# Forslag til nasjonale prøvar i engelsk 2025, for 8. trinn

## Del 1 - Utprøving september 2024

*INSTITUTT FOR FRAMANDSPRÅK*
*C. GROCOTT, H.HELNESS OG C. SVENSSON*
*Mars 2025*

UNIVERSITETET I BERGEN

# Innhaldsliste

**GENERELT OM NASJONAL PRØVE I ENGELSK FOR 8. TRINN ..... 3**

*Lesing og lytting i faget engelsk ..... 3*

*Kva måler prøvane på 8. trinn ..... 5*

*Nasjonal prøve i engelsk for 8. trinn og læreplanverket LK20/LK20S ..... 6*

*Om teksttypar og tema i prøvane på 8. trinn ..... 9*

**DEL 1 – UTPRØVING AV NP ENGELSK 8. TRINN HAUSTEN 2024 OG FORSLAG TIL PRØVAR 2025 ..... 11**

*1.1 Generell informasjon om utprøvinga til prøven i 2025 ..... 11*

*1.2 Forslag til prøvar 8. trinn 2025 ..... 14*

*1.3 DIF mellom gutar og jenter ..... 24*

*1.4 Oversikt over enkeltoppgåver 8. trinn ..... 26*

**DEL 2 – NASJONAL PRØVE I ENGELSK 8. TRINN 2025 ..... 31**

**2. Nasjonal prøve i engelsk 8. trinn høst 2025 - hovedresultat ..... 31**

*2.1 Gjennomføring 2025 ..... 31*

*2.2 Analyser av kohortprøven NPENG08 ..... 31*

2.2.1 Overordnet resultat ..... 32

2.2.2 Resultat på oppgavenivå ..... 34

Løsningsprosent (p-) og b-verdier ..... 37

Diskriminering – a-verdi ..... 37

Theta ..... 38

*2.3 Kjønn ..... 40*

*2.4 NP engelsk for 8. trinn 2025 – oppsummering ..... 44*

## Om rapporten

**Del 1 av rapporten gjer greie for arbeidet før 2025-utgåva av Nasjonal prøve i engelsk for 8. trinn blir gjennomført: utprøving av oppgåver og forslag til endeleg prøve. (Del 2 blir skrive etter at prøvane er gjennomførte hausten 2025 og summerer opp resultata frå den reelle prøvegjennomføringa.)**

Første del av denne delrapporten gir ein del generell informasjon om Nasjonal prøve i engelsk for 8. trinn (til dømes samanhengen mellom læreplan i engelsk og prøve, kva prøven måler, oppgåveformat og korleis oppgåvene i prøven blir kategorisert).

Sidan 2014 har det ved hjelp av dei nasjonale prøvane blitt undersøkt om ferdigheitene til elevane endrar seg over tid. Til og med 2021 har dette blitt gjort ved at eit utval elevar har tatt ein ankerprøve i engelsk, lesing og rekning, medan dei fleste har tatt hovudprøven som blir utprøvd og sett saman året før. Grunna nye læreplanar i 2020 og nytt konstrukt, starta trendmålingar på nytt i 2022, år null. Gjennomføringa i 2025 blir år tre.

Denne rapporten fokuserer på utprøving av oppgåver og forslag til Nasjonal prøve i engelsk for 8. trinn 2025. I samband med IRT-analysar er datamaterialet for alle oppgåvene som var med i 2024-gjennomføringa for 8. trinn lagt til grunn (kohort-, anker- og utprøvingsoppgåver).

## GENERELT OM NASJONAL PRØVE I ENGELSK FOR 8. TRINN

### Lesing og lytting i faget engelsk.

LK20 byggjer på eit kommunikativt syn på språklæring. Omgrepet kommunikativ kompetanse blir definert som evna til både å bruke grammatiske regler og forme setningar korrekt, men også å vite når det er passande å bruke språket på ein bestemt måte og til kven. Å kommunisere omfattar både det å motta og forstå språkleg informasjon, og det å produsere språk. Det blir ofte sagt at dei reseptive språkferdigheitene (lytte og lese) byggjer opp under dei produktive ferdigheitene (snakke og skrive). Språkleg resepsjon blir sett på som primært i forhold til språkproduksjon, som blir omtalt som sekundært. For å kunne snakke, må ein kunne

motta og tolke språkleg innputt (for eksempel ved å høyre), medan å skrive forutset det å kunne lese. Det betyr at for å utvikle språkkompetansen sin i engelsk, er det viktig å lytte til engelsk, lese engelsk og å bruke språket aktivt på ulike måtar og i ulike samanheng munnleg og skriftleg. Dette blir også understreka i kompetansemåla i engelskfaget som lister opp kva elevane bør/kan arbeide med og i kjerneelementet *kommunikasjon*. I tillegg framhevar Utdanningsdirektoratet sin definisjon av *grunnleggande ferdigheiter* at språkkompetanse består av ulike språklege element, både det å kunne lese, lytte, snakke og skrive. For at elevane skal utvikle den kommunikative kompetansen sin i engelsk, er det viktig at dei får høve til å bruke språket på mange forskjellige måtar i meiningsfulle situasjonar. Det same blir understreka i LK20, under **Fagets relevans og sentrale verdiar,** der engelskfaget blir omtalt på denne måten: «*Engelsk er et sentralt fag for kulturforståelse, kommunikasjon, danning og identitetsutvikling. Faget skal gi elevene et grunnlag for å kommunisere med andre lokalt og globalt, uavhengig av kulturell og språklig bakgrunn.*»

Engelsk er ikkje ein del av dei grunnleggande ferdigheitene, slik desse er definert i LK20. Derfor skil prøvane i engelsk seg frå dei andre nasjonale prøvane ved at dei tar utgangspunkt i det overordna kompetanseomgrepet, grunnleggande ferdigheiter og dei kompetansemåla som er relevante for lytting og lesing i faget engelsk.

I teksten om grunnleggande ferdigheiter i engelsk i LK20 under **Å kunne lese** står det følgande: *«Det innebærer å lese og finne informasjon i sammensatte tekster med konkurrerende budskap og å bruke lesestrategier til å forstå eksplisitt og implisitt informasjon. Utviklingen av leseferdigheter i engelsk går fra å eksperimentere med språklyder, stavemønstre og stavelser til å lese varierte og komplekse tekster med flyt og forståelse og i stadig større grad kunne reflektere over og vurdere ulike typer tekster kritisk.»*

I teksten om grunnleggande ferdigheiter i engelsk i LK20 under **Muntlige ferdigheter** står det følgande: *«Muntlige ferdigheter i engelsk er å skape mening gjennom å lytte, tale og samtale.»*

## Kva måler prøvane på 8. trinn

Nasjonale prøvar i engelsk skal måle i kva grad elevane sine ferdigheiter i engelsk er i samsvar med kompetansemåla i læreplanen for engelsk, knytt til lese- og lytteforståing, ord og uttrykk, og reglar for rettskriving, ordbøying og setningsstruktur.

På 8. trinn betyr dette mellom anna å:

- finne, forstå og bruke informasjon
- lytte til/lese og forstå ord og uttrykk i tilpassa og autentiske tekstar
- forstå ord, omgrep og hovudinnhold i tilpassa tekstar av varierande omfang frå forskjellige kjelder
- reflektere over innhald i tekstar
- forstå og bruke eit ordforråd og omgrep som dekker eigne interesser og aktuelle emne
- forstå betydinga av ord og uttrykk ut frå samanhengen dei er brukt i
- å forstå samanhengen i ein tekst
- følge reglar for rettskriving, ordbøying og setningsstruktur

Desse aspekta blir testa på ulike måtar (format) og i ulike kombinasjonar, og ikkje som ådskilde ferdigheiter. For eksempel må ein kunne forstå ord og omgrep for å kunne finne informasjon eller forstå hovudinnhaldet i en tekst. Ferdigheitene kan ikkje sjåast isolert.

Konstruktet for prøvane omfattar to hovuddimensjonar: 1) Nivået i teksten elevane må prosessere på for å løyse oppgåva og 2) Føremålet lesinga har (operasjonane oppgåva krev).

**Prosesseringsnivå**: Eleven kan finne svaret på ei einskild oppgåve enten i eit ord/uttrykk, ei setning eller ved å kople informasjon frå ulike deler av teksten.

**Føremålet** lesinga og lyttinga har: Ei oppgåve kan krevje at eleven kan finne informasjon og forstå detaljar, forstå hovudinnhald, tolke informasjon (trekke

slutningar), forstå samanheng i tekstar og oppfatte og bruke ord, stavemønster og grammatiske mønster.

Oppgåvene i prøven er utforma for å dekke alle desse aspekta. Det er likevel viktig å vere klar over at i mange tilfelle blir fleire ferdigheiter målt samtidig. I korte tekstar for dei yngste vil det til dømes ofte vere vanskeleg å skilje mellom å forstå hovudinnhald og detaljar. I slike tilfelle er det viktig å forstå detaljane (orda) for å forstå hovudinnhaldet.

## Nasjonal prøve i engelsk for 8. trinn og læreplanverket LK20/LK20S

Nasjonale prøvar i engelsk måler reseptive ferdigheiter og under referer vi til nokre viktige punkt i konstruktet:

I nasjonal prøve i engelsk blir delferdigheita **lytting** testa under **Munnlege ferdigheiter**, samt den grunnleggande ferdigheita **å kunne lese** i engelskfaget. Om **Munnlege ferdigheiter** seier LK20 at dette handlar om å «skape mening gjennom å lytte, tale og samtale». Lytting dannar grunnlaget for alle andre munnlege ferdigheiter i engelsk. I LK20 blir det sagt at ferdigheita **å kunne lese** inneber forståing og refleksjon, finne informasjon og å bruke lesestrategiar til å forstå eksplisitt og implisitt informasjon. Ferdigheita **å kunne skrive** blir ikkje testa direkte i nasjonal prøve, men lesing dannar grunnlaget også for denne ferdigheita. Dei fleste kompetansemåla i LK20 inneheld fleire av dei grunnleggande ferdigheitene samtidig.

**Kjerneelementa** står sentralt i nasjonal prøve i engelsk. Kjerneelementet ***Kommunikasjon*** omfattar ifylgje LK20 mellom anna «å skape mening med språk» og å «ta i bruk enkle strategier for å kommunisere muntlig og skriftlig i forskjellige situasjoner». Under blir nokre kompetansemål som er knytt til dette kjerneelementet beskrive, og samanhengen mellom desse og nasjonal prøve i engelsk.

**Etter 7. trinn**

- **bruke enkle strategiar i språklæring, tekstskaping og kommunikasjon**

I nasjonal prøve blir dette målet testa mellom anna gjennom oppgåver som inneber å finne informasjon, å tolke, å forstå hovudinnhald og å finne/forstå samanheng. Variasjonen i oppgåvene krev at elevane tar i bruk ulike strategiar for å løyse dei.

- **uttrykke seg forståeleg med eit variert ordforråd og høflegheitsuttrykk tilpassa mottakar og situasjon**
- **innleie, halde i gang og avslutte samtalar om eigne interesser og aktuelle emne**
- **skrive samanhengande tekstar, inkludert samansette, som gjenfortel, fortel, spør og uttrykker meiningar og interesser, tilpassa mottakar**

Desse kompetansemåla blir ikkje testa direkte i prøven, fordi dei omfattar dei produktive ferdigheitene. Elevane møter likevel i nasjonal prøve både eit variert ordforråd og høflegheitsuttrykk gjennom eit mangfald av samtalar og skriftlege tekstar, inkludert samansette. Dei blir testa i lese- og lytteforståing, som dannar grunnlaget for å oppnå kompetansemåla over.

Kjerneelementet ***Språklæring*** blir omtalt mellom anna slik i LK20: «Kunnskap om hvordan språklyder, ordforråd, og ord-, setningsstrukturer og tekststrukturer brukes, gir elevene valg og muligheter i kommunikasjon og samhandling.» Under blir nokre kompetansemål som kan knytast til dette kjerneelementet beskrive, og samanhengen mellom desse og nasjonal prøve i engelsk.

**Etter 7. trinn**

- **bruke enkle strategiar i språklæring, tekstskaping og kommunikasjon**

Dette kompetansemålet er like relevant her som under kjerneelementet kommunikasjon. I nasjonal prøve blir dette målet testa mellom anna gjennom å finne synonym, velje riktig form av ulike ord i ei gitt samanheng, og gjennom omskriving av setningar.

- **følge reglar for rettskriving, ordbøying og setningsstruktur**

Dette kompetansemålet blir testa direkte i nasjonale prøvar gjennom fleire ulike oppgåvetypar.

Om kjerneelementet ***Møte med engelskspråklige tekster*** seier LK20 «Språklæringen skjer i møte med engelskspråklige tekster», «Tekstbegrepet brukes i vid forstand», og «Gjennom å reflektere over, tolke og kritisk vurdere ulike typer engelskspråklige tekster skal elevene tilegne seg språk og kunnskap om kultur og samfunn.» Under

blir nokre kompetansemål som kan knytast til dette kjerneelementet beskrive, og samanhengen mellom desse og nasjonal prøve i engelsk.

**Etter 7. trinn:**

- **lytte til og forstå ord og uttrykk i tilpassa og autentiske tekstar**

Dette blir testa direkte gjennom ulike oppgåver i lyttedelen.

- **lese og vidareformidle innhald frå ulike typar tekstar, inkludert sjølvvalte tekstar**
- **lese og lytte til engelskspråklege sakprosatekstar og engelskspråkleg barne- og ungdomslitteratur og skrive og samtale om innhaldet**

For å oppnå desse kompetansemåla, er lese- og lytteforståing grunnleggande. Dette blir testa gjennom mange ulike oppgåvetypar i nasjonale prøvar.

- **utforske levemåtar og tradisjonar i ulike samfunn i den engelskspråklege verda og i Noreg og reflektere over identitet og kulturell tilhøyrigheit**

Nasjonal prøve ivaretar denne tematikken gjennom eit tekstutval med referansar til mange ulike stader og hendingar i den engelskspråklege verda.

Under teksten om ***underveisvurdering*** i LK20 står det slik: «Læreren skal gi veiledning om videre læring og tilpasse opplæringen slik at elevene kan bruke veiledningen for å utvikle leseferdigheter, skriveferdigheter og muntlige og digitale ferdigheter i faget.» Resultata frå nasjonal prøve og støtteressursar/PAS-prøvar dannar eit svært godt grunnlag for undervegsvurderinga av elevane sine grunnleggande ferdigheiter i faget innanfor dei reseptive ferdigheitene å lytte og å lese, som igjen dannar grunnlaget for dei produktive ferdigheitene å snakke og å skrive.

Oppgåvene og tekstane i nasjonale prøvar i engelsk kan med fordel brukast i oppfølgingsarbeidet og generelt i arbeidet med engelsk i klasserommet, med utgangspunkt i LK20.

## Om teksttypar og tema i prøvane på 8. trinn

Tekstane elevane skal lese på og lytte til i 8. trinnsprøven, er jamt over noko lengre og meir komplekse enn dei i prøven for 5. trinn. Nokre av oppgåvene i 8. trinnsprøven har også bilde, men færre enn i prøven for 5. trinn. Prøven inneheld tre relativt lange tekstar i lytteformata *Multiple choice long, Ordering* og *Matching*. I lesedelen er det to relativt lange tekstar, i formata *Who could say* og *Move paragraph*.

Tabell 8, som er ei oversikt over forslaget til prøven på 8. trinn, viser at teksttypane i lyttedelen er beskrivingar, monologar, dialogar, podcast, reklame og nyheitsmeldingar. Lesedelen inneheld beskrivingar, forteljingar og faktatekstar. *Click and drag* i lesedelen beskriv dyr, personar, stader eller situasjonar. Mange av tekstane har emne frå natur, historie og samfunn. Det finst også tekstar i prøven som omhandlar daglegdagse situasjonar og meir nære og personlege forhold.

Elevar som har øvd på å lese og lytte til tekstar med ulikt fokus og brukt ulike lese- og lyttestrategiar, bør ha eit godt utgangspunkt for å svare på oppgåvene i dei nasjonale prøvane i engelsk. På same måte burde prøvane eigna seg godt i etterkant, for å øve på å lese og lytte til tekstar av ulik lengd og kompleksitet.

## Om oppgåveformat

**Tabell 1a: Oppgåveformat i nasjonal prøve i engelsk for 8. trinn - lyttedelen**

| **Format Lytting** | **Lukka /opent format** | **Kva skal eleven gjere?** |
|---|---|---|
| **Click picture (CP)** | Lukka | Elevane blir presentert for fire bilde. Først lyttar dei til instruksjonen *Listen to the recording and click on the correct picture. You can listen to the recording twice.* Deretter lyttar dei til ein tekst og basert på innhaldet i denne klikkar dei på det bildet som passar til teksten. |
| **Multiple choice short (MCS)** | Lukka | Først lyttar elevane til instruksjonen *Listen to the recording. You can see the question on the screen while you listen. Click on the correct answer. You can listen to the recording twice.* Elevane lyttar deretter til ein kort lyttetekst og eit spørsmål. Dette er ei fleirvalsoppgåve |

| | | |
|---|---|---|
| | | der elevane svarer ved å velje eitt av fire alternative svar. |
| **Multiple choice long (MCL)** | Lukka | Først lyttar elevane til instruksjonen *Listen to the recording. You can see the questions on the screen while you listen. Answer the three questions. You can listen to the recording twice.* Elevane lyttar deretter til ein litt lengre tekst. Til oppgåva høyrer tre spørsmål. Dette er fleirvalsoppgåver der elevane svarer ved å velje eitt av fire alternative svar i kvart av spørsmåla. |
| **Ordering (OR)** | Opent | Først lyttar elevane til instruksjonen *Listen to the recording. Click on the key words and place them in the correct order. The first key word is given. You can move the key words while you listen. You can listen to the recording twice.* Elevane ser fem nøkkelord/frasar på venstre side av skjermen. Basert på innhaldet i teksten dei lyttar til, skal dei plassere nøkkelorda i riktig rekkefølge til høgre på skjermen. I tillegg er det eitt nøkkelord som allereie står fast øvst. |
| **Matching (MA)** | Lukka | Først lyttar elevane til instruksjonen *Listen to the recording. Give your answer by clicking four of the boxes in the table. You can listen to the recording twice.* Elevane ser ein tabell på skjermen med fire stikkord i kolonneoverskrifter, og fire stikkord i radene i første kolonna nedover til venstre. Basert på innhaldet i teksten dei lyttar til, skal dei matche dei ulike stikkorda, ved å krysse av fire av avkryssingsboksane. |

**Tabell 1b: Oppgåveformat i nasjonal prøve i engelsk for 8. trinn – lesedelen**

| **Format Lesing** | **Lukka /opent format** | **Kva skal eleven gjere?** |
|---|---|---|
| **Click and drag (C&D)** | Opent | Elevane blir presentert for eitt stort bilde og bilde av fire mindre ting utanfor det store bildet. I tillegg skal dei lese ein kort tekst som inneheld ein instruksjon. Oppgåva er å følgje instruksjonen og plassere ein av tinga på rett plass i det store bildet. |
| **Multiple choice short (MCS)** | Lukka | Elevane blir presentert for ein kort tekst og eit spørsmål med fire alternative svar. Elevane skal klikke på eitt av alternativa. |
| **Multiple choice medium (MCM)** | Lukka | Elevane blir presentert for ein middels lang tekst og to spørsmål. Kvart spørsmål har fire alternative svar. Elevane skal klikke på eitt av alternativa for kvart spørsmål. Dette er ein multimodal tekst, som i tillegg til |

| | | |
|---|---|---|
| | | hovudteksten inneheld eit bilde, og t.d. ei lita ordliste og/eller faktaboks. |
| **Click word (CW)** | Opent | Elevane blir presentert for ein kort tekst og ein instruksjon, t.d. *Click on the word that means almost the same as '…….'.* Oppgåva er å finne det ordet i teksten som betyr om lag det same og klikke på det. |
| **Who could say (WS)** | Opent | Elevane blir presentert for ein litt lengre tekst, der fleire personar blir introdusert, og fem ytringar. Kvar ytring representerer ein av personane i teksten. Oppgåva er å finne ut kven ytringa representerer og klikke på namnet til denne personen. |
| **Complete sentence (CS)** | Opent | Elevane blir presentert for to setningar. Den første setninga er fullstendig. I den andre setninga manglar det noko. Eleven skal skrive inn nokre ord i setninga. |
| **Write the word (WW)** | Opent | Elevane blir presentert for ein kort tekst med tre luker. Elevane skal skrive inn det ordet som manglar, i kvar av dei tre lukene. |
| **Move paragraph (MP)** | Opent | Elevane blir presentert for ein tekst med seks avsnitt. Det første avsnittet ligg fast. Dei andre avsnitta skal flyttast slik at teksten får ein logisk samanheng. Det er berre éi rekkefølge på avsnitta som er korrekt. |

Nasjonal prøve i engelsk måler reseptive ferdigheiter og har to delar: ein lyttedel og ein lesedel. I lyttedelen kan elevane høyre sjølve tekstfilene maksimalt to gonger, og dei har 60 minutt til saman for begge prøvedelane til å svare på alle oppgåvene. Dersom ein elev høyrer alle lydfilene i lyttedelen to gonger (inkludert instruksjonane som går igjen i kvar oppgåve innanfor eit format), tar lyttedelen ca. 30 minutt.

Alle oppgåvene i Nasjonal prøve i engelsk blir skåra automatisk. På Udir.no ligg eksempeloppgåver og Nasjonale prøvar frå 2023 og 2024, slik at elevane kan gjere seg kjende med oppgåveformata på førehand.

# DEL 1 – UTPRØVING AV NP ENGELSK 8. TRINN HAUSTEN 2024 OG FORSLAG TIL PRØVAR 2025

## 1.1 Generell informasjon om utprøvinga til prøven i 2025

I denne rapporten nemner vi tre typar oppgåver:

- Kohortoppgåver. Dette er oppgåvene som er inkludert i årets hovudprøve, NPENG 2024. Alle elevar tar desse oppgåvene.
- Ankeroppgåver. Dette er oppgåver som blir brukt i samband med trendmålingar. Ankeroppgåvene er dei same frå år til år, og ved hjelp av desse kan vi lenke data frå eitt år til eitt eller fleire andre år. Ankeroppgåvene vert ikkje publisert.
- Utprøvingsoppgåver. Dette er oppgåver som blir prøvd ut med tanke på den nasjonale prøven det komande året, dvs. kohortprøven.

Alle elevane tok alle kohortoppgåvene i hovudprøven/NPENG2024. I tillegg fekk kvar elev éi til fire ekstra oppgåver som anten var utprøvingsoppgåver til NPENG 2025 eller ankeroppgåver (til trendmåling). Tal på ekstraoppgåver varierte i forhold til vanskegrad, kva oppgåveformat og teksttypar elevane vart tildelt oppgåver i. Dersom dei fekk tildelt ein lang, (anteke) vanskelegare tekst, fekk dei t.d. éi oppgåve, og dersom dei fekk korte, enklare oppgåver, fekk dei fleire. Total arbeidsmengde for ekstraoppgåver skulle vere om lag det same for alle elevane. Elevane visste ikkje kva oppgåver som var «tilleggsoppgåver», og resultata frå utprøvingsoppgåvene vart heller ikkje rekna med i prøveresultatet totalt.

Det er fleire fordelar med å prøve ut nye oppgåver i samband med gjennomføringa av reelle prøvar. Utvalet som tek dei ekstra oppgåvene er representative for heile populasjonen (tilfeldig frå heile landet) og tidspunktet er det same som for reell gjennomføring. Dette tenkjer vi gjer resultata meir til å stole på.

Ein har tidlegare sett at oppgåvene vert lettare frå utprøving til reell gjennomføring, mest sannsynleg fordi elevane gjer ein større innsats i reell gjennomføring. Dette unngår ein når elevane får tilleggsoppgåver i sjølve NPENG-gjennomføringa.

I tillegg treng skulane heller ikkje å bruke ekstra tid på å vere med på utprøvingar av neste års prøve, og prøveutviklarane treng ikkje bruke tid på å kontakte skular rundt i landet og spørje om dei vil delta i utprøving. Tidlegare var det ein del skular som meldte at dei ofte vart trekt ut til å delta på utprøvingar for nasjonale prøvar (engelsk, lesing, rekning), og at det kunne opplevast belastande. Med dagens system vert utprøving av nye oppgåver svært mykje mindre

tidkrevjande for både skulane, prøveutviklarane og Udir, samstundes som resultatet er meir til å stole på.

**Analyse av data**

Data frå utprøvinga hausten 2024 er blitt analysert i *R* med IRT-metoden ved NPENG-faggruppa ved UiB og Utdanningsdirektoratet. Det er brukt Item Response Theory (IRT)-kalibrering med ein 2PL IRT-modell og med Generalized partial credit model (GPCM-modell) for polytome oppgåver.

UiB har brukt også Excel til å sette saman oppgåvene som er valt ut til forslag til NPENG2025, 8. trinn.

Analysane er basert på at heile datamaterialet for 8. trinn, det vil seie at data er frå analysar av alle oppgåvene samla som var med i gjennomføringa, totalt 179 oppgåver: 33 kohortoppgåver, 16 ankeroppgåver frå 2024 og 130 utprøvingsoppgåver. (Nokre av utprøvingsoppgåvene vart prøvd ut med tanke på å auke talet på ankeroppgåver i gjennomføringa i 2025, og det var difor nokre fleire utprøvingsoppgåver med i 2024 i forhold til i 2023.)

**Tidspunkt for utprøving**

September 2024

**Utprøvingsdesign**

Vi har brukt eit delvis lenka design. Alle elevar fekk 1 til 4 ekstra oppgåver (anker- og/eller utprøvingsoppgåver) under gjennomføring av NPENG 2024, 8. trinn.

**Tal på oppgåver som vart prøvd ut**

130 oppgåver

**Tal på elevar som deltok**

Dei same elevane som tok kohortoppgåvene i engelsk 8. trinn i 2024 deltok i utprøvinga: 60673.

Prøvedesignet var sett opp med eit mål om at kvar utprøvings- og ankeroppgåve skulle svarast på av minimum 600 elevar. Oppgåvene som er foreslått til NPENG2025 5. trinn er svart på av frå 588 til 1938 elevar.

**Samansetjing av 2025-prøven**

I arbeidet med å setje saman eit forslag til prøve for 2025, måtte vi balansere følgjande faktorar: lytteoppgåver, leseoppgåver, oppgåveformat, oppgåver totalt, føremål, tekstlengd, tema, løysingsprosent, eventuelle kjønnsforskjellar og diskriminering. Excel vart brukt til dette arbeidet.

## 1.2 Forslag til prøvar 8. trinn 2025

Vi har sett saman ein nasjonal prøve på 8. trinn med lese- og lytteoppgåver i engelsk i Udir sitt eksamens- og prøvesystem, EPS.

Prøven i 2025 vil innehalde stort sett dei same oppgåveformata og tal på oppgåver per format, som i 2024-prøven.

Tabell 2 under, viser tal på oppgåver per format for foreslått prøve for 8. trinn.

**Tabell 2: Oppgåver per format foreslått for NP 2025, 8. trinn**

| Format | Forslag til Nasjonal prøve 2025, 8. trinn |
|---|---|
| **Lytting: Click picture** | 3 oppgåver |
| **Lytting: Multiple choice short** | 8 oppgåver |
| **Lytting: Multiple choice long** | 1 (3) oppgåve * |
| **Lytting: Ordering** | 1 (3) oppgåve * |
| **Lytting: Matching** | 1 (4) oppgåve * |
| **Lesing: Click and drag** | 2 oppgåver |

| | |
|---|---|
| **Lesing: Multiple choice short** | 3 oppgåver |
| **Lesing: Multiple choice medium** | 1 (2) oppgåver * |
| **Lesing: Who could say** | 5 oppgåver |
| **Lesing: Click word** | 3 oppgåver |
| **Lesing: Complete sentence** | 3 oppgåver |
| **Lesing: Write words** | 1 (3) oppgåve * |
| **Lesing: Move paragraph** | 1 (3) oppgåve * |
| **Tal på oppgåver totalt** | **33 oppgåver (45)** |

* Desse oppgåvene er polytome, det vil seie at eleven må gi fleire svar for å svare fullstendig på oppgåva. Dersom vi tar *Write words*-formatet som eit døme, så møter elevane der ein kort tekst med tre luker. Elevane må skrive inn riktig svar i kvar av dei tre lukene.

Dei aller fleste oppgåveformata har oppgåver som blir skåra dikotomt (0/1). Elevane gir eitt svar på desse oppgåvene. I nokre av oppgåvene kan elevane gi fleire svar og oppnå fleire poeng; i polytome oppgåver. Maksimum poengsum frå utprøvinga er representert i parentesane i tabellane 2 til 6.

Tabell 3 nedanfor gir ei oversikt over vanskegrad (*b*-verdiar), løysingsprosent (*p*-verdiar) og diskrimineringsindeks (*a*-verdiar) i forslag til 2025-prøven for 8. trinn, med estimat frå *R* og Excel.

*b*-verdi viser til vanskegraden oppgåvene har (minusverdiar viser til lette oppgåver, og plussverdiar til vanskelege).

*p*-verdi viser til løysingsprosent på oppgåvene, altså proporsjon rette svar. Verdien 0,506 viser t.d. at 50,6 prosent av elevane svarte riktig på oppgåvene under utprøving.

*a*-verdi (diskriminering) viser til kor godt oppgåva skil mellom elevar som presterer godt og dårleg. *a*-verdien bør vere over 0,8.

**Tabell 3: Gjennomsnittleg vanskegrad og diskriminering, forslag til NP 8. trinn 2025**

| | **Nasjonal prøve 8. trinn** |
|---|---|
| **Tal på oppgåver** | 33 (45) |
| **Gjennomsnittleg *p*-verdi** | 0,506 |
| **Gjennomsnittleg *b*-verdi** | 0,017 |
| **Gjennomsnittleg *a*-verdi** | 1,44 |

*b*- og *p*-verdien i tabell 3 og 4 er gjennomsnittet for alle oppgåvene i prøven, inkludert snittet for polytome. For polytome oppgåver betyr det at t.d. *b*1, *b*2, *b*3 og *b*4 er lagt saman og deretter delt på tal på poeng. Dette snittet er utgangspunkt for det totale snittet i verdien ovanfor.

Dersom vi deler prøven inn i lytte- og leseoppgåver, får vi fylgjande tal:

**Tabell 4: Gjennomsnittleg diskriminering, løysingsprosent, og vanskegrad, lytting og lesing kvar for seg – forslag til NP2025 – 8. trinn**

| | **Lytting** | **Lesing** |
|---|---|---|
| **Tal på oppgåver** | 14 (21) | 19 (24) |
| **Gjennomsnittleg *a*-verdi** | 1,18 | 1,63 |
| **Gjennomsnittleg *p*-verdi** | 0,551 | 0,473 |

| | | |
|---|---|---|
| **Gjennomsnittleg *b*-verdi** | -0,133 | 0,127 |

Prøven måler reseptive ferdigheiter. Vi ser i tabell 4 at lytteoppgåvene i gjennomsnitt er noko lettare enn leseoppgåvene – gjennomsnittleg *p*-verdi for lytting er 0,551 mot 0,473 i lesing. Den samla *p*-verdien for alle oppgåvene, for reseptive ferdigheiter, er som vi ser i tabell 3 ovanfor 0,506.

Tabell 5 og 6 under viser data frå kohort- og ankeroppgåvene samla. Det er imidlertid viktig å hugse at elevane får ulike ankeroppgåver, 1 – 4 i tillegg til kohortoppgåvene (anker- og/eller pilotoppgåver), dvs. at ingen elevar får alle desse oppgåvene.

**Tabell 5: Gjennomsnittleg vanskegrad og diskriminering, forslag til NP 8. trinn 2025 inkludert ankeroppgåver**

| | **Nasjonal prøve 8. trinn (kohort og anker)** |
|---|---|
| **Tal på oppgåver** | 57(74) |
| **Gjennomsnittleg *p*-verdi** | 0,506 |
| **Gjennomsnittleg *b*-verdi** | 0,390 |
| **Gjennomsnittleg *a*-verdi** | 1,47 |

*b*- og *p*-verdien i tabell 5 og 6 er gjennomsnittet for alle oppgåvene i prøven (kohort- og ankeroppgåver), inkludert snittet for polytome på same måte som i tabell 3 og 4.

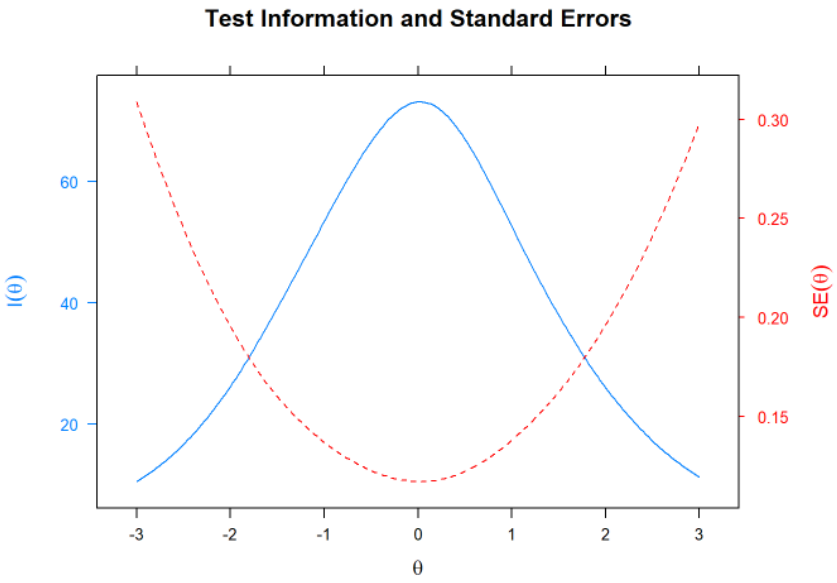
Dersom vi deler prøven inn i lytte- og leseoppgåver, får vi fylgjande tal:

**Tabell 6: Gjennomsnittleg diskriminering, løysingsprosent, og vanskegrad, lytting og lesing kvar for seg – forslag til NP2025 – 8. trinn, inkludert ankeroppgåver**

| | **Lytting (kohort og anker)** | **Lesing (kohort og anker)** |
|---|---|---|
| **Tal på oppgåver** | 24 (33) | 33 (14) |
| **Gjennomsnittleg *a*-verdi** | 1,19 | 1,67 |
| **Gjennomsnittleg *p*-verdi** | 0,55 | 0,474 |
| **Gjennomsnittleg *b*-verdi** | -0,182 | 0,108 |

Datagrunnlaget for figur 1 under er frå alle oppgåvene analysert samla. Figuren viser testinformasjonsfunksjon (TIF) i blå farge, og betinga målefeil (CSEM) i stipla rød line. Vi ser i figuren at prøven gir mest informasjon i området rundt 0.

**Figur 1: Testinformasjon og målefeil**




## Polytome oppgåver

I forslaget til hovudprøven for 8. trinn finst det i utgangpunktet sju polytome oppgåver, der elevane svarar på fleire oppgåver og får ein samleskår for kvar av desse oppgåvene. Tabell 7 gir ei oversikt over *a*- og *b*-verdiar for dei polytome oppgåvene under utprøving.

**Tabell 7: Fordeling av *a*- og *b*-verdiar for polytome oppgåver etter utprøving**

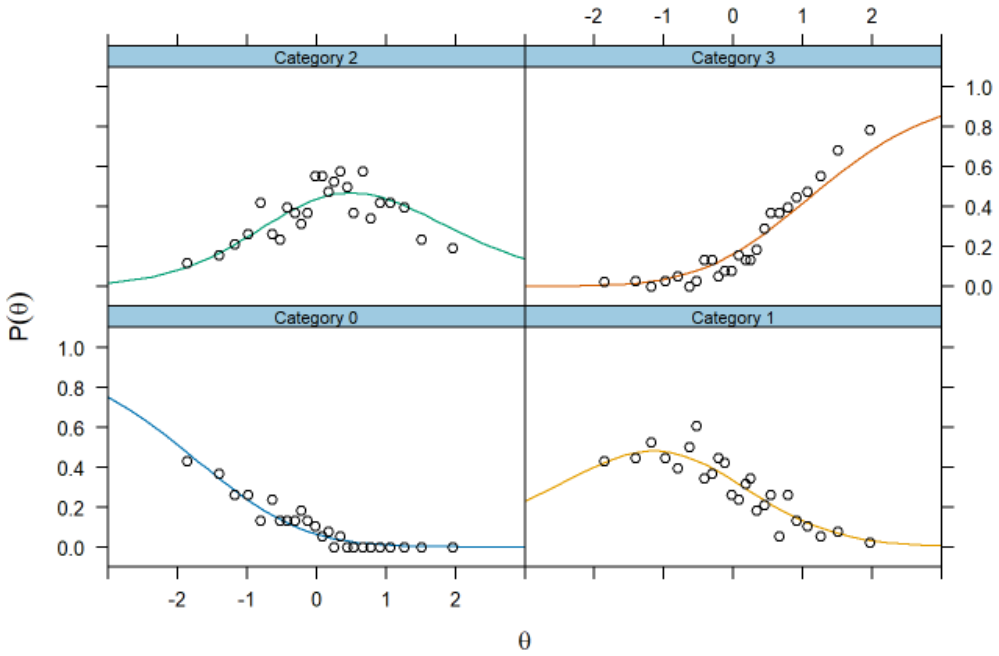
| Oppgåver | *a*-verdi | *b*-verdi<br>Eitt rett svar | *b*-verdi<br>To rette svar | *b*-verdi<br>Tre rette svar | *b*-verdi<br>Fire rette svar |
|---|---|---|---|---|---|
| ***Lytting*: MC long – Special Edition** | 0,937 | -1,74 | -0,295 | 1,051 | |
| ***Lytting*: Ordering – exchange program** | 0,616 | 1,459 | 4,666 | -1,456 | |
| ***Lytting*: Matching - Debatecast** | 0,37 | -1,593 | 0,928 | 3,587 | 4,807 |

| | | | | | |
|---|---|---|---|---|---|
| ***Lesing*: MC medium- Blood types** | 1,368 | -1,101 | -0,352 | | |
| ***Lesing:* Write words – dye jeans** | 1,272 | -1,131 | -0,493 | 0,701 | |
| ***Lesing:* Move paragraph – crop rotation** | 0,723 | 2,177 | 3,123 | 0,025 | |

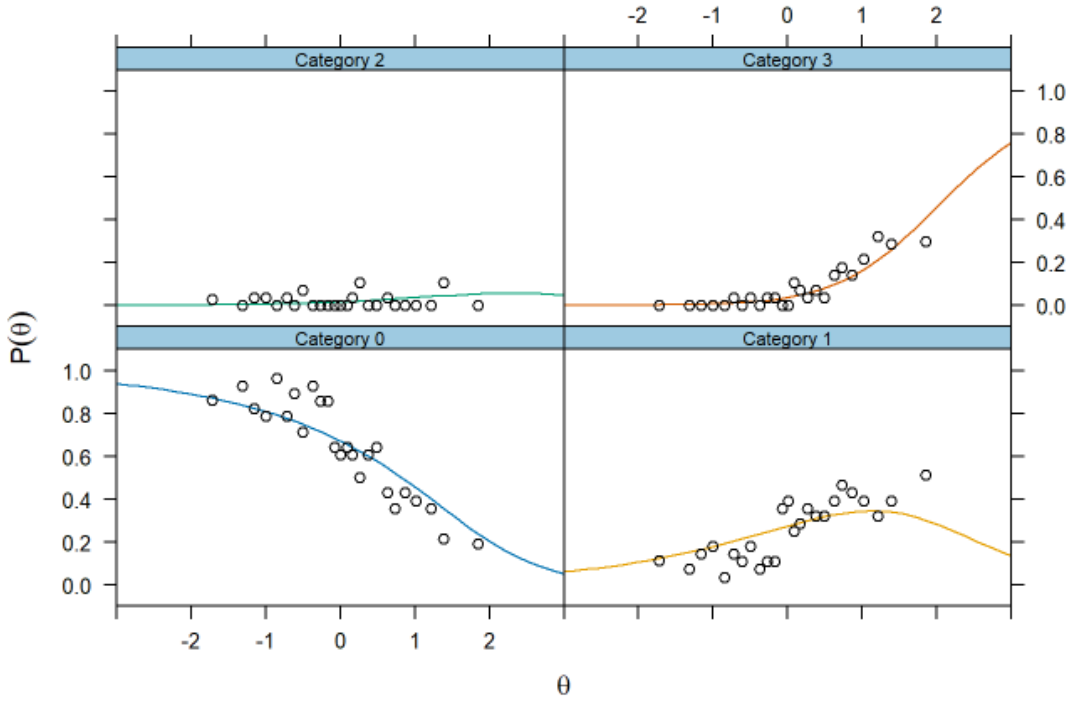
Figur 2 til 7 nedanfor viser IRT-bilde (*Item fit* eller elementtilpasning) av dei polytome oppgåvene, med informasjonskurve, responsfunksjon og empiriske plot. Figurane viser polytome oppgåver der elevane kan få t.d. opptil fire poeng (0, 1, 2, 3 eller 4 poeng.)

I t.d. Figur 2 under representerer X-aksen elevane sitt ferdigheitsnivå (theta), og Y-aksen sannsynet for å fylgje kvar av kurvene si responsfunksjon. Dei heil-lina kurvene viser sannsynet for at ein elev med ein gitt theta avgir korresponderande svarmønster for dei fire kategoriane. Category 0 gir null poeng, Category 1 gir eitt poeng, Category 2 gir to poeng og Category 3 gir tre poeng. Linene indikerer kvar oppgåva gir mest informasjon. Dei empiriske plotta (prikkane) i figuren viser faktisk fordeling av svara samanlikna med den predikerte fordelinga frå IRT.

**Figur 2: *Lytting*: MC long – Special Edition**



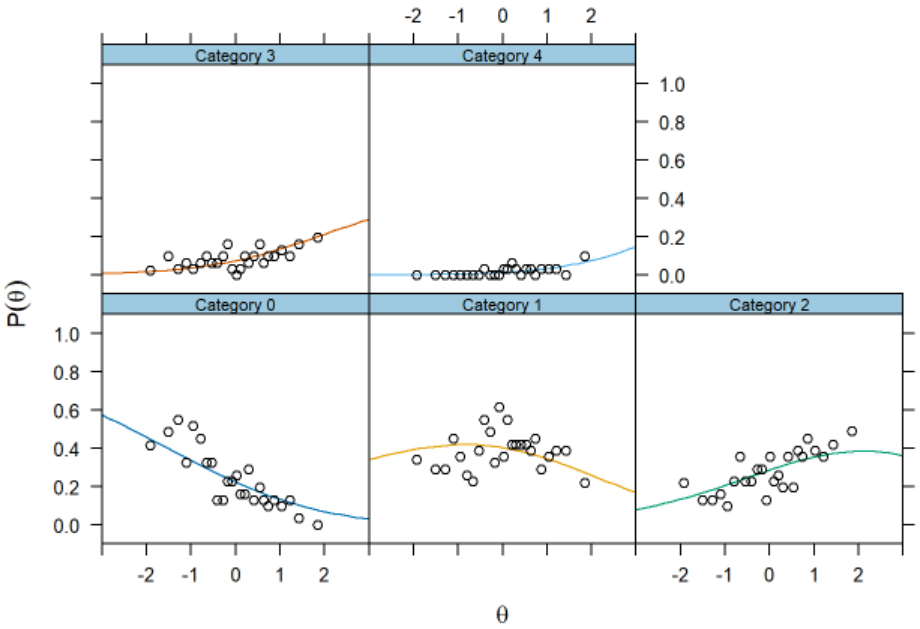
**Figur 3: *Lytting*: Ordering – exchange program**




Diskrimineringsverdien til oppgåvene skal i utgangspunktet vere over 0,8. Det kan likevel vere slik i enkelte tilfelle at IRT-analyser fungerer betre på dikotome oppgåver enn for polytome, m.a. fordi det i polytome oppgåver er fleire parameter å estimere for kvar oppgåve. Sidan polytome oppgåver gir fleire svaralternativ, er data meir komplekse og varierte. I oppgåva i Figur 3, er *a*-verdien 0,616, og altså noko under det som er ynskjeleg. Vi ser i figuren (og i tabell 7 ovanfor) at det svært vanskeleg å få to poeng, men lett å få tre poeng. Dette er fordi det i denne oppgåvetypen er ein intern avhengigheit. I oppgåva skal elevane plassere nøkkelord i riktig rekkefølge i forhold til innhaldet dei høyrer i lyttefila. Når dei har plassert dei første nøkkelorda riktig, vil det siste nøkkelordet gi seg sjølv. Det kan difor sjå ut som det er lettare å få tre poeng enn to.

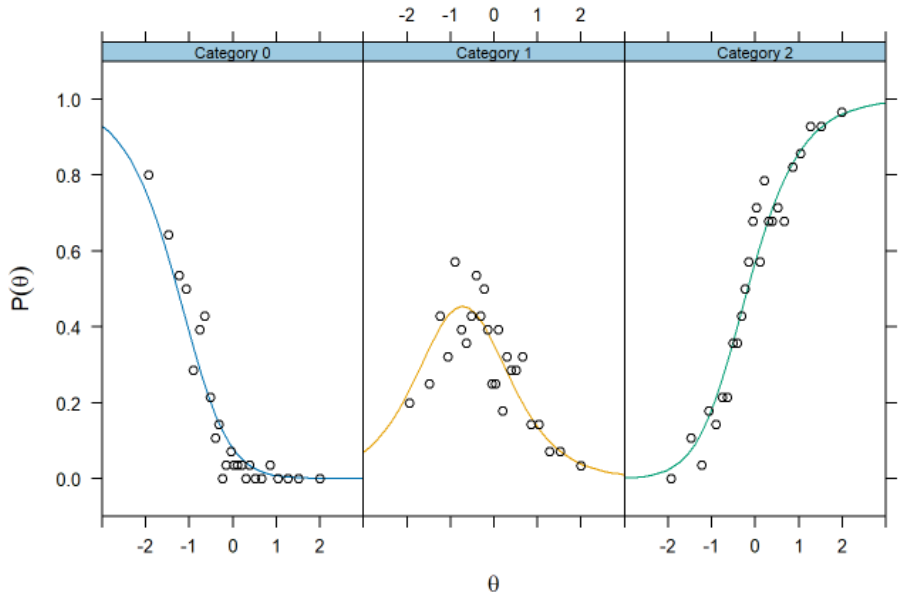
Vi har også tidlegare sett at denne oppgåvetypen oppfører seg ulikt andre oppgåver i analysar. Det er likevel viktig å inkludere denne då oppgåva måler at elevane kan forstå samanhengen i ein tekst, i samsvar med konstruktet til nasjonale prøver i engelsk. Det gir vidare eit signal til skular og elevar om at dette er viktig kompetanse, og noko dei kan jobbe med i klasserommet.

**Figur 4: *Lytting*: Matching - Debatecast**

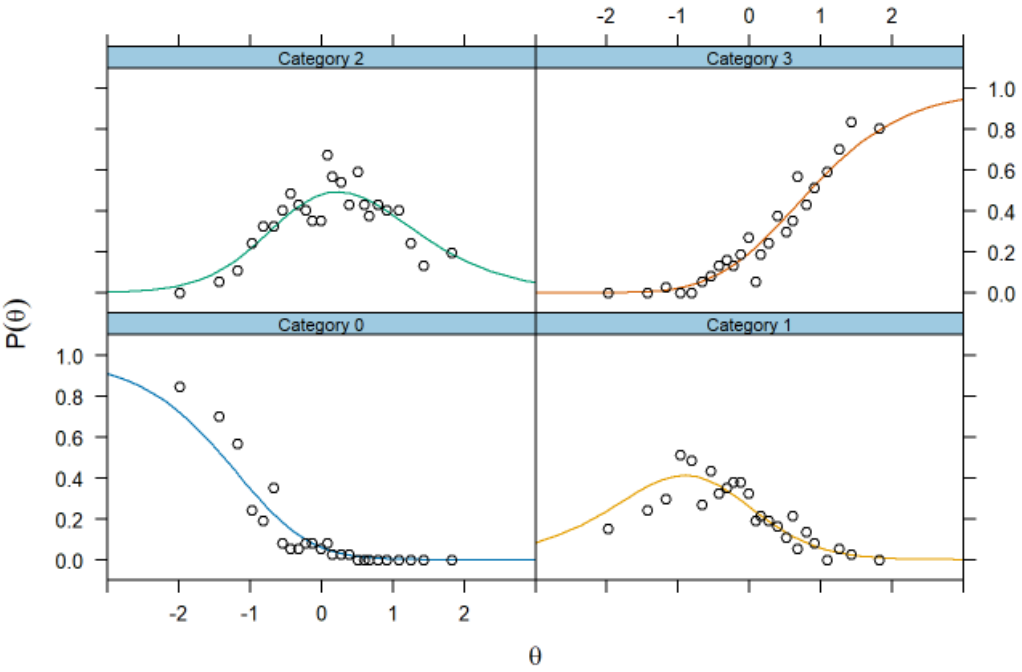



Oppgåva i Figur 4 viser også at *a*-verdien i denne oppgåva er under det som er ynskjeleg, nemleg 0,37. Vi vil likevel foreslå å inkludere denne oppgåva, fordi *b*-verdiane viser at det er ei stigande vanskegrad frå 0 poeng til 4 poeng: 0 poeng har *b*-verdi -1,593, 1 poeng har *b*-verdi 0,928, 2 poeng har *b*-verdi 3,587 og 3 poeng har *b*-verdi 4,807). Oppgåva er viktig pedagogisk, då elevane her høyrer ein lengre lyttetekst og skal deretter kople ulik informasjon dei høyrer i teksten. Vi har difor valt å prioritere det pedagogiske i denne samanhengen framføre det reint psykometriske.

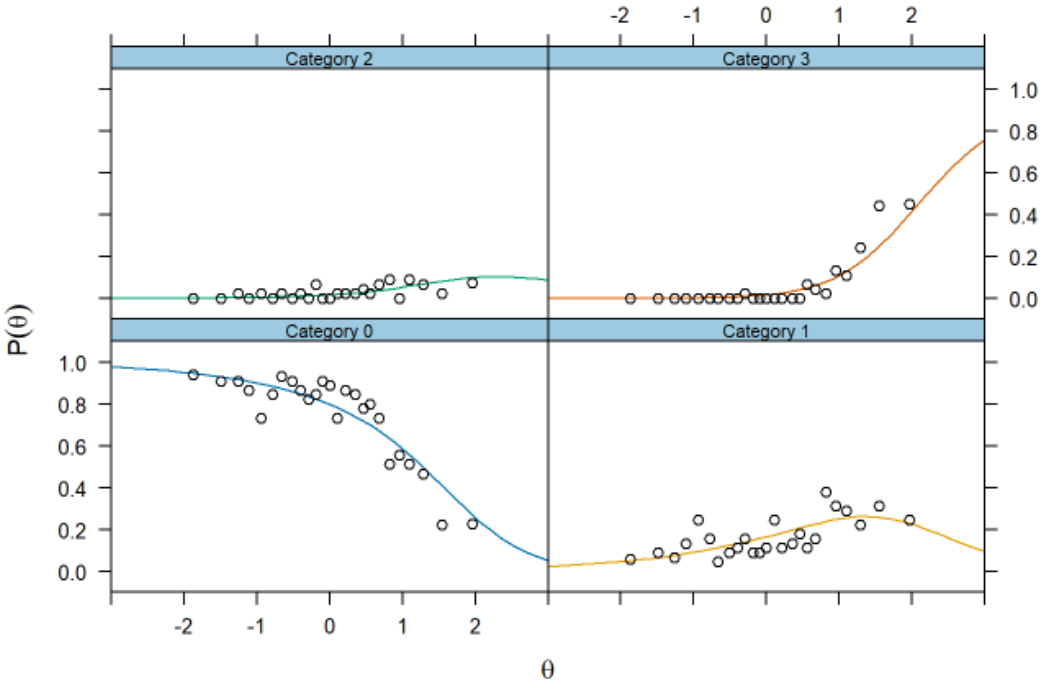
**Figur 5: *Lesing*: MC medium- Blood types**



**Figur 6: *Lesing:* Write words – dye jeans**




**Figur 7: *Lesing:* Move paragraph – crop rotation**



Oppgåva i figur 7 ovanfor har ein *a*-verdi på 0,723. Vi ser i figuren (og i tabell 7 ovanfor) at det på same måte som i oppgåva i figur 3 ovanfor, er vanskeleg å få eitt og to poeng, men lett å få tre poeng. Dette er fordi det også i denne oppgåvetypen er ein intern avhengigheit. Når elevane har plassert dei første avsnitta riktig, vil det siste svaret gi seg sjølv.

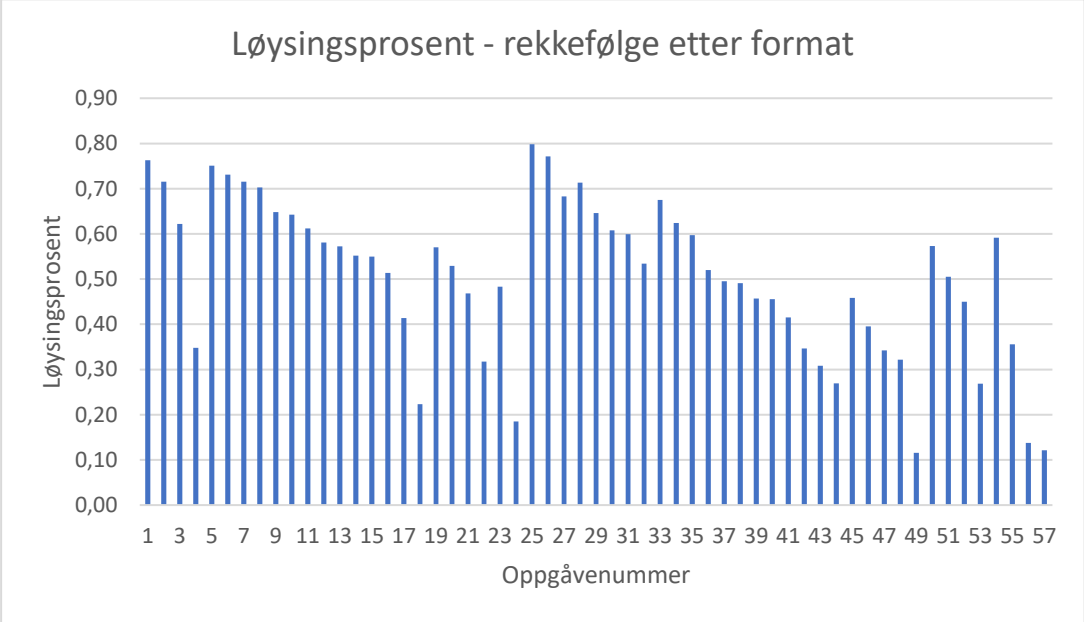
Vi inkluderer oppgåva for å måle at elevane kan forstå samanhengen i ein tekst, i samsvar med kontruktet til nasjonale prøver i engelsk.

### Løysingsprosent

Figur 8 under viser løysingsprosentane for dei ulike oppgåvene i prøven, alle kohort- og ankeroppgåver. Prøvesettet inkluderer som nemnt 24 ankeroppgåver (16 frå 2024 og forslag om 8 nye), men der alle elevar får alle kohortoppgåvene, kan kvar elev få maks fire ekstra oppgåver som nemnt tidlegare. Dette kan vere anker- og/eller utprøvingsoppgåver.

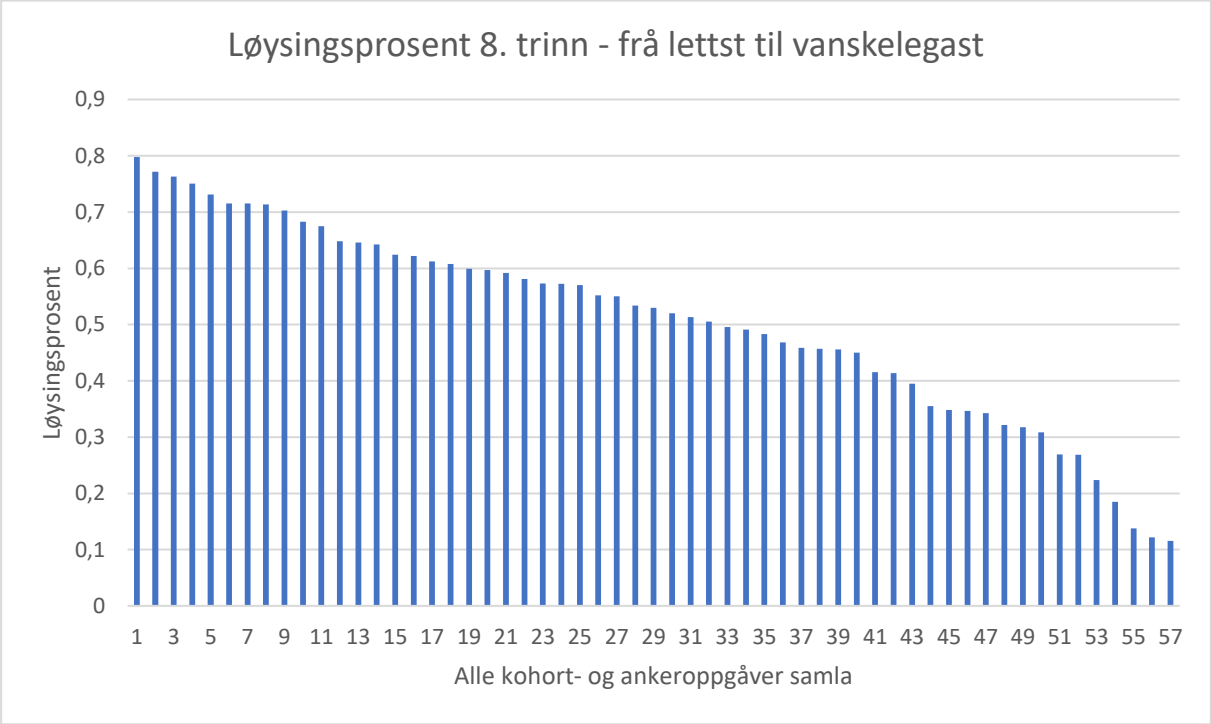
Det er ynskjeleg å inkludere oppgåver i prøven som varierer frå lette til vanskelege, slik at heile elevgruppa kan vise kva dei kan. Prøven inneheld fleire oppgåveformat, og alle oppgåver innanfor kvart format kjem etter kvarandre. Figuren viser korleis oppgåvene startar med lette i kvart format og deretter vert det i all hovudsak vanskelegare.

**Figur 8: Løysingsprosent i forslag til NP2025 8. trinn**



Figur 9 under viser løysingprosent på oppgåvene når desse er sortert frå høgast til lågast (og dette er ikkje same rekkefølge som i prøven). Vi ser i figuren at vanskegraden varierer frå den lettaste som 80 % av elevane har klart til den vanskelegaste som 12 % har klart. I tillegg til det grafen viser, er det t.d. fleire oppgåver i prøven der elevane kan få fleire poeng (polytome oppgåver). Grafen viser gjennomsnittskåre for desse. I éi av dei polytome oppgåvene er det få prosent av elevane som har fått full skåre: i *debatecast* har 2 % av elevane fått fire poeng og i *exchange program* har 8 % fått tre poeng (sjå tabell 8).

**Figur 9: Løysingsprosent i forslag til NP2025 8. trinn - sortert**




## 1.3 DIF mellom gutar og jenter

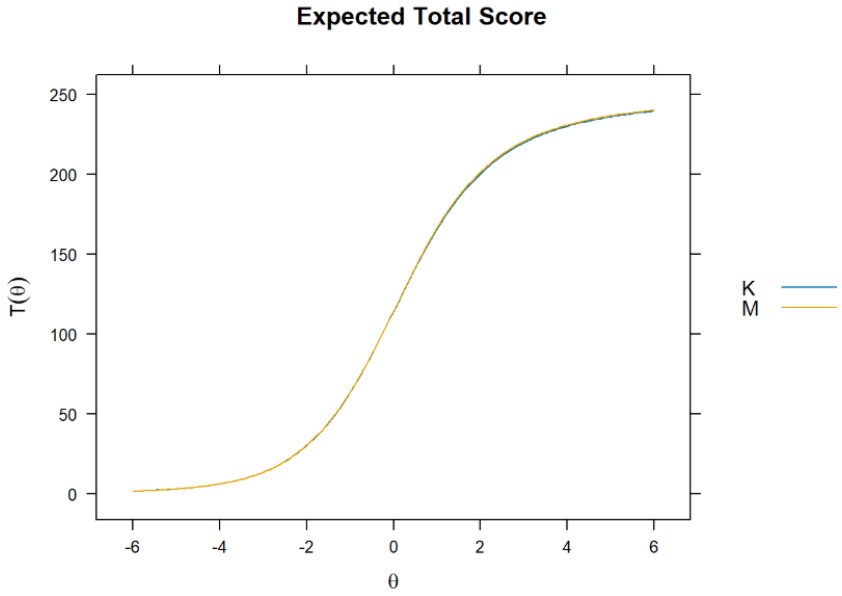
Det blei gjort ein DIF-analyse (*Differential Item Functioning*) i programmet *R* for å sjekke om det var forskjell på korleis gutar og jenter løyste oppgåvene, dvs. om det er oppgåver i prøven som favoriserer anten gutar eller jenter. Dersom ei oppgåve viser forskjellar (DIF) basert på kjønn, betyr det at gutar og jenter med lik dyktigheit/underliggande ferdigheit (ability) svarer på oppgåva systematisk forskjellig. Det kan tyde på at oppgåva favoriserer det eine kjønnet framfor det andre.

Ingen av oppgåvene foreslått for NPENG 2025 for 8. trinn inneheldt DIF for kjønn under utprøving.

Figur 10 og 11 under viser at det er liten forskjell mellom gutar og jenter for heile oppgåvesettet når det gjeld *Test characteristic curve* (TCC) og *Test information curve* (TIC).

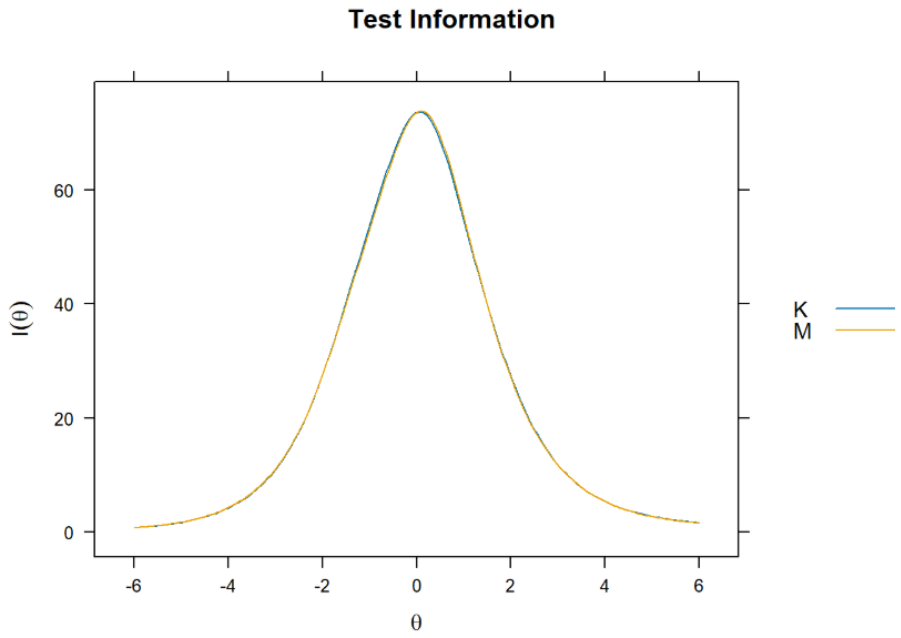
TCC i IRT-analysen viser samanhengen mellom ferdigheitsnivået (theta) til elevane og dei forventa poengsummane elevane vil få på prøven, for gutar og jenter.

**Figur 10: TCC - gutar og jenter for heile oppgåvesettet**




Figur 11 under viser *Test information curve* (TIC) mellom gutar og jenter for heile oppgåvesettet. Figuren skildrar kor mykje informasjon prøven gir om elevane sine ferdigheiter på dei ulike nivåa av ferdigheitsskalaen, for gutar og jenter.

**Figur 11: TIC - gutar og jenter for heile oppgåvesettet**




## 1.4 Oversikt over enkeltoppgåver 8. trinn

Tabell 8 under viser oversikt over alle oppgåvene som er foreslått til Nasjonal prøve i engelsk for 8. trinn 2025.

Tala i tabellen er frå *R*-analysar av alle oppgåvene analysert under eitt, frå gjennomføringa i 2024 for 8. trinn, 179 oppgåver totalt (kohort-, anker- og utprøvingsoppgåver).

Oppgåvenamna i tabellen er korta ned i forhold til namna i EPS (prøveadministrasjonssystemet til Udir). Rekkefølgje på oppgåvene kan bli endra i endeleg prøve.

LP i kolonna under står for Løysingsprosent.

**Tabell 8: Forslag til Nasjonal prøve i engelsk 8. trinn 2025 – oppgåve for oppgåve**

| Format | Oppgåvenamn | Tal på elevar deltatt | a-verdi | b | b1 | b2 | b3 | b4 | Snitt b-verdi | % svar per poeng - poly | Maks skåre | LP | P-verdi gj.snitt viss fleire poeng | Teksttype | Formål | Forståelses-nivå |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lytt-CP | house chores 2 | 1050 | 1.198 | -1.022 | | | | | -1.022 | | 1 | 0.716 | 0.716 | Dialog | Finne informasjon / forstå detaljar | Setning(ar) |
| Lytt-CP | Manual work2-women | 1123 | 0.92 | -0.605 | | | | | -0.605 | | 1 | 0.622 | 0.622 | Beskriving | Finne informasjon / forstå detaljar | Setning(ar) |
| Lytt-CP | men-Frank | 1059 | 1.23 | 0.666 | | | | | 0.666 | | 1 | 0.348 | 0.348 | Beskriving | Finne informasjon / forstå detaljar | Setning(ar) |
| Lytt-MCS | Creative | 628 | 1.599 | -0.99 | | | | | -0.990 | | 1 | 0.751 | 0.751 | Monolog | Tolke og forstå | Tekstkopling |
| Lytt-MCS | big things in NZ | 659 | 1.292 | -1.119 | | | | | -1.119 | | 1 | 0.731 | 0.731 | Reklame | Tolke og forstå | Tekstkopling |
| Lytt-MCS | space tourism | 610 | 1.226 | -0.629 | | | | | -0.629 | | 1 | 0.648 | 0.648 | Nyheit | Forstå hovudinnhald | Tekstkopling |
| Lytt-MCS | pros and cons-SchoolClosed | 588* | 1.586 | -0.422 | | | | | -0.422 | | 1 | 0.612 | 0.612 | Monolog | Forstå hovudinnhald | Tekstkopling |
| Lytt-MCS | TheParkingSpace | 594* | 0.997 | -0.396 | | | | | -0.396 | | 1 | 0.581 | 0.581 | Dialog | Forstå hovudinnhald | Tekstkopling |
| Lytt-MCS | DNA | 613 | 2.113 | -0.182 | | | | | -0.182 | | 1 | 0.573 | 0.573 | Monolog | Tolke og forstå | Tekstkopling |
| Lytt-MCS | PompeiiDiscovery | 616 | 1.19 | -0.214 | | | | | -0.214 | | 1 | 0.552 | 0.552 | Nyheit | Forstå hovudinnhald | Tekstkopling |
| Lytt-MCS | DogsEyeContact2 | 601 | 1.31 | -0.105 | | | | | -0.105 | | 1 | 0.514 | 0.514 | Nyheit | Forstå hovudinnhald | Tekstkopling |
| Lytt-MCL | Special Edition | 958 | 0.937 | | -1.74 | -0.295 | 1.051 | | -0.328 | 0=11%<br>1=30%<br>2=37%<br>3=22% | 3 | 1.711 | 0.570 | Monolog / podcast | 2 forstå detaljar / 1 forstå hovudinnhald | 2 setning(ar) / 1 tekstkopling |
| Lytt-MA | Debatecast | 795 | 0.37 | | -1.593 | 0.928 | 3.587 | 4.807 | 1.932 | 0=23%<br>1=38%<br>2=28%<br>3=9%<br>4=2% | 4 | 1.271 | 0.318 | Dialog / podcast | Finne informasjon / forstå detaljar | Tekstkopling |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lytt-OR | exchange program | 717 | 0.616 | | 1.459 | 4.666 | -1.456 | | 1.556 | 0=64%<br>1=26%<br>2=2%<br>3=8% | 3 | 0.556 | 0.185 | Monolog / infomøte | Forstå samanheng | Tekstkopling |
| Les-CD | NorwegianLandscape1 | 1920 | 1.195 | -1.302 | | | | | -1.302 | | 1 | 0.772 | 0.772 | Beskriving | Finne informasjon / forstå detaljar | Tekstkopling |
| Les-CD | Citysquareentrance2 | 1938 | 1.559 | -0.735 | | | | | -0.735 | | 1 | 0.683 | 0.683 | Beskriving | Tolke og forstå | Tekstkopling |
| Les-MCS | animal teeth | 741 | 1.227 | -0.589 | | | | | -0.589 | | 1 | 0.646 | 0.646 | Fakta /science | Forstå hovudinnhald | Setning(ar) |
| Les-MCS | dislike food- | 715 | 1.771 | -0.325 | | | | | -0.325 | | 1 | 0.608 | 0.608 | Fakta /science | Forstå hovudinnhald | Tekstkopling |
| Les-MCS | bear break in | 739 | 1.252 | -0.435 | | | | | -0.435 | | 1 | 0.599 | 0.599 | Nyheit | Forstå hovudinnhald | Tekstkopling |
| Les-MCM | Blood types | 704 | 1.368 | | -1.101 | -0.352 | | | -0.727 | 0=18%<br>1=30%<br>2=52% | 2 | 1.350 | 0.675 | Fakta / science | 1 forstå detaljar / 1 forstå hovudinnhald | 1 tekstkopling / 1 setning(ar) |
| Les-WS | ImpressivePeople11 | 679 | 2.184 | -0.3 | | | | | -0.300 | | 1 | 0.597 | 0.597 | Fakta / historie | Tolke og forstå | Setning(ar) |
| Les-WS | ImpressivePeople1 | 672 | 2.694 | 0.156 | | | | | 0.156 | | 1 | 0.457 | 0.457 | Fakta / historie | Tolke og forstå | Tekstkopling |
| Les-WS | ImpressivePeople13 | 675 | 1.823 | 0.219 | | | | | 0.219 | | 1 | 0.456 | 0.456 | Fakta / historie | Tolke og forstå | Setning(ar) |
| Les-WS | ImpressivePeople14 | 638 | 2.291 | 0.525 | | | | | 0.525 | | 1 | 0.347 | 0.347 | Fakta / historie | Finne informasjon / forstå detaljar | Tekstkopling |
| Les-WS | ImpressivePeople9 | 682 | 1.796 | 0.687 | | | | | 0.687 | | 1 | 0.309 | 0.309 | Fakta / historie | Tolke og forstå | Setning(ar) |
| Les-CW | lost cat | 877 | 2.188 | 0.154 | | | | | 0.154 | | 1 | 0.459 | 0.459 | Forteljing | Forstå avansert vokabular | Setning(ar) |
| Les-CW | CyclistJonas | 834 | 1.153 | 0.783 | | | | | 0.783 | | 1 | 0.322 | 0.322 | Forteljing | Forstå avansert vokabular | Setning(ar) |
| Les-CW | AncientHunters | 841 | 1.766 | 1.723 | | | | | 1.723 | | 1 | 0.116 | 0.116 | Fakta / science | Forstå avansert vokabular | Setning(ar) |
| Les-CS | Leftovers | 984 | 1.205 | 0.004 | | | | | 0.004 | | 1 | 0.505 | 0.505 | Grammatikk-setning | Bruke grammatiske strukturar | Setning(ar) |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Les-CS | Schoolwork | 982 | 2.081 | 0.191 | | | | | 0.191 | | 1 | 0.450 | 0.450 | Grammatikk-setning | Bruke grammatiske strukturar | Setning(ar) |
| Les-CS | FootbalPractice | 977 | 1.51 | 0.917 | | | | | 0.917 | | 1 | 0.269 | 0.269 | Grammatikk-setning | Bruke grammatiske strukturar | Setning(ar) |
| Les-WW | dye jeans | 943 | 1.272 | | -1.131 | -0.493 | 0.701 | | -0.308 | 0=15%<br>1=22%<br>2=35%<br>3=28% | 3 | 1.775 | 0.592 | Fakta / beskriving | Bruke grammatiske strukturar | Setning(ar) |
| Les-MP | crop rotation | 1140 | 0.723 | | 2.177 | 3.123 | 0.025 | | 1.775 | 0=74%<br>1=17%<br>2=3%<br>3=6% | 3 | 0.414 | 0.138 | Fakta / science | Forstå samanheng | Tekstkopling |

*Prøvedesignet var sett opp med eit mål om at alle utprøvings- og ankeroppgåvene skulle bli svart på av minimum 600 elevar (med eit utgangspunkt i at ca. 58 500 elevar deltok). I gjennomføringa var det 60673 elevar som deltok totalt. To oppgåver (sjå * i kolonna *Tal på elevar deltatt*) var likevel svart på av litt i underkant av 600 under gjennomføringa i 2024.

## Prøvedesign

Eit elevkull består vanlegvis av ca. 60 000 elevar. Elevane på 8. trinn skal svare på oppgåver i 13 seksjonar som inneheld kohort-, anker- og pilotoppgåver. Alle elevane svarar på alle kohortoppgåvene. I tillegg til kohortoppgåvene vert elevane tildelt éin tilfeldig underseksjon av 13 (med pilot- og ankeroppgåver) og i denne får kvar elev éi til fire ekstra oppgåver som anten er ankeroppgåver eller pilotoppgåver. Dersom dei t.d. er tildelt ein seksjon med lengre tekstar eller vanskelege oppgåver, får elevane i denne seksjonen berre éi oppgåve kvar. Er dei tildelt ein seksjon med lettare oppgåver, kan dei få fleire. I oppsettet for prøven har vi tatt utgangspunkt i 58 500 elevar. Kvar av dei 13 ekstra seksjonane vert då tildelt 4500 tilfeldige elevar (1/13). Slik sikrar vi at alle pilot- og ankeroppgåver vert svart på av minimum 600 elevar. Vidare er alle pilot- og ankeroppgåver sterkt lenka gjennom at *alle* elevane tek dei same kohortoppgåvene.

**Tabell 9: Eksempel på prøvedesign i EPS**

| | | |
|---|---|---|
| **Introduction** | Infoside | Alle elevar |
| | Infoside | Alle elevar |
| **Section 1 – Click picture** | | |
| **Section 1.1 - Listening Click picture** | Kohortoppgåver | Alle elevar |
| | | |
| **Section 1.2 - Listening Click picture** | Ankeroppgåve | Min. 600 elevar per oppgåve |
| (Elevar som får tildelt denne seksjonen får t.d. tre tilfeldige oppgåver kvar) | Pilotoppgåver 1 - 12 | |
| | | |
| **Section 2 – Listening Multiple choice short** | | |
| **Section 2.1 – Listening Multiple choice short** | Kohortoppgåver | Alle elevar |
| | | |
| **Section 2.2. – Listening Multiple choice short** | Ankeroppgåver | Min. 600 elevar per oppgåve |
| (Elevar som får tildelt denne seksjonen får t.d. fire tilfeldige oppgåver kvar) | Pilotoppgåver 1 - 27 | |
| | | |
| **Section 3 – 13 (andre format i prøven)** | Same inndelingar som seksjon 1 og 2 | |

# 2. Nasjonal prøve i engelsk 8. trinn høst 2025 - hovedresultat

## 2.1 Gjennomføring 2025

Den nasjonale prøven i engelsk for 8. trinn 2025, NPENG08 2025, ble gjennomført i tidsrommet 1. – 26. september 2025, i Utdanningsdirektoratets prøvegjennomføringssystem EPS – Eksamens- og prøvesystemet.

Totalt var det 60797 elever som gjennomførte prøven i september 2025, 30859 gutter og 29938 jenter. Prøven omfattet 33 kohortoppgaver, hvorav 27 var dikotome oppgaver og 6 var polytome.

Som forklart i Del 1 ble utprøving av oppgaver gjennomført samtidig med de nasjonale prøvene. Det ble gjort på den måten at alle elevene svarte på alle oppgavene i den nasjonale kohortprøven (2024), og i tillegg fikk de mellom én og fire ekstra oppgaver, enten ankeroppgaver og/eller oppgaver som ble prøvd ut med tanke på prøven i 2025.

Ankeroppgaver er oppgaver som går igjen fra år til år, og som gjør det mulig å måle trend, eventuell endring i engelskkompetanse hos elevene over tid. Mens svar på ankeroppgavene teller med i elevens endelige resultat, gjør ikke utprøvingsoppgavene det. Kohortoppgavene blir offentliggjort i ettertid. Dette gjelder ikke anker- og utprøvingsoppgavene.

I denne rapporten ser vi på resultat fra Nasjonal prøve i engelsk for 8. trinn 2025, kohortprøven.

## 2.2 Analyser av kohortprøven NPENG08

Vi har analysert data fra de nasjonale prøvene både ved hjelp av klassiske analyser i Excel og SPSS, og ved hjelp av IRT i analyseverktøyet R. Utdanningsdirektoratet har hjulpet oss med IRT-analyser (R). I de klassiske analysene vi har gjort, finner vi mål for reliabilitet, og vi kan si noe om forskjeller mellom gutter og jenter ved hjelp av t-tester. IRT analysene gir oss mål på theta (evne/kompetanse), vanskegrad (*b*-verdier) og diskriminering (*a*-verdier). Løsningsprosent (*p*-verdier) finner vi både ved hjelp av klassiske analyser og IRT-analyser.

De aller fleste oppgaveformatene har oppgaver som blir skåret dikotomt (0/1). Tre av lytteformatene inneholder oppgaver som er utviklet for polytom skåring (eleven svarer på flere deloppgaver og kan dermed få flere rette svar): *Ordering, Multiple choice long* og *Matching*. Tre av leseformatene inneholder oppgaver som er utviklet for polytom skåring: *Multiple Choice Medium*, *Write word* og *Move paragraph.* Hver av *Write word*-oppgavene, inneholder for eksempel tre luker der eleven skal skrive inn ett ord i hver luke.

Utdanningsdirektoratet bestemte at de ville slå sammen noen av svarkategoriene i to av de polytome oppgaveformatene (Matching og Ordering – begge to er lytteoppgaver) for at analysene skulle fungere bedre. Slik ble de polytome oppgaveformatene skåret i 2025:

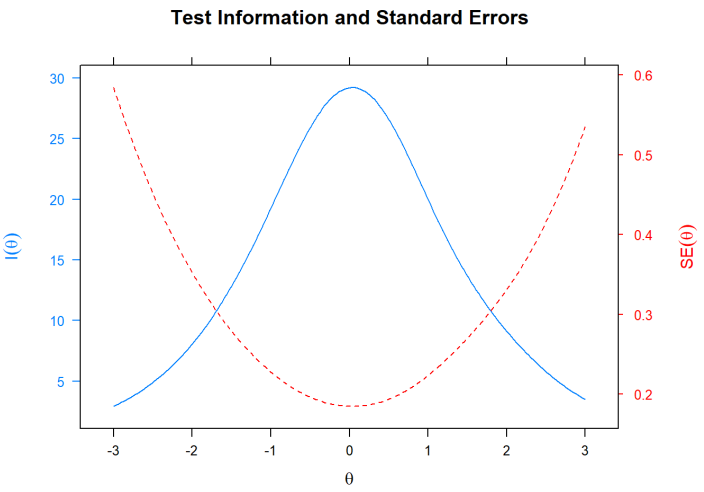
**Tabell 10. Omkoding/skåring polytome oppgaver NPENG08 2025**

| Oppgaveformat | Opprinnelege tal på svarkategorier | Svarkategorier etter omkoding som ligger til grunn for analysene |
|---|---|---|
| **Lytt Multiple choice long (MCL)** | 3 | 3 |
| **Lytt Order (OR)** | 3 | 2 |
| **Lytt Matching (MA)** | 3 | 2 |
| **Les Multiple choice medium (MCM)** | 2 | 2 |
| **Les Write Word (WW)** | 3 | 3 |
| **Les Move paragraph (MP)** | 3 | 3 |

### 2.2.1 Overordnet resultat

Figur 12 viser testinformasjon og måle(u)sikkerhet for prøven. Den blå linjen er informasjonskurven, og viser hvor godt prøven måler på forskjellige deler av ferdighetsskalaen, og den røde linjen viser måle(u)sikkerhet. Prøven gir mest sikker informasjon når theta (evnene/kompetansen til elevene) er i området -1 til +0,5. I dette området er det altså mest sikre målinger.

**Figur 12. Testinformasjon og standard måle(u)sikkerhet NPENG08 2025**




Tabell 11 viser statistikk for NPENG08 2025, nærmere bestemt antall oppgaver, mål på reliabilitet, gjennomsnittlige *p*-verdier og gjennomsnittlig klassisk mål på diskriminering (*a*).

**Tabell 11. Oppsummerende statistikk NPENG08 2025**

| | | | | |
|---|---|---|---|---|
| **Antall elever** | 60797 | | **Standardavvik** | 9,792 |
| **Antall oppgaver** | 33 (42) | | **Reliabilitet** | 0,904 |
| **Antall lytteoppgaver** | 14 (18) | | **Gjennomsnittlig løsningsprosent (*p*)** | 0,52 |
| **Antall leseoppgaver** | 19 (24) | | **Gjennomsnittlig *b*-verdi (dikotome oppgaver)** | -0,021 |
| **Gjennomsnittlig skår** | 23,03 | | **Gjennomsnittlig *a*-verdi** | 1,283 |

8. trinnsprøven er reliabel (0,904) og diskriminerer på en fornuftig måte mellom elevene (gjennomsnittlig *a*: 1,283). Prøven har en vanskegrad som er omtrent det vi sikter mot (gjennomsnittlig *p*- (0,52) og *b*-verdi (-0,021). Prøven diskriminerer også godt mellom elever på ulike ferdighetsnivåer. Lesedelen diskriminerer bedre enn lyttedelen. Det er vanlig å si at ved bruk av GPCM-modellen i IRT-analyse, så bør diskrimineringsmålet på individuelle oppgaver helst være over 0,8.

Tabell 12 fokuserer på lytting og lesing hver for seg og rapporterer mål for reliabilitet og gjennomsnittlige *p*-verdier.

**Tabell 12. Oppsummerende statistikk NPENG8 2025 lesing og lytting**

| NPE8 | Reliabilitet lytting | Reliabilitet lesing | Gjennomsnittleg *p*-verdi lytting | Gjennomsnittleg *p*-verdi lesing |
|---|---|---|---|---|
| **60797 elever** | 0,777 | 0,866 | 0,57 | 0,48 |

Tabell 12 viser at om vi ser på lytte- og leseoppgavene hver for seg, så er lesedelen mest reliabel med en reliabilitetskoeffesient på 0,866. Lyttedelen har en reliabilitetskoeffesient på 0,777. Reliabiliteten vil gå ned når antall oppgaver blir redusert, slik som er tilfelle når vi her deler prøven i to - lesing og lytting.

At lesedelen er mer reliabel enn lyttedelen kommer trolig av at teamet som utvikler den nasjonale prøven, har mer erfaring med å utvikle leseoppgaver. Tabell 12 viser at lyttedelen er lettere enn lesedelen.

Den gjennomsnittlige *p*-verdien etter utprøving var 0,51 (se tabell 3). Den reelle prøven hadde en gjennomsnittlig *p*-verdi på 0,52, vanskegraden for hele prøven har ikke endret seg mye fra utprøving.

### 2.2.2 Resultat på oppgavenivå

#### *Oversikt over alle oppgavene*

Tabell 13 gir en oversikt over oppgaveformat, hva da ulike oppgavene måler, pluss psykometriske mål som karakteriserer oppgavene i kohortprøven: løsningsprosent, *b*-verdier og *a*-verdier.

**Tabell 13. Løsningsprosent (*p*-verdier) og mål på diskriminering (*a*)**

| Rekke-følge | Lytt (L) Les (R) | Format | Oppgavenavn | Oppgaven måler | *p* | *b* | *a* |
|---|---|---|---|---|---|---|---|
| 1 | L | **CP** | House chores 2 | Finne informasjon / forstå detaljer | 0,59 | -0,438 | 0,944 |
| 2 | L | **CP** | Manual work 2 | Finne informasjon / forstå detaljer | 0,70 | -1,044 | 0,942 |
| 3 | L | **CP** | Frank | Finne informasjon / forstå detaljer | 0,43 | 0,321 | 1,135 |
| 4 | L | **MCS** | Big things in NZ | Tolke og forstå | 0,80 | -1,384 | 1,337 |
| 5 | L | **MCS** | The parking space | Forstå hovedinnhold | 0,57 | -0,382 | 0,814 |
| 6 | L | **MCS** | Creative | Tolke og forstå | 0,74 | -1,052 | 1,367 |
| 7 | L | **MCS** | DNA | Tolke og forstå | 0,57 | -0,276 | 1,569 |
| 8 | L | **MCS** | Dogs eye contact | Forstå hovedinnhold | 0,53 | -0,146 | 1,168 |
| 9 | L | **MCS** | Pompeii discovery | Forstå hovedinnhold | 0,55 | -0,221 | 1,028 |
| 10 | L | **MCS** | Space tourism | Forstå hovedinnhold | 0,67 | -0,801 | 1,143 |
| 11 | L | **MCS** | School closed | Forstå hovedinnhold | 0,56 | -0,281 | 1,177 |
| 12 | L | **MCL** | Special edition | 1 tolke og forstå, 2 forstå hovedinnhold | Snitt: 0,56<br>Full skår: 21,1% | b1 -1,594<br>b2 -0,447<br>b3 1,148 | 0,946 |
| 13 | L | **OR** | Exchange program | Forstå sammenheng | Snitt: 0,24<br>Full skår: 10,0% | b1 1,369<br>b2 2,202 | 0,607 |
| 14 | L | **MA** | Debatecase | Finne informasjon / forstå detaljer | Snitt: 0,42<br>Full skår: 2,0% | b1 -2,216<br>b2 5,346 | 0,750 |
| 15 | R | **CD** | City square entrance 2 | Tolke og forstå | 0,68 | -0,821 | 1,195 |
| 16 | R | **CD** | Norwegian landscape 1 | Finne informasjon / forstå detaljer | 0,79 | -1,381 | 1,229 |
| 17 | R | **MCS** | Animal teeth | Forstå hovedinnhold | 0,69 | -0,870 | 1,175 |
| 18 | R | **MCS** | Bear break in | Forstå hovedinnhold | 0,59 | -0,394 | 1,109 |
| 19 | R | **MCS** | Dislike food | Forstå hovedinnhold | 0,59 | -0,310 | 1,793 |
| 20 | R | **MCM** | Blood types | 1 forstå hovedinnhold / 1 finne informasjon | Snitt: 0,70<br>Full skår: 54,8% | b1 -1,240<br>b2 -0,402 | 1,285 |
| 21 | R | **CW** | Ancient hunters | Forstå avansert vokabular | 0,12 | 1,826 | 1,498 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 22 | R | **CW** | Cycling Jonas | Forstå avansert vokabular | 0,31 | 1,111 | 0,857 |
| 23 | R | **CW** | Lost cat | Forstå avansert vokabular | 0,45 | 0,171 | 1,718 |
| 24 | R | **WS** | Impressive people 11 | Tolke og forstå | 0,59 | -0,299 | 2,165 |
| 25 | R | **WS** | Impressive people 13 | Tolke og forstå | 0,43 | 0,276 | 1,596 |
| 26 | R | **WS** | Impressive people 14 | Finne informasjon / forstå detaljer | 0,36 | 0,460 | 2,136 |
| 27 | R | **WS** | Impressive people 1 | Tolke og forstå | 0,61 | -0,331 | 2,310 |
| 28 | R | **WS** | Impressive people 9 | Tolke og forstå | 0,34 | 0,744 | 1,104 |
| 29 | R | **CS** | Football practice | Bruke grammatiske strukturer | 0,29 | 0,830 | 1,558 |
| 30 | R | **CS** | Leftovers | Bruke grammatiske strukturer | 0,50 | 0,022 | 1,232 |
| 31 | R | **CS** | Schoolwork | Bruke grammatiske strukturer | 0,44 | 0,243 | 1,604 |
| 32 | R | **WW** | Dye jeans | Bruke grammatiske strukturer | Snitt: 0,59<br>Full skår: 31,7% | b1 -0,936<br>b2 -0,470<br>b3 0,404 | 1,140 |
| 33 | R | **MP** | Crop rotation | Forstå sammenheng | Snitt: 0,13<br>Full skår: 5,60% | b1 2,289<br>b2 3,157<br>b3 0,068 | 0,698 |

Tabell 13 gir en oversikt over alle oppgavene i kohortprøven med oppgaveformat, gjennomsnittlige *p*- og *b*-verdier og mål på diskriminering og *a*-verdier.

Tabellen viser at den letteste oppgaven (oppgave 4) svarte 80% av elevene rett på, mens den vanskeligste dikotome oppgaven (oppgave 21) svarte omtrent 12% rett på.  Den alle vanskeligste oppgaven er polytom, oppgave 14. Bare 2% fikk full skår på den.

### Løsningsprosent (p-) og b-verdier

Tabell 13 ovenfor viser at prøven inneholder flere lette dikotome oppgaver enn vanskelige. Det er i alt 27 oppgaver i prøven som blir skåret dikotomt. Tabellene 14 og 15 gir en oversikt over lette og vanskelige dikotome oppgaver i prøven.

**Tabell 14. Distribusjon av *p*-verdier**

| ***p*-verdi i området** | **Antall oppgaver** |
|---|---|
| 0,8 – 0,9 | 1 |
| 0,7 – 0,8 | 3 |
| 0,6 – 0,7 | 4 |
| 0,5 – 0,6 | 10 |
| 0,4 – 0,5 | 4 |
| 0,3 – 0,4 | 3 |
| 0,2 – 0,3 | 1 |
| 0,1 – 0,2 | 1 |

**Tabell 15. Distribusjon av *b*-verdier**

| ***b*-verdi i området** | **Antall oppgaver** |
|---|---|
| Under -1,5 | 0 |
| -1,0 - -1,5 | 4 |
| -0,5 - -1,0 | 3 |
| 0 - -0,5 | 10 |
| 0,5 – 0 | 6 |
| 1 – 0,5 | 2 |
| 1,5 - 1 | 1 |
| Over 1,5 | 1 |

Det er 18 dikotome oppgaver som flere enn 50% av elevene svarer rett på og 9 som færre enn 50% av elevene svarer rett på. 17 oppgaver har en *b*-verdi under 0, og 10 oppgaver har en *b*-verdi over 0.

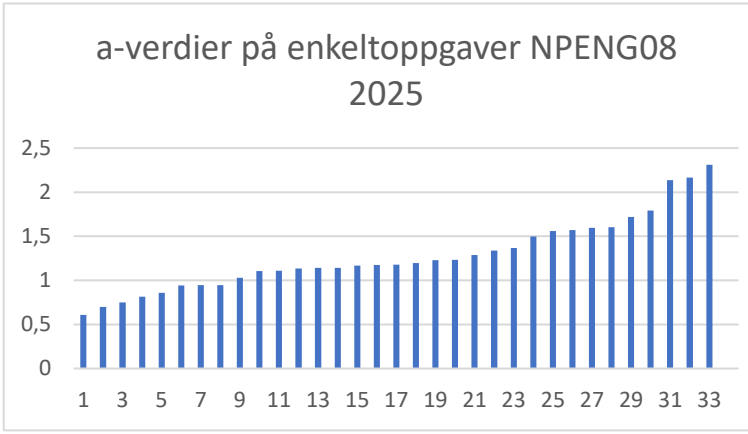
Det er seks oppgaver i kohortprøven som blir skåret polytomt, tre lytteoppgaver og tre leseoppgaver.

### Diskriminering – a-verdi

En oppgave som diskriminerer godt, skiller godt mellom elever på ulike kompetansenivå, for eksempel mellom elever på mestringsnivå 3 og de på mestringsnivå 4. Når vi analyserer resultatene i R, vil Utdanningsdirektoratet at vi bruker GPCM-modellen (Generalized Partial

Credit Model) der en ønsker at oppgavene skal ha et mål på diskriminering, en *a*-verdi, over 0,8. Tabell 13 viser *a*-verdien for hver av oppgavene i kohortprøven. Vi ser at de fleste oppgavene i kohortprøven har *a*-verdier over 0,8. Figur 13 og tabell 16 viser hvordan *a*-verdiene på oppgavene i kohortprøven fordeler seg.

**Figur 13. *a*-verdier på enkeltoppgaver NPENG08 2025**




**Tabell 16. Fordeling av *a*-verdier NPENG08 2025**

| ***a*-verdier i området** | **Antall oppgaver** |
|---|---|
| Over 2,0 | 3 |
| 1,5 – 2,0 | 6 |
| 1,0 – 1,5 | 16 |
| 0,8 – 1,0 | 5 |
| Under 0,8 | 3 |

Figur 13 og Tabell 16 viser at oppgavene i prøven diskriminerer stort sett godt. Tre av oppgavene har *a*-verdier i området under 0,8, mens 30 oppgaver har *a*-verdier over 0.8. De tre med *a*-verdier under 0,8 hadde *a*-verdier over 0,8 da de ble prøvd ut i 2024.

## Theta

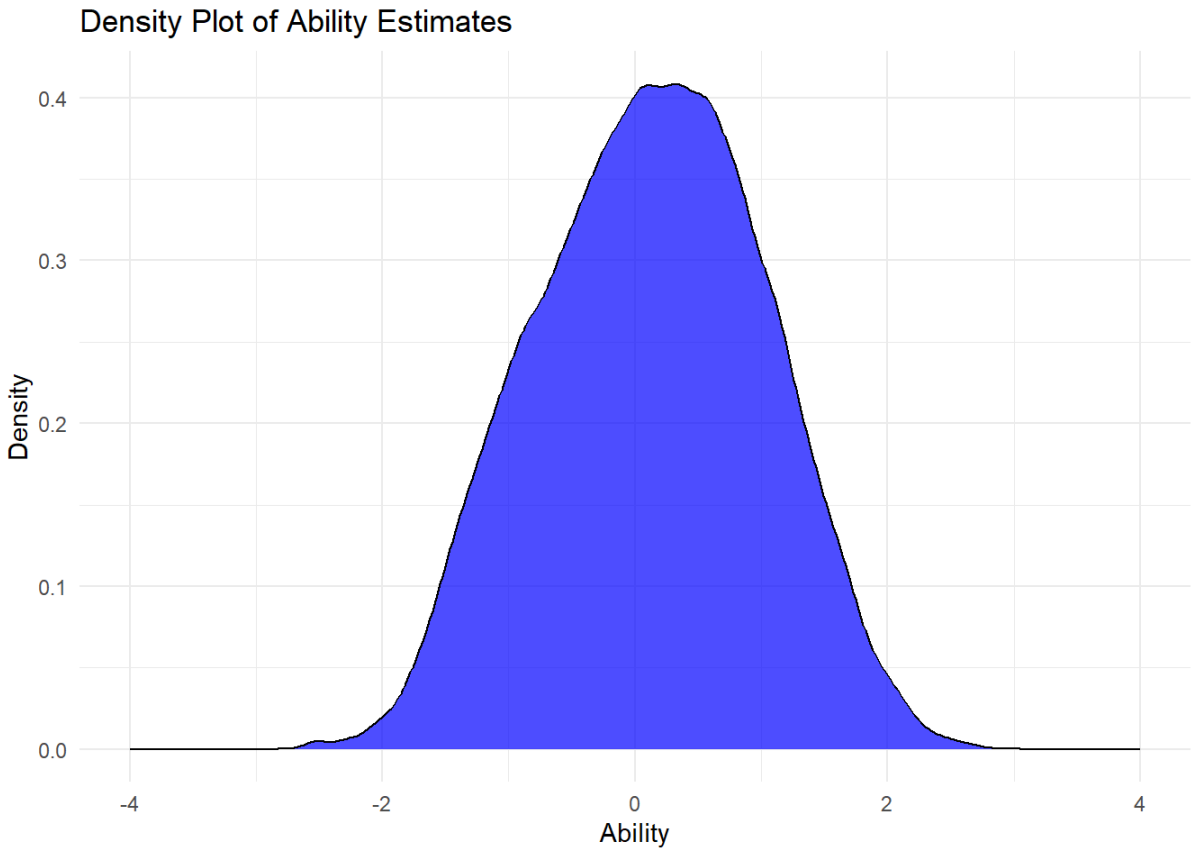
Tabell 17 og Figur 14 oppsummerer thetamål i 8. trinns populasjonen. Tabellen og figuren er fra analyser der ankeroppgavene er inkludert. Vi ser i Tabell 17 at distribusjonen av theta (ability) i 8. trinnsgruppen varierer fra -2,580 til +2,862.

**Tabell 17. Oppsummerende statistikk for theta 8. trinn 2025**

| **Elever** | **Gjennom-snitt** | **Standard-avvik** | **Minimum** | **Median** | **Max** |
|---|---|---|---|---|---|
| 60797 | 0,118 | 0,891 | -2,580 | 0,142 | 2,862 |

Variasjonen er framstilt grafisk i figur 14.

**Figur 14. Distribusjon av theta i 8. trinns-populasjonen**.




Vi ser at distribusjonen av theta i 8. trinnspopulasjonen er tilnærmet normalfordelt. Data fra prøven viser altså at det finnes elever som har svært lave ferdigheter i engelsk, og samtidig finnes det de med høye ferdigheter i engelsk. De aller fleste elevene har ferdigheter et eller annet sted mellom -1 og +1.

## 2.3 Kjønn

Tabell 18 gir en oversikt over forskjellen i løsningsprosent/*p*-verdi på enkeltoppgaver og kjønn.

**Tabell 18. Løsningsprosent – gutter og jenter NPENG08 2025**

| Rekke-følge | Lytt (L) Les (R) | Oppgavenavn | *p* | *p* gutter | *p* jenter | *Differanse gutter/jenter* |
|---|---|---|---|---|---|---|
| 1 | L | House chores 2 | 0,59 | 0,598 | 0,574 | 0,024 |
| 2 | L | Manual work 2 | 0,70 | 0,716 | 0,677 | 0,039 |
| 3 | L | Frank | 0,43 | 0,453 | 0,404 | 0,049 |
| 4 | L | Big things in NZ | 0,80 | 0,822 | 0,782 | 0,04 |
| 5 | L | The parking space | 0,57 | 0,572 | 0,563 | 0,009 |
| 6 | L | Creative | 0,74 | 0,748 | 0,741 | 0,007 |
| 7 | L | DNA | 0,57 | 0,601 | 0,548 | 0,053 |
| 8 | L | Dogs eye contact | 0,53 | 0,526 | 0,542 | -0,016 |
| 9 | L | Pompeii discovery | 0,55 | 0,601 | 0,492 | 0,109 |
| 10 | L | Space tourism | 0,67 | 0,716 | 0,631 | 0,085 |
| 11 | L | School closed | 0,56 | 0,585 | 0,544 | 0,041 |
| 12 | L | Special edition | Snitt: 0,56 | Snitt: 0,560 | Snitt: 0,560 | 0 |
| 13 | L | Exchange program | Snitt: 0,24 | Snitt: 0,213 | Snitt: 0,264 | Snittdifferanse: -0,051 |
| 14 | L | Debatecase | Snitt: 0,42 | Snitt: 0,428 | Snitt: 0,412 | Snittdifferanse: 0,016 |
| 15 | R | City square entrance 2 | 0,68 | 0,684 | 0,681 | 0,003 |
| 16 | R | Norwegian landscape 1 | 0,79 | 0,787 | 0,794 | -0,007 |
| 17 | R | Animal teeth | 0,69 | 0,687 | 0,696 | -0,009 |
| 18 | R | Bear break in | 0,59 | 0,616 | 0,557 | 0,059 |
| 19 | R | Dislike food | 0,59 | 0,585 | 0,594 | -0,009 |
| 20 | R | Blood types | Snitt: 0,70 | Snitt: 0,691 | Snitt: 0,703 | Snittdifferanse: -0,012 |
| 21 | R | Ancient hunters | 0,12 | 0,121 | 0,113 | 0,008 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 22 | R | Cycling Jonas | 0,31 | 0,332 | 0,278 | 0,054 |
| 23 | R | Lost cat | 0,45 | 0,449 | 0,459 | -0,01 |
| 24 | R | Impressive people 11 | 0,59 | 0,58 | 0,607 | -0,027 |
| 25 | R | Impressive people 13 | 0,43 | 0,44 | 0,414 | 0,026 |
| 26 | R | Impressive people 14 | 0,36 | 0,346 | 0,382 | -0,036 |
| 27 | R | Impressive people 1 | 0,61 | 0,616 | 0,594 | 0,022 |
| 28 | R | Impressive people 9 | 0,34 | 0,327 | 0,357 | -0,03 |
| 29 | R | Football practice | 0,29 | 0,3 | 0,279 | 0,021 |
| 30 | R | Leftovers | 0,50 | 0,491 | 0,50 | -0,009 |
| 31 | R | Schoolwork | 0,44 | 0,414 | 0,458 | -0,044 |
| 32 | R | Dye jeans | Snitt: 0,59 | Snitt: 0,568 | Snitt: 0,603 | Snittdifferanse: -0,035 |
| 33 | R | Crop rotation | Snitt: 0,13 | Snitt: 0,131 | Snitt: 0,127 | Snittdifferanse: 0,004 |

Differansen mellom gruppene finnes i kolonnen helt til høyre. De oppgavene som jentene gjør det best på har en negativ verdi. Vi ser at av de 33 oppgavene kohortprøven inneholder, gjør jentene det best på 13 av oppgavene, mens guttene gjør det best på 19 av oppgavene, med én oppgave med ingen forskjell. Jentene skårer høyest på flest leseoppgaver, mens guttene skårer høyest på flest lytteoppgaver. Det er ingen oppgaver med DIF i kohortprøven 2025.

Tabell 19 viser forskjeller mellom gutter og jenter i gjennomsnittlig totalskår, løsningsprosent og theta.

**Tabell 19. Forskjeller mellom gutter og jenter i gjennomsnittlig skår og løsningsprosent**

| Gjennomsnittlige estimat | | | | |
|---|---|---|---|---|
| | **Totalskår (42)** | **Lytting (18)** | **Lesing (24)** | **Løsnings-prosent** |
| Gutter | 21,15 | 9,90 | 11,25 | 0,524 |
| Jenter | 20,89 | 9,53 | 11,36 | 0,513 |

Tabell 19 viser at guttene skårer i gjennomsnitt litt høyere enn jentene totalt og på lyttedelen, og jentene skårer litt høyere på lesedelen, men forskjellene er marginale.

En t-test understreker også forskjeller mellom gutter og jenter i 8. trinns-populasjonen. Tabell 20 viser forskjellene mellom gutter og jenter basert på en t-test. Vi ser at gjennomsnittlig totalskår for guttene (21,15) er litt høyere enn den for jentene (20,89). Standardavviket er nesten likt i de to gruppene, litt høyere for jentene. Dette forteller oss at variasjonen i jentegruppen og guttegruppen er omtrent lik.

**Tabell 20. Forskjeller mellom jenter og gutter på 8. trinn – t-test**

| **Prøve** | **Kjønn** | **Gjennomsnittlig antall rette svar** | **n** | **Standardavvik** | ***t*** | **Sig 2-tailed** |
|---|---|---|---|---|---|---|
| **NPENG08 2025** | Gutter | 21,15 | 30859 | 8,856 | 3,656 | < 0,001 |
| | Jenter | 20,89 | 29938 | 9,055 | | |

Tabell 20 viser at med utgangspunkt i råskåren, er det en signifikant forskjell mellom gutter og jenter i favør av guttene. Når dette resultatet er signifikant, så betyr det at forskjellene med all sannsynlighet ikke er tilfeldige.

Om vi tar utgangspunkt i totalskåren, så skårer guttene litt høyere enn jentene. Denne forskjellen er signifikant på 0,001-nivå.

**Tabell 21. Forskjeller mellom gutter og jenter i lytting og lesing – t-test**

| Prøve | Kjønn | Gjennomsnittlig antall rette svar - lytting | n | Standardavvik | *t* | Sig 2-tailed |
|---|---|---|---|---|---|---|
| NPENG08 2025 | Gutter | 9,90 | 30859 | 3,671 | 12,035 | < 0,001 |
| | Jenter | 9,53 | 29938 | 3,909 | | |
| | | **Gjennomsnittlig antall rette svar - lesing** | **n** | **Standardavvik** | ***t*** | **Sig 2-tailed** |
| | Gutter | 11,25 | 30859 | 5,768 | -2,248 | 0,025 |
| | Jenter | 11,36 | 29938 | 5,681 | | |

Om vi kjører separate t-tester for lytting og lesing, ser vi at mens forskjellen mellom gutter og jenter er signifikant når det gjelder lytting (sig <0,001), så er den mindre signifikant for lesing (sig 0,025). Dette vil si at forskjellen mellom kjønnene sannsynligvis ikke er tilfeldig i begge tilfeller, men mer signifikant når det gjelder lytting.

## 2.4 NP engelsk for 8. trinn 2025 – oppsummering

Tidligere i rapporten har vi fokusert på hvordan Nasjonal prøve i engelsk for 8. trinn fungerte i 2025. Vi har sett på hovedresultat, ulike oppgaveformater, kjønnsforskjeller.

I all hovedsak har vi konkludert med at prøven ser ut til å fungere bra for målgruppene. Nærmere bestemt:

- Den nasjonale prøven i engelsk for 8. trinn 2025 måler lytting og lesing.
- Hovedprøven måler et utvalg av kompetansemål i læreplanen.
- Prøven er reliabel. Reliabilitetskoeffesienten er 0,904.
- Prøven diskriminerer godt mellom elevene. Gjennomsnittlig *a* for hovedprøven er 1,283.
- Prøven skiller godt mellom elever på ulike ferdighetssnivåer.
- Både de dikotome og polytome oppgavene ser ut til å fungere noenlunde greit.
- Guttene skårer høyere enn jentene totalt og på lyttedelen. Jentene skårer litt høyere enn guttene på lesedelen. Disse forskjellene er signifikante, altså sannsynligvis ikke tilfeldige.
- Forskjellene i vanskegrad fra utprøving til reell prøve er mindre enn det som har vært tilfelle før da vi prøvde ut oppgaver et par uker før de nasjonale prøvene ble gjennomført. Gjennomsnittlig *p*-verdi etter utprøving var 0,51 i prøven som ble foreslått. Data fra den nasjonale prøven viser en gjennomsnittlig *p*-verdi på 0,52. Det betyr at vanskegraden totalt er nesten den samme etter utprøving som i reell prøve.

Når en prøve skal tilpasses så ulike elevgrupper som de som representerer mestringsnivåene 1 - 5, må det bli slik at mange oppgaver er vanskelige for elever på mestringsnivå 1 samtidig som mange oppgaver er lette for elever på mestringsnivå 5. Likevel viser resultatet at prøven inneholdt oppgaver som utfordrer elever på alle mestringsnivåer.


UNIVERSITAS
BERGENSIS