

Rapport fra piloteringer og gjennomføring av den nasjonale prøven i lesing på 8. trinn 2023



Jostein Andresen Ryen, Solveig Aspelund, Anna Eriksen, Tove Stjern Frønes,
Tove Hårsaker, Eva K. Narvhus, Morten Skar og Cecilie Weyergang

Enhet for kvantitative utdanningsanalyser
Institutt for lærerutdanning og skoleforskning

Universitetet i Oslo

15.02.2024

Innhold

Innledning	4
Måling av leseforståelse i de nasjonale prøvene i lesing	4
Tekster og oppgaveformater	5
Leseprosesser og eksempeltekst	7
Generaliserbarhet	9
Analyser i rapporten	9
Nasjonal prøve i lesing for 8. trinn 2023	12
Prøveutvikling i perioden 2021–2023	12
Innholdet i prøven	12
Første pilotering	15
Innholdet i prøve A, B og C	15
Blanke svar i prøve A, B og C	18
Klassisk item-analyse av prøve A, B og C	20
IRT-analyse av prøve A, B og C	24
Oppgavene fordelt på mestringsnivåer	26
Uvalg av tekster til andre pilotering	29
Endringer etter første pilotering	31
Andre pilotering	33
Innholdet i prøve D	33
Blanke svar i prøve D	34
Klassisk item-analyse av prøve D	35
IRT-analyse av prøve D	37
Ekstrapilot	39
Endringer før ekstrapiloten	39
Klassisk item-analyse av ekstrapiloten	39
IRT-analyse av ekstrapiloten	40
Endringer etter siste pilotering	40
Forslag til endelig prøve 2023	41
Interessante leseutfordringer i forslaget til endelig prøve	42
IRT-analyse av prøveforslaget	44
Oppgaver med lav diskriminering og DIF	45
Oppgavene fordelt på mestringsnivåer	46
Testinformasjon og måleusikkerhet	47
Prøvens TRF (Test respons function)	47
Gjennomføring	49
Resultater for prøven som helhet	50
Kjønnsforskjeller	52

Klassisk item-analyse av prøven 2023	54
Oppgaver med lav diskriminering	56
Oppgavene fordelt på mestringsnivåer	57
Testinformasjon og måleusikkerhet	58
Prøvens TRF (Test respons function)	59
Referanser	60
Vedlegg 1: Beskrivelse av lesekompetanse på de fem ulike nivåene	61
Vedlegg 2: Oversikt over oppgaver og endringer	62
Vedlegg 3A: Spørreskjema til prøve A	71
Vedlegg 3B: Spørreskjema til prøve B	72
Vedlegg 3C: Spørreskjema til prøve C	73

Innledning

Denne rapporten beskriver og dokumenterer de lesefaglige og psykometriske vurderingene som ligger til grunn for utvikling av den nasjonale prøven i lesing på 8. trinn 2023, 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. Videre dokumenterer vi utviklings- og kvalitets-sikringsprosessen, før vi til slutt beskriver resultatene fra gjennomføringen av prøven i september 2023.

Måling av leseforståelse i de nasjonale prøvene i lesing

Å kunne lese er en grunnleggende ferdighet og en forutsetning for læring i alle fag. Lesing handler om å kunne forstå, bruke, reflektere over, vurdere og engasjere seg i tekster. I dagens informasjonssamfunn kreves det at vi leser for flere formål, at vi forholder oss til mange ulike typer tekster og at vi ofte må lese flere tekster som handler om det samme, for å tilegne oss ny kunnskap. Tekster har ulik utforming i forskjellige medier, og kan ved siden av verbalspråk også bestå av illustrasjoner, grafiske framstillinger, kart og andre visuelle elementer. Kunnskap om hva som kjennetegner ulike typer tekster og sjangre, er en viktig del av det å kunne lese, fordi sjangertrekk aktiviserer leserens forventninger og forkunnskaper om hvordan teksten skal leses og forstås.

Lesing er en sammensatt og komplisert aktivitet, som blant annet påvirkes av avkodingsferdigheter, lesehastighet, flyt, vokabular- og begrepsforståelse, samt kunnskap om tekststruktur og tekstens tema. God leseforståelse krever dessuten at lesere har evne til å identifisere forfatterens hensikt og vinkling, forstå hvilke bevis som brukes for å belegge påstander, og bestemme hvorvidt teksten er troverdig. I tillegg må lesere være bevisste på den påvirkningskraften tekster har i ulike kontekster. Lesere må ha kunnskap om hvilke strategier som er hensiktsmessige for å forstå ulike tekster, og tilpasse lesingen til det de leser og formålet med lesingen. Noen ganger leter lesere etter konkret informasjon i teksten, mens de andre ganger tolker, reflekterer eller foretar vurderinger av teksten (Utdanningsdirektoratet, 2022).

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 (Utdanningsdirektoratet 2020) og spesifiseres nærmere i «Rammeverk for grunnleggende ferdigheter» (Utdanningsdirektoratet 2017). Lesefaglige og psykometriske krav til de nasjonale leseprøvene defineres nærmere av rammeverket for nasjonale prøver (Utdanningsdirektoratet 2022), både generell del (del 1), om prøvenes innhold (del 2.4) og de tekniske kravene (del 3). 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, 2020):

- Å kunne lese i naturfag innebærer «å kunne forstå naturfaglige begreper, symboler, figurer og argumenter gjennom arbeid med naturfaglige tekster.»
- Å kunne lese i matematikk vil si «å sortere informasjon, analysere og vurdere form og innhold og sammenfatte informasjon i sammensatte tekster.»
- Å kunne lese i norsk innebærer «å kunne lese og reflektere over skjønnlitteratur og sakprosa, å beherske lesestrategier tilpasset formålet med lesingen og å kunne vurdere tekster kritisk. Lesing i norsk innebærer også å lese sammensatte tekster som kan inneholde skrift, bilder, tegninger, tall og andre uttrykksformer.»
- Å kunne lese i samfunnsfag innebærer «å utforske, tolke og reflektere kritisk over ulike historiske, geografiske og samfunnskunnskaplege kjelder.»

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.

Tekster og oppgaveformater

Tekstene i nasjonale prøver i lesing skal gjenspeile det tekstmangfoldet som elevene møter i de ulike fagene og i samfunnet ellers. Ingen av tekstene er skrevet spesielt for prøvene, men har stått på trykk i ulike bøker, tidsskrifter, brosjyrer, nettsider etc. Det er ikke plass til alle fag i hver prøve, men over tid er det et mål at tekster fra alle skolefagene er representert, sammen med tekster fra hverdagslivet forøvrig. Tekstene kan ha tematisk tilknytning til faget eller være strukturert på en måte som stiller krav til den fagspesifikke leseforståelsen. Tekster som inneholder grafiske framstillinger, kan for eksempel være kategorisert under matematikk, selv om tekstens tema kan knyttes til andre fag. De aller fleste tekstene er aktuelle i mer enn ett fag og kan brukes tverrfaglig i etterarbeidet med prøveresultatene. Prøvene skal inneholde både skjønnlitterære tekster og sakprosaetekster. De skjønnlitterære tekstene består hovedsakelig av sammenhengende verbaltekst, mens sakprosaetekstene ofte vil være sammensatt av

ulike elementer, som for eksempel illustrasjoner, tabeller, figurer, diagram, kart osv. I tråd med læreplanen kan prøvene inneholde tekster både på bokmål og nynorsk.

Til hver tekst er det mellom tre og åtte oppgaver som tar utgangspunkt i relevante måter å lese teksten på. Et oppgavesett berører leseutfordringer knyttet både til innholdet i teksten og særtrekkene i måten den er skrevet og bygd opp på. Oppgavene i prøvene legger blant annet vekt på meningsinnhold som er implisitt uttrykt, bortgjemt informasjon eller opplysninger som står i motsetning til det forventede. Eksempler på andre leseutfordringer i prøvene er å avgjøre hva som er formålet med teksten og hvem som er målgruppen, hvilke stemmer som kommer til uttrykk, eller hvilke roller ulike aktører har i teksten. Det stilles også krav til at elevene skal sammenligne innholdet i flere tekster som handler om det samme, og identifisere og håndtere motstridende informasjon.

Oppgaver til tekster i leseprøvene skal både fungere som forbilder for hvordan tekster kan leses og forstås, og ikke minst skal de oppleves som relevante målinger av elevenes leseferdigheter. Et oppgavesett vil ha en logisk innholdsmessig oppbygning og en spredning i vanskegrad. Det vil være både lette og vanskelige oppgaver til alle tekster.

I Eksamen- og prøvegjennomføringssystemet (EPS) legges teksten inn som delt innhold (shared stimuli) i hver oppgave. For å unngå at elevenes lesing blir styrt av at første oppgave dukker opp med en gang, og for å gi elevene et overblikk over teksten, blir teksten vist uten en tilknyttet oppgave ved første visning. Elevene kan selv velge om de vil lese teksten her, eller om de vil klikke seg videre og starte med oppgavene med en gang.

I EPS er det flere ulike oppgaveformater, og disse har vi gitt navn som beskriver hva oppgavene går ut på. I leseprøven på 8. trinn er det flest flervalgsoppgaver med et spørsmål og fire svaralternativer der bare ett svar er riktig. I tillegg er det tabell- og rekkefølgeoppgaver, klikk ord- og klikk bilde-oppgaver, og oppgaver der elevene skal fylle inn ord eller fullføre en setning. Tabell-oppgavene består av en tabell med påstander eller opplysninger som elevene skal ta stilling til ved å krysse av for eksempel «riktig» eller «galt» eller «nevnt i tekst 1», «nevnt i tekst 2» eller «nevnt i begge tekstene». Rekkefølge-oppgaver inneholder tre eller fire utvalgte hendelser som elevene blir bedt om å plassere i riktig rekkefølge, fra det som skjedde først til det som skjedde sist. Dette formatet er særlig aktuelt i tekster der kronologien er brutt. I klikk ord-oppgaver får elevene mulighet til å trykke på ett eller flere ord i en tekst, mens de i klikk bilde-oppgavene skal velge mellom markerte områder i et bilde. Fyll inn ord-oppgavene går ut på å fullføre én eller flere setninger ved å velge riktige ord i en nedtrekksmeny.

De nevnte oppgavetyperne blir skåret automatisk og er velegnet for leseprøver, både fordi elevens skriveferdigheter ikke påvirker resultatet, og fordi oppgavene vurderes uavhengig av lærerens skjønn. Disse oppgavene har imidlertid den begrensningen at de ikke kan brukes i tolknings- og refleksjonsoppgaver der flere svar er mulige, og de må av den grunn rettes mot elementer i teksten som kan besvares enten riktig eller feil. Åpne oppgaver kan derimot brukes for å la elevene vise hvordan de resonnerer og reflekterer, særlig i tilfeller der spørsmålet ikke bare har ett riktig svar. Ved å ha en åpen oppgave knyttet til hver tekst inviteres elevene til å uttrykke seg med egne ord, og lærerne får anledning til å bli kjent med tekstenes innhold når de vurderer elevsvarene. Lærernes kjennskap til tekster og oppgaver er

en forutsetning for godt etterarbeid i klasserommet og oppfølging av elevenes resultater. Det er med andre ord flere årsaker til at åpne oppgaver innlemmes i de nasjonale prøvene, selv om de utfordrer prøvenes reliabilitet ved at svarene i større eller mindre grad vurderes ulikt blant lærere. Vi legger stor vekt på å utarbeide en klar og tydelig vurderingsveiledning som gir minst mulig rom for tvil blant lærerne. Oppgavene inviterer dessuten kun til korte svar, gjerne 1–3 setninger, og ved vurdering skal det kun legges vekt på meningsinnholdet i svaret, ikke den språklige utformingen.

Når det gjelder oppgavetyperne som blir skåret automatisk, er det et mål at elevene skal få svaralternativer og påstander presentert i rekkefølger som *enten* gjenspeiler den logiske orden i teksten (tallrekker 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, og for komplekse oppgavetyper gir vi begrunnelser for poengsettingen støttet av analyseresultater.

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 elevenes evne til å finne informasjon, tolke og sammenholde informasjon eller å reflektere over og vurdere teksters form og innhold. Disse tre lesemåtene 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.

(Rammeverk for nasjonale prøver, pkt. 2.4)

Tre eksempeloppgaver fra teksten «Verdens kaldeste sted?» som var en del av leseprøven for 8. trinn i 2022, illustrerer de tre lesemåtene. Det riktige svaret på flervalgsoppgavene er markert med en sirkel.

Denne teksten er et utdrag fra en artikkel skrevet av Per Kristian Aale. Teksten er hentet fra nettavisen aftenposten.no i desember 2014. Bruk teksten når du svarer på oppgavene.

Verdens kaldeste sted?



Slik ser øyevipper ut i 50 minusgrader. Foto: Anastasia Gruzdeva (Instagram)

Dypt inne i Russlands største ødemark ligger Ojmjakon, en landsby i det nordøstlige Sibir. Her må befolkningen leve med vinter i ni måneder og en gjennomsnittlig minimumstemperatur i januar på minus 51 grader.

I sentrum av byen står en spesiell metallkonstruksjon fra sovjettiden, med hammer og sigd på toppen. Den minner om at Ojmjakon er verdens kaldeste bebodde sted. I 1926 registrerte vitenskapsmannen Sergej Obrutjev den ekstreme temperaturen minus 71,2 grader, ifølge Wikipedia.

Kulderekorder

Aldri siden er så lave temperaturer blitt registrert — hvis man ser bort fra ubebodde områder i Antarktis. Det må tilføyes at Obrutjev ikke hadde noe termometer som kunne måle en så lav temperatur og derfor kom til resultatet ved matematiske beregninger.

Den offisielle kulderekorden er fra 6. februar 1933, da det ble målt minus 67,7 grader. Ifølge Store norske leksikon deler Ojmjakon denne verdensrekorden med Verkhjansk, en annen landsby i Sibir. Dette er den

Ojmjakon



By med cirka 500 innbyggere i den østslaviske republikken Sakha. Blir sammen med Verkhjansk regnet som det kaldeste stedet i verden der det bor folk. Gjennomsnittlig minimums-temperatur i januar er – 51 °C, mens maksimums-temperaturen er –43 °C. Somrene er varme; i juli kommer temperaturen ofte over 30 °C. Sakha (tidligere Jakutia) er den største republikken i Russland i utstrekning. Hovedstaden er Jakutsk.

Kilde: Wikipedia

laveste temperaturen som er målt der det bor mennesker, og det er den laveste temperaturen som er målt på den nordlige halvkule. Kun Antarktis har målt enda lavere temperatur: minus 89,2 grader på Vostok-stasjonen 21. juli 1983.

Ojmjakon

Ojmjakon betyr «vann som ikke fryser» på jakutisk, med henvisning til byens varme kilde. Her er det så kaldt at det ikke gror noen ting, og befolkningen lever av reinsdyr- og hestekjøtt. I landsbyen har de skole, et postkontor, en bank og til og med en flystripe, selv om den bare er åpen i sommerhalvåret. Nå i romjulen er det sjeldent mildt, og ifølge yr.no ligger gradestokken og vipper rundt minus 40 grader.

Til tross for kulden har de fleste husene fremdeles utedo, noe Aftenposten fikk erfare da vi besøkte stedet for noen år siden. Selv når det er le for vinden, skal det mye til for å hale ut toalett-besøket i mer enn ett minutt.

Spyttkylser fryser til is

De færreste innbyggerne har termometer. Isteden har de andre måter å finne ut hvor kaldt det er, noe en lokal dame forklarte Aftenposten: Når det er varmere enn minus 50 grader, er det bare vanlig kaldt utendørs og ingen er opptatt av om det er minus 42 eller 48 grader.

Under minus 50 grader skjer det noe med luften. Ofte ligger det en kuldedis eller tåke over hele landsbyen som i de mildere tilfeller (minus 50–55 grader) forsvinner ut på formiddagen. På de kaldeste dagene, med 55–60 minusgrader, blir kuldedis liggende hele dagen. Kommer man under minus 60 grader, skjer det helt spesielle ting, som eksempelvis at ens egen pust blir som krystaller og lyder som knuste glasskår. Eller at plastposer brekker og spyttkylser fryser til is før de treffer bakken.

I kulde på rundt minus 50 grader fungerer ikke mobiltelefonene, og folk her er så redde for at bilene skal stoppe, at de går på tomgang hele dagen. Opprinnelig var dette et stoppested for gjettere som trengte vann til dyrene sine, men er nå et permanent bosted for rundt fem hundre innbyggere. Nærmeste bebodde sted er tre dagers kjøretur unna.



Ojmjakon, kuldens pol, står det på skiltet der den tidligere værstasjonen lå. Her ble kulderekorden på minus 67,7 grader målt i 1933.

Foto: Maxim Shemetov

1 Hva er den offisielle kulderekorden i Ojmjakon?

- A minus 51 grader
- ☒ B minus 67,7 grader
- C minus 71,2 grader
- D minus 89,2 grader

Leseren må finne informasjonen i teksten og skille mellom den offisielle rekorden og andre temperaturer som er oppgitt.

2 Hva er grunnen til at temperaturen som vitenskapsmannen Sergej Obrutjev registrerte i 1926, ikke er den offisielle kulderekorden i Ojmjakon, ifølge teksten?

- A Ojmjakon måtte dele temperaturrekorden fra 1926 med Verkhjansk.
- B Det var ingen vitner som kunne bekrefte Obrutjevs observasjon.
- C Det ble målt enda lavere temperatur i Ojmjakon noen år etterpå.
- ☒ D Obrutjev hadde ikke noe instrument å måle så lave temperaturer med.

Leseren må tolke og forstå at setningen «Det må tilføyes at Obrutjev ikke hadde noe termometer og derfor kom til resultatet ved matematiske beregninger», er en innvending mot det som står om den offisielle kulderekorden i avsnittet etter.

3 Hvorfor er kulderekordene i Ojmjakon og Verkhojansk interessante når det er målt lavere temperaturer andre steder i verden?

Leseren må reflektere over innholdet i teksten, der det flere steder står at Ojmjakon er verdens kaldeste bebodde sted, eller sted der det bor folk.

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 prøveforslaget for 8. trinn 2023 er i tråd med beskrivelser av lesing som grunnleggende ferdighet og kompetansemål i fagene etter 7. trinn, noe som blir bekreftet gjennom møter med lærere og andre eksterne kvalitetssikrere. Prøven inneholder både lette og vanskelige oppgaver, og måler dermed både elever med svake og sterke leseferdigheter. Prøvens innhold ser ikke ut til å favorisere verken gutter eller jenter i stor 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 hvorvidt vi måler det ønskede konstruktet på en konsistent måte. Det endelige prøveforslagets reliabilitetskoeffisient er 0,906 (Cronbachs alpha), som regnes for å være en høy og tilfredsstillende verdi (Nunnally & Bernstein, 1994).

Åpne oppgaver utfordrer som nevnt prøvenes reliabilitet ved at svarene kan vurderes ulikt blant lærere. I prøveutviklingsprosessen stilles det høye krav ($> 0,85$) til samsvaret mellom vurderere, og de erfaringene vi gjør underveis, blir brukt til å skrive en, så langt det er mulig, utvetydig vurderingsveiledning for lærere.

Prøvens validitet handler om i hvor stor grad vi klarer å måle det vi ønsker å måle, nemlig elevens leseferdighet. Den digitale prøveplattformen som er valgt, samt kravene til universell utforming, begrenser til en viss grad både hvilke tekster det er mulig å bruke, og hvordan tekstene kan utformes. Prøven gir likevel et godt mål på elevens leseferdighet, men det er viktig å både utforske og utfordre den digitale prøveplattformen slik at prøveopplevelsen for elevene blir best mulig og at leseferdigheten blir målt i tråd med beskrivelsene av lesing som grunnleggende ferdighet.

Analyser i rapporten

I tråd med Rammeverk for nasjonale prøver (Utdanningsdirektoratet, 2022) skal prøvene utformes med beste mulige standarder for prøvekonstruksjon, ved hjelp av egnede psykometriske metoder. Oppgaveutviklingen skal skje med klassiske mål for reliabilitet og oppgaveegenskaper. I tillegg skal enkeltoppgaver og hele prøven analyseres ved hjelp av item response theory (IRT). For 2023-prøven er klassisk analyse utført med programmet SPSS,

mens IRT-analysene er utført med henholdsvis XCalibre 4.2 etter første pilotering og R 4.2.2 og R-pakken «mirt» fra og med andre pilotering.

Ved hjelp av IRT kan man estimere oppgavens vanskegrad og beregne elevenes ferdighet uavhengig av hvilke oppgaver de besvarte, forutsatt at elevene i utvalget representerer hele ferdigheten. IRT-analysen leverer resultatet i form av theta-verdier (θ), som går fra omtrent -3 til +3 med et gjennomsnitt på 0. Oppgavens vanskegrad blir også uttrykt på samme skala. Etter en enkel omregningsnøkkel ($\theta * 10 + 50$) regnes theta-verdiene om til positive verdier, såkalte skalapoeng.

IRT et godt verktøy i prøveutviklingen fordi analysene av pilotresultatene estimerer hvordan oppgavene vil fungere på et større antall elever. Når det kommer til gjennomføringen av prøven, handler det ikke om å estimere resultater, men å analysere data for hele populasjonen, og da er klassisk analyse mer relevant enn IRT.

IRT-analysen leverer en sum av informasjonsverdier og målefeil for hele ferdighetsfordelingen i en prøve (TIF) og en Test Respons Funksjon (TRF). Sistnevnte er en beskrivelse av sammenhengen mellom elevenes estimerte ferdighet (theta) og andel riktige besvarte oppgaver i prøven. Vi bruker TRF til å sammenligne årets prøve med fjorårets.

I IRT-analysene av oppgavene i prøven brukes en to-parameter-modell (2PL). Modellen viser oppgavens vanskegrad (b) og oppgavens diskriminering (a og point biserial correlation). En oppgaves diskrimineringsgrad sier noe om i hvilken grad oppgaven skiller mellom elever med svake og sterke leseferdigheter. Diskrimineringsmålene viser korrelasjon mellom skår på hver enkelt oppgave og total skår på prøven. Oppgaver med høy diskriminering gir høy reliabilitet for prøven. En oppgave med lav diskriminering ($< 0,250$ point biserial eller $< 0,8$ a-parameter) kan være «bortkastet» ved ikke å skille godt mellom lesere med ulik ferdighet, men oppgaven kan likevel forsvares pedagogisk eller lesefaglig. For eksempel kan en svært lett oppgave som nesten alle lesere mestrer, gi en god og oppklarende introduksjon til teksten og oppgavesettet. Polytome oppgaver, som skåres med 0, 1, 2 eller flere poeng, gir som regel høyere diskriminering, fordi de gir mer informasjon. Dersom en prøve inneholder mange oppgaver med for høy diskriminering, kan det være en indikasjon på lav validitet ved at oppgavene er for like og dermed dekker konstruktet for dårlig. Tilsvarende kan en prøve med mange lavt diskriminerende oppgaver også være en indikasjon på lav validitet ved at oppgavene måler noe annet enn leseforståelse.

IRT-analysen gir også svar på om oppgavene, og dermed prøven som helhet, favoriserer én bestemt gruppe framfor en annen gjennom en analyse av Differential Item Functioning (DIF). For eksempel om en oppgave favoriserer jenter framfor gutter, gitt at representanter for begge grupper har samme underliggende ferdighet (Utdanningsdirektoratet 2022).

Resultatene fra klassisk analyse blir rapportert i form av prosentvis fordeling av elevsvar (p-verdier), gjennomsnittlig dyktighet til gruppene av elever som valgte de ulike svaralternativene, mål på diskriminering (point biserial correlation), og mål for reliabilitet/indre konsistens (Cronbachs alpha).

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).

Chronbach's alpha er et reliabilitetsmål som gir 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 0,7 og 0,9.

Nasjonal prøve i lesing for 8. trinn 2023

Prøveutvikling i perioden 2021–2023

For å få et så representativt utvalg som mulig, inviterer vi pilotskoler med ulikt antall elever fra hele landet. Skolene er tilfeldig utvalgt fra en liste gitt av Utdanningsdirektoratet, og hele klasser rekrutteres. Prøven for 2023 er pilotert i to omganger: første gang i februar 2022 og andre gang i august 2022. I tillegg ble én av tekstene pilotert for tredje gang i februar 2023. Det var elever på 8. trinn som deltok i piloteringene, og de var henholdsvis et halvt år eldre og like gamle 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 piloteringene.

Prøven er kvalitetssikret av eksterne lesefaglige eksperter ved Uppsala universitet, av psykometrisk ekspert ved NTNU og av en referansegruppe bestående av lærere på ungdomstrinnet.

Innholdet i prøven

Forslaget til leseprøve for 2023 består av 7 tekster med til sammen 42 oppgaver som måler elevenes leseforståelse, i tråd med læreplanen og kompetansemålene etter 5.–7. trinn. Tekstene er presentert i tabell 1 nedenfor.

Tabell 1: Oversikt over tekstene i forslaget til nasjonal prøve i lesing på 8. trinn 2023.

Tekst	Fagtilhørighet	Teksttype	Antall ord: 5686	Antall oppgaver: 42		
				Flervalg: 32	Åpne: 7	Andre: 3
<i>Tasmansk pungulv</i>	NAT, SAF	Multippel: Utdrag fra fagbok og nettartikkel	1040	5	1	1
<i>Stargate</i>	NOR	Skjønnlitterært utdrag	745	4	1	1
<i>Lysforurensing</i>	NAT, KHV, SAF	Argumenterende blogginnlegg	1050	6		
<i>Skulearbeid (sidemål)</i>	SAF, MAT	Grafikk	500	5	1	
<i>Snikfotografen</i>	KHV, SAF	Nettartikkel	983	4	1	1
<i>Den korte historien om egget</i>	MHE, NAT	Utdrag fra fagbok	867	4	1	
<i>Et skår i gleden?</i>	NOR, SAF	Reklame	500	4	2	

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, kunst og håndverk, matematikk og mat og helse. Tekstene er tildelt ulike plasser i prøven, i tråd med plan for teksttyper i ankeret og med tanke på at den første teksten skal være engasjerende og

gi flest mulig elever en god start på prøven. På samme måte er den siste teksten kort – noe som kan motivere elevene til fortsatt innsats, slik at vi får en god måling av leseferdigheten også mot slutten av prøven. Det er som vanlig én skjønnlitterær tekst og seks sakprosattekster med i utvalget. Flertallet av sakprosatekstene 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 12 finne-, 17 tolke- og 13 refleksjonsoppgaver.

Ved siden av at alle oppgaver skal tilfredsstillende lesefaglige og psykometriske krav, legges det stor vekt på at prøven som helhet representerer en bredde i fagtilhørighet, sjangre, tema og leseutfordringer. Den følgende beskrivelsen av leseutfordringer i prøven viser at årets prøve ivaretar alle disse målsettingene, og at prøven måler lesing som grunnleggende ferdighet i flere fag:

«**Tasmansk pungulv**» er en multippel tekst bestående av et utdrag fra fagboka *Dyrene som forsvant* (2017) av Line Renslebråten og en nyhetsartikkel fra *nettavisen.no*. Begge tekstene handler om den utdødde arten tasmansk pungulv, men innholdet og perspektivet er forskjellig. Lene Renslebråten's bok er skrevet for barn i alderen 9 til 12 år, og hun forteller historien om den siste kjente pungulven, Benjamin, som døde i en dyrehage i 1936. Hun beskriver også arten og forteller om bakgrunnen for at pungulven ble utryddet. Nyhetsartikkelen fra Nettavisen tar utgangspunkt i en rapport fra australske myndigheter, som spekulerer på om pungulven kanskje ikke er utryddet likevel. Oppgavene krever at elevene forstår årsaken til at pungulven ble utryddet, og at de kan sortere hendelsene i kronologisk rekkefølge. De må også se de to tekstene i sammenheng og reflektere over nettavisens kildebruk og hvilke svakheter som hefter ved de påståtte observasjonene og videobevisene av pungulv.

«**Stargate**» er første kapittel i den skjønnlitterære boka *Stargate – en julefortelling* (2021) av Ingvild H. Rishøi. Boka er ikke skrevet spesielt for barn og unge, men miljøet og de to karakterene i utdraget vil likevel være gjenkjennelig for elevene. Jeg-personen er ei jente som går på Tøyen skole, og i tekstutdraget innleder vaktmesteren på skolen en samtale med henne. Mellom linjene går det fram at vaktmesteren kjenner jenta, og han vet at faren hennes ikke har jobb. Vaktmesteren gir jenta mat og viser henne en plakate der det står at juletresele søkes. Han river ned plakaten og ber jenta gi den til faren sin slik at han kan søke på jobben. Oppgavene legger vekt på å tolke det som står mellom linjene og forstå at vaktmesteren har gode intensjoner og omtanke for jenta og familien.

«**Lysforurensing**» er en argumenterende tekst hentet fra bloggen til professor i bevaringsøkologi Anne Sverdrup-Thygeson ved Universitetet for miljø og biovitenskap. Forfatteren forklarer hvorfor feil lys til feil tid er et problem for dyre- og plantelivet, og hun viser til hva vi mennesker kan gjøre med det. Teksten er ikke skrevet for målgruppa, men temaet lys og utendørs lyssetting passer inn både i naturfag, samfunnsfag og kunst og håndverk. Lysforurensing er dessuten et problem som ikke har fått like mye oppmerksomhet som annen forurensing av naturen, og teksten er derfor egnet til å bevisstgjøre elevene. Oppgavene legger vekt på hva lysforurensing er, og hvordan problemet kan løses.

«**Skularbeid**» er et utdrag fra rapporten *Ungdata 2021*. Teksten er på sidemål, og temaet er skolearbeid og stress – noe som angår elevenes egen hverdag. Utdraget er satt sammen av en innledende verbaltekst om Ungdata-undersøkelsen og ulike grafiske framstillinger og tekstbokser. De grafiske elementene består av stolpe- og linjediagrammer og viser hva elevene i undersøkelsen har svart på spørsmål om hvor mye tid de bruker på lekser, og hvor ofte de blir stresset av dette. Resultatene vises for 2021 sammenlignet med 2018–20, samt for gutter og jenter på ulike klassetrinn og for ungdomsskoler og videregående skoler over flere år. Den største utfordringen er å finne fram til riktig diagram og lese av de resultatene som oppgavene ber om. Elevene må også forstå hensikten med teksten og med selve Ungdata-undersøkelsen, og vise at de kan reflektere over resultatene fra undersøkelsen.

«**Snikfotografen**» er en artikkel hentet fra *nrk.no*. I denne artikkelen fortelles historien om studenten Carl Størmer som på 1890-tallet fotograferte Henrik Ibsen og andre tilfeldig forbipasserende med et lite kamera som han hadde på innsiden av skjorta. Resultatet ble en rekke bilder av Ibsen og andre personer i mer hverdagslige situasjoner enn det som var vanlig å fange med kamera på den tiden. En av oppgavene belyser denne forskjellen, mens en annen oppgave utfordrer elevenes evne til å sortere hendelser i kronologisk rekkefølge. For å svare på oppgavene må elevene også forstå hvem Carl Størmer var og reflektere over hvem som uttaler seg i artikkelen. Temaet «snikfotografering» på 1890-tallet er morsomt og spennende i seg selv, men også interessant sett i lys av at det er et aktuelt og gjenkjennelig fenomen for elevene.

«**Den korte historien om egget**» er et kapittel hentet fra boka *Barnas aller beste kokebok* (2021) av Erik Røed ved Geitmyra matkultursenter. Som tittelen viser, er boka skrevet for en ung målgruppe. Det utvalgte kapitlet har overskriften «Den korte historien om egget» og innledes med spørsmålet «Har du tenkt over hva et hønseegg egentlig er?» Det er fem oppgaver til teksten, og i den første oppgaven må elevene reflektere over hva forfatteren mest sannsynlig ønsker å oppnå med å stille et innledende spørsmål. To av de andre oppgavene utfordrer også elevenes evne til å reflektere over forfatterens valg og tekstens utforming, mens de to siste oppgaven tester elevenes evne til å finne og tolke og forstå det faglige innholdet i teksten.

«**Et skår i gleden?**» er en helsides annonse for butikkjeden Jernia. Overskriften «Et skår i gleden?» står med store bokstaver øverst til venstre, og midt på siden er det et stort bilde av en knust tallerken som fanger leserens oppmerksomhet. Nederst er det en kort verbaltekst, etterfulgt av slagordet «Sammen kan vi gjøre en forskjell!» og logoen til Jernia. Verbalteksten har en dobbel agenda: på den ene siden inneholder den informasjon om hvorfor ødelagte drikkeglass, porselen og keramikk ikke skal gjenvinnes sammen med glass, og på den andre siden er det en oppfordring om å levere ødelagt kjøkkenutstyr til Jernia. Som et lite apropos til slutt står det at Jernia gir 20 prosent innbytterabatt på en ny vare. Oppgavene legger vekt på at elevene skal forstå at dette er en reklame, der Jernia framstiller seg selv som en miljøbevisst aktør, samtidig som de har til hensikt å øke salget.

Første pilotering

I den første piloteringen av prøven 2023 ble det prøvd ut 18 tekster fordelt på prøve A, B og C. To av de 18 tekstene tilhører 2022-prøven, og ble pilotert for tredje gang, for en ekstra kvalitetskontroll. De øvrige 16 tekstene og 125 oppgavene danner grunnlag for utvelgelse av tekster til prøven 2023. Pilotprøven ble gjennomført i den nye, digitale plattformen EPS, og prøvetiden var 90 minutter, som ved gjennomføring av prøven. 1003 elever fra 13 skoler i 6 fylker deltok. Skolene ble trukket ved tilfeldig og ikke-representativ utvalgsteknikk, og erfaringsmessig gir dette dekkende utvalg så lenge et høyt antall elever deltar fra både små og store skoler med god geografisk spredning. Prøven ble pilotert på 8. trinn i februar 2022, og det betyr at elevene var ca. et halvt år eldre enn de elevene som gjennomfører prøven like etter skolestart høsten 2023.

Tabell 2: Oversikt over antall elever i første pilotering

	Elever
Antall inviterte elever	4071 (fra 35 skoler)
Antall oppmeldte elever	1925 (fra 16 skoler)
Antall elever som gjennomførte (N)	Totalt: 1003 Prøve A: 342 (185 gutter, 157 jenter) Prøve B: 331 (173 gutter, 158 jenter) Prøve C: 341 (169 gutter, 172 jenter)
Antall skoler	13

Innholdet i prøve A, B og C

Tekstene fordeler seg på fag og teksttyper som vist i tabell 3, 4 og 5. For hver tekst er det her oppgitt en gjennomsnittlig p-verdi og kjønnsforskjell, basert på p-verdien til hver enkelt oppgave som er knyttet til den enkelte tekst. 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.

Antall ord som står oppgitt i tabellen, viser hvor mange ord hver enkelt tekst består av, inkludert oppgavene. Dette sier ikke noe om vanskegraden på tekstene, men gir en indikasjon på lengden på prøven og hvor mye elevene må lese. I alle de tre prøvene ligger antall ord nokså likt gjennomsnittet for nasjonal leseprøve på 8. trinn de siste 10 år, som er 5753 ord.

Tabell 3: Oversikt over tekstene i prøve A.

Tekst	Fagtilhørighet	Teksttype	Antall ord: 5532	Antall oppg.: 45			Gj. snittlig p-verdi	Kjønns- forskjell (j-g)*
				Flervalg: 36	Åpne: 8	Andre: 1		
<i>Et skår i gleden?</i>	Norsk, samfunnsfag	Reklame	327	4	1		0,48	0,05
<i>Stargate</i>	Norsk	Skjønnlitterært utdrag	848	7	1		0,70	0,09
<i>Lysforurensing</i>	Naturfag, kunst og håndverk, samfunnsfag	Argumenterende blogginnlegg	1029	5	1		0,55	-0,01
<i>Tasmansk pungulv</i>	Naturfag, samfunnsfag	Multippel: utdrag fra fagbok og nettartikkel	1018	5	1	1	0,51	-0,02
<i>Den korte historien om egget</i>	Mat og helse, naturfag	Utdrag fra fagbok	863	4	2		0,54	0,01
<i>Snikfoto grafen</i>	Kunst og håndverk, samfunnsfag	Nettartikkel	947	5	1		0,53	0,04
<i>Skolearbeid</i>	Samfunnsfag	Grafisk	500	6	1		0,50	0

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

Tabell 4: Oversikt over tekstene i prøve B.

Tekst	Fagtilhørighet	Teksttype	Antall ord: 5742	Antall oppg.: 47			Gj. snittlig p-verdi	Kjønns- forskjell (j-g)*
				Flervalg: 39	Åpne: 7	Andre: 1		
<i>Når gammel kjærlighet ruster</i>	Norsk, samfunnsfag	Reklame	347	5	1		0,58	0,03
<i>Stargate</i>	Norsk	Skjønnlitterært utdrag	848	7	1		0,70	0,08
<i>Dyreparker</i>	Naturfag, samfunnsfag	Nettartikkel, mening	1216	6	1		0,64	0,08
<i>Stripemysteriet</i>	Naturfag	Multippel	787	6	1		0,64	0,07
<i>Noe du trenger, mor?</i>	Norsk, samfunnsfag	Utdrag fra bok	881	5	1		0,59	0,15
<i>Salum Kashafali</i>	Kroppsoving, samfunnsfag	Nettartikkel	1163	5	1	1	0,47	0,05
<i>Klimavennlig mat</i>	Mat og helse, naturfag, samfunnsfag	Grafisk	500	5	1		0,49	0,07

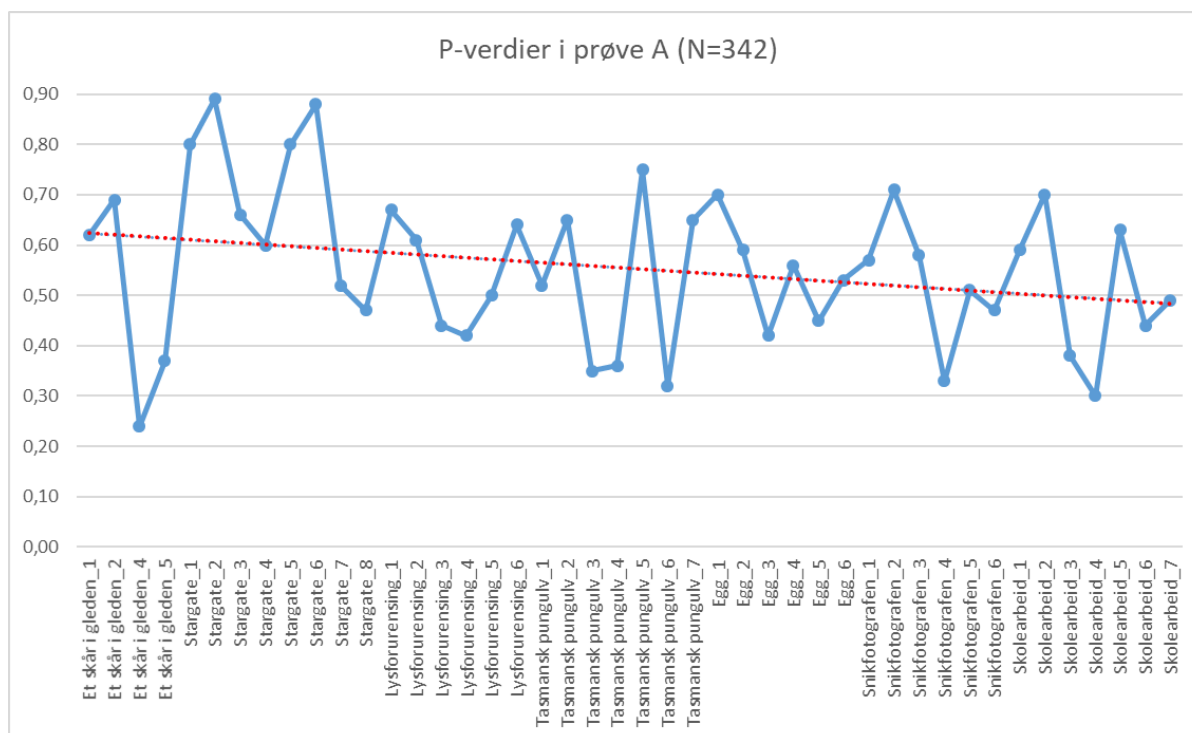
* Positive verdier er i jentenes favør, negative verdier i guttenes favør.

Tabell 5: Oversikt over tekstene i prøve C.

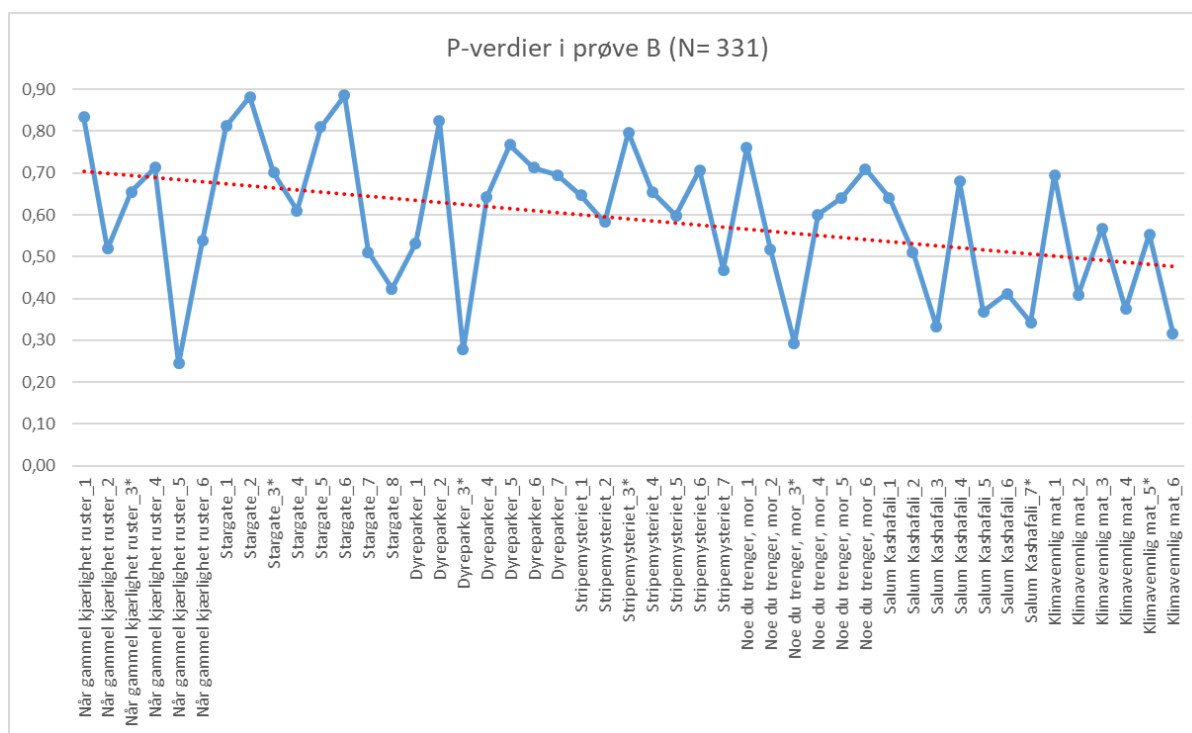
Tekst	Fagtilhørighet	Teksttype	Antall ord: 5020	Antall oppg.: 46			Gj. snittlig p-verdi	Kjønns- forskjell (j-g)*
				Flervalg: 36	Åpne: 8	Andre: 2		
<i>Et skår i gleden?</i>	Norsk, samfunnsfag	Reklame	327	4	1		0,47	0,08
<i>Stargate</i>	Norsk	Skjønnlitterært utdrag	848	7	1		0,71	0,13
<i>iHuman</i>	Kunst og håndverk, samfunnsfag	Multippel, mening	700	4	1	1	0,45	0,06
<i>Kraften</i>	Samfunnsfag, norsk, musikk	Utdrag fra bok	1230	5	2		0,47	0,07
<i>Sjøku</i>	Naturfag, samfunnsfag	Multippel	785	5	1	1	0,65	0,09
<i>Knøttfe og himmelreker</i>	Mat og helse, naturfag	Utdrag fra bok	786	6	1		0,50	0,06
<i>Vandring gjennom tiden</i>	Samfunnsfag	Grafisk	344	5	1		0,56	0,06

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

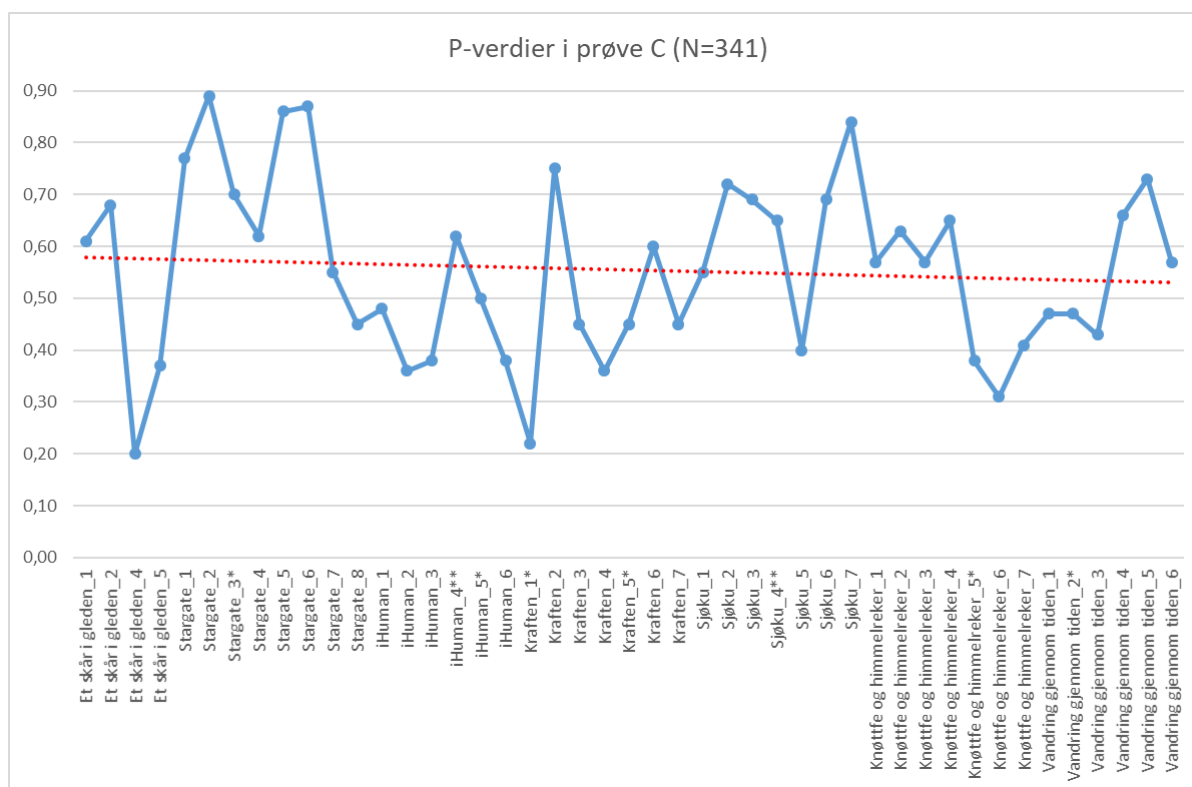
Figur 1, 2 og 3 viser p-verdien for hver av oppgavene i henholdsvis prøve A, B og C. Oppgavene er presentert i samme rekkefølge som i prøvene.



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.

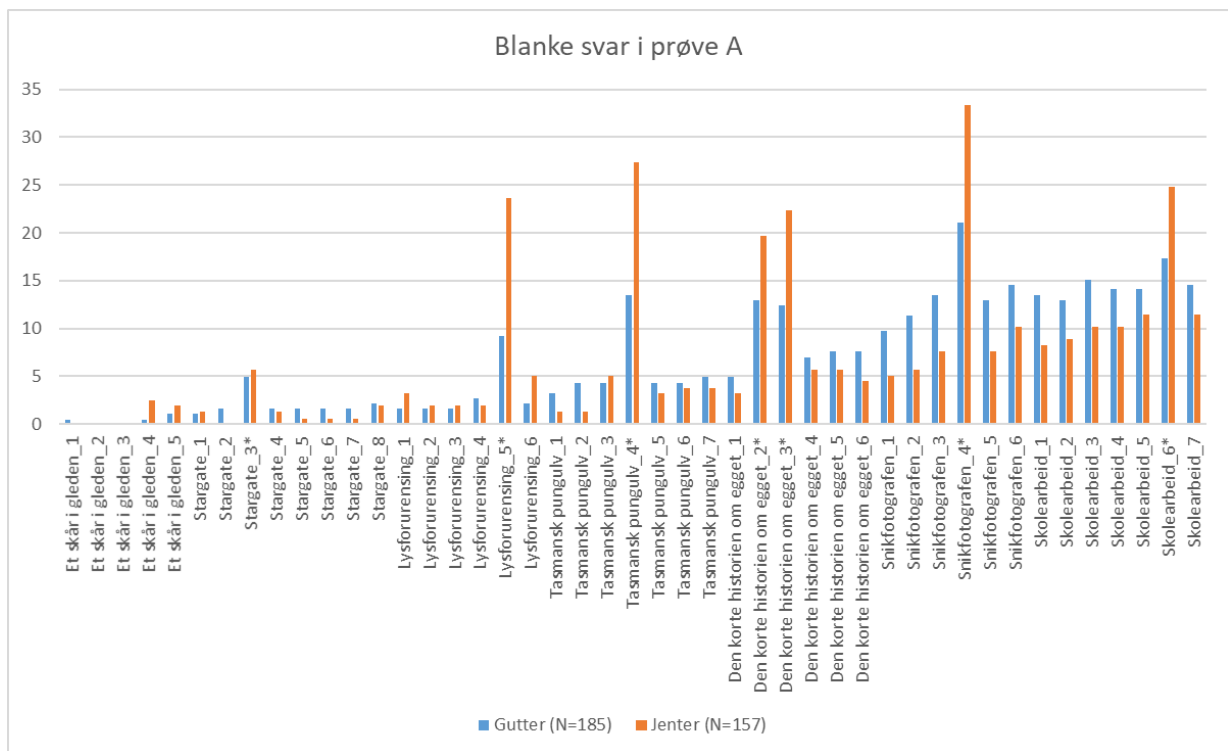


Figur 3: Andelen elever med riktig svar (p-verdi) på oppgavene i prøve C, organisert etter rekkefølgen i prøven. Oppgaven Et skår i gleden_3 er tatt ut av analysegrunnet.

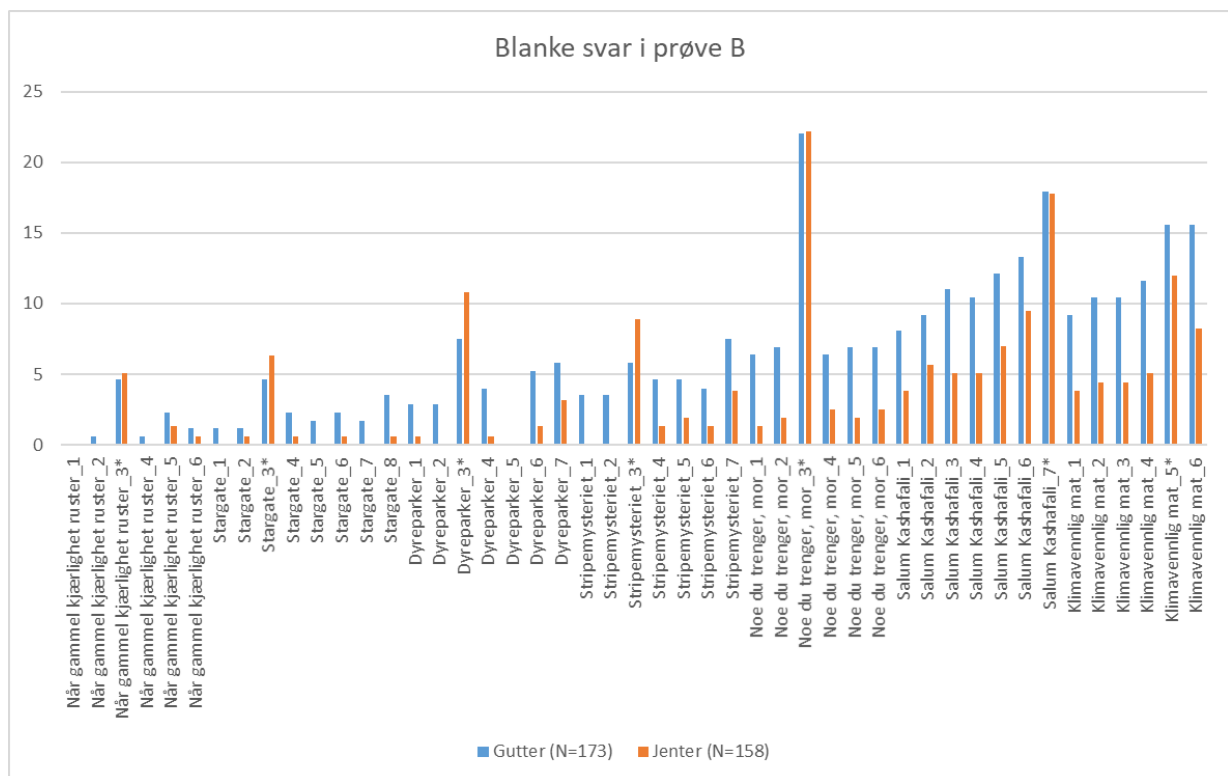
Figur 1, 2 og 3 viser at det er både lette og vanskelige oppgaver i alle de tre prøvene, og den røde trendlinjen viser at p-verdien blir lavere mot slutten, altså at det er færre elever som får til de siste oppgavene. Dette er forventet fordi lesere med svært svake leseferdigheter ikke har tid eller utholdenhet til å fullføre hele prøven.

Blanke svar i prøve A, B og C

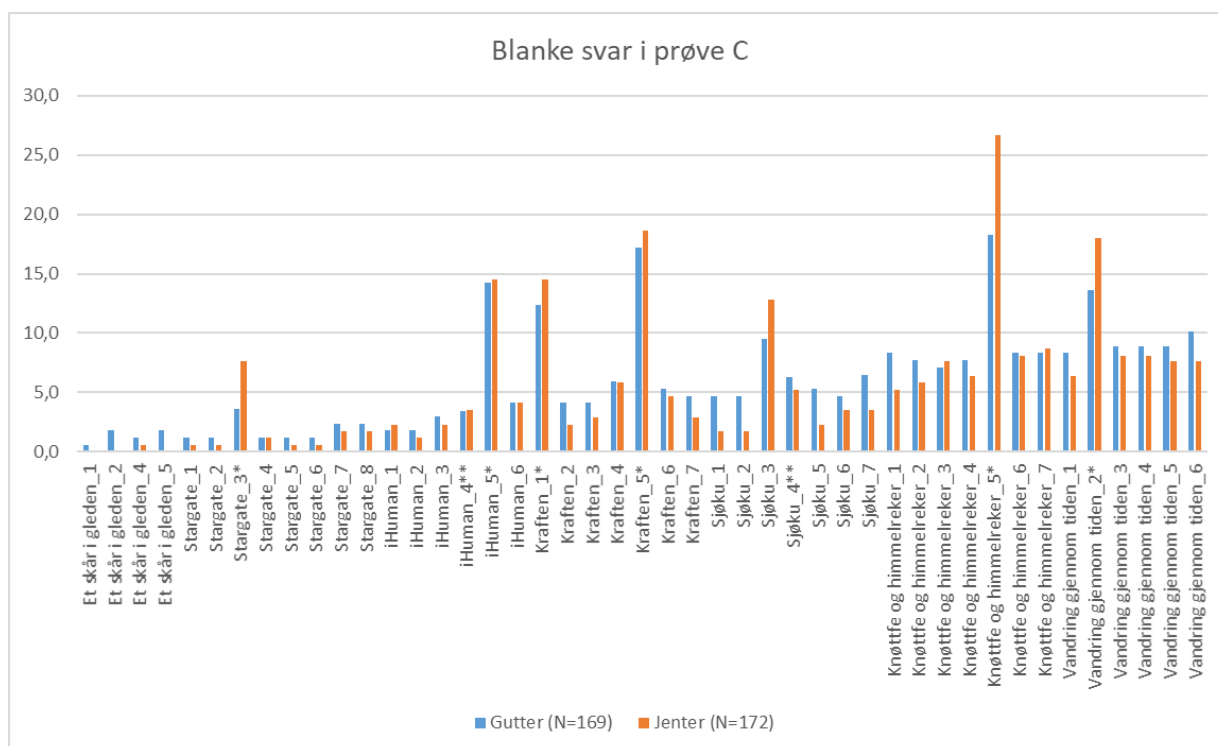
Figur 4, 5 og 6 nedenfor viser prosentandelen gutter og jenter som ikke har svart på hver oppgave i prøve A, B og C. Åpne oppgaver er markert med stjerne (*). For prøve A er den gjennomsnittlige andelen blanke svar 7,3 prosent, mens den for prøve B og C er henholdsvis 5,6 og 5,8 prosent. I figurene ser vi at de aller fleste søylene som stikker seg ut med en høyere prosentandel enn gjennomsnittet, representerer åpne oppgaver. Vi ser også at andelen blanke svar øker noe mot slutten i alle prøvene og at det er en større andel gutter enn jenter som har latt være å svare på mange av oppgavene i alle prøvene.



Figur 4: Prosentandel gutter og jenter som ikke svarte på hver oppgave i prøve A. Åpne oppgaver er merket (*).



Figur 5: Prosentandel gutter og jenter som ikke svarte på hver oppgave i prøve B. Åpne oppgaver er merket (*).



Figur 6: Prosentandel gutter og jenter som ikke svarte på hver oppgave i prøve C. Åpne oppgaver er merket (*).

Klassisk item-analyse av prøve A, B og C

Data fra både klassisk analyse og IRT-analyse vil være grunnlag for utvalg av tekster og oppgaver til andre pilot. Tabell 6, 7 og 8 viser resultater fra klassisk item-analyse av henholdsvis prøve A, B og C. Oppgavene er her gitt et løpenummer, og hver oppgave er tildelt en item-ID og et navn. Navnet viser 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 omplasseringer som eventuelt blir gjort mellom de ulike piloteringene.

Resultatene i tabellene viser prosentvis svarfordeling for hver oppgave og gjennomsnittlig antall poeng på prøven for elevene som har valgt de ulike svaralternativene. Kolonnen «diskriminering» viser en biserial korrelasjon, som ifølge fastsatte mål bør være over 0,250 for at oppgaven skal differensiere godt mellom lesere med svake og sterke leseferdigheter. De siste kolonnene viser de ulike oppgavens p-verdi og kjønnsforskjell, der positive verdier er i jentenes favør og negative verdier i guttenes favør.

I rekkefølgeoppgaven i prøve A måtte elevene plassere alle hendelser i riktig rekkefølge for å få poeng. Tilsvarende måtte alle kryss settes på riktig plass i tabell-oppgaven Sjøku_4 i prøve C. I iHuman_4, som også er en tabell, var det mulig å få to poeng, men denne oppgaven tilhører 2022-prøven.

Prøve A

Tabell 6: Resultater fra klassisk item-analyse, prøve A.

N = 342 (185 gutter, 157 jenter). Reliabilitet (Cronbachs alpha): 0,909.

Gjennomsnittlig p-verdi: 0,55 (gutter: 0,54/jenter: 0,57). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Nr.	Item-ID	Prøve A	Svarfordeling i % (frekvens):					Gj.snittlig poengsum på prøven per svaralternativ				Diskrim. (biseriell)	P-verdi	Kjønnsforskjell (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D				
			0 p.	1 p.	2 p.			0 p.	1 p.	2 p.					
1	L141S01	Et skår i gleden_1	5	62	16	17		16	27	21	21	0,308	0,62	0,06	
2	L141S02	Et skår i gleden_2	7	19	5	69		17	19	17	27	0,422	0,69	0,09	
3	L141S03	Et skår i gleden_3													Tatt ut
4	L141S04	Et skår i gleden_4	24	30	10	21	2	29	24	17	24	0,205	0,24	0,02	*
5	L141S05	Et skår i gleden_5	26	32	37	4	2	23	24	27	17	0,167	0,37	0,02	*
6	L142S01	Stargate_1	3	9	80	7	1	11	17	26	17	0,428	0,80	0,01	
7	L142S02	Stargate_2	89	1	4	5	1	26	15	13	17	0,353	0,89	0,06	
8	L142S03	Stargate_3	29	66			5	18	28			0,525	0,66	0,18	Åpen
9	L142S04	Stargate_4	32	3	4	60	2	22	11	17	27	0,311	0,60	0,10	
10	L142S05	Stargate_5	10	80	3	6	1	17	27	13	15	0,466	0,80	0,09	
11	L142S06	Stargate_6	4	2	5	88	1	13	10	19	26	0,385	0,88	0,07	
12	L142S07	Stargate_7	8	52	14	25	1	18	28	19	24	0,303	0,52	0,06	
13	L142S08	Stargate_8	10	11	47	31	2	19	16	27	25	0,267	0,47	0,15	
14	L143S01	Lysforurensing_1	18	9	4	67	2	21	13	15	28	0,490	0,67	0,05	
15	L143S02	Lysforurensing_2	19	6	61	12	2	21	17	27	23	0,292	0,61	-0,01	
16	L143S03	Lysforurensing_3	6	31	44	17	2	17	21	29	22	0,419	0,44	-0,08	
17	L143S04	Lysforurensing_4	30	42	7	19	2	24	27	16	24	0,231	0,42	-0,10	*
18	L143S05	Lysforurensing_5	34	50			16	20	29			0,515	0,50	0,05	Åpen
19	L143S06	Lysforurensing_6	64	20	7	5	4	28	20	20	15	0,413	0,64	0,03	
20	L144S01	Tasmansk pungulv_1	52	39	4	4	2	26	25	21	19	0,095	0,52	-0,01	*
21	L144S02	Tasmansk pungulv_2	10	16	6	65	3	19	20	14	28	0,502	0,65	-0,04	
22	L144S03	Tasmansk pungulv_3	61	35			5	22	31			0,484	0,35	-0,07	Rekkefølge
23	L144S04	Tasmansk pungulv_4	44	36				22	31			0,493	0,36	0,00	Åpen
24	L144S05	Tasmansk pungulv_5	75	12	3	7	4	28	16	17	17	0,550	0,75	-0,02	
25	L144S06	Tasmansk pungulv_6	35	32	7	22	4	25	27	14	26	0,144	0,32	-0,04	*
26	L144S07	Tasmansk pungulv_7	18	66	6	6	4	19	28	16	19	0,486	0,65	0,01	
27	L145S01	Den korte historien om egget_1	5	6	70	15	4	12	23	27	20	0,434	0,70	0,01	
28	L145S02	Den korte historien om egget_1	25	59			16	19	29			0,572	0,59	0,08	Åpen
29	L145S03	Den korte historien om egget_1	41	42			17	22	31			0,563	0,42	0,05	Åpen
30	L145S04	Den korte historien om egget_1	56	7	9	22	6	27	17	24	24	0,303	0,56	-0,10	
31	L145S05	Den korte historien om egget_1	9	45	34	5	7	18	29	23	19	0,413	0,45	0,00	
32	L145S06	Den korte historien om egget_1	9	53	14	18	6	19	27	23	24	0,308	0,53	0,02	
33	L146S01	Snikfotografen_1	57	16	12	7	8	29	20	21	18	0,526	0,57	0,07	
34	L146S02	Snikfotografen_2	7	71	9	4	9	17	28	18	16	0,586	0,71	0,04	
35	L146S03	Snikfotografen_3	58	10	18	3	11	29	17	21	22	0,493	0,58	0,04	
36	L146S04	Snikfotografen_4	40	33			27	23	32			0,567	0,33	-0,01	Åpen
37	L146S05	Snikfotografen_5	24	8	7	51	11	21	16	20	30	0,598	0,51	0,08	
38	L146S06	Snikfotografen_6	14	9	47	18	13	20	19	30	23	0,510	0,47	0,00	
39	L147S01	Skolearbeid_1	59	15	2	14	11	28	20	18	23	0,447	0,59	0,00	
40	L147S02	Skolearbeid_2	4	70	11	4	11	18	28	20	12	0,511	0,70	0,04	
41	L147S03	Skolearbeid_3	8	30	38	11	13	20	24	28	25	0,281	0,38	0,02	
42	L147S04	Skolearbeid_4	21	12	25	30	12	22	19	25	31	0,405	0,30	-0,03	
43	L147S05	Skolearbeid_5	8	11	63	6	13	19	21	28	20	0,430	0,63	-0,01	
44	L147S06	Skolearbeid_6	35	44			21	22	30			0,499	0,44	0,03	Åpen
45	L147S07	Skolearbeid_7	49	18	13	8	13	29	18	23	24	0,454	0,49	-0,03	

*Lav diskriminering (under 0,250).

Prøve B

Tabell 7: Resultater fra klassisk item-analyse, prøve B.

N = 331 (173 gutter, 158 jenter). Reliabilitet (Cronbachs alpha): 0,917.

Gjennomsnittlig p-verdi: 0,59 (gutter: 0,55/jenter: 0,63). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Nr.	Item –ID	Prøve B	Svarfordeling i % (frekvens):					Gj.snittlig poengsum på prøven per svaralternativ				Diskrim. (biseriell)	P-verdi	Kjønns-forskjell (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D				
			0 p.	1 p.	2 p.			0 p.	1 p.	2 p.					
1	L151S01	Når gammel kjærlighet rustet_1	1	2	14	83		11	11	20	30	0,388	0,83	0,04	
2	L151S02	Når gammel kjærlighet rustet_2	8	52	12	28		21	31	21	28	0,245	0,52	-0,03	*
3	L151S03	Når gammel kjærlighet rustet_3	30	66			5	22	31			0,402	0,66	0,04	Åpen
4	L151S04	Når gammel kjærlighet rustet_4	26	1	71	2		24	11	30	19	0,240	0,71	0,09	*
5	L151S05	Når gammel kjærlighet rustet_5	25	37	11	26	2	33	27	25	25	0,284	0,24	0,02	
6	L151S06	Når gammel kjærlighet rustet_6	31	7	54	8	1	25	21	31	20	0,347	0,54	0,01	
7	L142S01	Stargate_1	1	6	81	11	1	20	18	30	19	0,449	0,81	0,01	
8	L142S02	Stargate_2	88	1	3	7	1	29	11	14	17	0,442	0,88	0,09	
9	L142S03	Stargate_3	25	70			5	22	31			0,410	0,70	0,06	Åpen
10	L142S04	Stargate_4	27	5	5	61	2	25	19	16	31	0,378	0,61	0,10	
11	L142S05	Stargate_5	8	81	3	8	1	20	30	20	16	0,458	0,81	0,11	
12	L142S06	Stargate_6	3	1	6	89	2	14	13	19	29	0,402	0,89	0,05	
13	L142S07	Stargate_7	9	51	16	24	1	22	30	23	28	0,206	0,51	0,03	*
14	L142S08	Stargate_8	13	9	42	34	2	22	19	32	28	0,323	0,42	0,22	
15	L159S01	Dyreparker_1	8	53	21	17	2	19	30	25	31	0,190	0,53	0,11	*
16	L159S02	Dyreparker_2	4	6	83	7	2	15	17	30	20	0,496	0,82	-0,02	
17	L159S03	Dyreparker_3	63	28			9	26	36			0,456	0,28	0,18	Åpen
18	L159S04	Dyreparker_4	8	64	5	20	2	16	31	18	27	0,398	0,64	0,13	
19	L159S05	Dyreparker_5	77	8	6	7	2	30	18	22	21	0,452	0,77	0,03	
20	L159S06	Dyreparker_6	3	12	11	71	3	18	22	21	31	0,475	0,71	0,14	
21	L159S07	Dyreparker_7	9	6	11	70	5	21	23	21	31	0,484	0,69	0,01	
22	L154S01	Stripemysteriet_1	65	5	27	2	2	31	16	25	21	0,352	0,65	0,05	
23	L154S02	Stripemysteriet_2	3	4	58	33	2	12	15	30	28	0,205	0,58	0,00	*
24	L154S03	Stripemysteriet_3	13	80			7	17	30			0,517	0,80	0,06	Åpen
25	L154S04	Stripemysteriet_4	66	5	18	9	3	31	18	26	20	0,395	0,66	0,11	
26	L154S05	Stripemysteriet_5	60	12	16	9	3	32	23	25	18	0,460	0,60	0,05	
27	L154S06	Stripemysteriet_6	9	71	9	8	3	19	32	21	22	0,519	0,71	0,11	
28	L154S07	Stripemysteriet_7	23	8	47	16	6	25	18	33	26	0,463	0,47	0,11	
29	L160S01	Noe du trenger, mor_1	6	76	10	4	4	21	31	20	19	0,506	0,76	0,08	
30	L160S02	Noe du trenger, mor_2	12	52	15	18	5	28	32	25	22	0,375	0,52	0,15	
31	L160S03	Noe du trenger, mor_3	49	29			22	26	36			0,495	0,29	0,19	Åpen
32	L160S04	Noe du trenger, mor_4	11	14	60	11	5	21	21	32	24	0,520	0,60	0,16	
33	L160S05	Noe du trenger, mor_5	64	7	7	18	5	32	20	21	23	0,536	0,64	0,14	
34	L160S06	Noe du trenger, mor_6	13	3	8	71	5	20	17	20	32	0,606	0,71	0,16	
35	L161S01	Salum Kashafali_1	16	9	5	64	6	22	20	20	32	0,546	0,64	0,05	
36	L161S02	Salum Kashafali_2	51	11	6	24	8	33	20	19	25	0,561	0,51	0,04	
37	L161S03	Salum Kashafali_3	29	17	33	13	8	27	25	32	29	0,261	0,33	-0,02	
38	L161S04	Salum Kashafali_4	3	68	11	11	8	17	32	24	20	0,571	0,68	0,04	
39	L161S05	Salum Kashafali_5	37	10	12	31	10	33	26	24	27	0,332	0,37	0,08	
40	L161S06	Salum Kashafali_6	47	41			12	24	34			0,525	0,41	0,11	
41	L161S07	Salum Kashafali_7	48	34			18	26	35			0,484	0,34	0,06	Åpen
42	L162S01	Klimavennlig mat_1	70	11	5	8	7	31	23	19	19	0,510	0,69	0,03	
43	L162S02	Klimavennlig mat_2	8	17	27	41	8	26	24	27	32	0,303	0,41	0,06	
44	L162S03	Klimavennlig mat_3	17	12	6	57	8	26	21	19	32	0,437	0,57	-0,01	
45	L162S04	Klimavennlig mat_4	29	38	15	10	9	27	33	23	27	0,397	0,37	0,03	
46	L162S05	Klimavennlig mat_5	31	55			14	22	33			0,615	0,55	0,19	Åpen
47	L162S06	Klimavennlig mat_6	23	32	17	17	12	29	32	27	25	0,255	0,32	0,10	

Lav diskriminering (under 0,250).

Prøve C

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

N = 341 (169 gutter, 172 jenter). Reliabilitet (Cronbachs alpha): 0,907.

Gjennomsnittlig p-verdi: 0,55 (gutter: 0,52/jenter: 0,59). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Nr.	Item -ID	Prøve C	Svarfordeling i % (frekvens):					Gj.snittlig poengsum på prøven per svaralternativ				Diskrim. (biseriell)	P-verdi	Kjønns-forskjell (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D				
			0 p.	1 p.	2 p.			0 p.	1 p.	2 p.					
1	L141S01	Et skår i gleden_1	4	61	19	16		16	28	22	25	0,216	0,61	0,07	*
2	L141S02	Et skår i gleden_2	4	21	6	68	1	20	22	16	28	0,361	0,68	0,10	
3	L141S03	Et skår i gleden_3													Tatt ut
4	L141S04	Et skår i gleden_4	20	32	11	23	1	30	25	20	26	0,168	0,20	0,01	*
5	L141S05	Et skår i gleden_5	24	35	37	4	1	22	25	30	18	0,335	0,37	0,14	
6	L142S01	Stargate_1	1	9	77	11	1	13	22	28	17	0,415	0,77	0,16	
7	L142S02	Stargate_2	89	1	4	5	1	27	6	11	18	0,437	0,89	0,11	
8	L142S03	Stargate_3	25	70			6	18	29			0,546	0,70	0,13	Åpen
9	L142S04	Stargate_4	28	3	5	63	1	23	12	17	29	0,387	0,62	0,17	
10	L142S05	Stargate_5	4	86	3	7	1	17	27	17	15	0,430	0,86	0,11	
11	L142S06	Stargate_6	3	3	6	87	1	16	15	16	27	0,420	0,87	0,08	
12	L142S07	Stargate_7	8	55	16	20	2	21	28	22	25	0,254	0,55	0,12	
13	L142S08	Stargate_8	11	8	46	34	2	19	18	29	26	0,326	0,45	0,15	
14	L134S09	iHuman_1 (2022)	18	7	48	24	2	25	16	30	22	0,402	0,48	0,08	
15	L134S02	iHuman_2 (2022)	23	3	36	36	2	22	20	25	30	0,292	0,36	-0,07	
16	L134S04	iHuman_3 (2022)	12	36	38	12	3	22	24	31	20	0,426	0,38	0,08	
17	L134S05	iHuman_4 (2022)	10	54	35			14	25	29		0,388	0,62	-0,05	Tabell
		iHuman_4A	32	66			2	23	27						
		iHuman_4B	78	19			3	28	19						
		iHuman_4C	42	52			6	24	28						
18	L134S06	iHuman_5 (2022)	35	50			14	22	31			0,574	0,50	0,15	Åpen
19	L134S07	iHuman_6 (2022)	22	15	21	38	4	22	24	24	31	0,414	0,38	0,04	
20	L163S01	Kraften_1	65	22			14	26	29			0,165	0,22	-0,02	Åpen
21	L163S02	Kraften_2	10	4	75	7	3	17	21	28	22	0,430	0,75	0,11	
22	L163S03	Kraften_3	30	13	9	46	4	25	19	18	30	0,430	0,45	0,19	
23	L163S04	Kraften_4	10	38	36	10	6	23	26	29	23	0,252	0,36	0,02	
24	L163S05	Kraften_5	37	46			18	23	31			0,521	0,45	0,11	Åpen
25	L163S06	Kraften_6	60	10	20	5	5	29	19	22	19	0,467	0,60	0,03	
26	L163S07	Kraften_7	45	17	9	25	4	29	26	21	23	0,315	0,45	0,04	
27	L157S01	Sjøku_1	28	55	6	8	3	24	29	18	21	0,342	0,55	0,21	
28	L157S02	Sjøku_2	72	9	11	5	3	29	16	20	19	0,548	0,72	0,07	
29	L157S03	Sjøku_3	20	69			11	20	29			0,496	0,69	0,06	Åpen
30	L157S04	Sjøku_4	36	65				17	30			0,631	0,65	0,11	Tabell
		Sjøku_4A	81	2	5	8	4	28	15	16	19				
		Sjøku_4B	3	9	79	3	7	15	17	28	16				
		Sjøku_4C	77	3	5	9	7	29	13	15	18				
31	L157S05	Sjøku_5	8	39	40	9	4	22	24	30	22	0,345	0,40	-0,02	
32	L157S06	Sjøku_6	11	10	6	69	4	19	20	16	29	0,549	0,69	0,11	
33	L157S07	Sjøku_7	84	4	4	4	5	28	17	16	17	0,494	0,84	0,07	
34	L127S01	Knøttfe og himmelreker_1 (2022)	13	13	10	57	7	23	19	20	30	0,550	0,57	0,10	
35	L127S06	Knøttfe og himmelreker_2 (2022)	63	11	12	7	7	30	20	23	17	0,504	0,63	0,05	
36	L127S02	Knøttfe og himmelreker_3 (2022)	17	9	58	9	7	24	19	30	20	0,469	0,57	0,01	
37	L127S03	Knøttfe og himmelreker_4 (2022)	8	65	9	11	7	20	30	18	19	0,558	0,65	0,04	
38	L127S07	Knøttfe og himmelreker_5 (2022)	40	38			23	24	32			0,503	0,38	0,03	Åpen
39	L127S05	Knøttfe og himmelreker_6 (2022)	23	11	31	27	8	25	18	31	26	0,337	0,31	0,12	
40	L127S08	Knøttfe og himmelreker_7 (2022)	9	41	28	14	9	20	29	26	25	0,265	0,41	0,06	
41	L164S01	Vandring gjennom tiden_1	9	19	18	47	7	19	22	26	30	0,384	0,47	0,04	
42	L164S02	Vandring gjennom tiden_2	38	47			16	22	31			0,506	0,47	0,05	Åpen
43	L164S03	Vandring gjennom tiden_3	43	39	4	4	9	27	28	23	15	0,055	0,43	0,03	*
44	L164S04	Vandring gjennom tiden_4	19	66	4	3	9	25	28	15	16	0,336	0,66	0,08	
45	L164S05	Vandring gjennom tiden_5	7	6	6	73	8	19	18	20	29	0,478	0,73	0,05	
46	L164S06	Vandring gjennom tiden_6	17	58	7	9	9	22	29	20	26	0,372	0,57	0,13	

*Lav diskriminering (under 0,250)

IRT-analyse av prøve A, B og C

IRT-analysen av oppgavene i prøve A, B og C er utført ved hjelp av programvaren XCalibre 4.2. Tabell 9 viser parameterne for alle oppgavene i prøvene.

De tre prøvene har en felles tekst, «Stargate», og med hjelp av denne lenketeksten er prøvene samkalibrert slik at alle verdier er på samme skala. Lenketeksten lå på plass nummer to i alle prøvene ved pilotering.

I tabell 9 er prøvene adskilt ved hjelp av farger: prøve A er markert som grønn, prøve B er blå og C er rosa. Lenketeksten «Stargate» og «Et skår i gleden», som ble prøvd ut både i prøve A og C, er markert med gult og plassert til slutt i tabellen. For øvrig ligger tekstene og oppgavene i den samme rekkefølgen som de lå i ved pilotering ute på skolene.

Prøve ABC

Tabell 9: Parametre fra IRT-analyse av oppgavene i prøve A, B og C.

A: N=342 (185 gutter, 157 jenter). B: N=331 (173 gutter, 158 jenter). C: N=341 (169 gutter, 172 jenter).

Lav diskriminering (point biserial < 0,250 samt a < 0,5) er merket (*).

Oppg.	Item-ID	Oppgave	Diskriminering (point biserial)	a-verdi	b	DIF (disfavør)
1	L143S01	Lysforurensing_1	0,489	0,806	-0,704	
2	L143S02	Lysforurensing_2	0,291	0,486*	-0,604	
3	L143S03	Lysforurensing_3	0,418	0,668	0,254	
4	L143S04	Lysforurensing_4	0,230*	0,424*	0,515	
5	L143S05	Lysforurensing_5	0,513	0,828	-0,020	
6	L143S06	Lysforurensing_6	0,412	0,650	-0,680	
7	L144S01	Tasmansk pungulv_1	0,095*	0,331*	-0,143	
8	L144S02	Tasmansk pungulv_2	0,501	0,816	-0,609	
9	L144S03	Tasmansk pungulv_3	0,484	0,808	0,619	
10	L144S04	Tasmansk pungulv_4	0,492	0,857	0,552	
11	L144S05	Tasmansk pungulv_5	0,548	1,087	-0,949	
12	L144S06	Tasmansk pungulv_6	0,144*	0,383*	1,246	
13	L144S07	Tasmansk pungulv_7	0,485	0,787	-0,660	
14	L145S01	Den korte historien om egget_1	0,434	0,713	-0,924	
15	L145S02	Den korte historien om egget_2	0,571	1,003	-0,329	F
16	L145S03	Den korte historien om egget_3	0,561	0,977	0,285	
17	L145S04	Den korte historien om egget_4	0,302	0,493*	-0,350	
18	L145S05	Den korte historien om egget_5	0,413	0,648	0,218	
19	L145S06	Den korte historien om egget_6	0,307	0,505	-0,169	
20	L146S01	Snikfotografen_1	0,525	0,874	-0,281	
21	L146S02	Snikfotografen_2	0,585	1,124	-0,766	
22	L146S03	Snikfotografen_3	0,492	0,796	-0,331	
23	L146S04	Snikfotografen_4	0,564	1,088	0,595	
24	L146S05	Snikfotografen_5	0,597	1,071	-0,037	
25	L146S06	Snikfotografen_6	0,509	0,825	0,108	
26	L147S01	Skolearbeid_1	0,446	0,703	-0,394	
27	L147S02	Skolearbeid_2	0,510	0,884	-0,797	
28	L147S03	Skolearbeid_3	0,281	0,496*	0,657	
29	L147S04	Skolearbeid_4	0,403	0,692	0,919	
30	L147S05	Skolearbeid_5	0,429	0,670	-0,583	
31	L147S06	Skolearbeid_6	0,497	0,815	0,215	
32	L147S07	Skolearbeid_7	0,454	0,716	0,054	
33	L151S01	Når gammel kjærlighet rustet_1	0,387	0,741	-1,634	
34	L151S02	Når gammel kjærlighet rustet_2	0,243*	0,437*	-0,120	
35	L151S03	Når gammel kjærlighet rustet_3	0,400	0,638	-0,736	

Oppg.	Item-ID	Oppgave	Diskriminering (point biserial)	a-verdi	b	DIF (disfavør)
36	L151S04	Når gammel kjærlighet rustet_4	0,239*	0,476*	-1,268	
37	L151S05	Når gammel kjærlighet rustet_5	0,284	0,589	1,337	
38	L151S06	Når gammel kjærlighet rustet_6	0,345	0,556	-0,193	
39	L159S01	Dyreparker_1	0,190*	0,399*	-0,207	
40	L159S02	Dyreparker_2	0,495	1,008	-1,338	
41	L159S03	Dyreparker_3	0,456	0,861	0,904	
42	L159S04	Dyreparker_4	0,398	0,629	-0,682	
43	L159S05	Dyreparker_5	0,451	0,822	-1,158	
44	L159S06	Dyreparker_6	0,474	0,816	-0,894	
45	L159S07	Dyreparker_7	0,484	0,809	-0,814	
46	L154S01	Stripemysteriet_1	0,353	0,578	-0,737	
47	L154S02	Stripemysteriet_2	0,204*	0,401*	-0,538	
48	L154S03	Stripemysteriet_3	0,517	1,014	-1,190	
49	L154S04	Stripemysteriet_4	0,394	0,635	-0,738	
50	L154S05	Stripemysteriet_5	0,460	0,760	-0,412	
51	L154S06	Stripemysteriet_6	0,518	0,897	-0,824	
52	L154S07	Stripemysteriet_7	0,463	0,746	0,129	
53	L160S01	Noe du trenger, mor_1	0,505	0,920	-1,063	
54	L160S02	Noe du trenger, mor_2	0,373	0,611	-0,081	gutter
55	L160S03	Noe du trenger, mor_3	0,491	0,976	0,785	
56	L160S04	Noe du trenger, mor_4	0,519	0,861	-0,397	
57	L160S05	Noe du trenger, mor_5	0,535	0,908	-0,542	
58	L160S06	Noe du trenger, mor_6	0,604	1,190	-0,741	
59	L161S01	Salum Kashafali_1	0,546	0,936	-0,534	
60	L161S02	Salum Kashafali_2	0,561	0,980	-0,041	
61	L161S03	Salum Kashafali_3	0,262	0,496	0,941	
62	L161S04	Salum Kashafali_4	0,570	1,073	-0,652	
63	L161S05	Salum Kashafali_5	0,332	0,567	0,661	
64	L161S06	Salum Kashafali_6	0,525	0,933	0,330	
65	L161S07	Salum Kashafali_7	0,484	0,888	0,614	
66	L162S01	Klimavennlig mat_1	0,510	0,891	-0,774	
67	L162S02	Klimavennlig mat_2	0,304	0,536	0,476	
68	L162S03	Klimavennlig mat_3	0,437	0,682	-0,303	
69	L162S04	Klimavennlig mat_4	0,398	0,668	0,565	
70	L162S05	Klimavennlig mat_5	0,614	1,162	-0,181	
71	L162S06	Klimavennlig mat_6	0,256	0,495*	1,035	
72	L134S09	iHuman_1 (2022)	0,402	0,626	0,086	
73	L134S02	iHuman_2 (2022)	0,293	0,517	0,750	
74	L134S04	iHuman_3 (2022)	0,428	0,711	0,542	
75	L134S05	iHuman_4 (2022)	0,251	0,464*	0,902	
76	L134S06	iHuman_5 (2022)	0,574	0,996	-0,006	
77	L134S07	iHuman_6 (2022)	0,415	0,688	0,553	
78	L163S01	Kraften_1	0,163*	0,442*	1,866	
79	L163S02	Kraften_2	0,429	0,757	-1,095	gutter
80	L163S03	Kraften_3	0,431	0,687	0,210	
81	L163S04	Kraften_4	0,252	0,466*	0,826	
82	L163S05	Kraften_5	0,524	0,910	0,180	
83	L163S06	Kraften_6	0,465	0,747	-0,428	
84	L163S07	Kraften_7	0,315	0,518	0,268	
85	L157S01	Sjøku_1	0,344	0,552	-0,234	gutter
86	L157S02	Sjøku_2	0,546	1,065	-0,804	
87	L157S03	Sjøku_3	0,491	0,881	-0,739	
88	L157S04	Sjøku_4	0,631	1,360	-0,460	
89	L157S05	Sjøku_5	0,348	0,577	0,513	
90	L157S06	Sjøku_6	0,552	1,083	-0,669	
91	L157S07	Sjøku_7	0,490	1,111	-1,326	
92	L127S01	Knøttfe og himmelreker_1 (2022)	0,554	0,991	-0,250	
93	L127S06	Knøttfe og himmelreker_2 (2022)	0,505	0,870	-0,487	
94	L127S02	Knøttfe og himmelreker_3 (2022)	0,470	0,761	-0,299	
95	L127S03	Knøttfe og himmelreker_4 (2022)	0,560	1,045	-0,544	
96	L127S07	Knøttfe og himmelreker_5 (2022)	0,504	0,898	0,488	

Oppg.	Item-ID	Oppgave	Diskriminering (point biserial)	a-verdi	b	DIF (disfavør)
97	L127S05	Knøttfe og himmelreker_6 (2022)	0,337	0,582	0,975	
98	L127S08	Knøttfe og himmelreker_7 (2022)	0,266	0,458*	0,534	
99	L164S01	Vandring gjennom tiden_1	0,383	0,623	0,155	
100	L164S02	Vandring gjennom tiden_2	0,507	0,848	0,141	
101	L164S03	Vandring gjennom tiden_3	0,055*	0,311*	0,534	
102	L164S04	Vandring gjennom tiden_4	0,332	0,567	-0,822	
103	L164S05	Vandring gjennom tiden_5	0,475	0,824	-0,942	
104	L164S06	Vandring gjennom tiden_6	0,372	0,581	-0,356	
105	L141S01	Et skår i gleden_1	0,261	0,354*	-0,826	
106	L141S02	Et skår i gleden_2	0,391	0,568	-0,970	
Tatt ut	L141S03	Et skår i gleden_3				
107	L141S04	Et skår i gleden_4	0,186*	0,365*	2,178	
108	L141S05	Et skår i gleden_5	0,250	0,386*	0,897	
109	L142S01	Stargate_1	0,427	0,789	-1,353	
110	L142S02	Stargate_2	0,401	0,980	-1,791	gutter
111	L142S03	Stargate_3	0,488	0,832	-0,768	gutter
112	L142S04	Stargate_4	0,354	0,536	-0,584	gutter
113	L142S05	Stargate_5	0,441	0,888	-1,427	gutter
114	L142S06	Stargate_6	0,397	0,922	-1,769	
115	L142S07	Stargate_7	0,247*	0,385*	-0,173	
116	L142S08	Stargate_8	0,293	0,449*	0,299	gutter

I tabell 9 er oppgavenes diskrimineringsevne oppgitt både som et klassisk mål (point biserial) og en a-verdi. Oppgaver med diskriminering under 0,250 (point biserial) samt a-parameter under 0,5 er merket med stjerne. Oppgaver som er merket for begge diskrimineringsparameterne, er markert med grått, og de oppgavene som blir med videre til andre pilot, blir kommentert senere i denne rapporten.

Teksten «Stargate» har flere oppgaver med DIF i guttenes disfavør, men den er allikevel tatt med videre i prosessen. Begrunnelsen for dette er todelt. For det første ga referanselærerne våre tilbakemelding om at «Stargate» var den beste teksten (sammenliknet med «Kraften» og «Noe du trenger mor») å arbeid faglig med. Dette inntrykket ble forsterket gjennom møter med kvalitetssikrerne våre. For det andre mener vi at DIF-en i dette tilfellet kan være knyttet til at teksten har en jente som hovedperson, noe som vi av erfaring vet fører til store kjønnsforskjeller mellom gutter og jenter. Vi har opplevd å få kritikk for at de skjønnlitterære tekstene ofte er tilpasset guttelesere, og vi mener derfor at det er et poeng at vi noen ganger også bruker tekster med jenter som hovedperson, selv om dette kan favorisere jenter i noen grad. Vi kommer uansett til å følge med på DIF-analysene i det videre arbeidet med teksten, og søke å minimere utslag av DIF der det er mulig.

Oppgavene fordelt på mestringsnivåer

Tabell 10 viser oppgavene i de tre piloteringsprøvene fordelt på mestringsnivåer. Nivågrensene som er brukt, ble bestemt etter gjennomføringen av nasjonale leseprøver i 2016. Theta-verdiene som kommer ut av denne IRT-analysen, er sentrert på theta-fordelingen i piloteringsprøven og er ikke de samme som ved gjennomføringen. Piloteringen er altså ikke ankret på startpunktet av trenden i 2016. Det er derfor en viss usikkerhet knyttet til nivågrensene.

Etter gjennomføring av prøven i 2022 ble det satt nye nivågrenser, og disse vil brukes når resultatene etter andre pilot analyseres.

Tabell 10: Oppgavene i piloteringsprøvene, fordelt på mestringsnivåer.

Nivå	Løpenummer	Item-ID	Oppgave	Skalapoeng
Nivå 1	6	L142S02	Stargate_2	32,09
	10	L142S06	Stargate_6	32,31
	45	L151S01	Når gammel kjærlighet ruster_1	33,66
	9	L142S05	Stargate_5	35,73
	5	L142S01	Stargate_1	36,47
	52	L159S02	Dyreparker_2	36,62
	103	L157S07	Sjøku_7	36,74
	48	L151S04	Når gammel kjærlighet ruster_4	37,32
	60	L142S02	Stargate_2	32,09
Nivå 2	55	L154S03	Stripemysteriet_3	38,10
	91	L159S05	Dyreparker_5	38,42
	65	L163S02	Kraften_2	39,05
	2	L160S01	Noe du trenger, mor_1	39,37
	23	L141S02	Et skår i gleden_2	40,30
	115	L144S05	Tasmansk pungulv_5	40,51
	26	L164S05	Vandring gjennom tiden_5	40,58
	56	L145S01	Den korte historien om egget_1	40,76
	1	L159S06	Dyreparker_6	41,06
	63	L141S01	Et skår i gleden_1	41,74
	114	L154S06	Stripemysteriet_6	41,76
	57	L164S04	Vandring gjennom tiden_4	41,78
	98	L159S07	Dyreparker_7	41,86
	39	L157S02	Sjøku_2	41,96
	78	L147S02	Skolearbeid_2	42,03
	7	L162S01	Klimavennlig mat_1	42,26
	33	L142S03	Stargate_3	42,32
	70	L146S02	Snikfotografen_2	42,34
	99	L160S06	Noe du trenger, mor_6	42,59
	61	L157S03	Sjøku_3	42,61
	58	L154S04	Stripemysteriet_4	42,62
	47	L154S01	Stripemysteriet_1	42,63
	13	L151S03	Når gammel kjærlighet ruster_3	42,64
	54	L143S01	Lysforurensing_1	42,97
	18	L159S04	Dyreparker_4	43,18
	102	L143S06	Lysforurensing_6	43,20
	25	L157S06	Sjøku_6	43,31
	74	L144S07	Tasmansk pungulv_7	43,40
	20	L161S04	Salum Kashafali_4	43,48
	14	L144S02	Tasmansk pungulv_2	43,91
	8	L143S02	Lysforurensing_2	43,96
Nivå 3	42	L142S04	Stargate_4	44,16
	107	L147S05	Skolearbeid_5	44,17
	69	L127S03	Knøttfe og himmelreker_4	44,56
	59	L160S05	Noe du trenger, mor_5	44,58
	71	L154S02	Stripemysteriet_2	44,62
	105	L161S01	Salum Kashafali_1	44,66
	100	L127S06	Knøttfe og himmelreker_2	45,13
	95	L157S04	Sjøku_4	45,40
	62	L163S06	Kraften_6	45,72
	68	L154S05	Stripemysteriet_5	45,88
	38	L160S04	Noe du trenger, mor_4	46,03
	116	L147S01	Skolearbeid_1	46,06
	29	L164S06	Vandring gjennom tiden_6	46,44
	34	L145S04	Den korte historien om egget_4	46,50
	27	L146S03	Snikfotografen_3	46,69
	80	L145S02	Den korte historien om egget_2	46,71
	106	L162S03	Klimavennlig mat_3	46,97

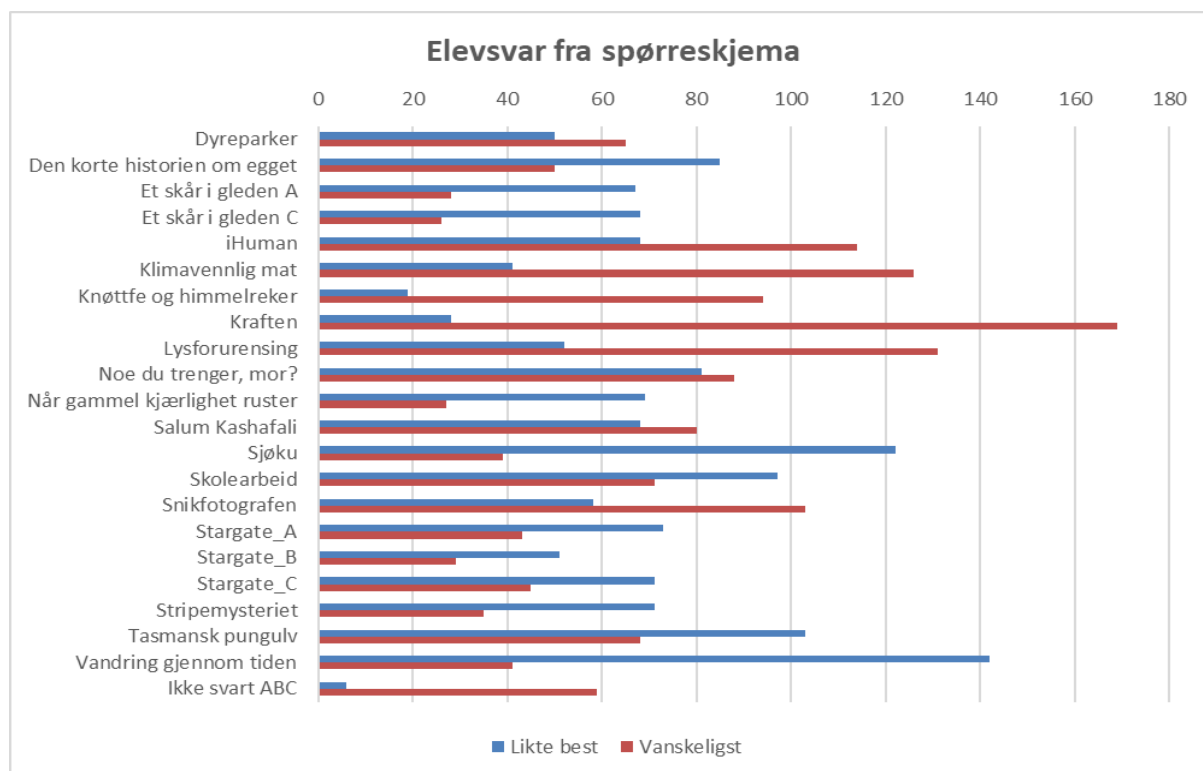
Nivå	Løpenummer	Item-ID	Oppgave	Skalapoeng
	32	L127S02	Knøttfe og himmelreker_3	47,01
	104	L146S01	Snikfotografen_1	47,19
	97	L127S01	Knøttfe og himmelreker_1	47,51
	51	L157S01	Sjøku_1	47,66
	50	L159S01	Dyreparker_1	47,94
	82	L151S06	Når gammel kjærlighet ruster_6	48,07
	11	L162S05	Klimavennlig mat_5	48,19
	31	L142S07	Stargate_7	48,27
	19	L145S06	Den korte historien om egget_6	48,31
	46	L144S01	Tasmansk pungulv_1	48,57
	66	L151S02	Når gammel kjærlighet ruster_2	48,80
	72	L160S02	Noe du trenger, mor_2	49,19
	36	L161S02	Salum Kashafali_2	49,59
	17	L146S05	Snikfotografen_5	49,63
	88	L143S05	Lysforurensing_5	49,80
	44	L134S06	iHuman_5	49,94
	84	L147S07	Skolearbeid_7	50,54
	37	L134S09	iHuman_1	50,86
	64	L146S06	Snikfotografen_6	51,08
	112	L154S07	Stripemysteriet_7	51,29
	111	L164S02	Vandring gjennom tiden_2	51,41
	94	L164S01	Vandring gjennom tiden_1	51,55
	92	L163S05	Kraften_5	51,80
	43	L163S03	Kraften_3	52,10
	30	L147S06	Skolearbeid_6	52,15
	15	L145S05	Den korte historien om egget_5	52,18
	96	L143S03	Lysforurensing_3	52,54
	28	L163S07	Kraften_7	52,68
	12	L145S03	Den korte historien om egget_3	52,85
	76	L142S08	Stargate_8	52,99
	79	L161S06	Salum Kashafali_6	53,30
	108	L162S02	Klimavennlig mat_2	54,76
	101	L127S07	Knøttfe og himmelreker_5	54,88
Nivå 4	16	L157S05	Sjøku_5	55,13
	113	L143S04	Lysforurensing_4	55,15
	110	L164S03	Vandring gjennom tiden_3	55,34
	86	L127S08	Knøttfe og himmelreker_7	55,34
	22	L134S04	iHuman_3	55,42
	89	L144S04	Tasmansk pungulv_4	55,52
	81	L134S07	iHuman_6	55,53
	35	L162S04	Klimavennlig mat_4	55,65
	77	L146S04	Snikfotografen_4	55,95
	21	L161S07	Salum Kashafali_7	56,14
	40	L144S03	Tasmansk pungulv_3	56,19
	75	L147S03	Skolearbeid_3	56,57
	85	L161S05	Salum Kashafali_5	56,61
	67	L134S02	iHuman_2	57,50
	93	L160S03	Noe du trenger, mor_3	57,85
	4	L163S04	Kraften_4	58,26
	87	L141S05	Et skår i gleden_5	58,97
	53	L134S05	iHuman_4	59,02
	41	L159S03	Dyreparker_3	59,04
	73	L147S04	Skolearbeid_4	59,19
	109	L161S03	Salum Kashafali_3	59,42
	83	L127S05	Knøttfe og himmelreker_6	59,75
	24	L162S06	Klimavennlig mat_6	60,36
	49	L144S06	Tasmansk pungulv_6	62,46
Nivå 5	90	L151S05	Når gammel kjærlighet ruster_5	63,37
	3	L163S01	Kraften_1	68,66
	6	L141S04	Et skår i gleden_4	71,78

Ideelt sett skal prøven måle tilfredsstillende og med et tilstrekkelig antall oppgaver i alle deler av skalaen, jevnt fordelt mellom mestringsnivåene. Dette har imidlertid vist seg å være vanskelig å få til, og som det går fram av tabell 10, er prøve A, B og C ikke noe unntak. Det er kun tre oppgaver på det øverste nivået, 24 oppgaver på det nest øverste nivået og 50 oppgaver på det midterste, nivå 3.

Utvalg av tekster til andre pilotering

Ved utvalg av tekster til andre pilotering ligger ikke bare resultatene og analysene fra første pilotering til grunn, men også tilbakemeldinger fra de eksterne, lesefaglige kvalitetssikrerne for nasjonale leseprøver ved Uppsala universitet, og innspill fra referansegruppen bestående av lærere på ungdomstrinnet. Kvalitetssikrerne fra Uppsala vurderte først og fremst kvaliteten på oppgaveformuleringene og svaralternativene, mens lærerne vurderte tekster og oppgaver med tanke på vanskegrad og egnethet for aldersgruppa samt potensiale for anvendelse og etterarbeid i klasserommet. Referanselærerne ga til dels svært positive tilbakemeldinger på tekstene.

Elevene som deltok på den første piloteringen, ga også tilbakemeldinger på hvilke tekster de likte best og hvilke de syntes var vanskeligst. Tilbakemeldingene ble innhentet ved hjelp av et enkelt spørreskjema (vedlegg 3A-C) som lå bakerst i prøven. Figur 7 viser hvor mange ganger en tekst ble nevnt som svar på spørsmålene «Hvilken tekst likte du best?» og «Hvilken tekst var vanskeligst?». Ettersom elevene kunne oppgi mer enn én tekst på begge spørsmålene, er det til sammen flere «stemmer» enn antall elever.



Figur 7: Antall ganger tekstene er nevnt som svar på spørsmålene «Hvilken tekst likte du best?» (blå søyler) og «Hvilken tekst var vanskeligst?» (røde søyler)

Etter en helthetsvurdering har vi kommet fram til at følgende tekster blir med videre til andre pilotering: «Et skår i gleden?», «Stargate», «Lysforurensing», «Skulearbeid», «Snikfotografen», «Tasmansk pungulv» og «Den korte historien om egget». Tekstene er nærmere beskrevet under avsnittet «Nasjonal leseprøve for 8. trinn 2023» i denne rapporten.

Nedenfor gir vi en kort begrunnelse for hvorfor de følgende ti tekstene ikke blir med videre: «Når gammel kjærlighet rustet», «Dyreparker», «Stripemysteriet», «Noe du trenger, mor?», «Salum Kashafali», «Klimavennlig mat», «Kraften», «Sjøku» og «Vandring gjennom tiden».

- «**Når gammel kjærlighet rustet**» er, i likhet med teksten «Et skår i gleden?», en helsides annonse for butikkjeden Jernia. Begge annonsene inneholder en kort, informerende verbaltekst og et stort bilde av henholdsvis en stekepanne og en knust tallerken, samt en overskrift som viser til et fast uttrykk. «Et skår i gleden?» ble vurdert til å ligge litt nærmere elevenes egne erfaringer og ble derfor valgt framfor teksten «Når gammel kjærlighet rustet».
- «**Dyreparker**» er en nyhetsartikkel fra nettavisen.no, der en MDG-politiker uttrykker at «dyreparker hører til på historiens skraphaug». Selv om teksten ikke er skrevet for målgruppa, er temaet kjent, og oppgavene var middels vanskelige. Elevene vurderte også teksten til å være middels vanskelig. «Dyreparker» har noen likhetstrekk med «Lysforurensing», som heller ikke er skrevet for målgruppa, men som handler om en viktig og sannsynligvis ikke fullt så kjent problemstilling. Det er ikke plass til flere enn én sakprosaetekst i denne sjangeren i prøven og valget falt på «Lysforurensing».
- «**Stripemysteriet**» handler om hvorfor sebraen har striper. Teksten er multippel og består av en magasinartikkel og et utdrag fra en sakprosaabok. Det er både sammenfallende og motstridende informasjon i de to tekstene, som dermed egner seg godt i leseprøven. Teksten «Tasmansk pungulv» har noen av de samme kvalitetene, og ut fra en helthetsvurdering ble denne valgt framfor «Stripemysteriet».
- «**Noe du trenger, mor?**» er et utdrag fra boka «Til de voksne» av Linn Skåber. Teksten er skrevet i en lett og morsom stil, men med et underliggende alvor som krever at leseren greier å skille mellom det jeg-personen sier og det hun tenker. Den passer derfor i leseprøven, men med bare én skjønnlitterær plass tilgjengelig, tapte den i denne omgang plassen mot et utdrag fra boka «Stargate» av Ingvill H. Rishøi.
- «**Salum Kashafali**» er en nyhetsartikkel fra august 2021, om Salum Kashafali som vant gull på 100-meter i paralympics. Teksten ble i første omgang vurdert som egnet, blant annet fordi den kombinerer fortellingen om Kashafalis dramatiske liv og den positive drivkraften som har ført til at han har oppnådd stor suksess som parautøver. Artikkelen er imidlertid lang og vil kreve mye omarbeiding for å fungere godt i prøven.
- «**Klimavennlig mat**» er en grafisk tekst som er oversatt fra engelsk. Den ble vurdert til å være mindre relevant og interessant for elevene enn den andre grafiske teksten i utvalget,

nemlig «Skularbeid», som er et utdrag fra Elevundersøkelsen.

- **«Kraften»** er et utdrag fra boka «Derfor må du vite at jeg er same» av artisten og vinneren av Stjernekamp Ella Marie Hætta Isaksen. Teksten er en personlig fortelling med referanser til samenes historie og fornorskingspolitikken. Dette er et relevant tema på ungdomstrinnet, men pilotresultatene og tilbakemeldingen fra elevene viser at dette tekstutdraget ble oppfattet som vanskelig, og kanskje egner det seg bedre som undervisningsmaterie i klasserommet enn i leseprøven.
- **«Sjøku»** er en multipl tekst bestående av et kapittel fra boka «Dyrene som forsvant» av Lene Renslebråten og en Wikipedia-artikkel. Denne teksten ble også prøvd ut på 5. trinn. Pilotresultatene viser at teksten ble lett på 8. trinn, og den tapte dessuten konkurransen om en plass i prøven med «Tasmansk pungulv» fra samme bok.
- **«Vandring gjennom tiden»** er en nokså avansert grafisk framstilling av menneskets vandring fra Afrika til de andre kontinentene opp gjennom millioner av år. Elevene rapporterte at de likte denne teksten godt, men det er noen problemer forbundet med fargekodene og en lite UU-vennlig framstilling som gjør at vi velger å ikke ha denne teksten med i prøven.

Endringer etter første pilotering

Nedenfor gir vi en oppsummering av endringer som ble gjort i tekstene og oppgavene som ble valgt ut til den endelige prøven, før de ble pilotert på nytt. For øvrig viser vi til vedlegg 2 som inneholder en detaljert oversikt over alle oppgavene med item-nummer og oppgave-ordlyden i de ulike piloteringene. Alle endringer er dokumentert ved hjelp av funksjonen «spor endringer» i Word, og tekstfilene er lagret i prøveutviklernes arkiver.

- **«Et skår i gleden»:** Oppgave_3 (L141S03) ble tatt ut av analysegrunnlaget i første pilotering, fordi det var stor variasjon i elevsvarene og vanskelig å kode dem på en konsistent måte. Oppgaven ble erstattet av en ny, åpen oppgave (L141S06). Flervalgsoppgaven L141S05, «Hva er hovedhensikten med denne teksten?» er tatt ut. En ny flervalgsoppgave (L141S07) og en ny åpen oppgave (L141S08) er lagt til.
- **«Stargate»:** Antall oppgaver er redusert fra åtte til seks ved at de fire oppgavene L142S01, -02, -03 og -06 er tatt ut og erstattet av to nye oppgaver (L142S09 og -10).
- **«Lysforurensing»:** Innledningen til teksten er forkortet med ca. 100 ord, og antall oppgaver er økt fra seks til sju. De tre flervalgsoppgavene L143S02, -03 og -04 er tatt ut, og fire nye flervalgsoppgaver, L143S07, -08, -09 og -10, er lagt til.
- **«Skularbeid»:** Oversatte teksten til fra bokmål til nynorsk og reduserte med én oppgave (L147S05). Endret formuleringen av den åpne oppgaven L147S06.

- **«Snikfotografen»:** Laget en ny oppgave, L146S07, som krever at elevene sorterer hendelsene i teksten i riktig rekkefølge.
- **«Tasmansk pungulv»:** Oppgave L144S05 ble tatt ut. En ny oppgave, L144S08, ble lagt til. Ellers ble det gjort små endringer i ordlyden i oppgave L144S01, -04 og -06 (se vedlegg 2).
- **«Den korte historien om egget»:** Det er kun gjort noen små endringer i ordlyden i oppgavene L145S02, -03 og -04 (se vedlegg 2).

Andre pilotering

De sju utvalgte tekstene ble satt sammen til en ny prøve, D. Denne prøven inneholdt 44 oppgaver, og tekstene ble prøvd ut i den rekkefølgen som er planlagt for den ordinære gjennomføringen. Prøvetiden var 90 minutter, og prøven ble gjennomført i digital plattform, i august 2022. Det var 583 elever fra 8 skoler i 5 ulike fylker som deltok.

Tabell 11: Oversikt over antall elever i andre pilot

	Elever
Antall inviterte elever	4394 (fra 39 skoler)
Antall oppmeldte elever	791 (fra 8 skoler)
Antall elever som gjennomførte (N)	Totalt: 583 (310 gutter, 273 jenter)
Antall skoler	8

Innholdet i prøve D

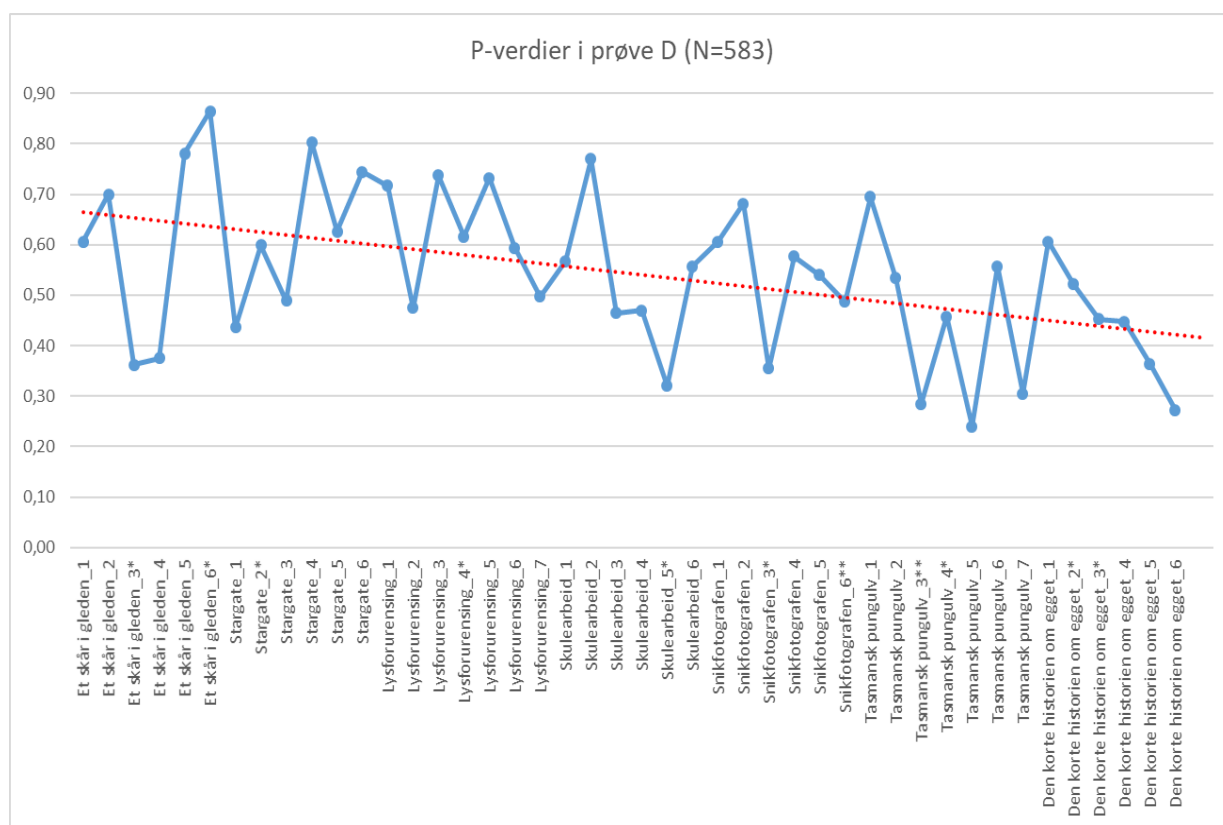
Tekstene fordeler seg på fag og teksttyper som vist i tabell 12. Kolonnen lengst til venstre viser hvilken prøve og plassering den enkelte tekst hadde i første pilot.

Tabell 12: Oversikt over tekstene i prøve D.

Plass i første pilot	Tekst	Fagtilhørighet	Teksttype	Antall ord: 5658	Antall oppgaver: 44			Gjennomsnittlig p-verdi	Kjønnsforskjell (j-g)*
					Flervalg: 33	Åpne: 9	Rekkefølge: 2		
A-1	<i>Et skår i gleden?</i>	NOR, SAF	Reklame	500	4	2		0,61	0,08
ABC-2	<i>Stargate</i>	NOR	Skjønnlitterært utdrag	745	5	1		0,62	0,14
A-3	<i>Lysforurensing</i>	NAT, SAF, KHV	Argumenterende blogginnlegg	1050	6	1		0,62	0,03
A-7	<i>Skularbeid (sidemål)</i>	SAF, MAT	Grafikk	500	5	1		0,53	0,04
A-6	<i>Snikfotografen</i>	KHV, SAF	Nettartikkel	983	4	1	1	0,54	0,09
A-4	<i>Tasmansk pungulv</i>	NAT, SAF	Multippel: fagbok og nettartikkel	1040	5	1	1	0,44	0,02
A-5	<i>Den korte historien om egget</i>	MHE, NAT	Utdrag fra fagbok	867	4	2		0,44	0,06

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

Figur 8 viser p-verdien for hver enkelt oppgave i prøve D.

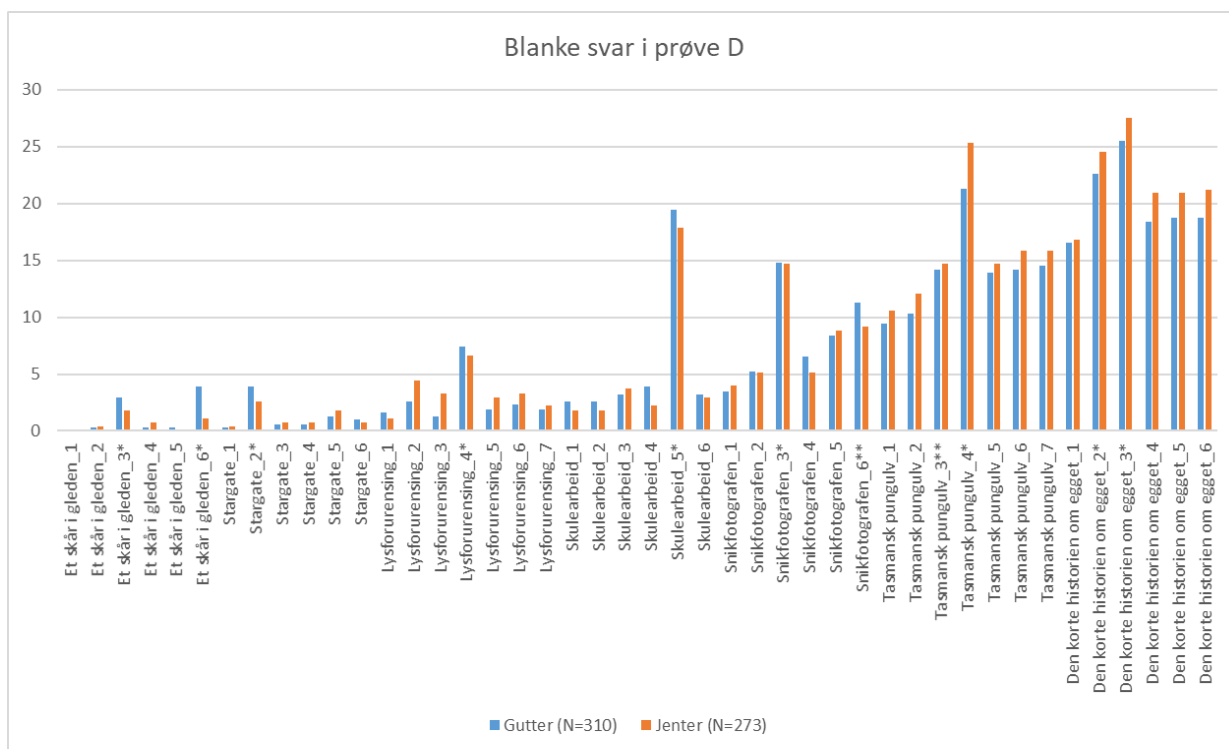


Figur 8: Andelen elever med riktig svar (p-verdi) på oppgavene i prøve D, organisert etter rekkefølgen i prøven.

Figur 8 viser at alle tekstene har både lette og vanskelige oppgaver, men også her viser trendlinjen en nedgang i p-verdi mot slutten av prøven. Dette er en vanlig tendens, og en del av forklaringen er at en stor andel elever lar være å svare på oppgaver langt bak i prøven – kanskje fordi de mister motivasjonen eller får for lite tid.

Blanke svar i prøve D

Figur 9 nedenfor viser prosentandelen gutter og jenter som ikke har svart på hver oppgave i prøve D. Åpne oppgaver er markert med stjerne stjerne (*).



Figur 9: Prosentandel gutter og jenter som ikke svarte på hver oppgave i prøve D. Åpne oppgaver er merket (*).

I prøve D er den gjennomsnittlige andelen blanke svar 7,8 prosent – 7,7 prosent for gutter og 8,2 for jenter. Dette er en forskjell fra første pilotering, der det var en større andel gutter enn jenter som hadde latt være å svare på oppgavene i alle de tre prøvene. For øvrig viser figur 9 det samme mønsteret som tidligere, nemlig at andelen blanke svar er stigende utover i prøven og høyest for de åpne oppgavene. I prøve A var den gjennomsnittlige andelen blanke svar 7,3 prosent, mens den for prøve B og C var henholdsvis 5,6 og 5,8 prosent. Ved ordinær gjennomføring de siste årene har andelen ligget på omtrent 4 prosent.

Klassisk item-analyse av prøve D

Tabell 13 viser resultater fra klassisk item-analyse av prøve D. Resultatene i tabellene 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, som ifølge fastsatte mål bør være over 0,250 for at oppgaven skal differensiere godt mellom lesere med svake og sterke leseferdigheter. De siste kolonnene viser de ulike oppgavenes p-verdi og kjønnsforskjell, der positive verdier er i jentenes favør og negative verdier i guttenes favør.

Tabell 13: Resultater fra klassisk item-analyse, prøve D.

N = 583 (310 gutter, 273 jenter). Reliabilitet (Cronbachs alpha): 0,906.

Gjennomsnittlig p-verdi: 0,54 (gutter: 0,51 / jenter: 0,58). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Nr.	Item-ID	Prøve D	Svarfordeling i % (frekvens):					Gj.snittlig poengsum på prøven per svaralternativ				Diskrim. (biserial)	P-verdi	Kjønnsforskjell (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D				
			0 p.	1 p.	2 p.			0 p.	1 p.	2 p.					
1	L141S01	Et skår i gleden_1	4	61	18	17		13	26	20	21	0,281	0,61	0,03	
2	L141S02	Et skår i gleden_2	7	17	6	70		16	20	17	26	0,333	0,70	0,03	
3	L141S06	Et skår i gleden_3	61	36			2	21	29			0,349	0,36	0,10	Åpen
4	L141S04	Et skår i gleden_4	38	32	16	14	1	28	22	17	24	0,302	0,38	0,07	
5	L141S07	Et skår i gleden_5	9	7	78	5		13	19	26	20	0,388	0,78	0,12	
6	L141S08	Et skår i gleden_6	13	84			3	14	26			0,455	0,84	0,14	Åpen
7	L142S09	Stargate_1	19	11	26	44		18	18	24	28	0,325	0,44	0,10	
8	L142S10	Stargate_2	37	60			3	20	27		11	0,346	0,60	0,16	Åpen
9	L142S04	Stargate_3	41	5	49	5	1	22	18	27	16	0,275	0,49	0,20	
10	L142S05	Stargate_4	6	80	3	10	1	16	26	15	17	0,399	0,80	0,15	
11	L142S08	Stargate_5	16	9	63	11	2	18	18	27	20	0,435	0,63	0,12	
12	L142S11	Stargate_6	8	7	9	74	1	16	17	22	26	0,334	0,74	0,13	
13	L143S01	Lysforurensing_1	16	6	5	72	1	19	13	16	27	0,452	0,72	0,04	
14	L143S09	Lysforurensing_2	23	48	14	12	3	25	27	16	19	0,323	0,48	-0,06	
15	L143S10	Lysforurensing_3	74	8	4	12	2	26	14	17	20	0,427	0,74	0,00	
16	L143S05	Lysforurensing_4	31	62			7	19	28			0,480	0,62	0,16	Åpen
17	L143S06	Lysforurensing_5	73	15	5	5	2	26	21	15	15	0,370	0,73	0,07	
18	L143S07	Lysforurensing_6	59	9	26	3	3	26	16	24	17	0,228	0,59	0,01	*
19	L143S08	Lysforurensing_7	12	19	17	50	2	18	23	21	27	0,334	0,50	-0,03	
20	L147S01	Skulearbeid_1 (sidemål)	57	13	3	26	2	27	20	9	23	0,285	0,57	0,05	
21	L147S02	Skulearbeid_2 (sidemål)	2	77	18	1	2	13	26	20	11	0,302	0,77	0,03	
22	L147S03	Skulearbeid_3 (sidemål)	7	31	47	12	3	20	21	27	24	0,262	0,46	0,03	
23	L147S04	Skulearbeid_4 (sidemål)	22	8	20	47	3	20	18	25	27	0,319	0,47	0,07	
24	L147S06	Skulearbeid_5 (sidemål)	49	32			19	22	31			0,469	0,32	0,04	Åpen
25	L147S07	Skulearbeid_6 (sidemål)	15	15	56	12	3	21	18	27	25	0,314	0,56	0,05	
26	L146S01	Snikfotografen_1	61	20	9	6	4	28	20	20	17	0,444	0,61	-0,02	
27	L146S03	Snikfotografen_2	68	7	12	8	5	27	19	18	20	0,445	0,68	0,10	
28	L146S04	Snikfotografen_3	50	36			15	21	31			0,547	0,36	0,23	Åpen
29	L146S05	Snikfotografen_4	26	6	5	58	6	18	16	21	29	0,569	0,58	0,07	
30	L146S06	Snikfotografen_5	15	10	54	13	9	19	17	29	22	0,522	0,54	0,10	
31	L146S07	Snikfotografen_6	41	49			10	21	29			0,486	0,49	0,09	Rekkefølge
32	L144S01	Tasmansk pungulv_1	70	10	6	4	10	27	20	17	17	0,506	0,69	0,04	
33	L144S02	Tasmansk pungulv_2	12	16	7	53	11	21	21	20	30	0,493	0,53	-0,02	
34	L144S03	Tasmansk pungulv_3	57	29			14	22	32			0,551	0,28	-0,03	Rekkefølge
35	L144S04	Tasmansk pungulv_4	31	46			23	21	32			0,635	0,46	0,01	Åpen
36	L144S05	Tasmansk pungulv_5	27	24	8	26	14	26	29	16	27	0,215	0,24	0,04	*
37	L144S06	Tasmansk pungulv_6	16	56	9	5	15	23	30	16	20	0,505	0,56	0,07	
38	L144S07	Tasmansk pungulv_7	13	22	20	31	15	22	23	26	31	0,382	0,31	0,00	
39	L145S01	Den korte historien om egget_1	2	5	61	16	17	17	20	29	22	0,404	0,61	0,06	
40	L145S02	Den korte historien om egget_2	24	52			24	21	31			0,553	0,52	0,14	Åpen
41	L145S03	Den korte historien om egget_3	28	45			26	21	32			0,584	0,45	0,09	Åpen
42	L145S04	Den korte historien om egget_4	45	7	5	24	20	30	18	20	25	0,411	0,45	0,01	
43	L145S05	Den korte historien om egget_5	9	36	29	6	20	19	31	25	22	0,396	0,36	0,03	
44	L145S06	Den korte historien om egget_6	19	27	18	16	20	22	31	26	25	0,352	0,27	0,02	

*Lav diskriminering (under 0,250).

IRT-analyse av prøve D

Oppgavene i prøve D er analysert ved hjelp av IRT, og tabellen nedenfor viser parameterne for alle oppgavene i prøven. Vi gjør oppmerksom på at IRT-analysen er utført ved hjelp av programvaren R. Dette er en forskjell fra første pilot, der analysene ble utført ved hjelp av Xcalibre. I tabell 14 er nedre grense for a-verdiene 0,8, og de er derfor ikke direkte sammenlignbare med a-verdiene etter første pilot (tabell 9).

Tabell 14: Parametre fra IRT-analyse av oppgavene i prøve D. Reliabilitet (Cronbachs Alpha): 0,906. N = 583 (273 gutter, 310 jenter). Lav diskriminering (point biserial < 0,250 og a < 0,8) er merket (*).

Nr.	Item-ID	Oppgave	Diskriminering (point biserial)	a-verdi	b	DIF (disfavør)
1	L141S01	Et skår i gleden_1	0,302	0,643*	-0,728	0
2	L141S02	Et skår i gleden_2	0,360	0,853	-1,143	0
3	L141S06	Et skår i gleden_3	0,374	0,856	0,766	0
4	L141S04	Et skår i gleden_4	0,321	0,701*	0,804	0
5	L141S07	Et skår i gleden_5	0,432	1,235	-1,312	0
6	L141S08	Et skår i gleden_6	0,483	1,759	-1,406	0,556
7	L142S09	Stargate_1	0,337	0,726	0,399	0
8	L142S10	Stargate_2	0,367	0,816	-0,56	0,673
9	L142S04	Stargate_3	0,292	0,604*	0,08	1,334
10	L142S05	Stargate_4	0,422	1,246	-1,439	0,721
11	L142S08	Stargate_5	0,470	1,174	-0,558	0,341
12	L142S11	Stargate_6	0,352	0,875	-1,412	0,659
13	L143S01	Lysforurensing_1	0,476	1,301	-0,939	0
14	L143S09	Lysforurensing_2	0,344	0,737	0,152	-0,64
15	L143S10	Lysforurensing_3	0,458	1,258	-1,063	0
16	L143S05	Lysforurensing_4	0,520	1,379	-0,464	0,378
17	L143S06	Lysforurensing_5	0,399	1,018	-1,191	0
18	L143S07	Lysforurensing_6	0,250	0,515*	-0,78	0
19	L143S08	Lysforurensing_7	0,366	0,798*	0,015	0
20	L147S01	Skulearbeid_1 (sidemål)	0,314	0,666*	-0,45	0
21	L147S02	Skulearbeid_2 (sidemål)	0,334	0,848	-1,632	0
22	L147S03	Skulearbeid_3 (sidemål)	0,286	0,592*	0,258	0
23	L147S04	Skulearbeid_4 (sidemål)	0,346	0,744*	0,182	0
24	L147S06	Skulearbeid_5 (sidemål)	0,502	1,363	0,742	0
25	L147S07	Skulearbeid_6 (sidemål)	0,339	0,728*	-0,354	0
26	L146S01	Snikfotografen_1	0,489	1,237	-0,45	-0,338
27	L146S03	Snikfotografen_2	0,484	1,285	-0,774	0
28	L146S04	Snikfotografen_3	0,593	1,816	0,511	0,589
29	L146S05	Snikfotografen_4	0,614	1,88	-0,261	0
30	L146S06	Snikfotografen_5	0,555	1,505	-0,152	0
31	L146S07	Snikfotografen_6	0,532	1,383	0,051	0
32	L144S01	Tasmansk.pungulv_1	0,560	1,707	-0,72	0
33	L144S02	Tasmansk.pungulv_2	0,550	1,479	-0,127	-0,348
34	L144S03	Tasmansk.pungulv_3	0,610	2,119	0,736	-0,302
35	L144S04	Tasmansk.pungulv_4	0,694	2,596	0,125	-0,221
36	L144S06	Tasmansk.pungulv_5	0,243*	0,577*	2,138	0
37	L144S07	Tasmansk.pungulv_6	0,567	1,575	-0,213	0
38	L144S08	Tasmansk.pungulv_7	0,432	1,092	0,931	0
39	L145S01	Den korte historien om egget_1	0,461	1,125	-0,477	0
40	L145S02	Den korte historien om egget_2	0,613	1,844	-0,074	0
41	L145S03	Den korte historien om egget_3	0,644	2,087	0,153	0
42	L145S04	Den korte historien om egget_4	0,458	1,094	0,239	0
43	L145S05	Den korte historien om egget_5	0,444	1,084	0,639	-0,218
44	L145S06	Den korte historien om egget_6	0,394	0,997	1,179	0

I tabell 14 er oppgaver med diskriminering (point biserial) under 0,250 og a-parameter under 0,8 merket med stjerne. En av oppgavene, Tasmansk pungulv_5, er merket for begge parametere og uthevet med grått. I denne oppgaven vil vi ta bort et overflødig ord i det riktige svaret, men ellers beholde svaralternativene som de er, fordi de er tydelig feil.

Den åpne oppgaven Stargate_2 diskriminerer innenfor grenseverdiene, men det var vanskelig å vurdere elevsvarene på en konsistent måte, og vi velger derfor å erstatte denne oppgaven med en helt ny oppgave som krever en ny pilotering. Endringene i denne oppgaven og Stargate_3 er nærmere beskrevet i avsnittet «Ekstrapilot» nedenfor.

Kolonnen til høyre i tabell 14 viser at det er noen oppgaver med DIF (Differential Item Functioning) for kjønn. Vi kommenterer oppgaver med DIF under «Forslag til endelig prøve» nedenfor.

Ekstrapilot

Den skjønnlitterære teksten «Stargate» ble pilotert for tredje og siste gang i februar 2023, som en del av den første piloteringen av tekster og oppgaver til 2024-prøven. Pilotprøvene ble distribuert til et tilfeldig utvalg elever, uten at utvalget er representativt. Oppgavene fra dette som vi har valgt å kalle en «ekstrapilot», vil bli analysert både ved hjelp av klassisk metode og IRT, og resultatet vil inngå i datagrunnlaget for det endelige prøveforslaget.

Endringer før ekstrapiloten

Før ekstrapiloten ble gjennomført, gjorde vi følgende endringer i oppgavene:

- Stargate_2 (L142S10) var en åpen oppgave der elevene skulle nevne to eksempler fra teksten som viser at vaktmesteren bryr seg om jenta. Ved vurdering av elevsvarene fra andre pilotering viste det seg vanskelig å avgrense riktige svar til kun to eksempler. Det ble for mange ulike svar som måtte godkjennes, og dermed ble oppgaven både mindre interessant og for lett, med en p-verdi på 0,60. For å teste om elevene forstår det viktige poenget i teksten, nemlig at vaktmesteren har gode hensikter, ble det laget en ny oppgave (L142S12) bestående av en tabell med fem adjektiv, og elevene ble bedt om å krysse av for om adjektivene «passer» eller «ikke passer» for å beskrive vaktmesteren.
- Stargate_3 (L142S04) ble endret fra flervalg til åpen oppgave. Ordlyden er den samme i begge oppgavene, men den åpne oppgaven ble skåret med 2 og 1 poeng.

Klassisk item-analyse av ekstrapiloten

Tabell 15 viser resultater fra klassisk item-analyse av oppgavene som var med i ekstrapiloten. «Stargate» ble plassert på plass nummer fire i pilotprøve B, og det var 348 elever som gjennomførte prøven.

Tabell 15: Data fra klassisk item-analyse av oppgavene i ekstrapilot. N=348 (175 gutter, 173 jenter). Sum svarfordeling kan avvike fra 100 pga. avrunding. *Lav diskriminering (under 0,250).

Nr.	Item-ID	Ekstrapilot	Svarfordeling i % (frekvens)					Gjennomsnittlig poengsum på prøven per svaralternativ				Diskrim. (biserial)	P-verdi	Kjønns-forskjell (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D				
			0p	1p	2p			0p	1p						
1	L142S09	Stargate_1	36	5	25	32	2	25	24	32	37	0,416	0,32	0,14	
2	L142S12	Stargate_2	56	41			3	23	38			0,549	0,41	0,13	Tabell
3	L142S04	Stargate_3	45	30	17		8	26	35	41	15	0,555	0,32	0,32	Åpen
4	L142S05	Stargate_4	16	68	6	7	3	21	35	20	23	0,597	0,68	0,15	
5	L142S08	Stargate_5	14	11	61	10	4	25	21	35	27	0,483	0,61	0,25	
6	L127S11	Stargate_6	10	13	12	62	4	24	23	25	35	0,496	0,62	0,14	

P-verden for tabell-oppgaven, Stargate_2, er regnet ut ved å gi 1 poeng til elever som har fått alt riktig. Elever som har én eller flere feil, har fått 0 poeng. P-verdien til den åpne, polytome

oppgaven Stargate_3 er regnet ut ved å legge sammen p-verdien for 2 poeng og 50 prosent av p-verdien for 1 poeng.

IRT-analyse av ekstrapiloten

Tabell 16: Parametre fra IRT-analyse av oppgavene til teksten «Stargate». N=348 (175 gutter, 173 jenter).

Nr.	Item-ID	Oppgave	Diskriminering (point biserial)	a-verdi	b	DIF (disfavør)
1	L142S09	Stargate_1	0,468	1,24	0,82	0
2	L142S12	Stargate_2	0,591	1,806	0,341	0
3	L142S04	Stargate_3	0,614	1,261	0,403	0
4	L142S05	Stargate_4	0,608	2,137	-0,541	0
5	L142S08	Stargate_5	0,501	1,348	-0,38	0,78
6	L127S11	Stargate_6	0,519	1,442	-0,41	0

Analysene viser at oppgavejusteringene har vært vellykkete og at oppgavene til teksten nå er ferdigbehandlet.

Endringer etter siste pilotering

I forslaget til ny leseprøve har vi foretatt noen endringer etter ekstrapiloten og den andre piloteringen, på bakgrunn av resultatene fra disse. Endringene er gjort for å sørge for god spredning i oppgavenes vanskegrad, og for 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. I tillegg er det et mål å opprettholde elevenes konsentrasjon og motivasjon gjennom prøven og redusere andelen ubesvarte oppgaver mot slutten. Av den grunn er følgende endringer gjort:

- Rekkefølgen på tekstene er endret slik at den korte reklameteksten «Et skår i gleden» kommer til slutt i prøven, mens den multiple og litt lange teksten «Tasmansk pungulv» kommer først. For øvrig er «Den korte historien om egget» flyttet fra siste til nest siste plass.
- Antall oppgaver er redusert fra 44 til 42 ved at Lysforurensing_4 (L143S05) og Den korte historien om egget_3 (L145S03) er tatt ut.

Andre og mindre språklige endringer i oppgavene kan ses i vedlegg 2 «Oversikt over oppgaver og endringer».

Forslag til endelig prøve 2023

Forslaget til endelig prøve inneholder de samme tekstene som i prøve D. Det er i alt 42 oppgaver, hvorav 32 flervalgsoppgaver, sju åpne oppgaver, én tabell og to rekkefølgeoppgaver. Alle oppgavene er kategorisert etter den leseprosessen som er mest framtreddende, og i antall er det 12 finne-, 17 tolke- og 13 refleksjonsoppgaver. Tabell 17 gir en oversikt over oppgavene med tilhørende p-verdier. Den gjennomsnittlige p-verdien for prøveforslaget er beregnet til 0,53.

Tabell 17: Oppgavene i forslaget til endelig prøve, p-verdi, riktig svar/oppgaveformat og leseprosess.

Nr.	Item-ID	Oppgave	P-verdi	Riktig svar	Leseprosess
1	L144S01	Tasmansk pungulv_1	0,69	A	Tolke
2	L144S02	Tasmansk pungulv_2	0,53	D	Tolke
3	L144S03	Tasmansk pungulv_3	0,28	C-D-B-A	Reflektere
4	L144S04	Tasmansk pungulv_4	0,46	Åpen	Tolke
5	L144S06	Tasmansk pungulv_5	0,24	B	Tolke
6	L144S07	Tasmansk pungulv_6	0,56	B	Tolke
7	L144S08	Tasmansk pungulv_7	0,31	D	Tolke
8	L142S09	Stargate_1	0,32	D	Tolke
9	L142S12	Stargate_2	0,41	A-B-A-B-B	Tolke
10	L142S08	Stargate_3	0,61	C	Reflektere
11	L142S05	Stargate_4	0,68	B	Tolke
12	L142S04	Stargate_5	0,32	Åpen	Tolke
13	L142S11	Stargate_6	0,62	D	Tolke
14	L143S01	Lysforurensing_1	0,72	D	Tolke
15	L143S09	Lysforurensing_2	0,48	B	Finne
16	L143S10	Lysforurensing_3	0,74	A	Tolke
17	L143S06	Lysforurensing_4	0,73	A	Finne
18	L143S07	Lysforurensing_5	0,59	A	Tolke
19	L143S08	Lysforurensing_6	0,50	D	Tolke
20	L147S01	Skularbeid_1 (sidemål)	0,57	A	Reflektere
21	L147S02	Skularbeid_2 (sidemål)	0,77	B	Finne
22	L147S03	Skularbeid_3 (sidemål)	0,46	C	Finne
23	L147S04	Skularbeid_4 (sidemål)	0,47	D	Reflektere
24	L147S06	Skularbeid_5 (sidemål)	0,32	Åpen	Reflektere
25	L147S07	Skularbeid_6 (sidemål)	0,56	C	Finne
26	L146S01	Snikfotografen_1	0,61	A	Tolke
27	L146S03	Snikfotografen_2	0,68	A	Finne
28	L146S04	Snikfotografen_3	0,36	Åpen	Finne
29	L146S05	Snikfotografen_4	0,58	D	Finne
30	L146S06	Snikfotografen_5	0,54	C	Reflektere
31	L146S07	Snikfotografen_6	0,49	B-D-C-A	Reflektere
32	L145S01	Den korte historien om egget_1	0,61	C	Reflektere
33	L145S02	Den korte historien om egget_2	0,52	Åpen	Finne
34	L145S04	Den korte historien om egget_3	0,45	A	Reflektere
35	L145S05	Den korte historien om egget_4	0,45	B	Tolke
36	L145S06	Den korte historien om egget_5	0,36	B	Reflektere
37	L141S01	Et skår i gleden_1	0,61	B	Reflektere
38	L141S02	Et skår i gleden_2	0,70	D	Finne
39	L141S07	Et skår i gleden_3	0,36	C	Reflektere
40	L141S06	Et skår i gleden_4	0,38	Åpen	Reflektere
41	L141S04	Et skår i gleden_5	0,78	A	Finne
42	L141S08	Et skår i gleden_6	0,84	Åpen	Finne
Gjennomsnittlig p-verdi:			0,53		

Interessante leseutfordringer i forslaget til endelig prøve

I det følgende viser vi et utdrag av støtteressursene som ligger ute i PAS – Prøver. Disse ressursene gir et lite innblikk i hvilken informasjon som man kan hente ut av prøvene, hvis man arbeider med resultatene og reflekterer over hva de kan fortelle om interessante og relevante leseutfordringer for elevene.

Lesetekst	Mestringsnivå 1	Mestringsnivå 2	Mestringsnivå 3	Mestringsnivå 4	Mestringsnivå 5
☒ <u>Tasmansk pungulv</u>		1 ✓	2 ✗ 4 ? 6 ✗	3 ✗ 7 ✗	5 ✗
☒ <u>Stargate</u>			9 ✓ 10 ? 11 ✓	8 ✓ 12 ✗	
☒ <u>Lysforurensing</u>		13 ✓ 15 ✓ 16 ✓ 17 ✗	14 ✓ 18 ?		
☒ <u>Skulearbeid</u>	20 ✓		19 ✗ 21 ✓ 22 ? 24 ✗	23 ?	
☒ <u>Snikfotografen</u>		26 ✓	25 ✓ 28 ✗ 29 ✗ 30 ✗	27 ?	
☒ <u>Den korte historien om egget</u>			31 ✓ 32 ✓ 33 ✓	34 ✓ 35 ✗	
Elevfanen: Lucas sine resultater		A ✓ A ✓		A ✗ A ✗	

Nora har fått til mange oppgaver til tekstene «Den korte historien om egget» og «Et skår i gleden?». Dette tyder på at Nora har mestret disse teksttypene. Til «Tasmansk pungulv» derimot har hun bare fått til én oppgave. Hun har også svart feil på flere av oppgavene til «Snikfotografen».

Lærerne trykker på «Tasmansk pungulv» for å se teksten. Når de leser den, forstår de hvorfor teksten har vært vanskelig for mange av elevene i klassen, ikke bare for Nora. De ønsker å se nærmere på oppgavene og går videre til oppgavevisningen.



Elevvisningen gir en oversikt over hvilke oppgaver hver enkelt elev har fått til, og hvilke mestringsnivå de ulike oppgavene er på. Ved å trykke på titlene til venstre, vises tekstene i prøven. Oppgavene vises ved å trykke på oppgavenumrene. Alle elevene har fått en ankertekst. Ankerteksten og oppgavene til denne er ikke synlige.

Oppgaveoversikt for BA

Oppgave %	Tekst %	Gruppens løsningsprosent %	Nasjonal løsningsprosent %	Avvik %	Mestringsnivå %	Formål %	Fag %	Teksttype %
4	Tasmansk pungulv	19%	46%	-27%	3	Tolke	Nat, Saf	Utdrag fra fagbok og nettartikkel
31	Snikfotografen	24%	49%	-25%	3	Reflektere	Khv, Saf	Nettartikkel
23	Skulearbeid	26%	47%	-21%	4	Reflektere	Saf, Mat	Grafikk
30	Snikfotografen	34%	54%	-20%	3	Reflektere	Khv, Saf	Nettartikkel
5	Tasmansk pungulv	4%	24%	-20%	5	Tolke	Nat, Saf	Utdrag fra fagbok og nettartikkel
36	Den korte historien om egget	17%	36%	-19%	4	Reflektere	Mhe, Nat	Utdrag fra fagbok
				-18%	4	Reflektere	Nat, Saf	Utdrag fra fagbok og nettartikkel

Oppgavefanen: Gruppen sin skår på ulike oppgavetyper

Lærerne ser videre på oppgavefanen. De ser hvordan elevene i klassen har prestert sammenliknet med den nasjonale løsningsprosenten. De sorterer etter både positivt og negativt avvik. Oppgaver med stort negativt avvik viser leseutfordringer som mange elever trenger å arbeide videre med. Lærerne bestemmer seg for å undersøke noen av disse oppgavene nærmere. Slik kan de få informasjon om hva som er utfordrende for elevene, og bruke dette til å tilpasse undervisningen i tiden framover.

- Oppgavevisningen gir oversikt over alle oppgaver i prøven. Her kan man se hvor vanskelige oppgavene har vært, og hvor stor andel av elever i klassen som har svart riktig. Avvik mellom elevgruppens resultater og det nasjonale gjennomsnittet vises i en egen kolonne.



Lærerne ser at to oppgaver til «Tasmansk pungulv» og «Snikfotografen» handler om å holde orden på den kronologiske rekkefølgen i en fortelling. Ut fra resultatene kan det virke som mange elever ikke forstår at framstillingen av hendelser i tekster følger en annen rekkefølge enn i virkeligheten.

I teksten om den tasmanske pungulven blir historien om hvordan det siste individet døde i 1936 fortalt aller først. Deretter følger fortellingen om hvorfor arten på 1800-tallet ble jaktet på og gradvis utryddet.

Teksten om snikfotografen Carl Størmer startet også med en interessevekkende og konkret fortelling om hvordan han tok bilder av Henrik Ibsen i 1890-årene før den bakenforliggende historien blir fortalt.

Lærerne bestemmer seg for å arbeide videre med kronologi og tekststruktur.

IRT-analyse av prøvoforslaget

Oppgavene i prøvoforslaget er analysert ved hjelp av programvaren R. Dataene som ligger til grunn for analysen av oppgavene til teksten «Stargate», er hentet fra ekstrapiloten. De andre dataene er hentet fra andre pilot (prøve D). Antall elever (N) er 583 i andre pilot og 348 i ekstrapiloten. Tabell 18 viser parameterne for alle oppgavene.

Tabell 18: Parametre fra IRT-analyse av oppgavene i prøvoforslaget 2023. Gjennomsnittlig p-verdi = 0,53
Cronbach's alpha er 0,906, og er basert på data fra andre pilotering, fordi prøvoforslaget er satt sammen av data fra andre pilot og ekstrapiloten. Lav diskriminering (point biserial < 0,250 og a < 0,8) er merket (*).

Oppgavenr	Item-ID	Oppgave	Diskriminering (point biserial)	a	b	DIF (disfavør)
1	L144S01	Tasmansk pungulv_1	0,570	1,735	-0,717	0
2	L144S02	Tasmansk pungulv_2	0,554	1,463	-0,129	-0,323
3	L144S03	Tasmansk pungulv_3	0,623	2,179	0,73	-0,31
4	L144S04	Tasmansk pungulv_4	0,706	2,64	0,125	-0,196
5	L144S06	Tasmansk pungulv_5	0,240*	0,562*	2,189	0
6	L144S07	Tasmansk pungulv_6	0,577	1,59	-0,213	0
7	L144S08	Tasmansk pungulv_7	0,440	1,104	0,925	0
8	L142S09	Stargate_1	0,484	1,134	0,826	0
9	L142S12	Stargate_2	0,596	1,503	0,332	0
10	L142S08	Stargate_3	0,517	1,173	-0,465	0,851
11	L142S05	Stargate_4	0,814	4,255	-0,491	0
12	L142S04	Stargate_5	0,792	2,028	0,202 / 1,052	0
13	L142S11	Stargate_6	0,705	2,188	-0,374	0
14	L143S01	Lysforurensing_1	0,455	1,195	-0,991	0
15	L143S09	Lysforurensing_2	0,343	0,726*	0,153	-0,614
16	L143S10	Lysforurensing_3	0,470	1,296	-1,046	0
17	L143S06	Lysforurensing_4	0,384	0,954	-1,249	0
18	L143S07	Lysforurensing_5	0,256	0,523*	-0,77	0
19	L143S08	Lysforurensing_6	0,362	0,775	0,015	0
20	L147S01	Skularbeid_1 (sidemål)	0,310	0,648	-0,461	0
21	L147S02	Skularbeid_2 (sidemål)	0,339	0,856	-1,621	0
22	L147S03	Skularbeid_3 (sidemål)	0,294	0,604	0,253	0
23	L147S04	Skularbeid_4 (sidemål)	0,350	0,744*	0,182	0
24	L147S06	Skularbeid_5 (sidemål)	0,507	1,356	0,744	0
25	L147S07	Skularbeid_6 (sidemål)	0,337	0,713*	-0,361	0
26	L146S01	Snikfotografen_1	0,504	1,274	-0,443	0
27	L146S03	Snikfotografen_2	0,491	1,297	-0,772	0
28	L146S04	Snikfotografen_3	0,597	1,791	0,515	0,657
29	L146S05	Snikfotografen_4	0,625	1,907	-0,261	0
30	L146S06	Snikfotografen_5	0,564	1,517	-0,153	0
31	L146S07	Snikfotografen_6	0,552	1,448	0,049	0
32	L145S01	Den korte historien om egget_1	0,472	1,149	-0,472	0
33	L145S02	Den korte historien om egget_2	0,625	1,869	-0,075	0
34	L145S04	Den korte historien om egget_3	0,472	1,123	0,235	0
35	L145S05	Den korte historien om egget_4	0,447	1,077	0,642	-0,213
36	L145S06	Den korte historien om egget_5	0,397	0,992	1,184	0
37	L141S01	Et skår i gleden_1	0,314	0,666*	-0,708	0
38	L141S02	Et skår i gleden_2	0,356	0,832	-1,166	0
39	L141S07	Et skår i gleden_3	0,367	0,823	0,79	0
40	L141S06	Et skår i gleden_4	0,323	0,697*	0,808	0
41	L141S04	Et skår i gleden_5	0,423	1,18	-1,353	0,487
42	L141S08	Et skår i gleden_6	0,471	1,658	-1,451	0,615

I tabell 18 er oppgaver med diskriminering (point biserial) under 0,250 og a-verdi under 0,8 merket med stjerne. Én oppgave diskriminerer under begge disse målene og er uthevet med grått. De oppgavene som er merket for begge parametre, er uthevet i grått. Oppgaver med

lave diskrimineringsverdier og oppgaver som har en DIF (Differential Item Functioning) for kjønn på mer enn 0,5 ESSD, vil bli kommentert nedenfor.

Oppgaver med lav diskriminering og DIF

Etter at vi har fjernet noen oppgaver etter andre pilotering og kjørt nye analyser på prøveforslaget, er det én oppgave som har for lav diskriminering på både klassisk mål og på a-verdien. Tasmansk pungulv_5 (L144S06) hadde lave diskrimineringsverdier både i første og andre pilotering, men det er blitt gjort endringer for å forsøke å bedre oppgavens diskrimineringsgrad. Også etter andre pilotering er det gjort endringer i oppgaveformuleringen som vi håper kan bedre verdiene. Dette er den eneste oppgaven på det høyeste mestringsnivået, så vi ønsker ikke å ta oppgaven ut av prøven, men heller reparere den.

Kolonnen til høyre i tabell 18 viser at det er noen oppgaver med DIF (Differential Item Functioning) for kjønn. Dette betyr at det er noe ved oppgaven som favoriserer det ene kjønn, gitt at begge har samme ferdighet. DIF skiller seg fra måling av kjønnsforskjell, som viser hvilket kjønn som presterer best totalt sett. DIF mellom 0,5 og 0,8 regnes som en moderat DIF, mens oppgaver som ligger over 0,8 har en stor DIF. Oppgaver som slår ut på en DIF-analyse med en effektstørrelse som er lavere enn 0,5, tar vi ikke hensyn til ettersom det er knyttet noe usikkerhet til DIF-analysene, og dette regnes som svært små utslag.

Det er ni oppgaver i prøveforslaget som slår ut på DIF-analysen. Av disse har fire oppgaver moderat utslag og én oppgave stort utslag. Oppgaven med stor DIF er knyttet til den skjønnlitterære teksten «Stargate», og er i jentenes favør. Som nevnt under analysene fra den første piloteringen, vet vi av erfaring at jenter ofte presterer bedre på skjønnlitterære tekster, og at denne effekten forsterkes når det er en jente i hovedrollen, slik det er i denne teksten. Antall oppgaver med DIF til denne teksten er redusert fra fem i første pilot til én i prøveforslaget, noe som tyder på at vi har klart å redusere utslagene av DIF. Antakelig vil denne teksten til en viss grad favorisere jenter, men vi mener at dette kan forsvares med tekstens kvaliteter og at vi får målt lesing av skjønnlitteratur på en god måte. Oppgaven med størst DIF handler dessuten om å forstå hvorfor noe i teksten er skrevet med kursivert skrift, altså et spørsmål om form og ikke selve handlingen i teksten.

De øvrige oppgavene med DIF har mellomstor DIF. Oppgaven Lysforurensing_2 har DIF i guttenes favør, men oppgavene Et skår i gleden_6 og Snikfotografen_3 har DIF i jentenes favør. Ettersom alle er knyttet til ulike tekster og i noen grad balanserer hverandre ut, avventer vi å ta dem ut av prøven. Det gjøres også endringer i flere av oppgavene før de havner i endelig prøve.

Oppgavene fordelt på mestringsnivåer

Tabell 19 viser oppgavene i forslag til endelig prøve, fordelt på mestringsnivåer. Nivå-grensene som er brukt, ble bestemt etter gjennomføringen av nasjonale leseprøver i 2022. 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. Det er derfor en viss usikkerhet knyttet til nivågrensene.

Tabell 19: Oppgavene i forslaget til endelig prøve, fordelt på mestringsnivåer

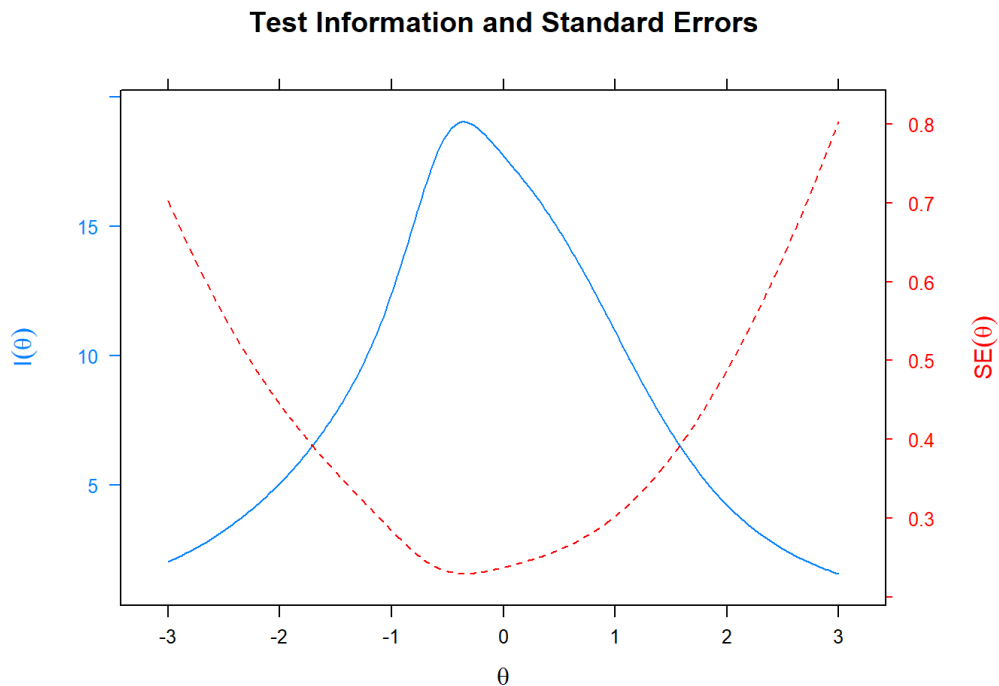
Nivå	Item-ID	Oppgave	Skalapoeng
Nivå 1	L147S02	Skulearbeid_2	33,79
	L141S08	Et skår i gleden_6	35,49
	L141S04	Et skår i gleden_5	36,47
Nivå 2	L143S06	Lysforurensing_4	37,51
	L141S02	Et skår i gleden_2	38,34
	L143S10	Lysforurensing_3	39,54
	L143S01	Lysforurensing_1	40,09
	L146S03	Snikfotografen_2	42,28
	L143S07	Lysforurensing_5	42,3
	L144S01	Tasmansk pungulv_1	42,83
	L141S01	Et skår i gleden_1	42,92
Nivå 3	L142S05	Stargate_4	45,09
	L145S01	Den korte historien om egget_1	45,28
	L142S08	Stargate_3	45,35
	L147S01	Skulearbeid_1	45,39
	L146S01	Snikfotografen_1	45,57
	L142S11	Stargate_6	46,26
	L147S07	Skulearbeid_6	46,39
	L146S05	Snikfotografen_4	47,39
	L144S07	Tasmansk pungulv_6	47,87
	L146S06	Snikfotografen_5	48,47
	L144S02	Tasmansk pungulv_2	48,71
	L145S02	Den korte historien om egget_2	49,25
	L143S08	Lysforurensing_6	50,15
	L146S07	Snikfotografen_6	50,49
	L144S04	Tasmansk pungulv_4	51,25
	L143S09	Lysforurensing_2	51,53
	L147S04	Skulearbeid_4	51,82
	L142S04	Stargate_5_1 p.	52,02
	L145S04	Den korte historien om egget_3	52,35
	L147S03	Skulearbeid_3	52,53
	L142S12	Stargate_2	53,32
Nivå 4	L146S04	Snikfotografen_3	55,15
	L145S05	Den korte historien om egget_4	56,42
	L144S03	Tasmansk pungulv_3	57,3
	L147S06	Skulearbeid_5	57,44
	L141S07	Et skår i gleden_3	57,9
	L141S06	Et skår i gleden_4	58,08
	L142S09	Stargate_1	58,26
	L144S08	Tasmansk pungulv_7	59,25
	L142S04	Stargate_5_2 p.	60,52
	L145S06	Den korte historien om egget_5	61,84
Nivå 5	L144S06	Tasmansk pungulv_5	71,89

Fordelingen av oppgaver på mestringsnivå viser det samme mønsteret som i tidligere prøver, der det er flest oppgaver på det midterste nivået. På de to nederste nivåene, nivå 1 og 2, er det til sammen elleve oppgaver. Til sammen er det også elleve oppgaver på de to øverste nivåene, men ti av disse er på nivå 4 og bare én er på nivå 5.

Testinformasjon og måleusikkerhet

IRT-analysen av prøveforslaget er basert på resultater fra andre pilotering og ekstrapiloten, men tre oppgaver er tatt ut. Det er også foretatt noen språklige justeringer i enkelte oppgaver som ligger til grunn for analysen av prøveforslaget. Med andre ord er dette kun et estimat, og vi regner med at IRT-analysen av prøven vil endre seg noe etter gjennomføringen.

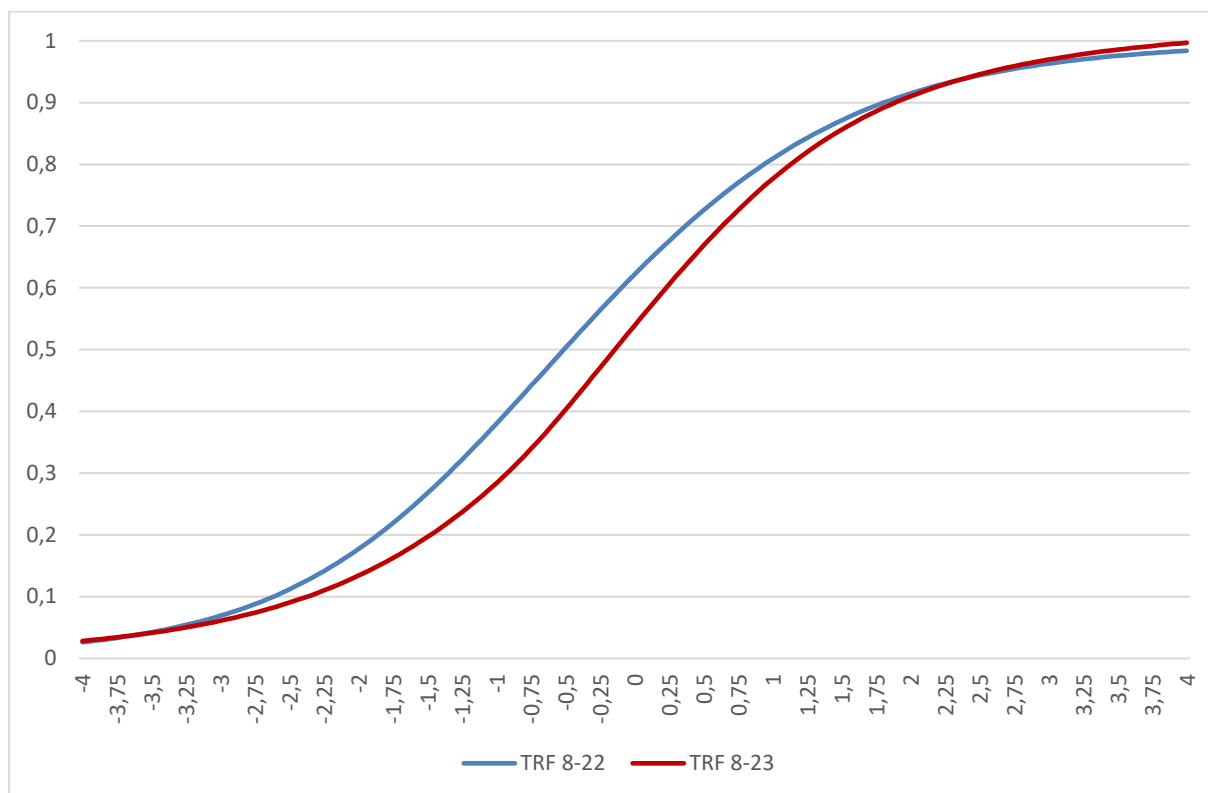
Figur 10 viser at testinformasjonen (TIF) for prøveforslaget er høyest for theta lik - 0,5 og at måleusikkerheten (CSEM) er minst i det samme ferdighetsområdet i prøven.



Figur 10: Informasjonsverdi og måleusikkerhet for prøveforslaget 2023.

Prøvens TRF (Test respons function)

Figur 11 viser testrespons-funksjonen (TRF) for prøveforslaget (rød linje), sammenlignet med 2022-prøven (blå linje). Ettersom 2022-prøven og prøveforslaget er analysert med ulike analyseverktøy, er ikke variasjonen i distribusjonen direkte sammenlignbar.



Figur 11: TRF for prøveforslaget 2023 (rød linje), sammenlignet med prøven 2022 (blå linje)

TRF er en beskrivelse av sammenhengen mellom estimert ferdighet (θ) og andel riktig besvarte oppgaver i prøven. Figur 11 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 noe forskjell mellom prøvene. Elever med ferdighet mellom -3 og 1,5 må løse færre oppgaver i prøveforslaget for å vise samme ferdighet, sammenliknet med 2022-prøven. Med andre ord forkorter 2023-prøveforslaget målingen i dette ferdighetsområdet sammenliknet med 2022-prøven. I de øvrige delene av skalaen er målingen tilnærmet lik.

Der er viktig å understreke at sammenligningen som her er gjort, er forbundet med noe usikkerhet, fordi 2022-prøven inneholder et sett av ankeroppgaver som ikke er inkludert i prøveforslaget 2023. I tillegg gjenspeiler ikke prøveforslaget de endringene som er gjort etter andre og tredje pilotering.

Gjennomføring

Tekstene i den nasjonale prøven i lesing for 8. trinn 2023 er presentert i tabell 1 på side 12 i denne rapporten. Prøven ble gjennomført i tidsrommet 4. september – 29. september 2023, i prøvegjennomføringssystemet PAS – Prøver. Totalt 61 980 elever på 8. trinn og 62 007 elever på 9. trinn gjennomførte leseprøven. Hver elev fikk en prøve med sju tekster.

På 8. trinn blir endringer i elevenes leseferdigheter målt over tid ved hjelp av ankertekster, det vil si tekster som blir brukt om igjen fra år til år. Det er i alt sju ankertekster, og i 2022 ble disse for første gang distribuert slik at alle elevene fikk en prøve bestående av én ankertekst og seks nye tekster. Med dette designet ble hver ankertekst distribuert til ca. 8 000 elever, mens hver av de nye tekstene ble distribuert til ca. 53 000 elever. Dermed har ingen elever gjennomført årets prøve i sin helhet.

Ankertekstene gjør det mulig å måle endring i elevers leseferdigheter over tid. Overgangen til nytt prøvegjennomføringssystem har ført til at målingen av utvikling over tid, trendmålingen, startet på nytt i 2022. Grensene for mestringsnivåene ble satt på nytt, med et nasjonalt gjennomsnitt på 50 skalapoeng. Resultatene fra gjennomføringen i 2023 er derfor ikke direkte sammenlignbare med resultater fra tidligere gjennomføringer.

Nedenfor presenterer vi resultatene fra de sju tekstene og 42 oppgavene som utgjorde den nye delen av prøven. Siden ankertekstene skal brukes om igjen, blir ikke resultatene fra disse omtalt i denne rapporten. For de andre tekstene blir det rapportert gjennomsnittlige p-verdier, det vil si prosentandel elever som fikk poeng på hver oppgave, på hvert årstrinn. P-verdiene blir oppgitt både samlet for alle tekstene og for hver enkelt oppgave. Der det er relevant, blir samsvar og ulikheter mellom årets prøve og tidligere prøver kommentert. Resultatene for 8. og 9. trinn blir sammenlignet enkelte steder, men fordi prøven primært er utviklet for elever på 8. trinn, og mønsteret er omtrent det samme når det gjelder elevenes relative styrker og svakheter på de to trinnene, presenteres som regel bare resultater for 8. trinn.

Tabell 20 under viser fordelingen av tekster fra årets prøve og ankertekster.

Tabell 20: Oversikt over fordelingen av tekster og prøveversjoner i årets gjennomføring

	Tekst 1	Tekst 2	Tekst 3	Tekst 4	Tekst 5	Tekst 6	Tekst 7
	N = 53 135	N = 44 255	N = 53 120	N = 61 980	N = 53 120	N = 53 121	N = 53 149
Prøve 1	Tasmansk pungulv	Stargate	Lysforurensing	Skulearbeid	Snikfotografen	Den korte historien om egget	Ankertekst 1
Prøve 2	Tasmansk pungulv	Stargate	Lysforurensing	Skulearbeid	Snikfotografen	Ankertekst 2	Et skår i gleden
Prøve 3	Tasmansk pungulv	Stargate	Lysforurensing	Skulearbeid	Ankertekst 3	Den korte historien om egget	Et skår i gleden
Prøve 4	Tasmansk pungulv	Stargate	Ankertekst 4	Skulearbeid	Snikfotografen	Den korte historien om egget	Et skår i gleden
Prøve 5	Tasmansk pungulv	Ankertekst 5	Lysforurensing	Skulearbeid	Snikfotografen	Den korte historien om egget	Et skår i gleden
Prøve 6	Ankertekst 6	Stargate	Lysforurensing	Skulearbeid	Snikfotografen	Den korte historien om egget	Et skår i gleden
Prøve 7	Tasmansk pungulv	Ankertekst 7 (pilot)	Lysforurensing	Skulearbeid	Snikfotografen	Den korte historien om egget	Et skår i gleden

Resultater for prøven som helhet

Tabell 21 inneholder en oversikt over alle oppgavene med beskrivelser av oppgaveformat og leseprosess, gjennomsnittlige resultater i form av p-verdier for alle elever på henholdsvis 8. og 9. trinn, samt forskjellen mellom de to trinnene. Tabellen viser også samlet p-verdi for hver av tekstene på 8. og 9. trinn, samt den gjennomsnittlige p-verdien for alle tekstene.

Tabell 21: Oppgaveformat, leseprosess og gjennomsnittlig p-verdi for hver oppgave på henholdsvis 8. og 9. trinn. Summene er avrundet til to desimaler i Excel.

Oppg.	Tekst/oppgave	Oppgaveformat	Prosess	8.trinn	9.trinn	Differanse (9-8)
1	Tasmansk pungulv_1	Flervalg	Tolke	0,81	0,86	0,05
2	Tasmansk pungulv_2	Flervalg	Tolke	0,69	0,75	0,06
3	Tasmansk pungulv_3	Rekkefølge	Reflektere	0,22	0,33	0,11
4	Tasmansk pungulv_4	Åpen	Tolke	0,56	0,67	0,11
5	Tasmansk pungulv_5	Flervalg	Tolke	0,24	0,29	0,05
6	Tasmansk pungulv_6	Flervalg	Tolke	0,71	0,78	0,07
7	Tasmansk pungulv_7	Flervalg	Tolke	0,38	0,48	0,10
Gjennomsnitt for Tasmansk pungulv				0,52	0,59	0,07
8	Stargate_1	Flervalg	Tolke	0,36	0,40	0,04
9	Stargate_2	Tabell	Tolke	0,48	0,58	0,10
10	Stargate_3	Flervalg	Reflektere	0,70	0,76	0,06
11	Stargate_4	Flervalg	Tolke	0,75	0,82	0,07
12	Stargate_5	Åpen	Tolke	0,40	0,48	0,08
13	Stargate_6	Flervalg	Tolke	0,70	0,77	0,07
Gjennomsnitt for Stargate				0,57	0,64	0,07
14	Lysforurensing_1	Flervalg	Tolke	0,70	0,78	0,08
15	Lysforurensing_2	Flervalg	Finne	0,50	0,57	0,07
16	Lysforurensing_3	Flervalg	Tolke	0,69	0,77	0,08
17	Lysforurensing_4	Flervalg	Finne	0,61	0,68	0,07
18	Lysforurensing_5	Flervalg	Tolke	0,70	0,78	0,08
19	Lysforurensing_6	Flervalg	Tolke	0,45	0,53	0,08
Gjennomsnitt for Lysforurensing				0,61	0,69	0,08
20	Skulearbeid_1	Flervalg	Reflektere	0,75	0,81	0,06
21	Skulearbeid_2	Flervalg	Finne	0,77	0,83	0,06
22	Skulearbeid_3	Flervalg	Finne	0,63	0,70	0,07
23	Skulearbeid_4	Flervalg	Reflektere	0,43	0,50	0,07
24	Skulearbeid_5	Åpen	Reflektere	0,30	0,45	0,15
25	Skulearbeid_6	Flervalg	Finne	0,55	0,62	0,07
Gjennomsnitt for Skulearbeid				0,57	0,65	0,08
26	Snikfoto grafen_1	Flervalg	Tolke	0,61	0,69	0,08
27	Snikfoto grafen_2	Flervalg	Finne	0,63	0,72	0,09
28	Snikfoto grafen_3	Åpen	Finne	0,25	0,38	0,13
29	Snikfoto grafen_4	Flervalg	Finne	0,56	0,68	0,12
30	Snikfoto grafen_5	Flervalg	Reflektere	0,53	0,64	0,11
31	Snikfoto grafen_6	Rekkefølge	Reflektere	0,40	0,50	0,10
Gjennomsnitt for Snikfoto grafen				0,50	0,60	0,07
32	Den korte historien om egget_1	Flervalg	Reflektere	0,66	0,73	0,07
33	Den korte historien om egget_2	Åpen	Finne	0,64	0,74	0,10
34	Den korte historien om egget_3	Flervalg	Reflektere	0,51	0,58	0,07
35	Den korte historien om egget_4	Flervalg	Tolke	0,42	0,51	0,09
36	Den korte historien om egget_5	Flervalg	Reflektere	0,41	0,46	0,05
Gjennomsnitt for Den korte historien om egget				0,53	0,60	0,07
37	Et skår i gleden_1	Flervalg	Reflektere	0,52	0,62	0,10
38	Et skår i gleden_2	Flervalg	Finne	0,54	0,67	0,13
39	Et skår i gleden_3	Flervalg	Reflektere	0,68	0,77	0,09
40	Et skår i gleden_4	Åpen	Reflektere	0,41	0,50	0,09
41	Et skår i gleden_5	Flervalg	Finne	0,30	0,39	0,09
42	Et skår i gleden_6	Åpen	Finne	0,68	0,78	0,10
Gjennomsnitt for Et skår i gleden?				0,52	0,62	0,10
Gjennomsnitt for hele prøven				0,54	0,63	0,08

Den gjennomsnittlige p-verdien for hele leseprøven er 0,54 på 8. trinn og 0,63 på 9. trinn. Det vil si at en gjennomsnittselev på 8. trinn klarer 54 prosent av oppgavene, mens en gjennomsnittselev på 9. trinn klarer 63 prosent av oppgavene. 196 av de 62 007 elevene på 9. trinn oppnådde maksimal poengsum på 43. Det er ikke mulig å finne det tilsvarende antallet på 8. trinn, fordi alle elevene fikk en ankertekst i prøven, noe som resulterte i sju ulike prøver med ulik maksimal poengsum.

Den gjennomsnittlige forskjellen i p-verdi mellom 8. og 9. trinn er 0,08. På tekstinivå ser vi at den største forskjellen er på 0,10 for reklamen «Et skår i gleden», som ligger til sist i prøven.

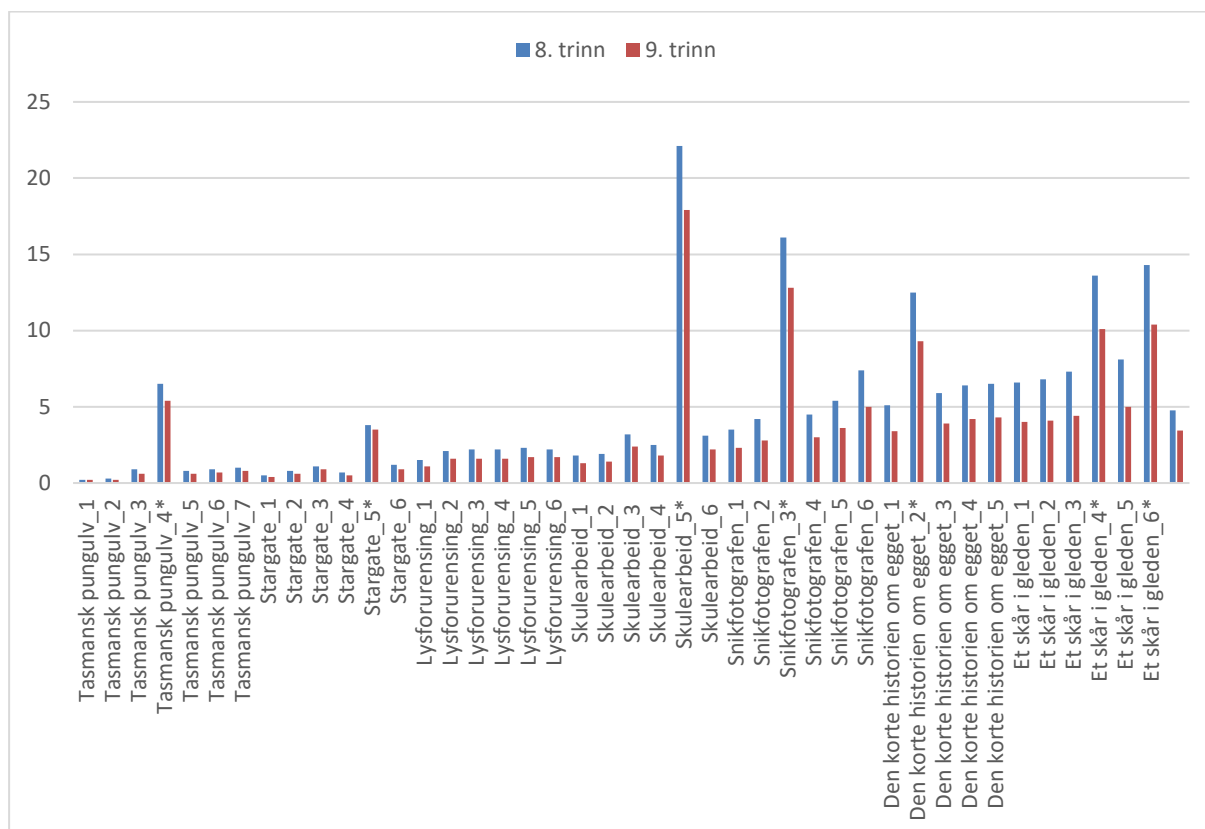
På oppgavenivå er det den åpne refleksjonsoppgaven Skulearbeid_5 som viser størst forskjell i prestasjoner mellom elevene på 8. og 9. trinn, med en forskjellen på 0,15. Der nest følger den åpne oppgaven Snikfotografen_3 og flervalgsoppgaven Et skår i gleden_2, hver med en forskjell på 0,13.

På 8. trinn er rekkefølge-oppgaven Tasmansk pungulv_3 den vanskeligste i prøven, med en p-verdi på 0,22. Oppgaven krever at elevene plasserer fire hendelser som er omtalt i teksten, i riktig rekkefølge. Dette har vist seg å være vanskelig for elevene, og i 2022 var det også en rekkefølge-oppgave som kom ut med den laveste p-verdien i prøven.

For øvrig har teksten Tasmansk pungulv både den vanskeligste og den letteste oppgaven i prøven. På 9. trinn er Tasmansk pungulv_5 den vanskeligste, mens den første oppgaven i prøven, Tasmansk pungulv_1, er lettest på begge trinn, med en p-verdi på henholdsvis 0,81 og 0,86.

Hvis vi sammenligner resultatet fra gjennomføringen med resultatet etter siste pilotering, ser vi at den gjennomsnittlige p-verdien på 8. trinn gikk opp fra 0,53 til 0,54. Dette er kun en liten, men likevel forventet endring, ettersom elevene sannsynligvis legger ned mer innsats under prøvegjennomføring enn under pilotering. Vi ser det samme når det gjelder andelen blanke svar, som i andre pilotering (prøve D) var på 7,8 prosent i gjennomsnitt, mens den etter gjennomføring av prøven er på 4,8 prosent på 8. trinn og 3,4 prosent på 9. trinn.

Figur 12 nedenfor viser andelen blanke svar for hver av oppgavene i prøvesettet på 8. og 9. trinn:



Figur 12: Prosentandel elever på 8. og 9. trinn som ikke har svart på hver enkelt oppgave i prøven. Åpne oppgaver er merket med stjerne (*).

I figuren ser vi at de aller fleste søylene som stikker seg ut med en høyere prosentandel enn gjennomsnittet, representerer åpne oppgaver. Vi ser også at andelen blanke svar øker mot slutten, og at dette gjelder i større grad for 8. trinn enn for 9. trinn.

Kjønnsforskjeller

Tabell 22 viser de gjennomsnittlige p-verdiene for jenter og gutter på de to trinnene, samt kjønnsforskjellene. Effektstørrelsene er beregnet som prosentandeler av standardavvikene.

Tabell 22: Kjønnsforskjeller i lesekompetanse, uttrykt som p-verdier og effektstørrelser

	8. trinn	9. trinn
Jenter	0,57	0,66
Gutter	0,52	0,60
Kjønnsforskjell (j-g)	0,05	0,06
Effektstørrelse i form av standardavvik	0,25	0,28

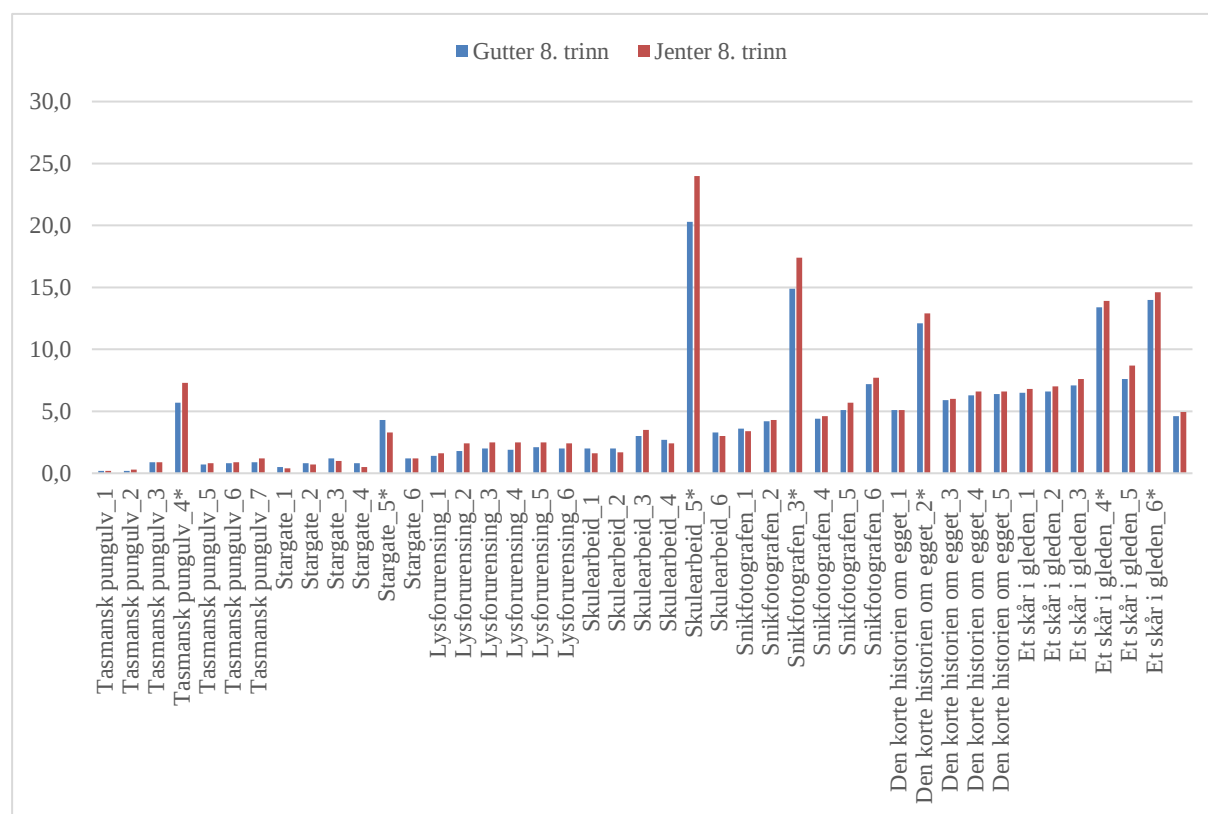
I likhet med tidligere gjennomføringer viser den nasjonale prøven i lesing for 2023 at jentene presterer bedre enn guttene. For årets prøve er den gjennomsnittlige kjønnsforskjellen i

jentenes favør 0,05 på 8. trinn og 0,06 på 9. trinn. Dette gir en effektstørrelse på henholdsvis 0,25 og 0,28, som er noe høyere enn tidligere år.

Når vi ser på oppgavene knyttet til hver enkelt tekst under ett på 8. trinn, er prestasjonsforskjellene mellom jenter og gutter størst for den skjønnlitterære teksten «Stargate». Kjønnsforskjellen går i jentenes favør og er på 0,11. De tre tekstene som viser minst prestasjonsforskjell mellom kjønnene, er «Den korte historien om egget», «Lysforurensing» og «Snikfotografen». For alle disse tekstene er kjønnsforskjellene i jentenes favør med henholdsvis 0,02; 0,04 og 0,01 i p-verdi. På 9. trinn er det «Stargate» som har størst kjønnsforskjell, mens Tasmansk pungulv har minst.

Andelen blanke svar er i gjennomsnitt 4,9 for jenter og 4,6 prosent for gutter på 8. trinn. Som ved tidligere gjennomføringer er andelen høyere for de åpne oppgavene enn for flervalgsoppgavene, og den er økende utover i prøvesettet.

Mønsteret for blanke svar fordelt på kjønn på 9. trinn er i all hovedsak det samme som for elevene på 8. trinn, med en gjennomsnittlig andel blanke svar på henholdsvis 3,3 prosent for gutter og 3,6 prosent for jenter. Figur 13 viser prosentandelen blanke svar for gutter og jenter på 8. trinn.



Figur 13: Prosentandel gutter og jenter på 8. trinn som ikke har svart på hver oppgave. Åpne oppgaver er merket med stjerne (*).

Klassisk item-analyse av prøven 2023

Tabell 23 og 24 viser resultater fra klassisk item-analyse av prøven på henholdsvis 8. og 9. trinn 2023. På grunn av det nye prøveoppsettet finnes det ikke lenger en N for årets prøve som helhet på 8. trinn, men vi vet at 61 980 elever gjennomførte leseprøven. Det er heller ikke mulig å gjøre reliabilitetsanalyse i SPSS. Se derfor diskrimineringsverdiene fra IRT-analysene i kapittelet «Oppgaver med lav diskriminering» nedenfor.

Tabell 23: Resultater fra klassisk item-analyse, endelig prøve, 8. trinn.

Gjennomsnittlig p-verdi: 0,54 (gutter: 0,52 / jenter: 0,57). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Nr	Item-ID	Oppgaver	Svarfordeling i % (frekvens)					Gj.snittlig poengsum på prøven per svaralt.				p-verdi	Kjønns-forskj. (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D			
			0p	1p	2p									
1	L144S01	Tasmansk pungulv_1	81	11	3	4	0	25	17	16	17	0,81	0,04	
2	L144S02	Tasmansk pungulv_2	9	17	5	69	0	17	19	16	26	0,69	-0,01	
3	L144S03	Tasmansk pungulv_3	77	22			1	21	30			0,22	-0,04	Rekkefølge
4	L144S04	Tasmansk pungulv_4	38	56			7	18	28			0,56	0,04	Åpen
5	L144S06	Tasmansk pungulv_5	37	24	5	33	1	23	26	14	24	0,24	0,05	
6	L144S07	Tasmansk pungulv_6	17	71	6	6	1	18	26	15	17	0,71	0,06	
7	L144S08	Tasmansk pungulv_7	21	21	19	38	1	21	20	22	27	0,38	-0,08	
8	L142S09	Stargate_1	25	7	32	36	1	18	18	24	28	0,36	0,11	
9	L142S12	Stargate_2	51	48			1	20	27			0,48	0,03	Tabell
10	L142S08	Stargate_3	13	8	70	8	1	18	16	26	19	0,70	0,12	
11	L142S05	Stargate_4	9	75	4	11	1	17	26	13	16	0,75	0,12	
12	L142S04	Stargate_5	38	36	22		4	20	24	30		0,40	0,13	Åpen
13	L142S11	Stargate_6	10	7	11	70	1	17	16	20	26	0,70	0,12	
14	L143S01	Lysforurensing_1	13	10	5	70	2	19	14	17	26	0,70	0,05	
15	L143S09	Lysforurensing_2	22	50	15	11	2	24	26	17	20	0,50	-0,01	
16	L143S10	Lysforurensing_3	7	11	69	11	2	17	16	26	18	0,69	0,03	
17	L143S06	Lysforurensing_4	8	22	61	7	2	17	20	26	15	0,61	0,09	
18	L143S07	Lysforurensing_5	71	12	11	5	2	26	17	17	18	0,70	0,04	
19	L143S08	Lysforurensing_6	13	21	19	45	2	18	22	20	27	0,45	0,03	
20	L147S01	Skulearbeid_1	75	14	3	5	2	25	19	14	17	0,75	0,09	
21	L147S02	Skulearbeid_2	3	77	16	2	2	14	25	19	14	0,77	0,03	
22	L147S03	Skulearbeid_3	12	64	16	6	3	20	25	20	20	0,63	0,05	
23	L147S04	Skulearbeid_4	26	8	20	43	3	20	17	23	27	0,43	0,07	
24	L147S06	Skulearbeid_5	48	30			22	21	30			0,30	0,09	Åpen
25	L147S07	Skulearbeid_6	14	18	56	10	3	20	18	27	22	0,55	0,09	
26	L146S01	Snikfotografen_1	61	17	10	8	4	27	18	20	18	0,61	0,03	
27	L146S03	Snikfotografen_2	63	8	17	8	4	26	18	19	20	0,63	0,06	
28	L146S04	Snikfotografen_3	59	25			16	22	31			0,25	0,08	Åpen
29	L146S05	Snikfotografen_4	27	7	5	56	5	19	14	18	28	0,56	0,07	
30	L146S06	Snikfotografen_5	15	10	53	17	5	18	17	27	21	0,53	0,04	
31	L146S07	Snikfotografen_6	52	40			7	20	29			0,40	0,02	Rekkefølge
32	L145S01	Den korte historien om egget_1	8	6	66	14	5	20	17	26	20	0,66	0,07	
33	L145S02	Den korte historien om egget_2	24	64			13	18	27			0,64	0,08	Åpen
34	L145S04	Den korte historien om egget_3	51	7	8	28	6	26	16	19	23	0,51	-0,01	
35	L145S05	Den korte historien om egget_4	10	42	35	6	6	17	28	22	19	0,42	0,00	
36	L145S06	Den korte historien om egget_5	15	42	21	15	7	19	26	25	22	0,41	-0,02	
37	L141S01	Et skår i gleden_1	8	52	17	16	7	16	27	22	22	0,52	0,05	
38	L141S02	Et skår i gleden_2	12	19	8	54	7	18	21	17	28	0,54	0,07	
39	L141S07	Et skår i gleden_3	12	8	68	6	7	17	17	27	19	0,68	0,08	
40	L141S06	Et skår i gleden_4	46	41			14	21	29			0,41	0,09	Åpen
41	L141S04	Et skår i gleden_5	30	30	18	14	8	28	23	20	24	0,30	0,04	
42	L141S08	Et skår i gleden_6	17	69			14	17	27			0,68	0,09	Åpen

Tabell 24: Resultater fra klassisk item-analyse, endelig prøve, 9. trinn.

Antall elever (N): 62 007 (31 618 gutter, 30 389 jenter). Reliabilitet (Cronbachs Alpha): 0,909.

Gjennomsnittlig p-verdi: 0,63 (gutter: 0,60 / jenter: 0,66). Sum svarfordeling kan avvike fra 100 pga. avrunding.

Nr	Item-ID	Oppgaver	Svarfordeling i % (frekvens)					Gj.snittlig poengsum på prøven per svaralt.				Diskrim. (biserial)	P-verdi	Kjønnsforskj. (j-g)	Merk
			A	B	C	D	Blank	A	B	C	D				
			0p	1p	2p										
1	L144S01	Tasmansk pungulv_1	86	8	3	3	0	28	20	18	19	0,306	0,86	0,03	
2	L144S02	Tasmansk pungulv_2	8	14	3	75	0	20	21	19	29	0,387	0,75	-0,01	
3	L144S03	Tasmansk pungulv_3	67	33			1	24	33			0,436	0,33	-0,04	Rekkefølge
4	L144S04	Tasmansk pungulv_4	28	67			1	21	30			0,512	0,67	0,04	Åpen
5	L144S06	Tasmansk pungulv_5	34	29	4	33	1	26	30	15	27	0,185	0,29	0,07	
6	L144S07	Tasmansk pungulv_6	13	78	4	4	1	20	29	16	18	0,435	0,78	0,05	
7	L144S08	Tasmansk pungulv_7	17	18	17	48	1	24	22	25	31	0,369	0,48	-0,09	
8	L142S09	Stargate_1	22	6	32	40	0	21	20	27	31	0,332	0,40	0,12	
9	L142S12	Stargate_2	42	58			1	23	30			0,383	0,58	0,03	Tabell
10	L142S08	Stargate_3	9	6	76	7	1	20	19	29	20	0,425	0,76	0,12	
11	L142S05	Stargate_4	7	82	3	8	1	18	29	14	18	0,479	0,82	0,11	
12	L142S04	Stargate_5	32	33	32		4	23	27	33		0,438	0,48	0,17	Åpen
13	L142S11	Stargate_6	8	6	8	77	1	19	19	22	29	0,403	0,77	0,11	
14	L143S01	Lysforurensing_1	10	6	4	78	1	21	15	19	29	0,466	0,78	0,05	
15	L143S09	Lysforurensing_2	22	58	10	9	2	26	30	19	22	0,303	0,57	0,00	
16	L143S10	Lysforurensing_3	5	8	77	8	2	18	17	29	21	0,478	0,77	0,02	
17	L143S06	Lysforurensing_4	7	19	69	5	2	20	23	30	16	0,430	0,68	0,09	
18	L143S07	Lysforurensing_5	78	9	8	4	2	29	18	19	20	0,469	0,78	0,05	
19	L143S08	Lysforurensing_6	9	20	17	53	2	20	25	23	31	0,406	0,53	0,05	
20	L147S01	Skulearbeid_1	82	11	2	3	1	29	21	14	18	0,394	0,81	0,08	
21	L147S02	Skulearbeid_2	2	83	12	2	1	16	29	21	15	0,395	0,83	0,03	
22	L147S03	Skulearbeid_3	10	70	13	4	2	23	29	23	22	0,317	0,70	0,05	
23	L147S04	Skulearbeid_4	20	7	21	50	2	22	19	26	31	0,361	0,50	0,06	
24	L147S06	Skulearbeid_5	37	45			18	23	33			0,553	0,45	0,13	Åpen
25	L147S07	Skulearbeid_6	12	12	62	11	2	23	19	30	25	0,382	0,62	0,10	
26	L146S01	Snikfotografen_1	69	14	9	6	2	30	21	22	20	0,457	0,69	0,03	
27	L146S03	Snikfotografen_2	72	6	13	7	3	30	19	20	23	0,438	0,72	0,07	
28	L146S04	Snikfotografen_3	49	38			13	24	33			0,523	0,38	0,09	Åpen
29	L146S05	Snikfotografen_4	20	5	5	68	3	20	15	20	31	0,586	0,68	0,07	
30	L146S06	Snikfotografen_5	11	7	64	14	4	21	19	30	23	0,487	0,64	0,05	
31	L146S07	Snikfotografen_6	45	50			5	23	32			0,530	0,50	0,01	Rekkefølge
32	L145S01	Den korte historien om egget_1	8	5	73	11	3	23	19	29	22	0,400	0,73	0,07	
33	L145S02	Den korte historien om egget_2	17	74			9	20	30			0,538	0,74	0,08	Åpen
34	L145S04	Den korte historien om egget_3	58	6	7	25	4	30	18	23	26	0,344	0,58	0,00	
35	L145S05	Den korte historien om egget_4	7	51	32	6	4	20	31	25	21	0,386	0,51	0,02	
36	L145S06	Den korte historien om egget_5	12	46	22	16	4	22	29	28	26	0,186	0,46	-0,02	
37	L141S01	Et skår i gleden_1	5	62	16	13	4	17	30	24	24	0,395	0,62	0,07	
38	L141S02	Et skår i gleden_2	8	15	6	67	4	19	23	18	30	0,515	0,67	0,09	
39	L141S07	Et skår i gleden_3	8	6	77	5	4	18	19	30	21	0,514	0,77	0,08	
40	L141S06	Et skår i gleden_4	40	50			10	24	32			0,479	0,50	0,10	Åpen
41	L141S04	Et skår i gleden_5	40	29	13	14	5	32	25	21	27	0,384	0,39	0,06	
42	L141S08	Et skår i gleden_6	11	78			10	18	30			0,559	0,78	0,08	Åpen

Oppgaver med lav diskriminering

Tabell 25 nedenfor viser diskrimineringsverdiene for alle oppgavene i prøven. Oppgavenes diskrimineringsevne er oppgitt både som et klassisk mål (point biserial) og en a-verdi. Oppgaver med diskriminering under 0,250 (point biserial) eller a-parameter under 0,8 er merket med stjerne.

Tabell 25: Oversikt over diskrimineringsverdier i prøven.

Oppg.	Item-ID	Oppgave	Diskriminering (point biserial)	a
1	L144S01	Tasmansk pungulv_1	0,498	0,920
2	L144S02	Tasmansk pungulv_2	0,541	0,980
3	L144S03	Tasmansk pungulv_3	0,606	1,250
4	L144S04	Tasmansk pungulv_4	0,679	1,430
5	L144S06	Tasmansk pungulv_5	0,236*	0,320*
6	L144S07	Tasmansk pungulv_6	0,598	1,170
7	L144S08	Tasmansk pungulv_7	0,475	0,790*
8	L142S09	Stargate_1	0,507	0,860
9	L142S12	Stargate_2	0,526	0,900
10	L142S08	Stargate_3	0,602	1,170
11	L142S05	Stargate_4	0,685	1,490
12	L142S04	Stargate_5	0,519	0,630*
13	L142S11	Stargate_6	0,557	1,030
14	L143S01	Lysforurensing_1	0,644	1,340
15	L143S09	Lysforurensing_2	0,399	0,610*
16	L143S10	Lysforurensing_3	0,676	1,460
17	L143S06	Lysforurensing_4	0,571	1,060
18	L143S07	Lysforurensing_5	0,659	1,400
19	L143S08	Lysforurensing_6	0,531	0,930
20	L147S01	Skularbeid_1	0,551	1,050
21	L147S02	Skularbeid_2	0,566	1,110
22	L147S03	Skularbeid_3	0,406	0,640*
23	L147S04	Skularbeid_4	0,484	0,800
24	L147S06	Skularbeid_5	0,694	1,540
25	L147S07	Skularbeid_6	0,540	0,950
26	L146S01	Snikfotografen_1	0,610	1,180
27	L146S03	Snikfotografen_2	0,560	1,030
28	L146S04	Snikfotografen_3	0,699	1,590
29	L146S05	Snikfotografen_4	0,764	1,880
30	L146S06	Snikfotografen_5	0,632	1,260
31	L146S07	Snikfotografen_6	0,680	1,450
32	L145S01	Den korte historien om egget_1	0,567	1,050
33	L145S02	Den korte historien om egget_2	0,721	1,660
34	L145S04	Den korte historien om egget_3	0,461	0,740*
35	L145S05	Den korte historien om egget_4	0,524	0,910
36	L145S06	Den korte historien om egget_5	0,300	0,410*
37	L141S01	Et skår i gleden_1	0,504	0,870
38	L141S02	Et skår i gleden_2	0,667	1,390
39	L141S07	Et skår i gleden_3	0,704	1,610
40	L141S06	Et skår i gleden_4	0,621	1,210
41	L141S04	Et skår i gleden_5	0,484	0,850
42	L141S08	Et skår i gleden_6	0,755	1,910

Det er én oppgave der begge diskrimineringsverdiene ligger under de nevnte grensene. Oppgave nummer 5, Tasmansk pungulv_5 (L144S06), har en diskriminering (point biserial) på 0,236, som ligger nær grenseverdien, men en a-verdi på 0,320 er alt for lav. Under hele prøveutviklingsprosessen har det vært gjort forsøk på å bedre oppgavens diskrimineringsverdier. Også etter andre pilotering ble det gjort endringer i oppgaveformuleringen som vi håpet ville bedre verdiene (jfr. vedlegg 2). Dessverre ser vi at de siste justeringene ikke bedret diskrimineringen og at oppgavene til slutt endte opp med for dårlige verdier. Vi vil likevel forsvare oppgaven da den stiller spørsmål med hvilke kilder forfatteren av teksten har brukt, noe som er svært relevant for å måle den kritiske leseferdighetene til elevene.

Oppgavene fordelt på mestringsnivåer

Tabell 26 viser oppgavene i den nasjonale prøven i lesing på 8. trinn fordelt på mestringsnivå. Det er satt nye grenser for mestringsnivåer i 2022, slik at resultatene fra gjennomføringen i 2023 ikke er direkte sammenlignbare med resultater fra tidligere gjennomføringer.

Antall oppgaver på hvert mestringsnivå er omtrent det samme som i estimatet etter siste pilotering. Som i estimatet er det bare en oppgave på det høyeste mestringsnivået. Det kan være flere grunner til at oppgaver flytter seg fra ett nivå til et annet. Det er alltid noe usikkerhet knyttet til mestringsnivåene i prøveforslaget ettersom disse ikke er samkalibrert med ankeret. Oppgaver som ligger nærme grensa mellom nivåer, kan være ekstra utsatt for å bytte nivå etter gjennomføring.

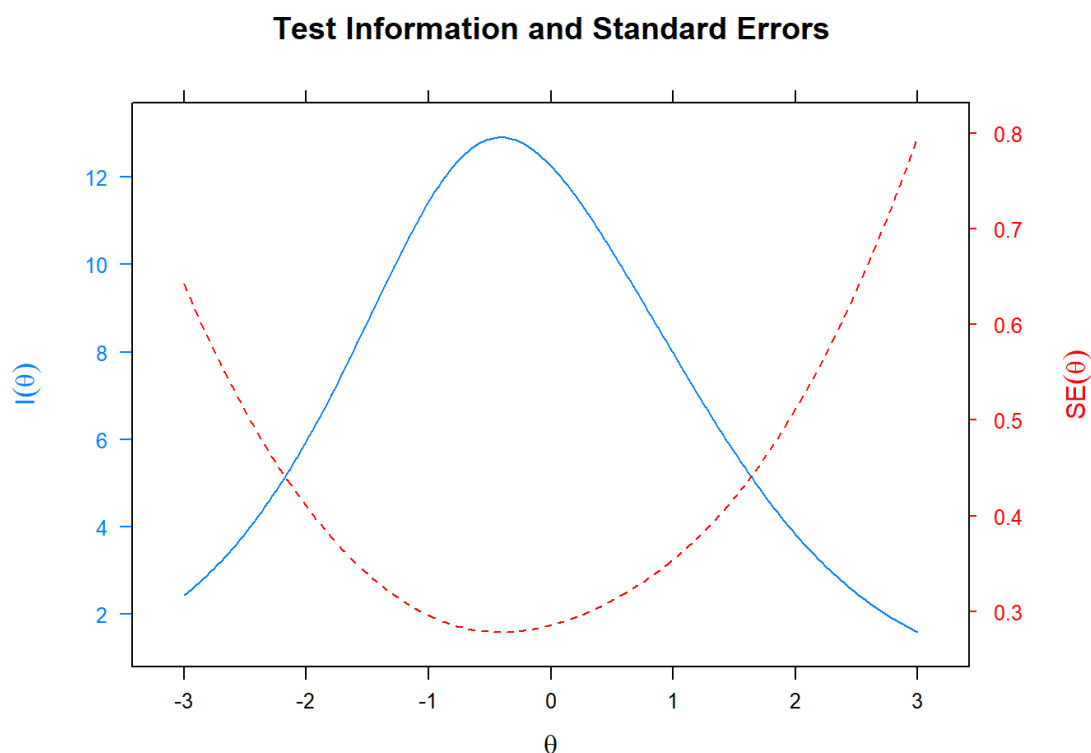
Tabell 26: Oppgavene i nasjonal prøve i lesing på 8. trinn 2023 fordelt på mestringsnivåer

Nivå	Item-ID	Oppgave	Skalapoeng
Nivå 1	L144S01	Tasmansk pungulv_1	30,2
	L147S02	Skulearbeid_2	35,3
	L147S01	Skulearbeid_1	35,7
Nivå 2	L142S05	Stargate_4	38,3
	L142S11	Stargate_6	38,5
	L144S02	Tasmansk pungulv_2	38,9
	L144S07	Tasmansk pungulv_6	39,1
	L147S03	Skulearbeid_3	39,2
	L142S08	Stargate_3	39,8
	L143S01	Lysforurensing_1	40,0
	L143S07	Lysforurensing_5	40,1
	L145S01	Den korte historien om egget_1	40,8
	L143S10	Lysforurensing_3	41,1
	L141S07	Et skår i gleden_3	41,9
	L141S08	Et skår i gleden_6	42,0
	L146S03	Snikfotografen_2	42,5
	L143S06	Lysforurensing_4	43,4
	L146S01	Snikfotografen_1	43,5
	L145S02	Den korte historien om egget_2	43,5
	L147S07	Skulearbeid_6	45,9
	L144S04	Tasmansk pungulv_4	46,3
	L146S05	Snikfotografen_4	46,4
	L141S02	Et skår i gleden_2	47,0

Nivå 3	L146S06	Snikfotografen_5	47,3
	L141S01	Et skår i gleden_1	47,6
	L145S04	Den korte historien om egget_3	47,9
	L143S09	Lysforurensing_2	48,7
	L142S04	Stargate_5_1p	49,4
	L142S12	Stargate_2	49,6
	L143S08	Lysforurensing_6	51,2
	L146S07	Snikfotografen_6	52,3
	L147S04	Skulearbeid_4	52,6
	L141S06	Et skår i gleden_4	52,6
	L145S05	Den korte historien om egget_4	53,0
Nivå 4	L144S08	Tasmansk pungulv_7	55,6
	L147S06	Skulearbeid_5	56,3
	L142S09	Stargate_1	56,6
	L145S06	Den korte historien om egget_5	57,6
	L146S04	Snikfotografen_3	58,4
	L142S04	Stargate_5_2p	60,3
	L141S04	Et skår i gleden_5	60,4
	L144S03	Tasmansk pungulv_3	61,5
Nivå 5	L144S06	Tasmansk pungulv_5	84,4

Testinformasjon og måleusikkerhet

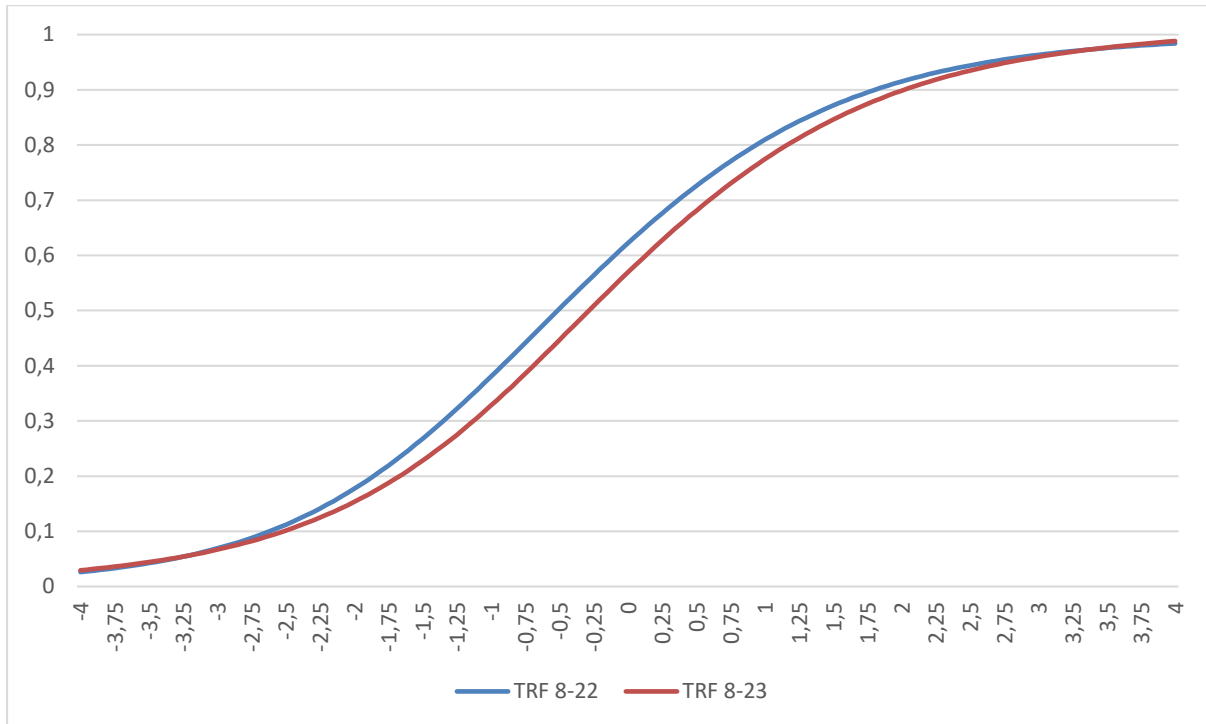
En IRT-analyse av den endelige prøven viser at testinformasjonen (TIF) som prøven leverer i sin helhet, er høyest for theta lik -0,4 og at måleusikkerheten (CSEM) er minst i det samme ferdighetsområdet i prøven. Figur 14 viser testinformasjon og måleusikkerhet for prøven.



Figur 14: Informasjonsverdi og måleusikkerhet for 2023-prøven

Prøvens TRF (Test respons function)

Figur 15 viser testrespons-funksjonen (TRF) for prøven 2023 (rød linje/1), sammenlignet med prøven 2022 (blå linje/0). Ettersom 2022-prøven og 2023-prøven er analysert med ulike analyseverktøy, er ikke variasjonen i distribusjonen direkte sammenlignbar.



Figur 15: TRF for prøven 2023, sammenlignet med prøven 2022

TRF er en beskrivelse av sammenhengen mellom estimert ferdighet (theta) og andel riktig besvarte oppgaver i prøven. Figur 15 viser med andre ord hvor stor andel av oppgavene en elev med en gitt ferdighet (theta) må klare for å få samme resultat på de to prøvene.

Det går fram av figuren at det er noe forskjell mellom prøvene. For elever med ferdighet mellom theta -4 og -3, og mellom 2,5 og 4, er det tilnærmet ingen forskjell, men mellom theta -3 og 2,5 ligger kurven for årets prøve under kurven for 22-prøven. Det betyr at elevene i dette ferdighetsområdet må få til en mindre andel oppgaver i årets prøve enn i prøven fra 2022 for å oppnå samme resultat.

Referanser

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

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

Utdanningsdirektoratet (2020). *Læreplanverket*. <https://www.udir.no/laring-og-trivsel/lareplanverket/>

Utdanningsdirektoratet (2017). Rammeverk for grunnleggende ferdigheter. Lastet ned fra: <https://www.udir.no/laring-og-trivsel/rammeverk/rammeverk-for-grunnleggende-ferdigheter/> [09.03.2023]

Utdanningsdirektoratet (2022). Rammeverk for nasjonale prøver. Lastet ned fra: <https://www.udir.no/eksamen-og-prover/prover/rammeverk-for-nasjonale-prover2/> [03.02.2023]

Vedlegg 1: Beskrivelse av lesekompetanse på de fem ulike nivåene

	FINNE	TOLKE	REFLEKTERE OG VURDERE
Nivå 1	lokalisere tydelig uttrykte informasjons-elementer i tekster med lite konkurrerende informasjon	trekke enkle slutninger eller oppfatte hovedtemaet i tekster når innholdet er tydelig uttrykt i tekstene	bruke personlig erfaring til å vurdere form eller innhold i tekster
Nivå 2	lokalisere tydelig uttrykte informasjons-elementer i tekster med noe konkurrerende informasjon	trekke slutninger om eller oppfatte hovedtemaet i tekster på bakgrunn av informasjon som ikke er tydelig uttrykt i tekstene	bruke personlig erfaring og/eller hverdagskunnskap til å vurdere form eller innhold i tekster
Nivå 3	lokalisere flere informasjonselementer fra ulike steder i tekster som inneholder klart konkurrerende informasjon	definere et ikke tydelig uttrykt tema i én eller flere tekster og/eller forstå sammenhenger mellom ulike deler av én eller flere tekster	bruke personlig erfaring og/eller formell kunnskap til å vurdere form eller innhold i tekster
Nivå 4	lokalisere og kombinere informasjon fra ulike steder i én eller flere tekster og vurdere hvilken informasjon som er relevant	forstå hvordan ikke tydelig uttrykte informasjonselementer i én eller flere tekster henger sammen, og/eller hvordan disse henger sammen med tekstene som helhet. Forstå meningsinnhold som står i motsetning til det forventede	vurdere form og/eller innhold i én eller flere tekster kritisk ved å sammenlikne, kontrastere eller kategorisere informasjon fra tekstene
Nivå 5	lokalisere og kombinere informasjon – ofte implisitt eller godt skjult – fra ulike steder i én eller flere tekster. Skille relevant informasjon fra sterkt konkurrerende informasjon	forstå tvetydigheter, meningsinnhold som står i motsetning til det forventede, eller meningsinnhold som er implisitt eller negativt uttrykt i én eller flere tekster	vurdere form og/eller innhold i én eller flere tekster kritisk og analytisk ved å utforme hypoteser på bakgrunn av egen kunnskap, sammenlikne, kontrastere eller kategorisere informasjon i tekstene

Vedlegg 2: Oversikt over oppgaver og endringer

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
Tasmansk pungulv				
L144S01	I det første avsnittet i tekst 1 vil forfatteren ...	Hva handler det første avsnittet i tekst 1 om?		
L144S02	Hva var den viktigste grunnen til at tasmansk pungulv ble utryddet, ifølge tekst 1?			
L144S03	Plasser hendelsene nedenfor i den rekkefølgen de faktisk skjedde, fra først (øverst) til sist (nederst).			
L144S04	Hva synes folk som bor på Tasmania, om pungulven i dag?	Nevn noe fra teksten som viser at de som bor på Tasmania, er stolte av pungulven.		
L144S05	Tekst 2 er en nyhetsartikkel. Hva er nyheten?	<i>Oppgaven er tatt ut.</i>		
L144S06	Hvor har journalisten som har skrevet tekst 2, hentet opplysningene sine fra?	Journalisten som har skrevet tekst 2, viser blant annet til en rapport fra tasmanske myndigheter. Hvor har journalisten ellers hentet opplysningene sine fra?		[...] har blant annet hentet opplysninger fra en rapport utgitt av tasmanske myndigheter. [...]

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L144S07	Den første teksten ble skrevet i 2017, den andre i 2019. Hva skjedde i mellomtiden?			Tekst 1 ble skrevet i 2017, og tekst 2 ble skrevet i 2019. Hva skjedde i mellomtiden? <i>Endret verbtider i svaralternativene.</i>
L144S08		Noen mener at det finnes videoer som beviser at pungulven eksisterer. Hva er problemet med disse videoene, ifølge teksten?		
Stargate				
				<i>Oppgave 3 og 5 byttet plass, (L142S04 og -08)</i>
L142S01	Hvor foregår handlingen i dette utdraget?	<i>Oppgaven er tatt ut.</i>		
L142S02	Hvorfor får jenta mat av vaktmesteren?	<i>Oppgaven er tatt ut.</i>		
L142S03	Gi et eksempel fra teksten som viser at jenta og vaktmesteren møter hverandre ofte.	<i>Oppgaven er tatt ut.</i>		
L142S09		I starten av teksten gjør jenta noe som egentlig ikke er lov. Hva er det?		I begynnelsen ...

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L142S10		Nevn to eksempler fra teksten som viser at vaktmesteren bryr seg om jenta.	<i>Oppgaven er tatt ut.</i>	
L142S12			Hvilke ord passer for å beskrive vaktmesteren? Sett kryss for «Passer» eller «Passer ikke».	
L142S04	På lyktestolpen henger det en plakat der noen søker etter folk som vil jobbe som juletreselger. Hva er den egentlige grunnen til at vaktmesteren fjerner hele plakaten?	Vaktmesteren fjerner plakaten der det står at juletreselger søkes. Hvorfor gjør han dette?		Hvorfor tar vaktmesteren ned hele plakaten og gir den til jenta?
L142S05	Hvem kan passe som juletreselger, ifølge vaktmesteren?			
L142S06	Hvilken påstand om vaktmesteren stemmer med innholdet i utdraget?	<i>Oppgaven er tatt ut.</i>		
L142S07	På hvilken måte viser vaktmesteren at han er en stor tenker?	<i>Oppgaven er tatt ut.</i>		
L142S08	Hvorfor er «Du er: pliktoppfyllende. Ansvarsfull. Glad i å være ute» skrevet med kursiv (skrå skrift)?			

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L142S11		Til slutt i samtalen med jenta sier vaktmesteren: «Nettopp. Du står her og ser på en stor tenker.» Hva vil han få fram med dette?		
Lysforurensing				
		<i>Innledningen til teksten er forkortet med ca. 100 ord.</i>		
L143S01	Hva er lysforurensing, ifølge teksten?			
L143S02	Forfatteren argumenterer for at ...	<i>Oppgaven er tatt ut.</i>		
L143S09		Hvem er det forfatteren mener må ta hovedansvaret for å redusere lysforurensingen?		
L143S03	Når omtrent fikk folk flest strøm i Norge?	<i>Oppgaven er tatt ut.</i>		
L143S04	Hvorfor nevner forfatteren Hammerfest og året 1891 i denne teksten?	<i>Oppgaven er tatt ut.</i>		
L143S10		I teksten står det: «De siste 70 åra har vi forandret natta på en måte som ikke har noe naturlig sidestykke». Hva mener forfatteren med dette?		<i>Byttet plass på svar A og C</i>

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L143S05	Hvorfor kan det være et problem at nyklekte havskilpadder blir påvirket av kunstig belysning?	Gi et eksempel på noe i teksten som viser at lysforurensing kan forstyrre naturen.		<i>Oppgaven er tatt ut.</i>
L143S06	Forfatteren bruker to lyssøyer i New York som eksempel på ...			<i>Byttet plass på svar A og C</i>
L143S07		Når forfatteren skriver at 80 prosent av verdens innbyggere – og 99 prosent av europeere – lever under en lysforurenset himmel, vil hun få fram ...		<i>Endret svaralternativ C.</i>
L143S08		Hva er hensikten med det siste avsnittet i teksten?		Hva er et viktig poeng i det siste avsnittet i teksten?
Skularbeid (sidemål)				
		<i>Teksten er oversatt fra bokmål til nynorsk.</i>		
L147S01	Hva er hovedhensikten med denne teksten?			
L147S02	Hvor mye tid er det mest vanlig å bruke på lekser per dag?			
L147S03	Hvor stor prosentandel av guttene på ungdomsskolen brukte minst én time på lekser hver dag i 2021?			<i>Endret svaralternativene (prosentandeler).</i>

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L147S04	Hva er hensikten med tekstboksen som har overskriften «Skolestress»?			Hva er poenget med tekstboksen som har overskriften «Skolestress»?
L147S05	Omtrent hvor stor prosentandel av jentene på 8. trinn brukte minst én time på lekser i 2021?	<i>Oppgaven er tatt ut.</i>		
L147S06	Andelen elever som oppgir at de blir stresset av skolearbeidet, er størst på 10. trinn og Vg3. Hva kan være årsaken til det?	<i>Bilde av Stian som sier:</i> Det er tydelig at elevene på 10. trinn og Vg3 blir påvirket av at de skal ha avsluttende eksamen og få vitnemål. Hvordan kan Stian begrunne denne påstanden ut fra resultatene i Ungdata-undersøkelsen?		[...] Hvordan kan Stian begrunne denne påstanden? Vis til konkrete resultater i Ungdata-undersøkelsen når du svarer.
L147S07	Hva er hensikten med Ungdata-undersøkelsen, ifølge teksten?			Hva er målet med Ungdata-undersøkelsen, ifølge teksten?
Snikfotografen				
L146S01	Hva er hensikten med de første avsnittene i teksten (før overskriften «Kulturhistorisk vendepunkt»)?			
L146S02	Hva gjorde Carl Størmer for å få folk til å se mot kameraet?	<i>Oppgaven er tatt ut.</i>		

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L146S03	Hvem mener at Størners bilder representerer et «kulturhistorisk vendepunkt»?			
L146S04	På hvilken måte er Carl Størners bilder forskjellig fra typiske bilder fra slutten av 1890-tallet, ifølge teksten?	[...], ifølge Robert Meyer?		På hvilken måte så folk annerledes ut på Carl Størners bilder sammenlignet med typiske bilder fra 1890-årene?
L146S05	Carl Størner er kjent som snikfotografen på Karl Johan. Hva annet er han kjent for, ifølge teksten?			
L146S06	Hvorfor er Alv Egeland intervjuet i denne teksten?			
L146S07		Plasser hendelsene nedenfor i den rekkefølgen de faktisk skjedde, fra først (øverst) til sist (nederst).		
Den korte historien om egget				
L145S01	Hva ønsker forfatteren mest sannsynlig å oppnå ved å stille et spørsmål først i teksten?			

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L145S02	Hvilke to viktige og sunne næringsstoffer finnes i eggeplommen og eggehviten?	Hvilke to viktige og sunne næringsstoffer finnes i eggeplommen og eggehviten, ifølge teksten?		
L145S03	Hva er grunnen til at de fleste hønseegg er hvite, ifølge teksten?	Hva er grunnen til at de fleste hønseegg er hvite nå for tiden, ifølge teksten?		<i>Oppgaven er tatt ut.</i>
L145S04	Hva vil forfatteren få fram ved å vise en matematisk formel?	Hva kan være en grunn til at forfatteren har med en matematisk formel i teksten?		
L145S05	Hvorfor er det en fordel å oppbevare egg i romtemperatur?			
L145S06	Hva er poenget med den nederste tekstboksen?			
Et skår i gleden?				
L141S01	Hva prøver Jernia å overbevise leseren om?			
L141S02	Hva skjer med drikkeglass, porselen og keramikk som leveres inn til Jernia?			Drikkeglass, porselen og keramikk som levers inn til Jernia, blir ...
L141S03	Nederst i teksten står det: «Sammen kan vi gjøre en forskjell». Hvilken forskjell er det snakk om?	<i>Oppgaven er tatt ut av analysegrunnlaget i første pilot, fordi det var stor variasjon i elevsvarene og dermed vanskelig å kode dem på en konsistent måte.</i>		

Item-ID	Første pilot (februar 2022)	Andre pilot (august 2022)	Ekstrapilot (februar 2023)	Prøveforslag
L141S06		Ifølge teksten vil Jernia få folk til å levere glass, porselen og keramikk til gjenvinning. Hva annet ønsker Jernia mest sannsynlig å oppnå med denne teksten?		Ifølge teksten er Jernia opptatt av miljøet og vil få folk til å ...
L141S04	Hvorfor står det en stjerne (*) etter ordet «vare» i teksten?	Etter ordet «vare» står det en stjerne (*). Hva viser denne stjernen til?		
L141S05	Hva er hovedhensikten med denne teksten?	<i>Oppgaven er tatt ut.</i>		
L141S07		I teksten står det: «Visste du at glass, porselen og keramikk ikke skal i containeren for glass og metall?» Hva kan være grunnen til at Jernia stiller dette spørsmålet?		
L141S08		Hvor har Jernia hentet opplysningene om gjenvinning og resirkulering fra?		

Vedlegg 3A: Spørreskjema til prøve A

Takk for at du gjennomførte prøven!

Nå vil vi som lager prøven, vite hva du syns om tekstene.

De tre oppgavene under får du ikke poeng for, og du kan krysse av for så få eller mange tekster du vil.

Hvilken tekst likte du best? Du kan krysse av for flere tekster.

- ☐ Et skår i gleden?
- ☐ Stargate
- ☐ Lysforurensing
- ☐ Tasmansk pungulv
- ☐ Den korte historien om egget
- ☐ Studenten som fanget Ibsen med skjult kamera
- ☐ Skolearbeid

Hvilken tekst syns du var vanskelig? Du kan krysse av for flere tekster.

- ☐ Et skår i gleden?
- ☐ Stargate
- ☐ Lysforurensing
- ☐ Tasmansk pungulv
- ☐ Den korte historien om egget
- ☐ Studenten som fanget Ibsen med skjult kamera
- ☐ Skolearbeid

Fikk du god nok tid til å svare på alle oppgavene?

- ☐ Ja
- ☐ Nei

Vedlegg 3B: Spørreskjema til prøve B

Takk for at du gjennomførte prøven!

Nå vil vi som lager prøven, vite hva du syns om tekstene.

De tre oppgavene under får du ikke poeng for, og du kan krysse av for så få eller mange tekster du vil.

Hvilken tekst likte du best? Du kan krysse av for flere tekster.

- ☐ Når gammel kjærlighet ruster
- ☐ Stargate
- ☐ Dyreparker
- ☐ Stripemysteriet
- ☐ Noe du trenger, mor?
- ☐ Salum Kashafali
- ☐ Klimavennlig mat

Hvilken tekst syns du var vanskelig? Du kan krysse av for flere tekster.

- ☐ Når gammel kjærlighet ruster
- ☐ Stargate
- ☐ Dyreparker
- ☐ Stripemysteriet
- ☐ Noe du trenger, mor?
- ☐ Salum Kashafali
- ☐ Klimavennlig mat

Fikk du god nok tid til å svare på alle oppgavene?

- ☐ Ja
- ☐ Nei

Vedlegg 3C: Spørreskjema til prøve C

Takk for at du gjennomførte prøven!

Nå vil vi som lager prøven, vite hva du syns om tekstene.

De tre oppgavene under får du ikke poeng for, og du kan krysse av for så få eller mange tekster du vil.

Hvilken tekst likte du best? Du kan krysse av for flere tekster.

- ☐ Et skår i gleden?
- ☐ Stargate
- ☐ iHuman
- ☐ Kraften
- ☐ Sjøku
- ☐ Knøttfe og himmelreker
- ☐ Vandring gjennom tiden

Hvilken tekst syns du var vanskelig? Du kan krysse av for flere tekster.

- ☐ Et skår i gleden?
- ☐ Stargate
- ☐ iHuman
- ☐ Kraften
- ☐ Sjøku
- ☐ Knøttfe og himmelreker
- ☐ Vandring gjennom tiden

Fikk du god nok tid til å svare på alle oppgavene?

- ☐ Ja
- ☐ Nei