Norge sin første pengeseddel?

- A 700
- B 995
- C 1695
- D 1816

Aspekt: Finne informasjon i teksten.

Nasjonal løsningsprosent: 48 %

Fag: Samfunnsfag

For å svare riktig på denne oppgaven må elevene lese grundig og skille sterkt konkurrerende informasjon fra informasjon som er relevant for oppgaven.

Svaret på oppgaven ligger i tekstboksen, og sterke lesere vil kunne svare ved først å letelese seg fram til riktig sted og deretter lese tekstboksen nøye.

Alle de fire tallene i svaralternativene finnes i teksten og er dermed plausible svar på spørsmålet. Tallene står dessuten i nærheten av det riktige årstallet, og dette kompliserer oppgaven, som krever at elevene leser konsentrert for ikke å plukke ut feil informasjon. Tallet 700 viser til hvor mange år det gikk fra en av de første myntene ble produsert, til vi fikk vår første seddel. Tallet 995 viser til hvilket år den eldste mynten vi kjenner til, er fra. Tallet 1695 viser til når den første norske seddelen utkom, mens 1816 viser til året da Norge fikk et nasjonalt system for å trykke penger og bruke sedler.

Oppgaven er forholdsvis vanskelig, og det kan være flere grunner til det. For det første vet vi at mange elever har problemer med å lokalisere informasjon i sammensatte tekster, og at de ikke leser tekstbokser eller bildetekster, men forholder seg kun til hovedteksten. For det andre kan det for mange elever være krevende å skille ut relevante opplysninger i en tekst som er oppramsende og informasjonstett.

IRT-analyse

Forslaget til endelig prøve er analysert ved hjelp av IRT, og tabell 11 viser parametrene for alle oppgavene i prøveforslaget.

Tabell 11: Parametre fra IRT-analyse av oppgavene i prøveforslaget 2018. N = 1561 (781 gutter, 780 jenter). Gjennomsnittlig p-verdi = 0,65. Alpha for hele prøven = 0,914. Gjennomsnittlig theta = - 0,021 (SD 0,901). Lav diskriminering (point biserial < 0,250, og/eller a-parameter < 0,5) er merket (*).

Oppg.	Item-ID	Oppgave	P-verdi	Diskrim. (p-bis)	a	b	Flagg	DIF (bias mot)
1	L047S01	George_1	0,729	0,342	0,489	-1,412		
2	L047S02	George_2	0,640	0,393	0,584	-0,735		
3	L047S04	George_3	0,771	0,377	0,611	-1,475		
4	L047S05	George_4	0,476	0,370	0,530	0,123		
5	L047S07	George_5	0,639	0,277	0,381 *	-0,985		
6	L047S06	George_6	0,717	0,410	0,639	-1,107		
7	L047S09	George_7	0,852	0,357	0,659	-1,993		
8	L046S03	Skyggen_1	0,828	0,382	0,689	-1,754		
9	L046S04a-d	Skyggen_2	0,686	0,417	0,646	-0,929		
10	L046S05	Skyggen_3	0,827	0,512	1,145	-1,368		
11	L046S07	Skyggen_4	0,737	0,480	0,870	-1,038		
12	L046S08	Skyggen_5	0,704	0,428	0,692	-0,990		Gutter
13	L043S01	Penger_1	0,578	0,457	0,705	-0,365		
14	L043S02	Penger_2	0,518	0,358	0,516	-0,108		
15	L043S03	Penger_3	0,565	0,406	0,597	-0,337		
16	L043S04	Penger_4	0,461	0,320	0,473	0,220		
17	L043S05	Penger_5	0,701	0,484	0,824	-0,886		
18	L043S06	Penger_6	0,470	0,417	0,654	0,130		
19	L051S01	Eplepai_1	0,914	0,351	0,839	-2,282		
20	L051S02	Eplepai_2	0,619	0,410	0,605	-0,609		
21	L051S03	Eplepai_3	0,664	0,210 *	0,329 *	-1,322		
22	L051S04	Eplepai_4	0,842	0,401	0,791	-1,721		
23	L051S07	Eplepai_5	0,763	0,503	0,939	-1,128		
24	L053S01	Røyking_1	0,451	0,237 *	0,522	0,255		
25	L053S03	Røyking_2	0,213	0,162 *	0,506	1,772		
26	L053S04	Røyking_3	0,775	0,256	0,668	-1,421		
27	L053S05	Røyking_4	0,787	0,292	0,771	-1,381		
28	L053S06	Røyking_5	0,245	0,225 *	0,582	1,390		
29	L053S08	Røyking_6	0,618	0,244 *	0,558	-0,633		

I tabell 11 er oppgavenes diskrimineringsevne oppgitt både som et klassisk mål (point biserial) og en a-verdi. Det klassiske målet kan avvike fra det som tidligere er vist i tabell 8 og 9, fordi IRT-analysen er basert på data fra pilotering av to ulike prøver, og det brukes ulike diskrimineringsmål. Teksten *Røyking* ble prøvd ut sammen med tekster som ble prøvd ut til den nasjonale prøven for 5. trinn i 2019. Data knyttet til denne teksten ble samkalibrert med data fra tekstene til andre pilotering (uten data fra teksten *Farger*) for å plassere alle item-ene på samme skala. Oppgaver med diskriminering under 0,250 (p-bis) eller a<0,5 er merket med stjerne, og vil bli kommentert i avsnittene nedenfor. Det samme gjelder oppgaver med DIF (Differential Item Functioning) for kjønn. Ingen av oppgavene ble flagget (F).

Tabell 12 viser den overordnede «model fit» for forslaget til endelig prøve. Ingen av oppgavene i forslaget har signifikant «model misfit».

Tabell 12: Overordnet «model fit» for prøveforslaget.

Test	Items	Chi-square	df	p	-2LL
Full Test	29	1286,822	377	0,000	26655

Oppgaver med lav diskriminering

Fire av oppgavene med lav diskriminering (under 0,250) er knyttet til teksten *Røyking*. Tekstene fra prøve C ble pilotert på 1129 elever, mens *Røyking* ble pilotert på 433 elever, der teksten ble pilotert på første plass i prøvestrukturen. Samkalibreringen har endret diskrimineringstillene for denne teksten, og flere av item-ene har dårligere diskrimineringsverdier enn de hadde i den opprinnelige piloteringen. Vi vil likevel argumentere for at teksten og oppgavene må inkluderes i den endelige prøven for 2018. For det første har to av oppgavene, «Røyking_2» og «Røyking_5», p-verdier på henholdsvis 0,21 og 0,25. Erfaringsmessig vet vi at oppgaver med lav p-verdi kan diskriminere dårlig. En kvalitativ vurdering av disse oppgavene fastslår at det er ingen tvil om hvilket svaralternativ som er riktig, og distraktorene er tydelig feil. For det andre ser vi at a-verdiene for alle oppgavene knyttet til denne teksten er over 0,5. Det er svært ønskelig å bruke teksten *Røyking*, siden den inneholder oppgaver på det øverste mestringsnivået.

Oppgaven «Eplepai_3» har en diskrimineringsevne på 0,210 som er under det fastsatte målet på 0,250. Oppgaven ser slik ut:

3	Står ingrediensene i oppskriften i noen bestemt rekkefølge? (L051S03)
A	ja, de står i den rekkefølgen de skal brukes
B	ja, først de tørre ingrediensene og så de våte
C	ja, de viktigste ingrediensene står til slutt
D	nei, ingrediensene står i tilfeldig rekkefølge

Vi vil argumentere for at oppgaven må beholdes i den endelige prøven av to grunner. For det første tester denne oppgaven en viktig del av leseforståelsen knyttet til sjangeren oppskrift. For det andre måler oppgaven elevens evne til å reflektere over tekstens form. Ingrediensene i en oppskrift skrevet for barn er ofte ordnet slik vi ser i denne teksten. Rekkefølgen er ikke tilfeldig, men ingrediensene står i den rekkefølgen de skal brukes (svaralternativ A). Elevene kan bruke sine forkunnskaper om oppskrifter for å løse oppgaven, men de kan også lese seg fram til svaret ved å tolke informasjon i teksten og bruke egne erfaringer med ingrediensene i oppskriften.

Oppgaven «Ensomme George_5» diskriminerer også litt dårlig. I denne oppgaven har vi gjort en mindre endring i en av distraktorene, noe som kan endre hvor godt oppgaven diskriminerer.

Oppgaver med DIF

En DIF-test avslører om oppgaven potensielt favoriserer det ene kjønn over det andre gitt at elever med ulikt kjønn har samme ferdighet. En slik analyse skiller seg fra en tradisjonell måling av kjønnsforskjell som måler hvilket kjønn som presterer best totalt sett.

Det er en oppgave med DIF (Differential Item Functioning) for kjønn i forslaget til prøve, og denne oppgaven er knyttet til teksten *Skyggen*. Oppgaven har DIF i disfavør av gutter.

Oppgaven «Skyggen_5» ser slik ut:

5 Hvilken setning oppsummerer best budskapet i eventyret? (L046S08)

- A Stol på deg selv, ikke be om råd fra andre.
- B Svaret på alle verdens problemer kan du finne i bøker.
- C Jo smartere du er, desto vanskeligere blir det å sove.
- D All verdens kunnskap løser ikke alltid et enkelt problem.

Det er flere grunner til at denne oppgaven må beholdes i det endelige prøveforslaget. For det første fungerer oppgaven godt etter andre statiske måleverdier som for eksempel diskriminering. For det andre måler denne oppgaven en sentral del av leseforståelsen knyttet til eventyret. Det er en viktig lesekompetanse å kunne trekke ut budskapet eller hovedpoenget i en tekst. For det tredje er dette en skjønnlitterær tekst som har en fast plass i konstruktet og som vi vet av erfaring at jentene gjør det bedre enn guttene på. Alt i alt er resultatene for denne type oppgaver knyttet til skjønnlitterære tekster i tråd med det forskning har vist om kjønnsforskjeller i lesing.

Oppgavene fordelt på mestringsnivåer

Tabell 13 viser oppgavene i forslag til endelig prøve, fordelt på mestringsnivåer. Nivå-grensene som er brukt, ble bestemt etter hovedgjennomføringen av nasjonale leseprøver i 2016. Theta-verdiene som kommer ut av denne IRT-analysen, er sentrert på theta-fordelingen i piloteringen og er ikke de samme som ved selve gjennomføringen. Piloteringen er altså ikke ankret på startpunktet av trenden i 2016. Det er derfor en viss usikkerhet knyttet til nivågrensene.

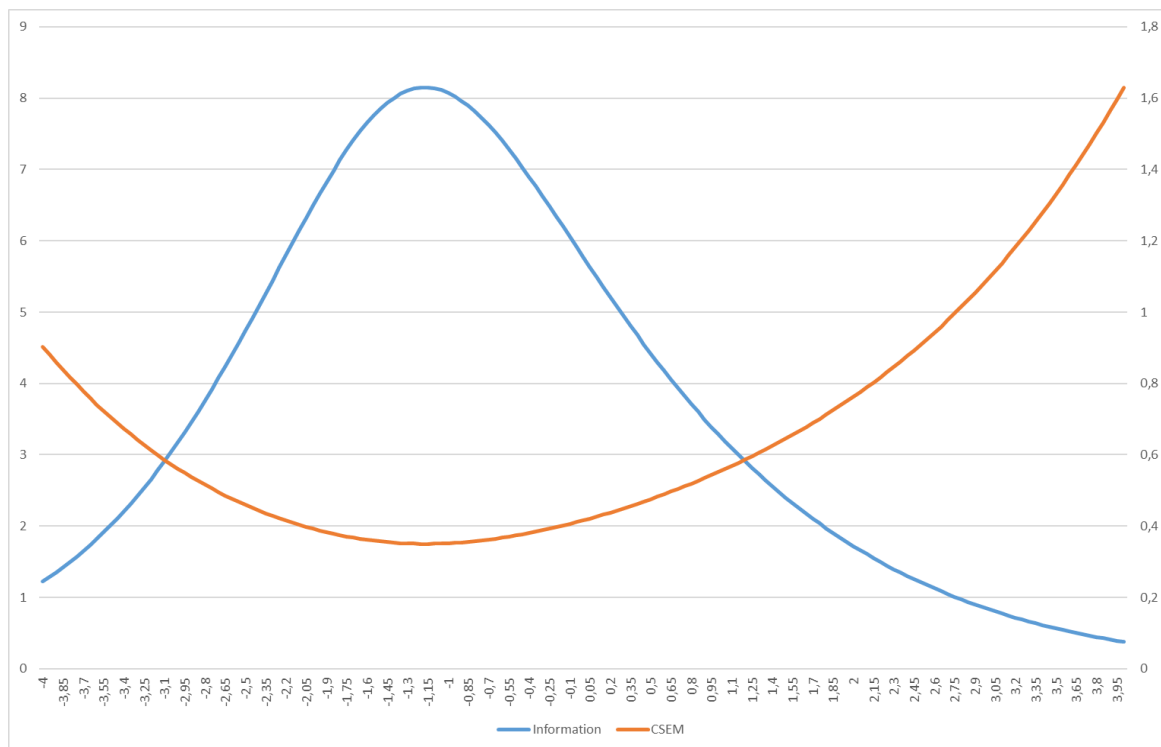
Ideelt sett skal prøven måle likt i alle deler av skalaen, og antallet oppgaver bør med andre ord være det samme innenfor alle mestringsnivåene. Dette har imidlertid vist seg å være vanskelig å få til, og i prøveforslaget for 2018 ser vi det samme som tidligere: Det er bare to oppgaver på det øverste nivået, og det er også færre oppgaver enn ønskelig på nivå to.

Tabell 13: Oppgavene i forslaget til endelig prøve, fordelt på mestringsnivåer.

Nivå	Item-ID	Oppgave	Skalapoeng
Nivå 1	L051S01	Eplepai_1	27,178
	L047S09	George_7	30,073
	L046S03	Skyggen_1	32,460
	L051S04	Eplepai_4	32,790
	L047S04	George_3	35,252
	L053S04	Røyking_3	35,788
	L047S01	George_1	35,881
	L053S05	Røyking_4	36,186
	L046S05	Skyggen_3	36,320
	L051S03	Eplepai_3	36,777
	L051S07	Eplepai_5	38,723
	L047S06	George_6	38,928
	L046S07	Skyggen_4	39,624
	L046S08	Skyggen_5	40,098
	L047S07	George_5	40,150
	L046S04a-d	Skyggen_2	40,707
	L043S05	Penger_5	41,138
	L047S02	George_2	42,645
Nivå 2	L053S08	Røyking_6	43,668
	L051S02	Eplepai_2	43,908
	L043S01	Penger_1	46,351
	L043S03	Penger_3	46,634
	L043S02	Penger_2	48,923
	L047S05	George_4	51,234
	L043S06	Penger_6	51,296
	L043S04	Penger_4	52,199
	L053S01	Røyking_1	52,547
Nivå 3	L053S06	Røyking_5	63,901
	L053S03	Røyking_2	67,719

Testinformasjon og måleusikkerhet

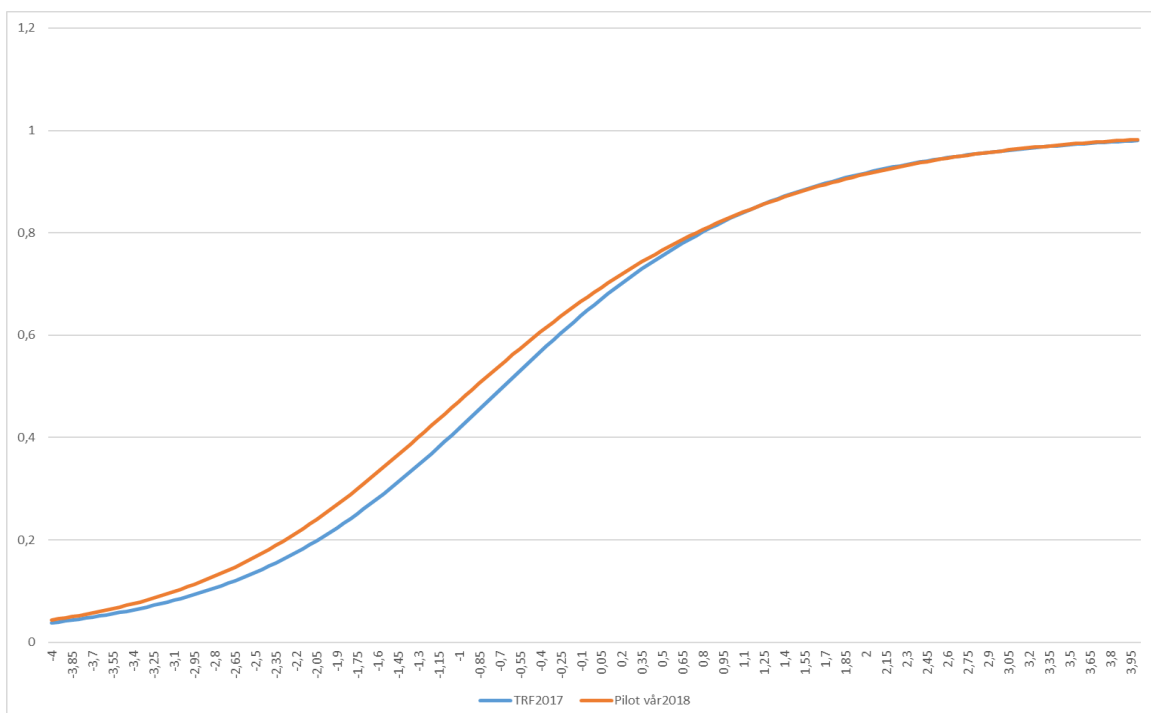
En IRT-analyse av det endelige prøveforslaget, med resultatet fra andre pilotering, viser at testinformasjonen fra prøven (TIF) som prøven leverer i sin helhet, er høyest på theta lik - 1,15, og at måleusikkerheten (CSEM) er minst i det samme ferdighetsområdet i prøven. Figur 4 viser testinformasjon og måleusikkerhet for forslaget til endelig prøve.



Figur 4: Informasjonsverdi og måleusikkerhet for prøveforslaget 2018

Prøvens TRF (Test response function)

Figur 5 viser test respons funksjon (TRF) for det nye prøveforslaget (rød linje), sammenlignet med prøven 2017 (blå linje).



Figur 5: TRF for prøveforslaget 2018, sammenlignet med prøven 2017

TRF er en beskrivelse av sammenhengen mellom estimert ferdighet (θ) og andel riktig besvarte oppgaver i prøven. Figur 5 viser med andre ord hvor stor andel av oppgavene en elev med en gitt ferdighet (θ) må klare for å få samme resultat på de to prøvene. Det går fram av figuren at det er en liten forskjell mellom prøvene, men forskjellen er ikke lik alle steder på ferdighetsskalaen: For å oppnå samme resultat må elever med ferdighet lavere enn 0,5 løse flere oppgaver i prøveforslaget enn de måtte i 2017-prøven.

Det er viktig å understreke at sammenligningen som her er gjort, er forbundet med noe usikkerhet, fordi 2017-prøven inneholder et sett av ankeroppgaver som ikke er inkludert i prøveforslaget 2018.

5. Hovedgjennomføring

Den nasjonale prøven i lesing for 5. trinn ble gjennomført i tidsrommet 15. oktober – 02. november 2018, i prøvegjennomføringssystemet PAS-Prøver. Totalt 55436 elever gjennomførte den ordinære leseprøven, som besto av fem tekster og 29 oppgaver. I tillegg ble tre ankerprøvesett bestående av tre ordinære tekster og to ankertekster tilfeldig distribuert til totalt 4527 av elevene på 5. trinn. Ankertekstene gjør det mulig å måle endring i elevers leseferdigheter over tid og vil bli brukt om igjen i neste års prøver.

Dette kapitlet beskriver resultatene til samtlige elever som gjennomførte den *ordinære* leseprøven på 5. trinn (55 436 elever). Resultatene blir presentert som gjennomsnittlige p-verdier for hver av oppgavene i prøven og for jenter/gutter, samt den samlede p-verdien for hele prøven.

Oppgavene blir også kategorisert etter ulike kriterier på tekst- og oppgavenivå, og det blir gitt resultater totalt og for jenter og gutter basert på disse kriteriene. Forskjellene mellom kjønnene og mellom ulike oppgaveformater og leseaspekter blir beskrevet som differanser i p-verdi. For prøven som helhet blir kjønnsforskjellene også rapportert som differanser målt med standardavvik som enhet.

Tabell 14 gir en oversikt over tekstene i den endelige prøven.

Tabell 14: Fagtilhørighet og teksttyper for tekstene i 2018-prøven:

Tekstnavn i rapport	Original tittel	Fagtilhørighet	Teksttype	Antall oppgaver: 29 *	Antall ord per tekst
<i>George</i>	«Ensomme George»	Naturfag	Fagboktekst	7	808
<i>Røyking</i>	«Jeg hater lukten av røyk!»	Samfunnsfag, norsk	Avisartikkel	6	1152
<i>Skyggen</i>	«Skyggen»	Norsk	Eventyr	5	1116
<i>Penger</i>	«Hva er penger?»	Samfunnsfag	Fagboktekst	6	689
<i>Eplepai</i>	«Eplepai»	Mat og helse	Oppskrift	5	546

* 28 av oppgavene i prøven er flervalgsoppgaver, 1 oppgave er en sammensatt flervalgsoppgave

Resultater for prøven som helhet

Tabell 15 nedenfor inneholder en oversikt over alle oppgavene i prøven med beskrivelser av oppgaveformat og leseaspekt, samt gjennomsnittlige resultater i form av p-verdier for jenter og gutter, samt differansen mellom disse. Tabellen viser også samlet p-verdi for hver av tekstene, samt den gjennomsnittlige p-verdien for hele prøven.

Tabell 15: Leseaspekt, oppgaveformat og gjennomsnittlige p-verdier for hver av oppgavene i prøven. (Summene er avrundet til to desimaler i Excel)

Oppg.	Tekst	Aspekt	Oppgave-format	Alle	Jenter	Gutter	Kjønns-forskj. (j-g)
1	George_1	Finne	Flervalg	0,67	0,68	0,66	0,01
2	George_2	Tolke	Flervalg	0,56	0,57	0,54	0,03
3	George_3	Finne	Flervalg	0,70	0,71	0,68	0,03
4	George_4	Tolke	Flervalg	0,47	0,45	0,49	-0,04
5	George_5	Reflektere	Flervalg	0,48	0,51	0,46	0,04
6	George_6	Finne	Flervalg	0,68	0,70	0,67	0,02
7	George_7	Tolke	Flervalg	0,73	0,76	0,70	0,06
Gjennomsnitt for George				0,61	0,62	0,60	0,02
8	Røyking_1	Tolke	Flervalg	0,42	0,46	0,38	0,07
9	Røyking_2	Tolke	Flervalg	0,27	0,27	0,27	0,01
10	Røyking_3	Finne	Flervalg	0,64	0,68	0,61	0,07
11	Røyking_4	Reflektere	Flervalg	0,65	0,69	0,61	0,07
12	Røyking_5	Tolke	Flervalg	0,26	0,28	0,24	0,04
13	Røyking_6	Reflektere	Flervalg	0,47	0,51	0,43	0,09
Gjennomsnitt for Røyking				0,45	0,48	0,42	0,06
14	Skyggen_1	Tolke	Flervalg	0,72	0,74	0,71	0,03
15	Skyggen_2	Finne	Paring	0,67	0,72	0,62	0,10
16	Skyggen_3	Tolke	Flervalg	0,76	0,80	0,73	0,06
17	Skyggen_4	Reflektere	Flervalg	0,68	0,72	0,65	0,07
18	Skyggen_5	Tolke	Flervalg	0,60	0,64	0,57	0,06
Gjennomsnitt for Skyggen				0,69	0,72	0,66	0,07
19	Penger_1	Tolke	Flervalg	0,54	0,54	0,54	0,00
20	Penger_2	Finne	Flervalg	0,56	0,54	0,57	-0,03
21	Penger_3	Tolke	Flervalg	0,56	0,58	0,53	0,05
22	Penger_4	Finne	Flervalg	0,47	0,47	0,48	-0,01
23	Penger_5	Tolke	Flervalg	0,64	0,65	0,63	0,02
24	Penger_6	Reflektere	Flervalg	0,45	0,45	0,44	0,01
Gjennomsnitt for Penger?				0,54	0,54	0,53	0,01
25	Eplepai_1	Finne	Flervalg	0,91	0,92	0,90	0,02
26	Eplepai_2	Tolke	Flervalg	0,65	0,68	0,61	0,07
27	Eplepai_3	Reflektere	Flervalg	0,65	0,67	0,63	0,03
28	Eplepai_4	Finne	Flervalg	0,82	0,84	0,79	0,05
29	Eplepai_5	Finne	Flervalg	0,72	0,74	0,70	0,03
Gjennomsnitt for Eplepai				0,75	0,77	0,73	0,04
Gjennomsnitt for hele prøven				0,60	0,62	0,58	0,04

Den gjennomsnittlige p-verdien for hele leseprøven er 0,6. Det vil si at en gjennomsnittselev på 5. trinn klarer 60 prosent av oppgavene. 398 av de 55 436 elevene på 5. trinn oppnådde maksimal poengsum på 29 poeng.

Prøven har en spredning i vanskegrad, både totalt og innen hver tekstenhet. Den vanskeligste oppgaven i 2018-prøven er «Røyking_5», med en gjennomsnittlig p-verdi på 0,26. Den letteste oppgaven i prøvesettet er «Eplepai_1», med en gjennomsnittlig p-verdi på 0,90.

Fire av de fem vanskeligste oppgavene i prøvesettet er knyttet til teksten *Røyking*, noe som var årsaken til at denne teksten ble tatt inn som erstatning for den piloterte teksten *Farger*.

Kjønnsforskjeller

I likhet med tidligere gjennomføringer viser den nasjonale prøven i lesing for 2018 at jentene presterer bedre enn guttene.

Tabell 16: Kjønnsforskjeller i lesekompetanse, uttrykt som p-verdier og effektstørrelse

Kjønnsforskjell	5. trinn
Jenter	0,62
Gutter	0,58
Kjønnsforskjell (jenter-gutter)	0,04
Effektstørrelser i form av standardavvik	0,18

Tabell 16 viser de gjennomsnittlige p-verdiene for jenter og gutter, samt kjønnsforskjell uttrykt som differanse mellom disse. Effektstørrelsen er beregnet som prosentandeler av standardavvikene. Kjønnsforskjellene i leseprestasjoner er på nivå med forskjellene i 2017-prøven.

Klassisk item-analyse av prøven 2018

Tabell 17 viser resultater fra klassisk item-analyse av den nasjonale prøven i lesing 2018.

Tabell 17: Resultater fra klassisk item-analyse, endelig prøve, 5. trinn.

Antall elever (N): 55436 (27948 gutter, 27488 jenter). Reliabilitet (Cronbachs Alpha): 0,860.

Gjennomsnittlig p-verdi: 0,60 (gutter: 0,58 / jenter: 0,62). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Oppg.	Item-ID	Tekst- og oppgavenavn	Svarfordeling i %: (frekvens)					Gj.snittlig poengsum på prøven for de som har valgt svaralt.					Diskim. (biserial)	P- verdi	Kjønns- forskj. (j-g)*	Komm.
		Flervalgsoppgaver	A	B	C	D	Blank	A	B	C	D	Blank				
			0	1				0	1							
1	L047S01	George_1	23	4	6	67	0	15	11	11	19	7	0,378	0,67	0,01	
2	L047S02	George_2	56	15	17	12	0	20	15	15	13	7	0,377	0,56	0,03	
3	L047S04	George_3	9	70	11	9	0	12	19	13	12	6	0,440	0,70	0,03	
4	L047S05	George_4	17	32	47	5	0	14	16	20	12	6	0,370	0,47	-0,04	
5	L047S07	George_5	16	24	11	48	0	14	17	15	19	7	0,243	0,48	0,04	**
6	L047S06	George_6	9	68	10	12	0	14	19	14	13	7	0,373	0,68	0,02	
7	L047S09	George_7	11	5	11	73	1	14	12	13	19	8	0,379	0,73	0,06	
8	L046S03	Røyking_1	46	5	42	8	0	16	14	20	15	4	0,326	0,42	0,07	
9	L046S04a-d	Røyking_2	23	26	27	24	1	14	17	21	18	6	0,281	0,27	0,01	
10	L046S05	Røyking_3	12	15	65	8	1	14	14	19	13	7	0,391	0,64	0,07	
11	L046S07	Røyking_4	12	4	65	18	1	14	10	20	14	7	0,437	0,65	0,07	
12	L046S08	Røyking_5	26	22	9	42	1	21	17	16	16	7	0,254	0,26	0,04	
13	L043S01	Røyking_6	28	8	17	47	1	14	13	15	21	6	0,467	0,47	0,09	
14	L043S02	Skyggen_1	7	10	10	72	1	10	14	13	19	6	0,424	0,72	0,03	
15	L043S03	Skyggen_2	33	67				13	20				0,465	0,67	0,10	***
16	L043S04	Skyggen_3	76	7	8	7	1	19	11	12	11	6	0,538	0,76	0,06	
17	L043S05	Skyggen_4	14	68	7	9	1	13	20	11	12	6	0,530	0,68	0,07	
18	L043S06	Skyggen_5	22	8	8	60	2	15	12	13	20	7	0,434	0,60	0,06	
19	L051S01	Penger_1	14	54	17	14	2	13	20	16	14	7	0,454	0,54	0,00	
20	L051S02	Penger_2	26	22	11	10	2	20	16	15	14	6	0,319	0,56	-0,03	
21	L051S03	Penger_3	16	14	13	56	2	13	15	14	20	7	0,445	0,56	0,05	
22	L051S04	Penger_4	11	25	47	15	2	14	16	20	15	7	0,313	0,47	-0,01	
23	L051S07	Penger_5	64	14	8	12	2	20	13	14	14	8	0,465	0,64	0,02	
24	L053S01	Penger_6	31	45	12	10	3	16	20	14	15	8	0,358	0,45	0,01	
25	L053S03	Eplepai_1	1	4	91	3	2	9	14	18	11	7	0,289	0,91	0,02	
26	L053S04	Eplepai_2	12	16	4	65	2	12	16	12	20	8	0,415	0,65	0,07	
27	L053S05	Eplepai_3	65	8	7	18	3	19	14	14	17	8	0,216	0,65	0,03	**
28	L053S06	Eplepai_4	3	82	10	3	3	12	19	13	10	8	0,409	0,82	0,05	
29	L053S08	Eplepai_5	72	6	11	9	3	19	12	13	13	8	0,478	0,72	0,03	

*Positive verdier i jentenes favør, negative i guttenes favør*** Lave diskrimineringsverdier (<0,250)

*** Sammensatt flervalgsoppgave

Andelen ubesvarte oppgaver er i gjennomsnitt 1,27 prosent – altså nokså lav. Økningen ubesvarte øker heller ikke nevneverdig ut over i prøven, og det er bare tre prosent som ikke har besvart de siste oppgavene i prøven. Det ser også ut til at andelen blanke svar er nokså jevnt fordelt mellom jentene og guttene på 5. trinn.

I estimatet for prøven hadde flere av oppgavene knyttet til teksten *Røyking* lav diskriminering. Etter gjennomføring ser vi at disse oppgavene har fungert tilfredsstillende. To av oppgavene i prøvesettet – «George_5» og «Eplepai_3» har diskrimineringsverdier under 0,250 etter hovedgjennomføringen. I kapitlet «Oppgaver med lav diskriminering» argumenterte vi for å beholde «Eplepai_3», selv om oppgaven diskriminerte litt dårlig. Under pilotering hadde «George_5» a-verdi under 0,5, men den klassiske diskrimineringen (point biserial) var på 0,277, noe som er over det fastsatte minstemålet. Etter gjennomføringen har diskrimineringsverdien falt under minstemålet, men disse endringene er marginale, og oppgaven forsvaret sin

plass i prøven. I likhet med «Eplepai_3» måler «George_5» elevens evne til å reflektere over formelle trekk ved teksten, og det er viktig å inkludere slike oppgavetyper i leseprøvene.

DIF-analyser av oppgavene i prøven 2018

IRT-analyser som verktøy i utviklingen av prøvene, har vist seg nyttig, men når de samme analysene anvendes på hele populasjonsdata, reiser det seg noen prinsipielle spørsmål. IRT-analysene er følsomme overfor antall elever (N), og i mange tilfeller blir mange oppgaver flagget fordi modellene som brukes ikke lenger passer når antallet elever blir så stort. Dette blir blant annet tydelig i DIF-analyser av resultatene fra den endelige gjennomføringen. Mantel Haenszel-testen, som brukes for å bestemme hvilke oppgaver som har DIF, er svært følsom overfor antall elever som er med i beregningsgrunnlaget. En DIF-test av resultatene fra hele prøven viser at over 20 oppgaver har DIF i disfavør et av kjønnene. Vi har også gjort tre randomiserte uttrekk fra populasjonsfilen, et på 1500 elever, et på 850 elever og et på 400 elever. Antall oppgaver med DIF endrer seg i de tre uttrekkene. I utvalget på 1500 elever, får fem oppgaver DIF. I utvalget på 850 elever får fire oppgaver DIF (men bare to av dem er identiske med oppgavene som fikk DIF i utvalget av 1500 elever). I utvalget på 400 elever har ingen oppgaver DIF. Slik vi ser det, vil en fornuftig bruk av testen kreve at to hensyn veies mot hverandre: Utvalget må være lite, slik at antall elever ikke får for stor betydning, samtidig som hele ferdigheten må være representert.

En DIF-analyse kan ha relevans i piloteringsfasen, da testen kan avsløre tekster og oppgaver som potensielt favoriserer det ene kjønn over det andre (gitt at elever med ulikt kjønn har samme ferdighet). Etter gjennomføring må denne testen brukes med fornuft, og vi etterlyser klare retningslinjer for hvilket utvalg som bør brukes og et klart mål for hva denne testen skal gi av informasjon etter den endelige gjennomføringen.

DIF-testen reiser også spørsmål om piloteringsrutinene er gode nok. Resultatene viser at det er vanskelig å få et representativt utvalg elever til å ta prøven. For eksempel kan det være vanskelig å få elever med høyest leseferdighet inkludert i pilotutvalgene. En mulig løsning på problemet kan være å inkludere piloteringen av neste års prøve i gjennomføringen av den nasjonale prøven.

Oppgavene fordelt på mestringsnivåer

Tabell 18 viser oppgavene i den nasjonale prøven i lesing fordelt på mestringsnivå. Sammenlignet med estimatet for den nasjonale prøven som baserte seg på piloteringsresultatene, ser vi at en del oppgaver har endret plassering, og i noen tilfeller nivå, men i hovedsak gir estimatet et ganske godt bilde av hvilke oppgaver som har vært lette og hvilke som har vært vanskelige. For eksempel er de samme oppgavene plassert på det høyeste mestringsnivået i estimatet og i den endelige prøven.

Tabell 18: Oppgavene i den nasjonale prøven i lesing på 5. trinn 2018 fordelt på mestringsnivå.

Nivå	Oppgave	Item-ID	Skalapoeng
Nivå 1	Elepai_1	L051S01	27,578
	Elepai_4	L051S04	36,14
	Elepai_3	L051S03	37,164
	Ensomme George_7	L047S09	38,578
	Skyggen_1	L046S03	39,825
	Ensomme George_6	L047S06	40,307
	Skyggen_3	L046S05	40,438
	Elepai_5	L051S07	40,885
	Ensomme George_1	L047S01	41,196
	Ensomme George_3	L047S04	41,279
	Skyggen_2	L046S04a-d	42,587
	Jeg hater lukten av røyk!_3	L053S04	42,595
	Elepai_2	L051S02	42,809
Nivå 2	Skyggen_4	L046S07	43,047
	Jeg hater lukten av røyk!_4	L053S05	43,049
	Penger_5	L043S05	43,856
	Skyggen_5	L046S08	45,129
	Penger_2	L043S02	46,323
	Ensomme George_2	L047S02	46,735
	Penger_3	L043S03	47,049
	Penger_1	L043S01	47,952
	Ensomme George_5	L047S07	50,582
	Jeg hater lukten av røyk!_6	L053S08	50,7
	Ensomme George_4	L047S05	51,043
	Penger_4	L043S04	51,074
	Penger_6	L043S06	52,065
	Jeg hater lukten av røyk!_1	L053S01	53,944
Nivå 3	Jeg hater lukten av røyk!_2	L053S03	64,174
	Jeg hater lukten av røyk!_5	L053S06	65,182

Vedlegg 1: Beskrivelser av lesekompetanse på de fem ulike nivåene

	FINNE	TOLKE	REFLEKTERE OG VURDERE
Nivå 1	lokalisere tydelig uttrykte elementer i en tekst med lite konkurrerende informasjon	trekke enkle slutninger eller oppfatte hovedtemaet i en tekst når innholdet er tydelig uttrykt i teksten	bruke personlige meninger til å kommentere form eller innhold i en tekst
Nivå 2	lokalisere tydelig uttrykte elementer i en tekst med klart konkurrerende informasjon	oppfatte hovedtemaet og forstå sammenhenger som ikke er tydelig uttrykt i teksten	identifisere formelle trekk ved tekster og ta stilling til eller vurdere tekstens innhold
Nivå 3	skille sterkt konkurrerende informasjon fra informasjon som er relevant for oppgaven	forstå motsetningsfylt innhold og komplekse sammenhenger i teksten	bruke kunnskap om språk og tekst til å identifisere og forholde seg til mer komplekse trekk ved tekstens form og innhold

Referanser

Silverman, D. (2011). *Interpreting qualitative data. Methods for analysing talk, text and interaction* (3. utg.). Thousand Oaks: SAGE Publications.

Utdanningsdirektoratet (2013). *Kunnskapsløftet. Læreplan for grunnskolen og videregående opplæring*.

Utdanningsdirektoratet (2017a). Rammeverk for grunnleggende ferdigheter. Lastet ned fra: <https://www.udir.no/laring-og-trivsel/lareplanverket/grunnleggende-ferdigheter/rammeverk-for-grunnleggende-ferdigheter/> [01.02.2019]

Utdanningsdirektoratet (2017b). Rammeverk for nasjonale prøver. Lastet ned fra <https://www.udir.no/eksamen-og-prover/prover/rammeverk-for-nasjonale-prover/> [25.04.18]

Utdanningsdirektoratet (2018). Metodegrunnlag for nasjonale prøver. Lastet ned fra <https://www.udir.no/globalassets/filer/vurdering/nasjonaleprover/metodegrunnlag-for-nasjonale-prover-august-2018.pdf> [09.02.19]