



UNIVERSITETET I BERGEN

Nasjonale prøvar i engelsk for 8. trinn 2023

Rapport del 1

Utprøving september 2022 - Forslag til prøvar (mars 2023)

Rapport del 2

Gjennomføring av prøvar september 2023

Rapport april 2024

Hildegunn Lahlum Helness, Eli Moe og team

Institutt for framandspråk

Innhold

.....	1
OM RAPPORTEN	3
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
Om itemformat	9
DEL 1 – UTPRØVING AV NP ENGELSK 8. TRINN HAUSTEN 2022 OG FORSLAG TIL PRØVAR 2023	13
1.1 Generell informasjon om utprøvinga til prøven i 2023.....	13
1.2 Forslag til prøvar 8. trinn 2023	15
1.3 Kva måler prøvane på 8. trinn?.....	19
1.4 Oversikt over enkeltitem 8. trinn	20
DEL 2 – NASJONALE PRØVAR I ENGELSK 8. TRINN 2023	24
2. Nasjonal prøve i engelsk 8. trinn haust 2023 - hovudresultat	24
2.1 Gjennomføring 2023.....	24
2.2 Analysar av kohortprøven NPENG08.....	24
2.2.1 Overordna resultat.....	25
2.2.2 Resultat på oppgåvenivå	27
2.3 Kjønn	34
2.4 NP engelsk for 8. trinn 2023 – oppsummering	38

OM RAPPORTEN

Del 1 av rapporten gjer greie for arbeidet før 2023-utgåva av Nasjonal prøve i engelsk for 8. trinn blir gjennomført: utprøving av item og forslag til endeleg prøve. Del 2 er skriva etter at prøvane er gjennomførte hausten 2023 og summerer opp resultatane 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 målar, itemformat og korleis itema 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 éin.

Denne rapporten fokuserer på utprøving av item i 2022 og Nasjonal prøve i engelsk for 8. trinn 2023. I samband med IRT-analysar er datamaterialet for alle itema som var med i gjennomføringa for 8. trinn lagt til grunn.

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 reglar 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 meningsfulle 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 hovudinnhald i tilpassa tekstar av varierende omfang frå forskjellige kjelder
- reflektere over innhald i tekstar
- forstå og bruke eit ordforråd og omgrep som dekker egne interesser og aktuelle emne
- forstå tydinga av ord og uttrykk ut frå samanhengen dei er brukt i
- følge reglar for rettskriving, ordbøying og setningsstruktur

Desse aspekta blir testa på ulike måtar (format) og i ulike kombinasjonar, og ikkje som ådskilte ferdigheiter. For eksempel må ein kunne forstå ord og omgrep for å kunne finne informasjon eller forstå hovudinnhaldet i ein 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 refererer 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 strategiar 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 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 sammenhengen 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 sakprosa- og engelskspråklege 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 tilhøyrighet**

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 vidare læring og tilpasse opplæring 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 ferdigheter 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 6, som er ei oversikt over forslaget til prøven på 8. trinn, viser at teksttypene i lyttedelen er beskrivingar, monologar, kunngjeringar, dialogar, podcast og nyheitsmeldingar. Lesedelen inneheld beskrivingar, forteljingar og faktatekstar. *Click and drag* i lesedelen beskriv 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.

Elevane 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 itemformat

Tabell 1a: Itemformat i den nasjonale prøven 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 <i>Listen to the recording and click on the correct picture. You can listen to the recording twice.</i> 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 <i>Listen to the recording. You can see the question on the screen while you listen. Click on the correct answer. You can listen to</i>

		<i>the recording twice.</i> 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 <i>Listen to the recording. You can see the questions on the screen while you listen. Answer the <u>three</u> questions. You can listen to the recording twice.</i> 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 <i>Listen to the recording. Click on the key words and place them in the order they are mentioned. The first key word is given. You can move the key words while you listen. You can listen to the recording twice.</i> 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 <i>Listen to the recording. Give your answer by clicking <u>three</u> of the boxes in the table. You can listen to the recording twice.</i> Elevane ser ein tabell på skjermen med stikkord i kolonneoverskrifter, og med små bilde i første kolonna nedover til venstre. Basert på innhaldet i teksten dei lyttar til, skal dei matche stikkorda med bildene, ved å krysse av tre av avkryssingsboksane.

Tabell 1b: Itemformat i den nasjonale prøven 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 eit stort bilde og bilde av fire eller fem 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 evt. ei liste eller faktaboks.
Click word (CW)	Opent	Elevane blir presentert for ein kort tekst og ein instruksjon, t.d. <i>Click on the word that means almost the same as 'fascinated'</i> . 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 fire 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)	Lukka	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 ei 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 i 2022, slik at elevane kan gjere seg kjende med oppgåveformata på førehand.

DEL 1 – UTPRØVING AV NP ENGELSK 8. TRINN HAUSTEN 2022 OG FORSLAG TIL PRØVAR 2023

1.1 Generell informasjon om utprøvinga til prøven i 2023

Hausten 2022 vart alle oppgåvene prøvd ut i eit nytt design for fyrste gong, i samband med reell gjennomføring av Nasjonale prøvar 2022. Alle elevane på 8. trinn tok alle oppgåvene i NPENG 2022 (kalla *kohort*-oppgåver). I tillegg fekk kvar elev ei til tre ekstra oppgåver som anten var utprøvingsoppgåver til NPENG 2023 eller ankeroppgåver (til trendmåling). 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 itema vert lettare frå utprøving til reell gjennomføring, mest sannsynleg fordi elevane gjer ein større innsats i reell gjennomføring. I rapporten etter gjennomføring i 2023 vil me samanlikne data frå utprøvinga med reell gjennomføring, for å sjå om det er mindre gap mellom utprøving og gjennomføring.

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 2022 er blitt analysert i R med IRT-metoden, av Utdanningsdirektoratet.

UiB har brukt Excel til samansetning av oppgåvene som er valt ut til forslag til NPENG2023, 8. trinn.

Heile datamaterialet for 8. trinn vart analysert under eitt, det vil seie at data er frå analysar av alle itema samla som var med i gjennomføringa; 180 item (36 *kohortoppgåver*, 17 *ankeroppgåver* og 126 *utprøvingsoppgåver*).

Tidspunkt for utprøving

September 2022

Utprøvingsdesign

Vi har brukt eit delvis lenka design. Alle elevar fekk 1 til 3 ekstra oppgåver (anker- eller utprøvingsoppgåver) under gjennomføring av NPENG 2022, 8. trinn.

Tal på item som vart prøvd ut

126 item

Tal på elevar som deltok

Dei same elevane som tok kohortoppgåvene i engelsk 8. trinn i 2022: 56 525 (28769 gutar og 27756 jenter).

Analyse av data:

Data frå gjennomføringa er analysert ved hjelp av *R* og Excel. Det er brukt Item Response Theory (IRT)-kalibrering med ein 2PL IRT-modell for dikotome oppgåver og med Generalised partial credit model (GPCM) for polytome oppgåver.

Samansetjing av 2023-prøven

I arbeidet med å setje saman eit forslag til prøve for 2023, måtte vi balansere følgjande faktorar: lytteitem, leseitem, itemformat, item totalt, føremål, tekstlengd, tema, løysingsprosent, eventuelle kjønnsforskjellar, og diskriminering.

1.2 Forslag til prøvar 8. trinn 2023

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

Prøven i 2023 vil innehalde stort sett dei same itemformata og tal på oppgåver per format, som i 2022-prøven. Unnataket er at prøven i 2023 har eit nytt format i lytting, *Multiple choice long*, med ein lengre lyttetekst som gir inntil 3 poeng. Elles viser tabell 2 under, tal på item per format for foreslått prøve for 8. trinn, og eventuelle forskjellar i høve 2022-prøvene.

Tabell 2: Item per format foreslått for NP 2023, 8. trinn

Format	Forslag til Nasjonal prøve 2023 8. trinn
Lytting: Click picture	3 item (2 i 2022)
Lytting: Multiple choice short	9 item (12 i 2022)
Lytting: Multiple choice long	1 (3) item* (nytt format)
Lytting: Ordering	1 (3) item*
Lytting: Matching	1 (3) item*
Lesing: Click and drag	2 item
Lesing: Multiple choice short	3 item
Lesing: Multiple choice medium	2 (4) item*
Lesing: Who could say	4 item
Lesing: Click word	3 item
Lesing: Complete sentence	2 item
Lesing: Write words	1 (3) item*
Lesing: Move paragraph	1 (3) item*
Tal på item totalt	33 item (45)

* Disse 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.

Vi ser i Tabell 2 at det er med tre oppgåver mindre i *Multiple choice short*-formatet i lytting i 2023 enn i 2022. Grunnen til dette er at vi i 2023 har inkludert ei ny oppgåve i eit nytt format, *Multiple choice long*. For at lyttedelen av prøven ikkje totalt sett skal ta lengre tid i år enn i fjor, tok vi difor ut tre kortare oppgåver.

Dei aller fleste oppgaveformata har item 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 er representert i parentesane i Tabell 2, 3 og 5.

Tabell 3 nedanfor gir ei oversikt over vanskegrad (b -verdiar), løysingsprosent (p -verdiar) og diskrimineringsindeks (a -verdiar) i forslag til 2023-prøven for 8. trinn, med estimat frå R .

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 0,514 viser t.d. at 51,4 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 her vere over 0,8.

Gjennomsnittleg løysingsprosent (p -verdiar) for heile prøven som er foreslått til NP2023 er noko høgre enn i fjor. Tilsvarende er vanskegraden (b -verdiar) noko mindre. Dette er basert på ei forventning om at oppgåvene som er med i prøven i 2023 ikkje blir lettare under reell gjennomføring enn under utprøving, då utprøvinga var ein del av sjølv gjennomføringa av NP2022.

Tabell 3: Gjennomsnittleg vanskegrad og diskriminering, forslag til NP 8. trinn 2023

	Nasjonal prøve 8. trinn
Tal på item	33 (45)*
Gjennomsnittleg p -verdi	0,514
Gjennomsnittleg b -verdi	-0,049
Gjennomsnittleg a -verdi	1,5

*Talet i parentes viser maks poengsum

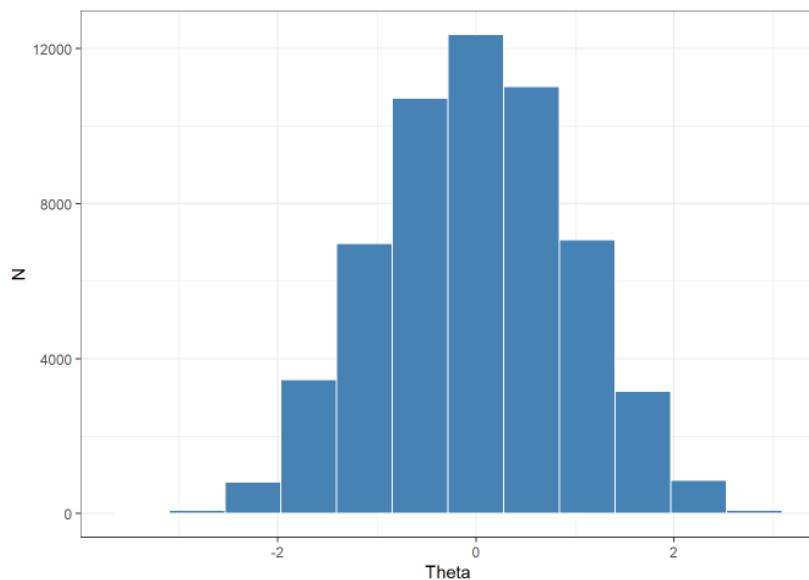
I hovudprøven for 8. trinn finst det sju polytome oppgåver, der elevane svarar på fleire item og får ein samleskår for kvar av desse oppgåvene. Tabell 4 gir ei oversikt over b -verdiar for dei polytome oppgåvene.

Tabell 4: Fordeling av b -verdier for polytome oppgåver

Itemformat	Eitt rett svar	To rette svar	Tre rette svar
<i>Lytting</i> : MC long – Transport changes	-1,301	0,386	2,303
<i>Lytting</i> : Ordering – Competition	-0,891	0,094	0,919
<i>Lytting</i> : Matching - Podcast opinions	-0,686	0,86	1,936
<i>Lesing</i> : MC medium- Surfing goat	-1,609	0,326	-
<i>Lesing</i> : MC medium – Meteorite	-0,493	1,745	-
<i>Lesing</i> : Write words – Silent film stars	-0,447	0,322	1,32
<i>Lesing</i> : Move paragraph – Gaming Mum	0,926	1,738	2,063

Figur 1 under viser fordelinga av θ -estimat for alle itema (lytting og lesing) som var med i gjennomføringa i 2022 (kohort-, anker- og utprøvingsoppgåver).

θ er eit uttrykk for elevane sine ferdigheiter/evner/kompetanse (ability). Minusverdier viser til lågare ferdigheit, og plussverdier viser til høgare. Vi ser at dei fleste itema har θ rundt null.

Figur 1: Thetaestimat frå gjennomføring 2022

Figur 2 under viser testinformasjon for alle kalibrerte item i sjølve prøven. Test information function (TIF) viser grad av informasjon i prøven for kvart θ -nivå.

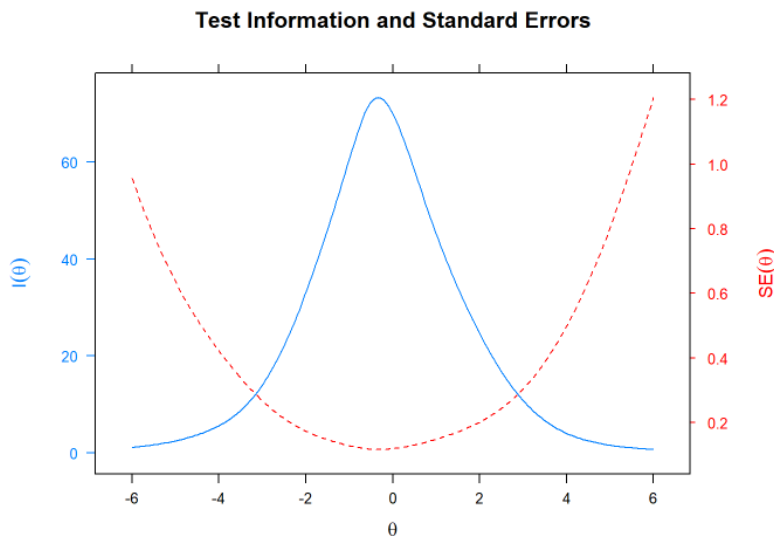
Figur 2: Test Information and Standard errors

Figure 5 displays the test information function (TIF) in blue with conditional standard errors of measurement (CSEM) in the dashed red line.

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

Tabell 5: gjennomsnittleg diskriminering, løysingsprosent, og vanskegrad, lytting og lesing kvar for seg

	Lytting	Lesing
Tal på item	15 (21)	18 (24)
Gjennomsnittleg diskriminering (α -verdi)	1,12	1,81
Gjennomsnittleg p -verdi	0,546	0,486
Gjennomsnittleg b -verdi	-0,219	0,093

*Tala i parentes viser maks poengsum

Prøven måler reseptive ferdigheiter. Vi ser i Tabell 5 at lytteoppgåvene i gjennomsnitt er noko lettare enn leseoppgåvene – gjennomsnittleg p -verdi for lytting er 0,546 mot 0,486 i lesing. Analysar frå piloteringa viser at fleire av lytteoppgåvene er noko lettare enn ønska, spesielt bildeoppgåvene. Kan hende er elevane også blitt flinkare i lytting. Vi jobbar framleis med å lage nok vanskelege lytteitem, mellom anna har vi lagt til eitt nytt format, *Multiple choice long*, der elevane kan få opptil tre poeng. Sjølv om gjennomsnittet for lytting er noko høgre enn ynskjeleg, ventar vi ikkje at dette skal stige frå utprøving til pilotering, som nemnt tidlegare. Den samla p -verdien for alle oppgåvene, for reseptive ferdigheiter, er som vi ser i Tabell 3 ovanfor 0,514.

1.3 Kva måler prøvane på 8. trinn?

Lista under viser korleis oppgåvene i forslaget til 2023-prøvene fordeler seg over dei aspekta som prøvane skal måle.

I enkelte av oppgåvene må elevane gi meir enn eitt svar (polytome oppgåver) og tala i parentes viser til kor mange rette svar det er mogleg å få totalt innanfor kvart aspekt.

Lesedelen i nasjonale prøvar i engelsk for 8. trinn måler:

Finne informasjon: 4 (5) item

Lesedelen inneheld fire item der elevane skal finne informasjon. I eitt av itema kan elevane finne svaret i eitt ord eller uttrykk. I tre tilfelle kan elevane finne svaret i ei setning.

Forstå hovudinnhald: 4 (5) item

Lesedelen inneheld fire item der elevane skal forstå hovudinnhald. I alle itema kan elevane finne svaret ved å kople informasjon frå ulike deler av teksten.

Tolke og forstå: 3 item

Lesedelen inneheld tre item der elevane til ei viss grad må tolke teksten. I alle itema finst svaret i ei setning.

Forstå avansert vokabular: 3 item

Lesedelen inneheld tre item der eleven er avhengig av å forstå relativt avansert vokabular for å finne svaret. I slike oppgåver kan eleven forstå kva ordet betyr ut frå samanhengen i teksten.

Bruke ord, stave- og grammatiske mønster: 3 (5) item

Lesedelen inneheld tre item der eleven skal følge enkle reglar for rettskriving og setningsstruktur. I alle tilfella må dei forstå setningane for å svare riktig.

Forstå samanheng i ein tekst: 1 item (3)

Lesedelen inneheld eitt item der elevane må vise at dei forstår samanhengen i eit tekst. Dette er ei *Move paragraph*-oppgåve der elevane kan oppnå tre poeng. Dei må kople informasjon frå fleire stader i teksten for å svare riktig.

Lyttedelen i nasjonale prøvar i engelsk for 8. trinn måler:

Forstå detaljar: 4 (8) item

Lyttedelen inneheld fire item der elevane skal forstå detaljar. I to av itema kan elevane finne

svaret i ei setning. I to tilfelle kan dette gjerast ved å kople informasjon frå ulike deler av teksten.

Forstå hovudinnhald: 9 item

Lyttedelen inneheld ni item der elevane skal forstå hovudinnhald. I alle tilfella kan dette gjerast ved å kople informasjon frå ulike deler av teksten.

Tolke og forstå: 1 item

Lyttedelen inneheld eitt item der eleven må tolke innhaldet i teksten for å svare riktig. Svaret finn dei i enkle ord.

Forstå samanheng: 1 (3) item

Lyttedelen inneheld éi oppgåve der elevane må forstå samanhengen i teksten, og svare ved å plassere fleire nøkkelord i rett rekkefølge i forhold til teksten dei høyrer. Dei må kople informasjon frå ulike deler av teksten.

Elevar som har øvd på å lese og lytte med ulikt fokus og med bruk av 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 tekstar av ulik lengd og kompleksitet.

1.4 Oversikt over enkeltitem 8. trinn

Tabell 6 under viser oversikt over alle itema som er foreslått til Nasjonal prøve i engelsk for 8. trinn 2023.

Tala i tabellen er frå *R*-analysar av alle itema analysert under eitt, frå gjennomføringa i 2022 for 8. trinn, 180 item totalt (kohort-, anker- og utprøvingsoppgåver).

Tabell 6: Forslag til Nasjonal prøve i engelsk 8. trinn 2023 – item for item

Rekkefølge*	Format	Oppgavenamn	Tal på elevar svart	a-verdi	b-verdi	% elev-svar per poeng	Maks skåre **	Løysingsprosent	Løysingsprosent jenter	Løysingsprosent gutar	Forskjell jenter / gutar	Teksttype	Formål	Forståelsesnivå
1	Lytting-CP	House chores 1	945	1,102	-1,08		1	0,723	0,717	0,729	-0,01	Beskriving	Forstå detaljar	Setning
2	Lytting-CP	Kids occupation photographer	983	0,968	-1,701		1	0,792	0,734	0,85	-0,12	Beskriving	Forstå detaljar	Tekstkopling
3	Lytting-CP	Activity 1-juggler	932	2,233	-1,468		1	0,881	0,86	0,902	-0,04	Beskriving	Forstå detaljar	Setning
4	Lytting-MCS	Goals on fire	313	1,233	-0,339		1	0,567	0,513	0,62	-0,11	Nyheitsmelding	Forstå hovudinnhald	Tekstkopling
5	Lytting-MCS	Teasing	325	1,133	0,522		1	0,371	0,337	0,404	-0,07	Dialog	Tolke og forstå	Vokabular
6	Lytting-MCS	Salmon sushi	320	1,117	0,02		1	0,496	0,521	0,471	0,05	Monolog	Forstå hovudinnhald	Tekstkopling
7	Lytting-MCS	London car free	-	1,024	-0,623		1	0,631	0,648	0,613	0,04	Nyheitsmelding	Forstå hovudinnhald	Tekstkopling
8	Lytting-MCS	Sour sweets	313	1,005	0,873		1	0,339	0,333	0,344	-0,01	Monolog	Forstå hovudinnhald	Tekstkopling
9	Lytting-MCS	Embarrasment	298	0,995	-0,566		1	0,616	0,662	0,57	0,09	Monolog	Forstå hovudinnhald	Tekstkopling
10	Lytting-MCS	Phone call work	318	0,925	-0,771		1	0,630	0,654	0,605	0,05	Dialog	Forstå hovudinnhald	Tekstkopling
11	Lytting-MCS	Presentation Shampoo	291	0,876	-0,21		1	0,527	0,48	0,573	-0,09	Monolog	Forstå hovudinnhald	Tekstkopling
12	Lytting-MCS	Debate	333	0,841	0,858		1	0,329	0,344	0,313	0,03	Kunngjering	Forstå hovudinnhald	Tekstkopling

13	Lytting-MCL	Transport changes	1122	1,129	Snitt 0,463	0=24% 1=36% 2=30% 3=11%	3	Snitt 0,424 (full skåre: 11%)	1,29	1,253	0,04	Dialog	1 forstå hovudinnhald/2 forstå detaljar	1 tekstkopling / 2 setning
14	Lytting-OR	Competition	1089	1,206	Snitt 0,041	0=30% 1=22% 2=18% 3=30%	3	Snitt 0,494 (full skåre: 30%)	1,321	1,645	-0,32	Monolog	Forstå samanheng	Tekstkopling
15	Lytting-MA	Podcast opinions	1106	0,946	Snitt 0,703	0=37% 1=30% 2=15% 3=17%	3	Snitt 0,375 (full skåre: 17%)	1,162	1,088	0,07	Podcast	Forstå detaljar	Tekstkopling
16	Lesing-CD	Sneakers	1663	1,45	-0,049		1	0,503	0,52	0,486	0,03	Beskriving	Finne informasjon	Tekstkopling
17	Lesing-CD	Restaurant2 pizza	1622	1,298	-0,055		1	0,512	0,556	0,468	0,09	Beskriving	Finne informasjon	Tekstkopling
18	Lesing-MCS	Embarrass dog 1	651	1,614	-1,353		1	0,830	0,847	0,813	0,03	Forteljing	Forstå hovudinnhald	Tekstkopling
19	Lesing-MCS	Snake selfie accident	659	1,309	-1,049		1	0,732	0,742	0,721	0,02	Fakta	Forstå hovudinnhald	Tekstkopling
20	Lesing-MCS	Raising money	648	2,08	-0,951		1	0,767	0,767	0,767	0,00	Forteljing	Forstå hovudinnhald	Tekstkopling
21	Lesing-MCM	Surfing goat	662	1,485	Snitt -0,642	0=15% 1=44% 2=41%	2	Snitt 0,631 (full skåre: 41%)	1,26	1,263	0,00	Fakta	1 forstå hovudinnhald / 1 finne informasjon	1 tekstkopling / 1 setning
22	Lesing-MCM	Meteorite	643	1,217	Snitt 0,626	0=39% 1=46% 2=15%	2	Snitt 0,380 (full skåre: 15%)	0,703	0,815	-0,11	Fakta	1 forstå hovudinnhald / 1 finne informasjon	1 tekstkopling / 1 setning
23	Lesing-CW	Bus hound1	335	2,297	0,125		1	0,452	0,411	0,492	-0,08	Fakta	Forstå avansert vokabular	Setning
24	Lesing-CW	Thousands of chocolate bars1	328	1,976	0,144		1	0,469	0,454	0,483	-0,03	Fakta	Forstå avansert vokabular	Setning
25	Lesing-CW	Helen Sherman	314	1,693	0,362		1	0,376	0,394	0,358	0,04	Fakta	Forstå avansert vokabular	Setning

26	Lesing-WS	Hero-High2-Angela	558	3558,407	-0,478		1	0,654	0,673	0,635	0,04	Forteljing	Tolke og forstå	Setning
27	Lesing-WS	Hero-High11-Mrs-Gibbs	593	3,251	-0,454		1	0,640	0,62	0,659	-0,04	Forteljing	Finne informasjon	Tekstkopling
28	Lesing-WS	Hero-High5-Linda	562	1,962	0,568		1	0,352	0,393	0,31	0,08	Forteljing	Tolke og forstå	Setning
29	Lesing-WS	Hero-High9-Claire	603	1,254	0,974		1	0,293	0,283	0,303	-0,02	Forteljing	Tolke og forstå	Setning
30	Lesing-CS	Passive must be done	710	1,656	-0,027		1	0,507	0,501	0,512	-0,01	Grammatikksetning	Bruke grammatiske strukturar	Setning
31	Lesing-CS	Direct-indirect 6-could not	727	1,585	1,961		1	0,091	0,101	0,08	0,02	Grammatikksetning	Bruke grammatiske strukturar	Setning
32	Lesing-WW	Silent film stars	1050	1,704	Snitt 0,398	0=38% 1=21% 2=23% 3=18%	3	Snitt 0,401 (full skåre: 18%)	1,219	1,184	0,04	Fakta	Bruke grammatiske strukturar	Setning
33	Lesing-MP	Gaming mum	1106	1,411	Snitt 1,576	0=72% 1=14% 2=4% 3=10%	3	Snitt 0,171 (full skåre: 10%)	0,53	0,496	0,03	Forteljing	Forstå samanheng	Tekstkopling

*Rekkefølge på itema kan bli endra i endeleg prøve.

**Dei grønne felta viser polytome oppgåver.

I kolonna *p-verdi* for polytome oppgåver (grønne celler), er det i tillegg til gjennomsnittleg *p-verdi* for heile oppgåva, også oppgitt kor mange prosent av elevane som fekk full skåre for oppgåva under utprøving, t.d. ca. 10 % elevar har oppnådd 3 poeng i oppgåve 33 (*GamingMum*).

DEL 2 – NASJONALE PRØVAR I ENGELSK 8. TRINN 2023

2. Nasjonal prøve i engelsk 8. trinn haust 2023 - hovudresultat

2.1 Gjennomføring 2023

Den nasjonale prøven i engelsk for 8. trinn 2023, NPENG08 2023, vart gjennomført i tidsrommet 4. – 29. september 2023, i prøvegjennomføringssystemet EPS – Eksamens- og prøvesystemet.

Totalt var det 62071 elevar som gjennomførte prøven i september 2023, 31488 gutar og 30583 jenter. Prøven omfatta 33 kohortoppgåver, 27 dikotome oppgåver og 6 polytome.

Som forklart i Del 1 vart utprøving av oppgåver gjennomført samtidig som dei nasjonale prøvane vart gjennomførte. Dette vart gjort på den måten at alle elevane svarte på alle oppgåvene i den nasjonale kohortprøven (2023), og i tillegg fekk dei mellom ei og fire ekstra oppgåver, enten ankeroppgåver eller oppgåver som vart pilotert med tanke på prøven i 2024.

Ankeroppgåver er oppgåver som går igjen frå år til år, og som gjer det mogeleg å måle trend, eventuell endring i engelskkompetanse hos elevane over tid. Medan svar på ankeroppgåvene tel med i elevens endelege resultat, gjer ikkje utprøvingsoppgåvene det. Kohortoppgåvene blir offentleggjort i ettertid. Dette gjeld ikkje anker- og utprøvingsoppgåvene.

I denne rapporten ser vi på resultat frå Nasjonal prøve i engelsk for 8. trinn 2023, kohortprøven.

2.2 Analysar av kohortprøven NPENG08

Vi har analysert data frå dei nasjonale prøvane både ved hjelp av klassiske analysar i Excel og SPSS, og ved hjelp av IRT (R). Utdanningsdirektoratet har hjelpt oss med IRT-analysar (R). I dei klassiske analysane vi har gjort, finn vi mål for reliabilitet, og vi kan seie noko om skilnader mellom gutar og jenter ved hjelp av t-testar. IRT analysane gir oss mål på theta (evne/kompetanse), vanskegrad (*b*-verdiar) og diskriminering (*a*-verdiar). Løysingsprosent (*p*-verdiar) finn vi både ved hjelp av klassiske analysar og IRT-analysar.

Dei aller fleste oppgåveformata har item som blir skåra dikotomt (0/1). Tre av lytteformata inneheld oppgåver som er utvikla for polytom skåring (eleven svarar på fleire deloppgåver og

kan dermed få fleire rette svar): *Ordering*, *Multiple choice long* og *Matching*. Tre av leseformata inneheld item som er utvikla for polytom skåring: *Multiple Choice Medium*, *Write words* og *Move paragraph*. Kvar av *Write words*-oppgåvene, til dømes, inneheld tre luker der eleven skal skrive inn eitt ord i kvar luke.

Utdanningsdirektoratet bestemte at dei ville slå saman nokre av svarkategoriene i tre av dei polytome oppgåveformata (to lytteformat og eitt leseformat) for at oppgåvene skulle fungere betre. Slik vart dei polytome oppgåveformata skåra i 2023:

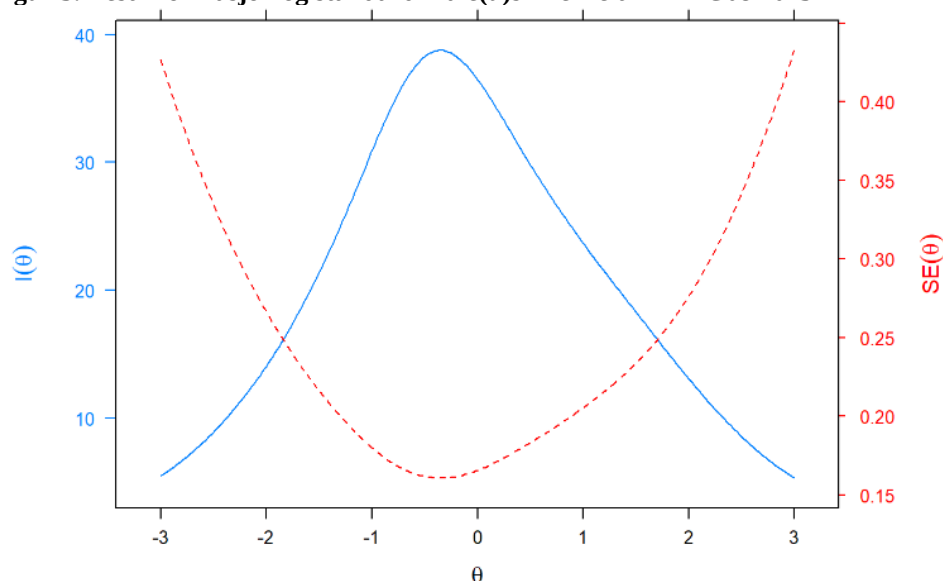
Tabell 8. Omkoding/skåring polytome oppgåver NPENG08 2023

Itemformat	Opprinnelege tal på svarkategoriar	Svarkategoriar etter omkoding som ligg til grunn for analysane
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	2

2.2.1 Overordna resultat

Figur 3 viser testinformasjon og målesikkerheit (eller måleusikkerheit) for prøven. Den blå lina viser theta (evne/kompetanse) i populasjonen av 8. trinnselevar, og den raude lina viser måle(u)sikkerheit. Prøven gir mest sikker informasjon når theta (evnene/kompetansen til elevane) er i området -1 til +0,5. I dette området er altså målingane sikrast.

Figur 3. Testinformasjon og standard måle(u)sikkerheit NPENG08 2023



Tabell 9 viser statistikk for NPENG08 2023, nærare bestemt tal på oppgåver, mål på reliabilitet, gjennomsnittlege p -verdiar og gjennomsnittleg klassisk mål på diskriminering (a).

Tabell 9. Oppsummerande statistikk NPENG08 2023

Tal på elevar	62071		Standardavvik	9,047
Tal på oppgåver	33 (42)		Reliabilitet	0,903
Tal på lytteoppgåver	15 (19)		Gjennomsnittleg løysingsprosent (p)	0,512
Tal på leseoppgåver	18 (23)		Gjennomsnittleg b -verdi (dikotome oppgåver)	-0,212
Gjennomsnittleg skår	21,52		Gjennomsnittleg a -verdi	1,31

8. trinnsprøven er reliabel (0,903) og diskriminerer på ein fornuftig måte mellom elevane (gjennomsnittleg a : 1,31). Prøven har ein vanskegrad som er omtrent det vi siktar mot (gjennomsnittleg p - (0,512) og b -verdi (-0,212)). Prøven diskriminerer også godt mellom elevar på ulike ferdigheitsnivå. Gjennomsnittleg a -verdi er 1,31. Lesedelen diskriminerer betre enn lyttedelen. Det er vanleg å seie at ved bruk av GPCM-modellen i IRT-analyse, så bør diskrimineringsmålet på individuelle oppgåver helst vere over 0,8.

Tabell 10 fokuserer på lytting og lesing kvar for seg og rapporterer mål for reliabilitet og gjennomsnittlege p -verdiar.

Tabell 10. Oppsummerande statistikk NPENG08 2023 lesing og lytting

NPENG08	Reliabilitet lytting	Reliabilitet lesing	Gjennomsnittleg p -verdi lytting	Gjennomsnittleg p -verdi lesing
62071elevar	0,759	0,873	0,528	0,500

Tabell 10 viser at om vi ser på lytte- og leseoppgåvene kvar for seg, så er lesedelen mest reliabel med ein reliabilitetskoeffesient på 0,873. Lyttedelen har ein reliabilitetskoeffesient på 0,759. Reliabiliteten vil gå ned når tal på oppgåver blir redusert. Tabell 11 viser at mens fire av lytteoppgåvene har a -verdiar under 0,8, så har lesedelen ingen med a -verdi under 0,8.

At lesedelen er meir reliabel enn lyttedelen kjem truleg av at teamet som utviklar den nasjonale prøven har meir erfaring med å utvikle leseoppgåver. Hausten 2022 var første gong prøven inneheldt lytteoppgåver. Tabell 10 viser at lyttedelen er litt lettare enn lesedelen, men skilnaden er ikkje så stor.

Den gjennomsnittlege p -verdien etter piloteringa var 0,514 (sjå tabell 3). Den reelle prøven hadde ein gjennomsnittleg p -verdi på 0,512. Vanskegraden for heile prøven har så å seia ikkje endra seg frå utprøvinga. Tidlegare då utprøving av oppgåver skjedde eit par veker før den nasjonale prøven vart gjennomført, var alltid den reelle prøven mellom 5 og 10 prosentpoeng

lettare enn utprøvinga indikerte. Tabell 10 viser at lyttedelen er litt lettare enn lesedelen. Dette var også tilfelle etter utprøvinga.

2.2.2 Resultat på oppgåvenivå

Oversikt over alle oppgåvene

Tabell 11 gir ei oversikt over oppgåveformat, kva dei ulike oppgåvene måler, pluss psykometriske mål som karakteriserer oppgåvene i kohortprøven: løysingsprosent, *b*-verdiar og *a*-verdiar.

Tabell 11. Løysingsprosent (*p*-verdiar), vanskegrad (*b*-verdiar) og mål på diskriminering (*a*)

Rekkefølge	Lytt (L) Les (R)	Format	Itemnamn	Oppgåva måler	<i>p</i>	<i>b</i>	<i>a</i>
1	L	CP	House chores 1	Forstå detaljer	0,709	-1,041	1,001
2	L	CP	Kids occupation photographer	Forstå detaljer	0,753	-1,550	0,799
3	L	MCS	Activity 1-juggler	Forstå detaljer	0,882	-1,603	1,874
4	L	MCS	Goals on fire	Forstå hovudinnhald	0,551	-0,253	0,842
5	L	MCS	Teasing	Tolke og forstå	0,433	0,347	1,039
6	L	MCS	Salmon sushi	Forstå hovudinnhald	0,491	0,076	0,986
7	L	MCS	London car free	Forstå hovudinnhald	0,647	-0,641	1,141
8	L	MCS	Sour sweets	Forstå hovudinnhald	0,349	0,773	1,013
9	L	MCS	Embarrasment	Forstå hovudinnhald	0,592	-0,530	0,744
10	L	MCS	Phone call work	Forstå hovudinnhald	0,701	-1,090	0,888
11	L	MCS	Presentation Shampoo	Forstå hovudinnhald	0,518	-0,052	1,097
12	L	MCL	Debate	Forstå hovudinnhald	0,341	1,061	0,710
13	L	MCL	Transport changes	1 forstå hovudinnhald/2 forstå detaljar	Snitt: 0,440 Full skår: 11,0%	b1 -1,244 b2 0,446 b3 1,954	0,802
14	L	OR	Competition	Forstå samanheng	Snitt: 0,458 Full skår: 23,8%	b1 -0,585 b2 1,076	0,969
15	L	MA	Podcast opinions	Forstå detaljar	Snitt: 0,412 Full skår: 15,6%	b1 -0,691 b2 2,000	0,702
16	L	CD	Sneakers	Finne informasjon	0,515	-0,026	1,363
17	R	CD	Restaurant2 pizza	Finne informasjon	0,537	-0,123	1,251
18	R	MCS	Embarrass dog 1	Forstå hovudinnhald	0,840	-1,526	1,437
19	R	MCS	Snake selfie accident	Forstå hovudinnhald	0,735	-1,096	1,133
20	R	MCS	Raising money	Forstå hovudinnhald	0,775	-0,962	2,061

21	R	MCM	Surfing goat	1 forstå hovudinnhald / 1 finne informasjon	Snitt: 0,653 Full skår: 43,2%	b1 -1,657 b2 0,200	1,210
22	R	MCM	Meteorite	1 forstå hovudinnhald / 1 finne informasjon	Snitt: 0,359 Full skår: 17,6%	b1 -0,226 b2 1,449	0,974
23	R	CW	Bus hound1	Forstå avansert vokabular	0,551	-0,158	1,572
24	R	CW	Thousands of chocolate bars1	Forstå avansert vokabular	0,525	-0,054	1,732
25	R	CW	Helen Sherman	Forstå avansert vokabular	0,330	0,677	1,579
26	R	Who	Hero-High2-Angela	Tolke og forstå	0,634	-0,370	2,820
27	R	Who	Hero-High11-Mrs-Gibbs	Finne informasjon	0,664	-0,454	3,093
28	R	Who	Hero-High5-Linda	Tolke og forstå	0,308	0,736	1,731
29	R	Who	Hero-High9-Claire	Tolke og forstå	0,396	0,454	1,348
30	R	CS	Passive must be done	Bruke grammatiske strukturar	0,583	-0,255	1,807
31	R	CS	Direct-indirect 6-could not	Bruke grammatiske strukturar	0,104	2,159	1,273
32	R	WW	Silent film stars	Bruke grammatiske strukturar	Snitt: 0,479 Full skår: 21,1%	B1 -0,441 b2 -0,073 b3 0,929	1,176
33	R	MP	Gaming mum	Forstå samanheng	Snitt: 0,241 Full skår: 13,6%	b1 1,088 b2 1,175	1,115

Tabell 11 gir ei oversikt over alle oppgåvene i kohortprøven og gir ei oversikt over oppgåveformat, gjennomsnittlege *p*- og *b*-verdiar og mål på diskriminering, *a*-verdi.

Tabellen viser at den lettaste oppgåva (oppg. nr 3) svarte 88% av elevane rett på, mens den vanskelegaste oppgåva (oppg. nr. 31) svarte omlag 10% rett på.

Løysingsprosent (p) og b -verdiar

Tabell 11 viser at prøven inneheld fleire lette dikotome oppgåver enn vanskelege. Det er i alt 26 oppgåver i prøven som blir skåra dikotomt. Tabellane 12 og 13 gir ei oversikt over lette og vanskelege dikotome oppgåver i prøven.

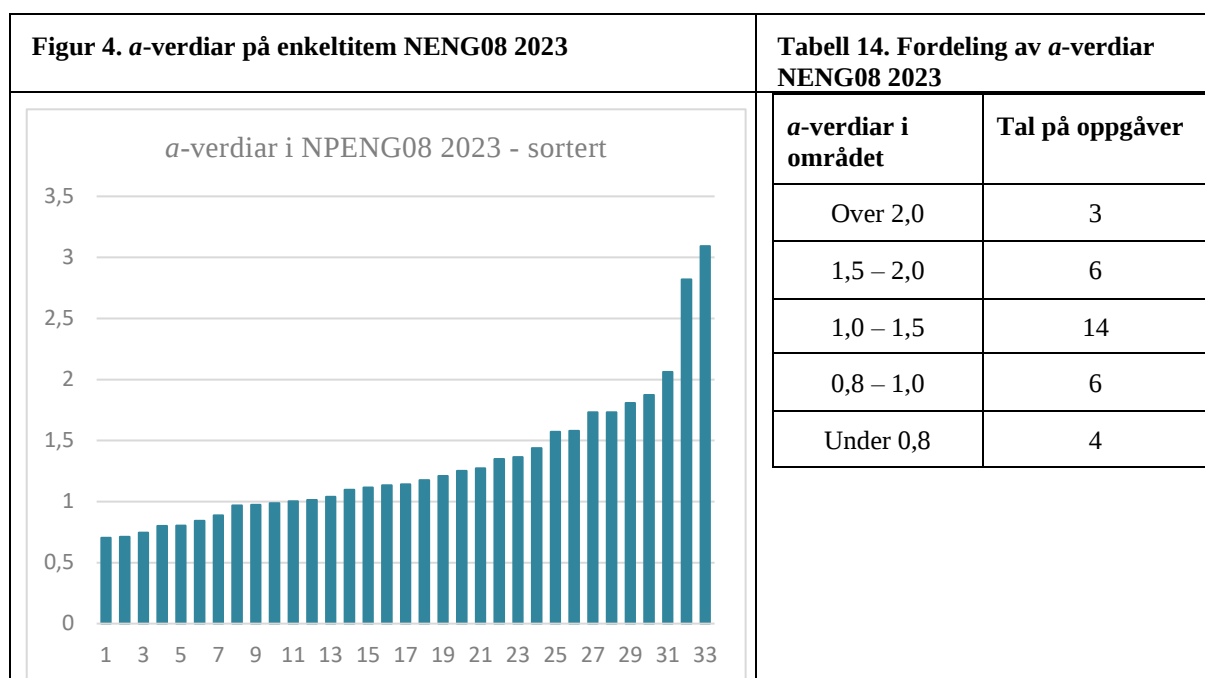
Tabell 12. Distribusjon av p -verdiar		Tabell 13. Distribusjon av b -verdiar	
p -verdi i området	Tal på oppgåver	b -verdi i området	Tal på oppgåver
0,8 til 0,9	2	Under -1,5	3
0,7 til 0,8	5	-1,0 til -1,5	3
0,6 til 0,7	3	-0,5 til -1,0	3
0,5 til 0,6	8	0 til -0,5	9
0,4 til 0,5	2	0,5 til 0	3
0,3 til 0,4	5	1 til 0,5	3
0,2 til 0,3	0	1,5 til 1	1
0,1 til 0,2	1	Over 1,5	1

Tabellen viser at det er ei litt skeiv fordeling mellom lette og vanskelege oppgåver. Det er 18 oppgåver som fleire enn 50% av elevane svarar rett på og åtte oppgaver som færre enn 50% av elevane svarar rett på. 18 oppgåver har ein p -verdi som er høgare enn 0,5 og ein b -verdi under 0. Åtte oppgåver har ein p -verdi under 0,5 og b -verdi over 0.

Det er 7 oppgåver i kohortprøven som blir skåra polytomt, 3 lytteoppgåver og 4 leseoppgåver.

Diskriminering – a -verdi

Ei oppgåve som diskriminerer godt, skil godt mellom elevar på ulike kompetansenivå, for eksempel mellom elevar på meistringsnivå 3 og dei på meistringsnivå 4. Når vi analyserer resultata i R, vil Utdanningsdirektoratet at vi brukar GPCM (Generalised Partial Credit Model) der ein ynskjer at itema skal ha eit mål på diskriminering, ein a -verdi, over 0,8. Tabell 11 viser a -verdien for kvart av itema i kohortprøven. Vi ser at dei fleste itema i kohortprøven har a -verdiar over 0,8. Figur 4 og tabell 14 viser korleis a -verdiane på oppgåvene i kohortprøven fordeler seg.



Figur 4 og Tabell 14 viser at oppgåvene i prøven diskriminerer stort sett godt. Fem av oppgåvene har α -verdier i området under 0,8. 29 oppgåver har α -verdi over 0,8. Fire oppgåver har α -verdi under 0,8. Disse fire er alle lytteoppgåver. Då oppgåvene vart prøvd ut i 2022, var α -verdien over 0,8.

Theta

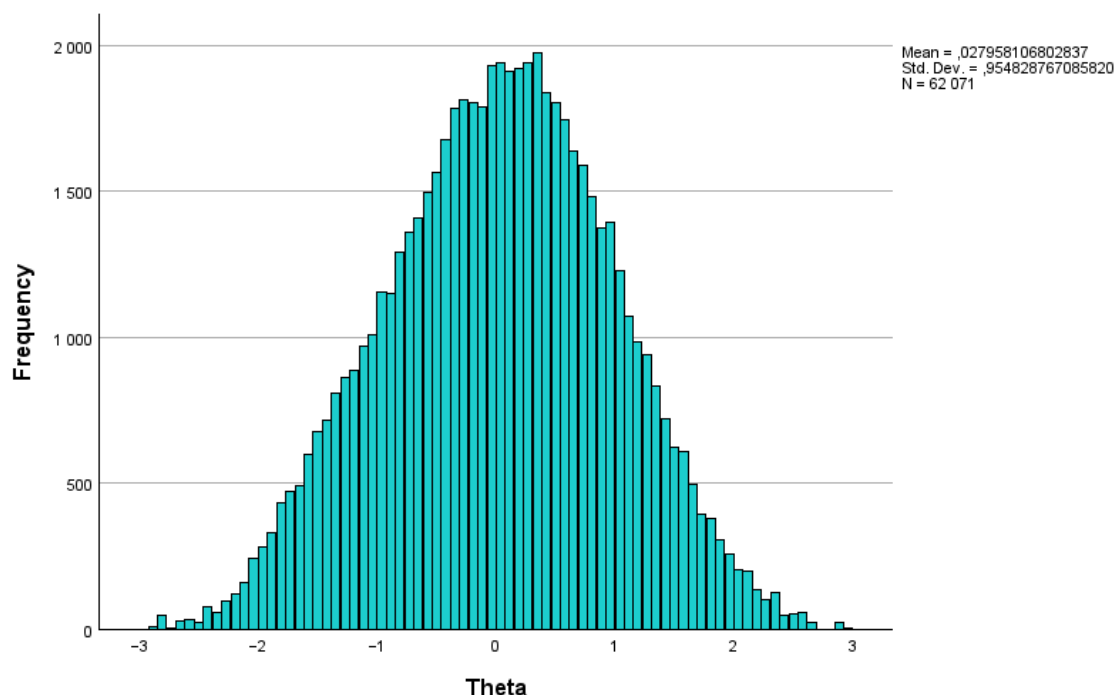
Tabell 15 og Figur 5 oppsummerer thetamål i 8. trinn populasjonen. Alle figurar og tabellar er frå analysar der ankeritema er inkludert. Vi ser i Tabell 15 at distribusjonen av theta (ability) i 8. trinnsgruppa varierer frå -2,879 til +2,957.

Tabell 15. Oppsummerande statistikk for theta 8. trinn 2023

Elevar	Gjennomsnitt	Standardavvik	Minimum	Median	Max
62071	0,279	0,955	-2,879	0,053	2,957

Variasjonen er framstilt grafisk i figur 5.

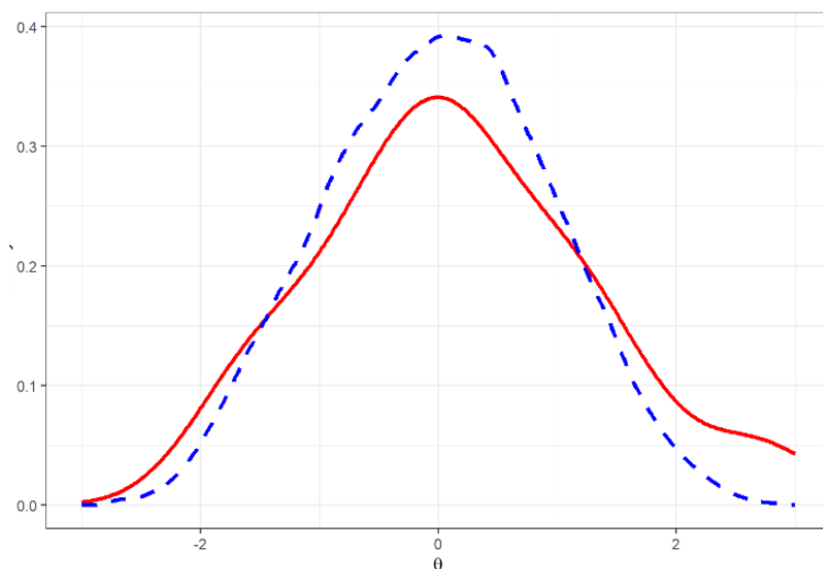
Figur 5. Distribusjon av theta i 8. trinnspopulasjonen.



Vi ser at distribusjonen av theta i 8. trinnspopulasjonen er tilnærma normalfordelt. Data frå prøven viser altså at det finst elevar som har svært låge ferdigheiter i engelsk, og samtidig finst det dei med høge ferdigheiter i engelsk. Dei aller fleste elevane har ferdigheiter ein eller annan stad mellom -1 og + 1.

Figur 6 viser korleis theta og vanskegrad på oppgåvene fordeler seg i populasjonen av 8. trinnselevar. Den blå stipla linea viser theta, mens den raude linea viser vanskegrad på oppgåvene.

Figur 6. Theta og vanskegrad NPENG08 2023



Dersom oppgåvene i prøven utfordrar elevar på alle ferdigheitsnivå, bør det vera nokolunde overlapp mellom theta og vanskegrad på oppgåvene. Det betyr i så fall at prøven er tilpassa elevar på ulike ferdigheitsnivå. Figur 6 viser at både theta og vanskegrad på oppgåvene er tilnærma normalfordelt. Det same rapporterer også tabellane 12 og 13.

2.3 Kjønn

Tabell 16 gir ei oversikt over skilnaden i løysingsprosent/ p -verdi på enkeltoppgåver og kjønn.

Tabell 16. Løysingsprosent – gutar og jenter NPENG08 2023

Rekkefølge	Lytt (L) Les (R)	Oppgåvenamn	p	p gutar	p jenter	Differanse gutar/jenter
1	L	House chores 1	0,709	0,708	0,710	-0,002
2	L	Kids occupation photographer	0,753	0,820	0,684	0,220
3	L	Activity 1-juggler	0,882	0,889	0,875	0,014
4	L	Goals on fire	0,551	0,563	0,539	0,024
5	L	Teasing	0,433	0,450	0,416	0,034
6	L	Salmon sushi	0,491	0,516	0,464	0,052
7	L	London car free	0,647	0,669	0,625	0,044
8	L	Sour sweets	0,349	0,362	0,337	0,025
9	L	Embarrasment	0,592	0,560	0,624	0,064
10	L	Phone call work	0,701	0,690	0,713	-0,023
11	L	Presentation Shampoo	0,518	0,575	0,459	0,116
12	L	Debate	0,341	0,312	0,370	-0,058
13	L	Transport changes	Snitt: 0,440 Full skår: 11,0%	Snitt: 0,431 Full skår: 10,8%	Snitt: 0,449 Full skår: 11,2%	Snittdifferanse: -0,018
14	L	Competition	Snitt: 0,458 Full skår: 23,8%	Snitt: 0,532 Full skår: 30,1%	Snitt: 0,382 Full skår: 17,9%	Snittdifferanse: 0,150
15	L	Podcast opinions	Snitt: 0,412 Full skår: 16,4%	Snitt: 0,414 Full skår: 15,4%	Snitt: 0,410 Full skår: 17,8%	Snittdifferanse: 0,004
16	L	Sneakers	0,515	0,497	0,533	-0,036
17	R	Restaurant2 pizza	0,537	0,497	0,577	-0,080
18	R	Embarrass dog 1	0,840	0,816	0,865	-0,049
19	R	Snake selfie accident	0,735	0,732	0,738	-0,006
20	R	Raising money	0,775	0,781	0,769	0,012

21	R	Surfing goat	Snitt: 0,653 Full skår: 43,2%	Snitt: 0,639 Full skår: 40,8%	Snitt: 0,667 Full skår: 45,8%	Snittdifferanse: -0,028
22	R	Meteorite	Snitt: 0,389 Full skår: 17,6%	Snitt: 0,413 Full skår: 20,1%	Snitt: 0,363 Full skår: 15%	Snittdifferanse: 0,050
23	R	Bus hound1	0,551	0,617	0,484	0,133
24	R	Thousands of chocolate bars1	0,525	0,551	0,497	0,054
25	R	Helen Sherman	0,330	0,318	0,343	-0,025
26	R	Hero-High2-Angela	0,634	0,610	0,658	-0,048
27	R	Hero-High11-Mrs-Gibbs	0,664	0,649	0,680	-0,031
28	R	Hero-High5-Linda	0,308	0,267	0,350	-0,050
29	R	Hero-High9-Claire	0,396	0,405	0,386	0,019
30	R	Passive must be done	0,583	0,599	0,566	0,033
31	R	Direct-indirect 6-could not	0,104	0,093	0,115	-0,022
32	R	Silent film stars	Snitt: 0,479 Full skår: 21,1%	Snitt: 0,464 Full skår: 19,4%	Snitt: 0,495 Full skår: 22,8%	Snittdifferanse: -0,031
33	R	Gaming mum	Snitt: 0,241 Full skår: 13,6%	Snitt: 0,205 Full skår: 11,1%	Snitt: 0,276 Full skår: 16,2%	Snittdifferanse: -0,071

Differansen mellom gruppene finst i kolonna heilt til høgre. Dei oppgåvene som jentene gjer det best på har ein negativ verdi. Vi ser at av dei 33 oppgåvene kohortprøven inneheld, gjer jentene det best på 16 av oppgåvene, mens gutane gjer det best på 17 av oppgåvene. Jentene skårar høgre på flest leseoppgåver, mens gutane skårar høgst på flest lytteoppgåver.

Tabell 17 viser skilnader mellom gutar og jenter i gjennomsnittleg totalskår, løysingsprosent og theta.

Tabell 17. Skilnader mellom gutar og jenter i gjennomsnittleg skår, løysingsprosent og theta

Gjennomsnittlege estimat					
	Totalskår (42)	Lytting (19) (råskår)	Lesing (23) (råskår)	Løysingsprosent	Theta
Gutar	21,64	10,30	11,34	0,515	0,0308
Jenter	21,41	9,75	11,66	0,510	0,0250

Tabell 17 viser at gutane skårar i gjennomsnitt høgare enn jentene totalt og på lyttedelen. Jentene skårar høgare på lesedelen. Samtidig gjennomsnittleg theta er litt høgare i gutegruppa enn i jentegruppa, men denne skilnaden er ikkje så stor.

Ein t-test er ein type signifikanstest som undersøker skilnader mellom gjennomsnitt i ulike grupper. I vårt tilfelle har vi eit utgangspunkt, ein null-hypotese, der vi går utifrå at det gjennomsnittlege resultatet for jenter og gutar er tilnærma likt eller ikkje. Ein t-test kan undersøke om dette stemmer. Eit signifikant resultat på t-testen vil seie at skilnaden mellom gruppene sannsynlegvis ikkje er tilfeldig, mens eit ikkje-signifikant resultat oppsummerer at skilnaden sannsynlegvis er tilfeldig. Ein t-test understrekar skilnader mellom gutar og jenter i 8. trinnspopulasjonen. Dette kjem fram av Tabell 18. Vi ser at gjennomsnittleg totalskår for gutane (21,64) er litt høgare enn den for jentene (21,41). Standardavviket er nesten likt i dei to gruppene, litt høgare for jentene. Dette fortel oss at variasjonen i jentegruppa og gutegruppa er omtrent lik.

Tabell 18. Skilnader mellom jenter og gutar på 8. trinn – t-test

Prøve	Kjønn	Gjennomsnittleg tal på rette svar	n	Standardavvik	t	Sig 2-tailed
NPENG08 2023	Gutar	21,64	31488	8,931	3,103	< 0,002
	Jenter	21,41	30583	9,163		
		Gjennomsnittleg theta	n	Standardavvik	t	Sig 2-tailed
	Gutar	0,0308	31488	0,9410	0,748	<0,454
	Jenter	0,0250	30583	0,9689		

Tabell 18 viser at med utgangspunkt i råskåren, er det ein signifikant skilnad mellom gutar og jenter i favør av gutane i lytting, og i favør av jentene i lesing (tabell 19). Når dette resultatet er signifikant, så betyr det at skilnadene sannsynlegvis ikkje er tilfeldige. Samtidig ser vi at om vi baserer t-testen på theta, så er skilnadene mellom gutar og jenter små og sannsynlegvis tilfeldige.

Altså; med utgangspunkt i gjennomsnittleg råskår, så er det svært liten skilnad mellom gutar og jenter (21,64 og 21,41). Om vi ser på resultatet av t-testane, er skilnadene signifikante på 0,002-nivå, altså ikkje tilfeldige, mens resultatet sannsynlegvis er tilfeldig om vi baserer analysen på theta. Likevel, med så store datasett som vi opererer med her, kan trivielle skilnader gi eit signifikant resultat, så vi må vere forsiktige når vi tolkar resultatet. Dermed treng det ikkje bety så mykje at skilnaden mellom gutar og jenter i råskår, er signifikant. Uansett så understrekar dette at skilnaden mellom gutar og jenter på åttande trinn er minimale i 2023.

Tabell 19 viser eit litt meir nyansert bilde. Her har vi køyrt t-testar på data frå lytting og lesing kvar for seg. Her ser vi at skilnadene mellom gutar og jenter er signifikant i favør av gutane når det gjeld lytting, mens skilnaden er signifikant i favør av jentene i lesing.

Tabell 19. Skilnader mellom gutar og jenter i lytting og lesing – t-test

Prøve	Kjønn	Gjennomsnittleg tal på rette svar - lytting	n	Standardavvik	t	Sig 2-tailed
NPENG08 2023	Gutar	10,30	31488	3,873	17,658	< 0,001
	Jenter	9,75	30583	3,955		
		Gjennomsnittleg tal på rette svar - lesing	n	Standardavvik	t	Sig 2-tailed
	Gutar	11,34	31488	5,616	-7,220	< 0,001
	Jenter	11,66	30583	5,754		

2.4 NP engelsk for 8. trinn 2023 – oppsummering

Tidlegare i rapporten har vi fokusert på korleis Nasjonal prøve i engelsk for 8. trinn fungerte i 2023. Vi har sett på hovudresultat, ulike itemformat, kjønnskilnader, men ikkje noko om kva elevar på ulike meistringsnivå kan. I all hovudsak har vi konkludert med at prøven ser ut til å fungere bra for målgruppa. Nærare bestemt

- Den nasjonale prøven i engelsk for 8. trinn 2023 måler lytting og lesing. Det er andre gong lytteoppgåver er inkludert i nasjonale prøvar i engelsk.
- Hovudprøven måler eit utval av kompetansemål i læreplanen.
- Prøven er reliabel. Reliabilitetskoeffesienten er 0,903.
- Prøven diskriminerer godt mellom elevane. Gjennomsnittleg α for hovudprøven er 1,31. Fire av oppgåvene har α under 0,8.
- Prøven skil relativt godt mellom elevar på ulike ferdigheitssnivå.
- Både dei dikotome og polytome oppgåvene ser ut til å fungere nokolunde greitt.
- Gutane skårar høgare enn jentene totalt, og denne skilnaden er med utgangspunkt i råskår, men ikkje om vi legg theta til grunn.
- Gutane skårar høgare enn jentene på lyttedelen, mens jentene skårar litt høgare enn gutane på lesedelen. Desse siste skilnadene er signifikante, altså sannsynlegvis ikkje tilfeldige.
- Skilnaden i vanskegrad frå pilotering til reell prøve er mindre enn det som har vore tilfelle dei siste åra. Gjennomsnittleg p -verdi etter pilotering var 0,514 i prøven som vart foreslått. Data frå den nasjonale prøven viser ein gjennomsnittleg p -verdi på 0,512. Det betyr at vanskegraden totalt er den same etter utprøving som i reell prøve.

Når ein prøve skal tilpassast så ulike elevgrupper som dei som representerer meistringsnivåa 1 - 5, må det bli slik at mange oppgåver er vanskelege for elevar på meistringsnivå 1 og mange oppgåver er lette for elevar på meistringsnivå 5. Likevel viser resultatet at prøven inneheldt oppgåver som utfordrar elevar på alle meistringsnivå.

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 (i 2023). 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 me tatt utgangspunkt i 58 500 elevar. Kvar av dei 13 ekstra seksjonane vert då tildelt 4500 tilfeldige elevar (1/13). Slik sikrar me 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.

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	