

NYE PRØVER MED LÆRINGSSTØTTENDE FORMÅL

19.06.2026

Delleveranse 2: konseptforslag fra tre lærergrupper og Udirs
innledende vurderinger av disse forslagene



Innhold

Forord	5
Sammendrag	6
1. Innledning	7
1.1 Om oppdraget.....	7
1.2 Hvordan forstå bestillingen om «læringsstøttende prøver»	10
1.3 Prøvenes formål og avveining av ulike hensyn.....	11
2. Prosjekt Nye prøver.....	13
2.1 Forprosjekt.....	14
2.2 Utvikling av prosjektet	14
2.3 Involvering.....	14
2.4 Innhenting av brukerinnsikt	15
3. Forankring av arbeidet i eksisterende kunnskapsgrunnlag	16
3.1 Støtte til lærernes faglig-pedagogiske arbeid.....	16
3.2 Tydeliggjøre formål og innhold	17
3.3 Støtte til å forstå resultater og å sette disse i sammenheng	17
3.4 Støtte til oppfølging av elevenes progresjon	17
3.5 Tilrettelegging for en sammensatt elevgruppe	18
3.6 Resultatvisninger og støtteressurser	18
3.7 Behov i den samiske skolen.....	19
4. Forslag fra arbeidsgruppene	19
4.1 Fellestrekk i konseptene for engelsk, matematikk og lesing	20
4.2 Engelskgruppens forslag og anbefalinger	23
4.3 Matematikkgruppens forslag og anbefalinger	28
4.4 Lesegruppens forslag og anbefalinger	33
4.5 Vurdering av endringer i prøvenes konstrukter	40
5. Sentrale utredningsspørsmål og veivalg.....	43
5.1 Kortere prøver	44

5.2	Blokkadaptive prøver	44
5.3	Vurdere resultater mot standard, norm eller gjennomsnitt	45
5.4	Følge opp elevenes faglige utvikling gjennom frivillige prøver	46
5.5	Analyser av oppgavekvalitet, elevferdigheter og trender.....	47
5.6	Prosessdata kan gi bedre innsikt	49
5.7	Tilgjengelige resultater og støtteressurser	50
5.8	Mer handlingsrettet informasjon fra prøvene.....	51
5.9	Behov for støtte til oppfølging	52
6.	Konsekvenser av forslagene for videre digitalisering	54
6.1	Prøveadministrasjonssystemet (PAS)	54
6.2	Eksamen- og Prøvesystemet (EPS).....	55
6.3	Hva er mulig på kort, mellomlang og lang sikt	55
6.4	Våre forslag til oppfølging og videre arbeid.....	57
7.	Informasjon om læringsutbytte for elever i den samiske skolen.....	58
7.1	Behov for forbedret kvalitativ informasjon fra samiske prøver	58
7.2	Behov for prøver som er utviklet på egne premisser	59
7.3	Behov for relevante sammenligningsgrunnlag.....	59
7.4	Behov for læremidler på samisk.....	59
7.5	Oppsummering.....	60
8.	Nye verktøy på 1.trinn	60
8.1	Pågående prosesser og endringer av relevans for arbeidet.....	60
8.2	Endringer for å gjøre dagens prøver mer fleksible	61
8.3	Behovsbeskrivelser i kunnskapsgrunnlaget	62
8.4	Mulige endringer på lengre sikt.....	65
8.5	Veien videre, nye verktøy til bruk i begynneropplæringen.....	65
9.	Anbefalinger om videre arbeid, nye prøver og nye verktøy.....	66
9.1	Vurdering av hvordan vi kan følge opp forslagene.....	66
9.2	Kommentarer og tilbakemelding fra organisasjonene i skolesektoren	68

9.3 Plan for videre arbeid.....	69
Kilder	72
Vedlegg	75
Vedlegg 1: Eksempler på hvordan resultater blir presentert i dagens løsning.....	75
Vedlegg 2: Innspill fra Statped til delleveranse 2.....	78

Forord

I den første delleveransen til Kunnskapsdepartementet beskrev vi hvordan vi ønsker å arbeide med oppdraget. Vi la til grunn at prøvene må utvikles i samarbeid med lærere og skoler for å bli gode verktøy for praksisfeltet. Dette er også i tråd med oppdragsbrevet, der vi blir bedt om å innhente synspunkter og erfaringer fra brukergruppene. I denne rapporten beskriver vi hvordan vi har fulgt opp dette, og hvilke anbefalinger vi har fått fra praksisfeltet om utvikling av nye prøver.

I den første delleveransen skrev vi at vi ønsker å utvikle løsninger for de læringsstøttende prøvene gjennom en åpen og utforskende prosess. En «smidig» og innovativ tilnærming skal både øke bruksverdien i skolen og bidra til at de nye prøvene får riktig teknisk og faglig kvalitet. En slik tilnærming tar også høyde for at det tar tid å utvikle gode digitale løsninger.

Dagens prøver er utformet for å gi presis informasjon om noen utvalgte ferdigheter. Med dette som utgangspunkt presenteres indikatorer som utgangspunkt for dialog om kvalitet på ulike nivåer. Til grunn for oppdraget ligger forståelsen av at dagens prøver i for liten grad støtter lærerens arbeid med elevenes læring. En sentral del av rapporten er derfor en presentasjon av rådene vi har fått fra lærerne som denne våren har utredet hvordan nye prøver kan bli nyttige verktøy i deres hverdag.

I leveransen drøfter vi flere steder hvordan kvantitativ informasjon fra prøver kan kombineres med vurderinger fra lærerne basert på flere kilder, og at dette blir del av det helhetlige arbeidet med elevenes læring. Dette ønsker vi å utforske nærmere i det videre arbeidet.

Arbeidsgruppene som har utredet konsepter for nye prøver har arbeidet med et åpent mandat. Dette har vi sett som en forutsetning for at vi skal få nye ideer til hvordan vi kan utforme disse verktøyene. Det betyr samtidig at disse forslagene må utredes nærmere. Vi ønsker å gå raskt i gang med å teste ut ideene gjennom utprøvinger. Samtidig må vi involvere fagmiljøene i universitetets- og høyskolesektoren som har spisskompetanse på denne typen prøveutvikling for å få deres vurderinger av hvordan vi kan lage prøver med god faglig kvalitet som samtidig møter ønskene fra lærerne.

Ambisjonen med leveransen er ikke å gi et endelig svar på hvordan de nye prøvene bør utformes, men å peke vei videre. Vi har lagt vekt på tett dialog med organisasjonene i skolesektoren om dette oppdraget. Vi er opptatt av å finne løsninger som mange kan stille seg bak, og som kan løfte det faglige arbeidet i skolen til det beste for elevenes læring. Vi oppfatter at det er støtte til at vi i dette arbeidet først konsentrerer oss om hvordan prøvene kan bli gode verktøy i skolen, og at vi så utreder nærmere hvilken informasjon kommunene kan få om elevenes læring for å kunne ivareta sitt ansvar. Vi mener forslagene også vil kunne gi god informasjon til kommunene, samtidig som læringsresultater ikke brukes som utgangspunkt for rangering av- og konkurranse mellom skoler. Dette vil vi utrede parallelt med at vi prøver ut nye konsepter med utgangspunkt i forslagene fra lærergruppene.

Sammendrag

Denne rapporten er andre delleveranse i Utdanningsdirektoratets arbeid med å utvikle nye prøver med læringsstøttende formål og nye fleksible verktøy for begynneropplæringen. Hovedmålet for arbeidet er å gjøre prøvene mer nyttige for lærere i det daglige arbeidet med elevenes læring, samtidig som de inngår i et nytt system for kvalitetsutvikling i skolen.

I rapporten presenterer vi forslag fra tre lærergrupper innen engelsk, matematikk og lesing som har utredet hvordan nye prøver kan bli gode verktøy til bruk i skolen. Utgangspunktet for anbefalingene er at dagens prøver må bli mer relevante for lærernes pedagogiske arbeid, samtidig som de fortsatt skal holde høy testfaglig kvalitet. Gruppene foreslår at de nye prøvene i større grad skal gi rask, forståelig og handlingsrettet informasjon om elevenes styrker, utfordringer og utvikling over tid.

Lærergruppene foreslår blant annet kortere og mer motiverende prøver. De har også ulike forslag til hvordan blokkadaptive løsninger kan gi elevene en bedre prøveopplevelse, samtidig som prøvene gir bedre informasjon om alle elever, også dem som befinner seg i ytterkantene av ferdighetsskalaen. Gruppene ønsker også bedre resultatvisninger, tydeligere støtteressurser, mulighet for å måle faglig utvikling og frivillige moduler. Vi er enige i at skolene må få bedre støtte, både i form av andre resultatvisninger og bedre støtteressurser, og også at disse prøvene må bli del av et helhetlig arbeid med elevenes læring. Noen forbedringer kan vi gå i gang med relativt raskt. Endringer i utforming av prøvene og i brukerflyten vil kreve mer arbeid i de tekniske systemene, og dermed ta mer tid.

I rapporten drøfter vi faglige, tekniske og økonomiske konsekvenser, samt konsekvenser for personvern. Vi gjør imidlertid ikke en fullstendig konsekvensutredning av de ulike forslagene. Formålet med rapporten er å peke retning for det videre arbeidet og skape enighet rundt noen målsettinger som kan ligge til grunn for den smidige utviklingsprosessen vi skisserte i den første delleveransen.

Rapporten omtaler arbeidet med nye verktøy for 1. trinn, som skal erstatte dagens kartleggingsprøver. Her redegjør vi for endringer vi har gjennomført, vi presenterer noen behovsvurderinger med utgangspunkt i foreliggende kunnskap, og vi peker videre retning for denne delen av arbeidet.

Rapporten drøfter også behov for egne løsninger for samisk skole. Det er behov for prøver, resultatvisninger og støtteressurser som er utviklet på samiske språk og premisser. En mindre elevpopulasjon gjør at disse prøvene ikke kan analyseres på samme måte som de andre prøvene. Vi vil utforske mulighetene for å kunne gi rikere og bedre informasjon fra også disse prøvene som del av arbeidet.

Til slutt anbefaler rapporten at Udir på kort sikt tester ut forbedringer i resultatvisning og støtteressurser, mens mer omfattende endringer i prøvedesign og systemstøtte må utvikles over tid. Arbeidet skal skje gradvis, i tett samarbeid med lærere, fagmiljøer og organisasjonene i skolesektoren. Dette vil bidra til at løsningene er i tråd med ambisjonene for oppdraget, og at utviklingen samtidig skjer innenfor tilgjengelige rammer.

1. Innledning

1.1 Om oppdraget

Kunnskapsdepartementet (KD) har gitt Utdanningsdirektoratet (Udir) i oppdrag å utvikle nye prøver med læringsstøttende formål og nye fleksible verktøy til bruk i den tidlige lese- og regneopplæringen (OB 2025-018¹). Oppdraget er en oppfølging av Meld. St. 34 (2023-2024). *En mer praktisk skole. Bedre læring, motivasjon og trivsel på 5. – 10. trinn*. Vi ser også oppdraget i sammenheng med de sentrale målene for direktoratet slik de er formulert i KDs tildelingsbrev til Udir. Det betyr at disse verktøyene skal bidra til bedre oppfølging av elever som har behov for tilrettelegging tidlig i skoleløpet, og til at flere elever lærer og trives. Arbeidet har også relevans for målsettingen om forbedret tverrfaglig samarbeid, og målet om et mer likeverdig opplæringstilbud uavhengig av hvor eleven bor.

Udir oversendte delleveranse 2 til Kunnskapsdepartementet 1.juni og varslet samtidig at vi ønsket å sende en justert versjon etter at vi har fått kommentarer fra organisasjonene i skolesektoren. Deres kommentarer ligger ved denne justerte versjonen, og vi har supplert drøftingen i kapittel 9 i lys av deres kommentarer.

Nye læringsstøttende prøver skal erstatte nasjonale prøver

Dagens nasjonale prøver skal gi informasjon til lokale og nasjonale myndigheter som grunnlag for å utvikle kvaliteten på opplæringen. Prøvene skal samtidig støtte lærerne i deres arbeid med elevenes læring (Meld. St. 34 (2023-2024)). I oppdragsbrev 2025-018 ber departementet Udir utvikle prøver med et *tydelig læringsstøttende formål* og skriver at de nye prøvene skal være *gode pedagogiske verktøy for lærerne til å støtte elevenes læring*. Samtidig forventes det også at disse prøvene skal være et *sentralt element i et nytt kvalitetsutviklingssystem*, og at de skal kunne *danne grunnlag for dialog, tillit og samarbeid om kvalitetsutvikling i skolen* (OB 2025-018).

Vi forstår «et tydelig læringsstøttende formål» slik at bruksverdien for skolene skal stå sentralt, og at lærernes arbeid med elevenes læring skal være omdreiningspunktet. Samtidig skal de nye prøvene også bidra til systematisk arbeid med elevenes læring både i og rundt skolen. En føring for oppdraget er at resultater ikke skal presenteres slik at det legger til rette for rangering eller konkurranse mellom skoler. Vi skal altså arbeide for å *forsterke* nytten av verktøyene for lærere, og *endre* måten informasjon fra prøvene brukes utenfor skolen.

Arbeidet med nye prøvene skal ses i sammenheng med ambisjonen om å etablere et nytt helhetlig system for å arbeide med kvalitetsutvikling i skolen. Det betyr at disse prøvene må understøtte de sentrale målene og prinsippene for et slikt system, slik det er beskrevet i Meld. St. 34. De nye prøvene må også ses i sammenheng med andre verktøy som skolene bruker i sitt helhetlige arbeid med kvalitet. Vi vil bidra til at skolene kan få støtte til å arbeide med trivsel og læring som gjensidig avhengige og forsterkende faktorer. Det er vanskelig å

¹ Tillegg nr. 16 til tildelingsbrev til Utdanningsdirektoratet for 2025 med oppdrag 2025-018 om å utvikle nye prøver med læringsstøttende formål

lære om man ikke trives, og det er vanskelig å trives om man ikke lærer. Gode prøver skal bidra til at flere opplever mestring og tilhørighet både faglig og sosialt.

Nye verktøy og prøver til bruk i begynneropplæringen

I oppdraget bes Udir avvikle dagens kartleggingsprøver på 1.trinn og erstatte kartleggingsprøvene med mer fleksible verktøy. Disse verktøyene skal hjelpe lærerne med å følge opp elever som trenger ekstra støtte i den tidlige lese- og regneopplæringen.

Kartleggingsprøvene på 1. trinn er frivillige for skolene å gjennomføre. Fra og med våren 2026 har Udir lagt til rette for at skolene har økt grad av fleksibilitet for både når de ønsker å gjennomføre en slik kartlegging, og hvilke elever de mener det kan være nyttig for. Oppdragsbrevet spesifiserer ikke hvilke forventninger som stilles til de nye fleksible verktøyene. Vi vil derfor kartlegge hvilke behov skolene har for støtte, og om det er andre typer verktøy enn dagens kartleggingsprøver som kan være mer hensiktsmessig. Dette drøfter vi nærmere i kapittel 8.

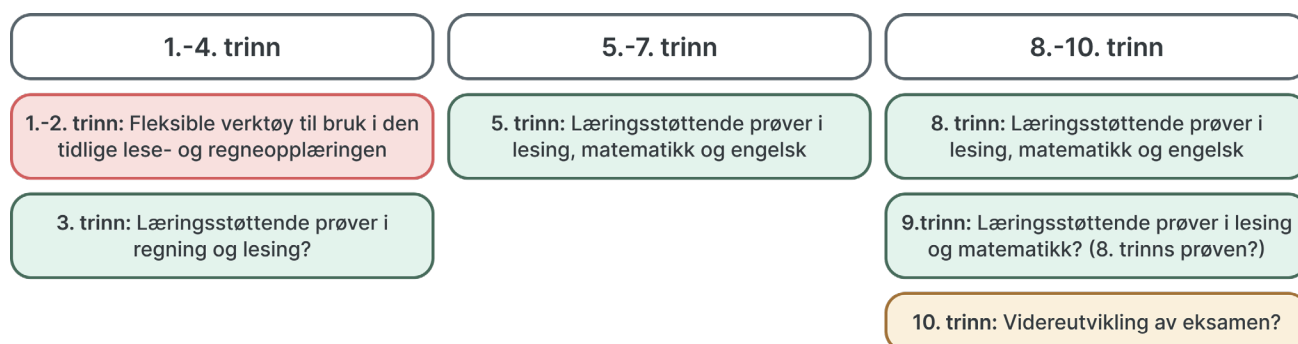
I oppdragsbrevet bes vi også vurdere om en ny prøve med læringsstøttende formål kan erstatte dagens kartleggingsprøver på 3.trinn. En slik endring vil bety at skolene får informasjon om alle elever, og ikke kun elever i oppfølgingsområdet. Samtidig vil det bety at prøvene må endres. Dagens prøver er utformet for å gi presis informasjon om elever som trenger ekstra oppfølging. Arbeid med nye verktøy og prøver til bruk i begynneropplæringen må ses i sammenheng med planlagte endringer i læreplanen for de første skoleårene.

Identifisere behov og mulige konsepter tett på praksis

Hovedaktiviteten våren 2026 har vært å organisere et utredningsløp i samarbeid med lærere fra hele landet. Disse lærerne har med utgangspunkt i sine erfaringer fra klasserommet utviklet ulike ideer og forslag som de mener vil være nyttige i det pedagogiske arbeidet. Underveis i arbeidet har lærerne diskutert med og innhentet synspunkter fra fagmiljøene i universitets- og høyskolesektoren som utvikler dagens prøver. Vi presenterer deres forslag, og vi drøfter hvordan vi kan gå videre med disse forslagene. Utredningen sammen med lærerne har hatt til hensikt å få opp nye ideer og bedre forståelse av behovene i det daglige arbeidet med elevenes læring. I den første leveransen fra desember 2025 argumenterte vi for at dette oppdraget best kan løses gjennom hypotesebaserte utprøvinger tett på praksisfeltet. I neste fase av arbeidet ønsker vi derfor å aktivt prøve ut ulike deler av forslagene fra disse arbeidsgruppene for å undersøke hvordan dette fungerer i praksis.

Helhet og progresjon i grunnskoleløpet

Oppdraget inneholder flere delspørsmål og berører hele grunnskoleløpet. I tillegg til å skulle foreslå nye verktøy til bruk i begynneropplæringen og nye læringsstøttende prøver på flere trinn i grunnskolen, bes vi også vurdere mulig ny innretning på eksamen på 10.trinn. Oppdraget kan med andre ord ha påvirkning på hele grunnskoleløpet. Vi mener det er en fordel å se disse vurderingssituasjonene i sammenheng. Det vil være gunstig om elevene både kan kjenne seg igjen i aktivitetene de møter på ulike tidspunkt i grunnskoleløpet, samtidig som de opplever en naturlig progresjon i hvilke faglige ferdigheter de får vist frem.



Figur 1: Leveranser og mulige endringer som følge av oppdraget, fordelt på trinn

Ulike typer spørsmål må utredes

Det er flere typer spørsmål som må utredes for å kunne besvare oppdraget. Vi må avklare *faglige spørsmål*. Dette inkluderer spørsmål om hvordan vi forstår de faglige ferdighetene som skal undersøkes. Et spørsmål er for eksempel hva som blir konsekvensen av å endre fra prøver i regning til prøver i matematikk. Oppdraget berører flere *prinsipielle spørsmål* om hvordan disse verktøyene skal inngå i det systematiske arbeidet med kvalitet. Her er det ulike synspunkter. Det er ikke hensiktsmessig å endre disse verktøyene for ofte. Det er derfor viktig å finne løsninger som flere kan samle seg om og som kan vare over tid. Nye verktøy og nye prøver skal for en stor del realiseres i digitale systemer. Det betyr at det er mange *tekniske spørsmål* som må avklares. Vi må blant annet vurdere hva vi kan få til med de IT-systemene vi har nå, og hva som krever annen og ny form for digitalisering. Disse vurderingene har stor betydning for de økonomiske aspektene ved arbeidet. Sist, men kanskje viktigst, er det mange *praktiske spørsmål* som må avklares. Den sentrale suksessfaktoren er hvordan disse verktøyene kan fungere godt i skolen.

Prioriteringer og avgrensninger

Vi drøfter altså først og fremst hvordan nye prøver og verktøy kan være gode verktøy for skolene. Vurderingene fra lærergruppene gir et godt utgangspunkt for å både identifisere sentrale behov og mulige måter å møte disse behovene på. Samtidig vil vi legge til rette for at organisasjonene i skolesektoren og praksisfeltet kan gjøre seg kjent med disse forslagene, og få mulighet til å gi tilbakemeldinger på dem. Udir må også utrede nærmere konsekvensene av disse forslagene.

Oppdragsbrevet inneholder flere utredningspunkter som ikke vil bli drøftet i denne leveransen. Vi vil komme tilbake til spørsmålet om hvilken informasjon lokale skolemyndigheter kan få fra et nytt kvalitetsutviklingssystem og nye prøver, uten at det påvirker utformingen til de nye prøvene og begrenser det læringsstøttende formålet de skal ha. Det samme gjelder spørsmålet om hvilken informasjon nye prøver og et nytt kvalitetsutviklingssystem kan gi til forskningsformål. I oppdragsbrevet står det at regjeringen fortsatt vil ha eksamen på 10. trinn, og at regjeringen ønsker å vurdere om eksamen på 10. trinn kan gi mer standardisert informasjon om utviklingen i elevenes læringsresultater over tid enn i dag. Udir vil utrede dette spørsmålet med sikte på å legge fram ulike alternativer for hvordan en slik endring eventuelt kan gjennomføres.

Nye prøver som støtter helhetlig arbeid med læring og trivsel

Overordnet del av læreplanverket er tydelig på at lokale skolemyndigheter, skoleledere, lærere og andre ansatte har et felles ansvar for å legge til rette for god utvikling i skolen. Skolen skal være et profesjonsfaglig fellesskap der lærere, ledere og andre ansatte samarbeider om å utvikle kvaliteten i skolen. Arbeidet skal understøtte samarbeid, dialog og tillit i skolen og til skolen. Skolens verdigrunnlag og brede mandat om danning og utdanning, slik det er beskrevet i overordnet del av læreplanverket, skal ligge til grunn for arbeidet med kvalitetsutvikling.

Arbeidet med de nye prøvene skal altså ses i sammenheng med ambisjonen om å styrke det systematiske arbeidet med kvalitet i skolen. Det betyr at de nye prøvene må fungere godt sammen med andre verktøy som skal bidra til helhetlig kvalitetsarbeid. For elevene skal prøvene bidra til læring, trivsel og utvikling, i tråd med målene i læreplanverket.

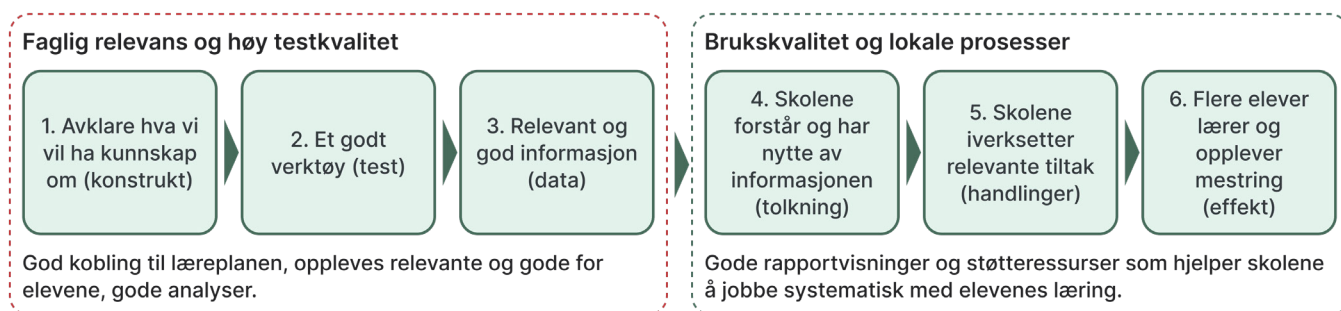
KD har gitt Udir i oppdrag å utvikle en ny skolemiljøundersøkelse som kan erstatte dagens elevundersøkelse. Disse oppdragene må ses i sammenheng, og vi må utrede hvordan vi kan legge til rette for at informasjon fra disse verktøyene kan kombineres. Udir vil oversende sine forslag til utforming av en ny skolemiljøundersøkelse i september 2026. Her vil også koblingen til de nye prøvene som del av en helhetlig tilnærming til kvalitetsarbeid bli drøftet. Her vil det være naturlig å også komme nærmere inn på hvordan disse kildene kan knyttes sammen med skolene og kommunenes øvrige kvalitets- og utviklingsarbeid.

1.2 Hvordan forstå bestillingen om «læringsstøttende prøver»

I delleveranse 1 knyttet vi det læringsstøttende til begrepet formative prøver, og skrev at vi ser dette som «et instrument eller en prosess der bevis for læring samles inn, vurderes og brukes av lærere i undervisningen for å støtte elevers videre læring».

«Bevis for læring» handler om at prøvene må gi informasjon vi kan stole på. «Bruk for å støtte læring» betyr at det ikke er tilstrekkelig at prøvene i seg selv gir pålitelig informasjon, men at denne informasjonen også er relevant for arbeidet med elevenes læring. Det læringsstøttende formålet handler om at prøver utformes slik at de fremmer faglig forståelse, gjør sentrale kriterier i det som vurderes tydelig for både lærere og elever, og muliggjør meningsfulle tilbakemeldinger (Carless, 2007).

I figur 2 har vi tegnet opp det vi oppfatter er sentrale antagelser bak en læringsstøttende prøve (bearbeidet fra Nordenbo et al. (2009)). Rundt kausalkjeden ligger faktorer som kan forsterke eller svekke koblingen mellom de ulike stegene i modellen, og dermed også utbyttet av prøven. Modellen fanger ikke opp alle faktorer som har betydning, men kan tjene som utgangspunkt for en strukturert samtale om hvilke forbedringer vi kan ta sikte på, og hvordan disse henger sammen.



Figur 2: hva gjør en prøve «læringsstøttende»?

I figur 2 markerer vi to typer forbedringer som begge er nødvendige for at prøvene skal fungere læringsstøttende. Innenfor den røde stiplede linjen vises forbedringer knyttet til prøvenes konstrukt og faglige begrunnelser, kvaliteten på selve prøven og hvilke data den gir. Innenfor den grønne stiplede linjen vises prosessene rundt prøvene og skolenes forutsetninger for å nyttiggjøre seg av dem. En drøfting av reliabilitet og validitet for de nye prøvene må, slik vi ser det, legges en slik forståelse til grunn. Faktorene innenfor den røde stiplede linjen kan vi påvirke direkte. Faktorene innenfor den grønne stiplede linjen er lokale forhold som vi kan påvirke indirekte ved å legge til rette for- og støtte god praksis.

Konstrukt

Et konstrukt er de egenskapene, kompetansene eller ferdighetene som vi forsøker å måle i en prøve (American Educational Research Association [AERA] et al., 2014). Et konstrukt kan være definert basert på en læreplan eller i et rammeverk. I de nasjonale prøvene i lesing er for eksempel konstruktet beskrivelsen av lesing som grunnleggende ferdighet, slik det er definert i rammeverket for grunnleggende ferdigheter.

Validitet

Validitet handler om gyldigheten til resultatene fra en prøve og i hvilken grad en prøve måler det vi forsøker å måle (AERA et al., 2014). Validiteten svekkes hvis konstruktet ikke måles på en god nok måte, eller hvis noe annet enn konstruktet måles i prøven. Figuren over angir at validitet også må ses opp mot formålet med prøven.

Reliabilitet

Reliabilitet handler om i hvilken grad vi kan forvente like resultater dersom en prøve gjennomføres på nytt under samme forutsetninger (AERA et al., 2014). Resultatene fra prøvene må være reliable og pålitelige for at tolkningen av resultatene skal være gyldige, altså valide.

1.3 Prøvenes formål og avveining av ulike hensyn

Kvalitetsutviklingsutvalget peker på avveining av ulike hensyn og kompromisser man må inngå for å møte ambisjonen om å gi *lærere relevant, umiddelbar og utviklingsorientert informasjon om elevene til det faglig-pedagogiske arbeidet*. Et av dilemmaene handler om hvordan både enkeltelementer og systemet som helhet kan støtte ulike aktører i tråd med de rollene og ansvaret de har for skolen.

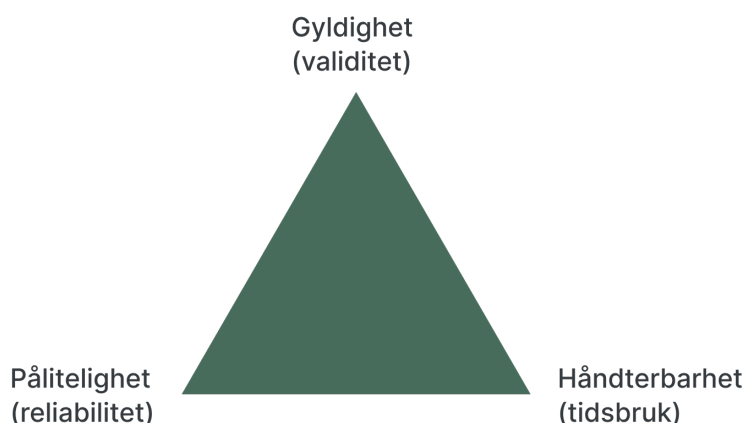
Utfordringen med «det doble formålet» presenteres som en konflikt mellom informasjon til styring og informasjon til utvikling. I realiteten er dette ofte integrerte prosesser, og det foregår både styring og utvikling på alle nivåer. Derfor er det mer presist å spørre hvilken informasjon ulike aktører trenger for å ivareta sine oppgaver på en god måte.

Gyldighet, pålitelighet og håndterbarhet

Kvalitetsutviklingsutvalget peker på at ulike aktører har behov for ulike type informasjon. Utvalget legger til grunn at styringsperspektivet, og dermed også kontrolldimensjonen, har vært for dominerende. Videre peker utvalget på at dette har hatt betydning for hvilke prioriteringer som er gjort i arbeidet med dagens prøver.

Dagens kartleggingsprøver og nasjonale prøver har høy testteoretisk kvalitet. Det betyr at informasjonen fra prøvene har høy gyldighet og pålitelighet knyttet til det prøvene måler. Dagens prøver er altså et uttrykk for at pålitelighet er vektlagt svært høyt. Begrunnelsen for oppdraget om å utvikle nye prøver ligger i at relevans, forstått som bruksverdi for lærer, ikke har blitt vektlagt tilstrekkelig.

Utvalget peker på en tredje faktor som også har betydning for utformingen av disse prøvene og systemet som helhet, nemlig håndterbarhet og omfang. I figur 3, har vi satt opp håndterbarhet som en rammefaktor på linje med gyldighet og pålitelighet. Vi er enig med utvalget i at en prøve alltid vil måtte ses som et kompromiss mellom disse tre hensynene.



Figur 3: ulike hensyn når man utvikler en prøve

Nye prøver, ny prioritering

I dagens prøver er altså hensynet til målepresisjon vektet høyt. De nasjonale prøvene skal gi så presis informasjon som mulig om noen utvalgte ferdigheter basert på en gjennomføring på inntil 60 eller 90 minutter. Dagens prøver har utviklingstid på to år, og denne tiden brukes blant annet til å prøve ut oppgavene for å sikre at de fungerer som intendert. Etter hver gjennomføring utarbeider fagmiljøene en teknisk rapport hvor de gjør rede for oppgaveutviklingen.

For at nye prøver og fleksible verktøy skal bli mer nyttige for skolene, må de gi rikere informasjon om elevenes læring og det må bli tydeligere hvordan de kan følge opp resultatene. Med utgangspunkt i modellen kan vi da si at hensynet til gyldighet, forstått som opplevd relevans, økes. Det blir da også interessant å drøfte hvor mye tid elevene skal bruke på denne type kartlegginger, og hvilke krav vi skal stille til presisjon i den informasjonen vi presenterer. Dette drøfter vi mer utfyllende i kapittel 5.

Utstrakt bruk av andre prøver og undersøkelser i norsk skole

Kvalitetsutviklingsutvalget kommenterer at mange skoler bruker andre prøver og undersøkelser enn de nasjonale myndigheter pålegger skoler- og skoleeier gjennom kvalitetsvurderingssystemet. Utvalget drøfter også hva dette kan ses som et uttrykk for. Mange verktøy som brukes i dag er dårlig dokumentert, og har varierende kvalitet (Arnesen et al., 2017). Dette kommenterer vi nærmere i kapittel 8.

Oppdragsbrevet (OB 2025-018) angir at det samlede omfanget av obligatoriske tester og prøver ikke skal øke. Vi oppfatter at formuleringen peker på prøver som blir initiert fra statlige myndigheter. Samtidig tas mange beslutninger på skole- og kommunenivå om bruk av denne typen verktøy. Noen kommuner utvikler egne prøver, mens andre benytter verktøy utviklet av andre aktører. Gode statlige prøver og verktøy gjør at kommunene i mindre grad må utvikle eller kjøpe egne verktøy, og det kan bidra til å heve kvaliteten på kartleggingsarbeidet i skolene. Vi mener derfor at en diskusjon om omfanget av denne typen kartlegginger og prøver bør ta utgangspunkt i et helhetlig blikk på kartlegging i skolen, og ikke kun prøver eller verktøy utviklet av statlige myndigheter.

Prøvenes rolle i arbeidet med elevenes læring

Vi ønsker at de nye prøvene skal gi rikere informasjon, samtidig som denne informasjonen skal være pålitelig og presis. Vi vil heller ikke at elevene skal bruke for mye tid på denne typen kartlegginger. Vi viste over til Kvalitetsutviklingsutvalgets drøfting av ulike hensyn og kompromisser man må inngå for å møte ambisjonen om å gi *lærere relevant, umiddelbar og utviklingsorientert informasjon om elevene til det faglig-pedagogiske arbeidet*. Forslagene fra arbeidsgruppene som har utredet mulige måter å utforme nye prøver på, viser hvordan disse lærerne ser for seg å finne et nytt kompromiss mellom disse hensynene.

Det faglig-pedagogiske arbeidet inkluderer både generell tilrettelegging for alle (§ 11-1), og oppfølging av elever med behov for særskilt tilrettelegging (§11-3). Dagens nasjonale prøver er utformet for å gi informasjon om alle elever som grunnlag for underveisvurdering og utvikling av kvaliteten. Kartleggingsprøvene er utformet for å gi informasjon om elever som har behov for ekstra støtte i den tidlige lese- og regneopplæringen. Begge verktøyene forutsetter videre kartlegging av elever som har behov for særskilt tilrettelegging.

Kartleggingsprøvene på 1.trinn skal erstattes med fleksible verktøy som kan støtte lærerne i å følge opp elever som har behov for det. Det betyr at vi vil legge økt vekt på støtte lærerens oppfølging av elever som ikke følger forventet progresjon i begynneropplæringen.

2. Prosjekt Nye prøver

For å løse oppdraget med utvikling av nye prøver med læringsstøttende formål har Udir satt ned et eget prosjekt med tverrfaglige ressurser. Prosjektet har blant annet kompetanse innen prøveutvikling, testteori, analyse, systemutvikling, tjenestedesign, involveringsarbeid og prosjektledelse. Prosjektet er bemannet for å kunne løse utredningsoppgavene knyttet til delleveranse 2 og 3 i oppdraget. Utviklingen av nye verktøy vil kreve økte ressurser. Udir arbeider med å styrke forutsetningene for å kunne løse denne delen av oppdraget.

2.1 Forprosjekt

Våren 2025 etablerte Udir et forprosjekt for å forberede arbeidet med oppdraget. Forarbeidet våren 2025 førte til at Udir raskt kunne sette i gang med konkrete tiltak rett etter at oppdragsbrevet ble mottatt i juni 2025. Et eksempel på et slikt tiltak er at prosjektet raskt kontaktet statsforvaltere for å komme i kontakt med aktuelle skoler for å innhente brukerinnsett til arbeidet. Forprosjektet samlet og kartla også eksisterende kunnskap på prøvefeltet, og det ble utarbeidet en plan for å sikre sammenheng mot andre prosesser i Udir og involvering av eksterne aktører tidlig i arbeidet.

2.2 Utvikling av prosjektet

I juni 2025 opprettet Udir et prosjekt som rapporterte til linjeorganisasjonen. Over tid har prosjektet blitt utvidet med ulike roller. I tillegg kommer representanter fra andre team som er viktig å inkludere i arbeidet, slik som analyseteamet i Udir. Et viktig moment er at prosjektet har tverrfaglig kompetanse og har evnen til å håndtere ulike fagområder som kreves for utvikling av nye prøver, og tjenesteutvikling generelt.

I delleveranse 1 orienterte vi om at vi ønsker å arbeide med dette med en innovativ og smidig tilnærming. Dette innebærer at prosjektet utvikler innsikt knyttet til spesifikke spørsmål i arbeidet. Denne innsikten supplerer eksisterende kunnskap på prøvefeltet. Ved å snakke direkte med de som skal bruke de nye prøvene og hvordan de benytter disse verktøyene i pedagogisk arbeid med egne elever, kommer prosjektet og Udir tettere på praksisfeltet. Det er de målgruppene som skal anvende de nye prøver som kan beskrive behov som nye prøver skal dekke fra et praksisnært perspektiv.

2.3 Involvering

I tråd med føringen i oppdragsbrevet har prosjektet lagt vekt på å informere om arbeidet underveis, og aktivt involvere de som blir berørt av utviklingsarbeidet. Det har blitt etablert arenaer hvor både interne og eksterne målgrupper har kunnet gi sine innspill til pågående prosjektaktiviteter og veivalg.

Udir har informert om arbeidet og innhentet tilbakemeldinger i en egen møteserie hvor både ledelsen i Udir og lederne for organisasjonene i utdanningssektoren har deltatt. Arbeidet er også blitt drøftet i referansegruppen for vurdering hvor disse organisasjonene er representert. I møtene er arbeidet med nye prøver blitt drøftet i lys av arbeidet med ny skolemiljøundersøkelse og et helhetlig kvalitetsutviklingssystem.

Prosjektet har blant annet informert om arbeidet i egne møter med blant annet Elevorganisasjonen, i ulike nettverk organisert av KS, og i møte med statsforvalterne. Statped har gitt innspill til rapportutkast og til plan for gjennomføring av oppdraget. Formålet med disse aktivitetene har vært at sentrale aktører har fått løpende informasjon om progresjon og målsettinger i arbeidet, og at de har fått mulighet til å dele sine perspektiver og gi innspill til arbeidet.

Prosjektet har opprettet en hjemmeside for arbeidet på udir.no hvor man kan finne oppdatert informasjon om status for arbeidet.

2.4 Innhenting av brukerinnsikt

Oppdragsbrev 2025-018 angir at Udir aktivt skal innhente synspunkt og erfaringer fra brukergruppene i utviklingsarbeidet. Prosjektet har etablert samarbeid med tolv strategiske samarbeidsskoler. Vi benytter også informasjon fra undersøkelser som gjennomføres i etterkant av dagens nasjonale prøver og kartleggingsprøver. Disse undersøkelsene gir oss spesifikke tilbakemeldinger på hvordan disse prøvene brukes, og hvordan de kan forbedres. Under utdyper vi hvordan vi har arbeidet med å innhente disse tilbakemeldingene.

Brukerinnsikt gjennom dybdeintervjuer med skoler

De strategiske samarbeidsskolene er altså valgt ut i dialog med statsforvalterne. Vi har lagt vekt på å få med ulike typer skoler, og vi har en god blanding av barneskoler, kombinert 1-10 skoler og ungdomsskoler. Skolene er geografisk spredt og varierer i størrelse og elevgrunnlag. Utvalget er begrenset, og vi er derfor varsomme med å tolke deres svar som gyldig for alle skoler.

Dialogen med disse skolene gir oss bedre innsikt i behovsbeskrivelser og eksisterende praksis. Samtalene med skolene gir oss et mer detaljert og praksisnært bilde av hvordan prøvene, tilhørende løsninger og ressurser faktisk brukes og oppleves i skolehverdagen. Innsikten er ment å være et supplement til eksisterende kunnskapsgrunnlag på prøvemarket, og den utdyper og konkretiserer behov og ønsker på et brukernivå. Systematisering av skolenes erfaringer er nyttig når vi skal utvikle disse tjenestene. Temaene omhandlet blant annet hvordan informasjon fra prøvene inngår i undervisnings- og oppfølgingsarbeidet, erfaringer med bruk av Prøveadministrasjonssystemet (PAS) og hvilke former for støtteressurser som oppleves eller kan oppleves som nyttige i praksis.

Det er gjennomført to runder med brukerinnsikt. Den første runden ble gjennomført høsten 2025, og omfattet besøk ved alle de tolv strategiske samarbeidsskolene. Formålet var å få nærmere innsikt i erfaringer med dagens nasjonale prøver og behov knyttet til utvikling av nye prøver med læringsstøttende formål. Den andre runden ble gjennomført våren 2026 og var avgrenset til samarbeidsskoler med 1. trinn. Denne runden var spesifikt knyttet til den delen av oppdraget som gjelder å erstatte dagens kartleggingsprøver på 1.trinn med forbedrede og mer fleksible verktøy. Vi ønsker også å bruke disse skolene for å få tilbakemelding på ideer og konsepter i en tidlig fase.

Spørreundersøkelser til lærere

Etter hver gjennomføring av nasjonale prøver og kartleggingsprøver sendes det ut en spørreundersøkelse til lærere og administrerende roller som har vært involvert i prøvegjennomføringen. I undersøkelsen skilles det mellom de nasjonale prøvene og kartleggingsprøvene på 1. og 3. trinn. Høsten 2025 mottok Udir til sammen 3798 besvarelser. 1 187 av disse var svar fra skoleadministratorer og gjaldt alle prøvetyper. 1 765 gjaldt nasjonale prøver besvart av lærere, og 846 gjaldt kartleggingsprøvene besvart av lærere.

I disse undersøkelsene får vi tilbakemeldinger og vurderinger knyttet til blant annet forarbeid, gjennomføring og oppfølging av prøvene, samt bruk av Prøveadministrasjonssystemet (PAS). Undersøkelsene er frivillige og anonyme. Resultatene gir indikasjoner på erfaringer og mønstre blant respondentene med utgangspunkt i deres

faktiske bruk. Dette gir oss konkrete innspill til hvordan denne typen tjenester kan forbedres, for eksempel med tanke på forenklinger og redusert tidsbruk for lærere.

3. Forankring av arbeidet i eksisterende kunnskapsgrunnlag

Vurderinger av de nasjonale prøvene og kartleggingsprøvene (heretter også kalt prøvene) bygger på et bredt og omfattende kunnskapsgrunnlag. Både NOU 2023:1 *Kvalitetsvurdering og kvalitetsutvikling i skolen – Et kunnskapsgrunnlag*, NOU 2023:27 *Et nytt system for kvalitetsutvikling – for elevenes læring og trivsel* og Meld. St. 34 (2023–2024) *En mer praktisk skole*, gir gode beskrivelser av hvordan dagens prøver og andre elementer i kvalitetsvurderingssystemet fungerer i praksis, og av utfordringer knyttet til disse.

Dokumentene peker samlet på at dagens nasjonale prøver i varierende grad oppleves som nyttige for skolene, og at det er spenninger mellom ulike formål. Det beskrives utfordringer knyttet til relevans for lærernes faglig-pedagogiske arbeid, balansen mellom læringsstøtte og styringsinformasjon. Lærerne trenger tilgang på lett tilgjengelig informasjon som kan brukes i undervisningsarbeid. Samtidig understreker kunnskapsgrunnlaget det at prøvene fortsatt er en sentral informasjonskilde for både lærere, skoleledere og skoleeiere, og en del av grunnlaget for å følge elevenes læringsutbytte og utviklingen i skolen.

Dette overordnede bildet utdypes også i Udirs tidligere rapport *Prøver i fagfornyelsen – hvilke behov skal prøvene dekke?* (Utdanningsdirektoratet [Udir], 2020) (Heretter kalt prøver i fagfornyelsen), som samler erfaringer fra evalueringer, forskning og brukerinnsikt over tid. I tillegg utdyper brukerinnsikten som er gjennomført i arbeidet med dette oppdraget allerede kjente behovsbeskrivelser.

Gjennomgang av kunnskapsgrunnlaget viser mange felles trekk i behovsbildet, samtidig som ulike kilder drøfter disse verktøyene på ulike måter. Vi samler her noe av det vi oppfatter som sentrale behovsbeskrivelser i kunnskapsgrunnlaget.

3.1 Støtte til lærernes faglig-pedagogiske arbeid

I NOU 2023: 27 vises det til funn om at lærere i mindre grad enn skoleledere og skoleeiere opplever informasjonen fra prøvene som nyttig i det daglige arbeidet med elevenes læring. Lærere peker blant annet på at det kan være vanskelig å se sammenhengen mellom resultater fra prøvene og egen opplæringspraksis, og at ferdighetsområdene fra prøvene ikke alltid gir tilstrekkelig støtte for tilpasning av undervisningen. Dette bidrar til at mange skoler supplerer de nasjonale prøvene med andre verktøy for å få mer inngående informasjon. Behovet for at lærerne må få umiddelbare og konkrete tilbakemeldinger om elevenes læring understrekes også.

Prøver i fagfornyelsen (Udir, 2020) peker på at lærere i hovedsak har behov for informasjon som kan brukes til planlegging av undervisning, vurdering av elevs faglige ståsted og oppfølging av enkeltelever og elevgrupper. Dette innebærer støtte til underveisvurdering gjennom resultater som forteller detaljert om elevenes prestasjonsnivå som legger et godt grunnlag for tilbakemeldinger og oppfølging.

For at prøver skal oppleves som nyttige, er det nødvendig at resultatene skal kunne sees i sammenheng med lokal kontekst og være tilgjengelige på et tidspunkt og i et format som gjør dem handlingsrelevante.

3.2 Tydeliggjøre formål og innhold

Prøver i fagfornyelsen viser at det er stor variasjon i hvordan skoler tolker prøvenes hensikt (Udir, 2020). Tilsvarende funn i prosjektets brukerinnsikt tyder på at flere skoler tolker de nasjonale prøvene som rene fagprøver, mens andre ser dem som prøver i grunnleggende ferdigheter. I flere av tilbakemeldingene blir for eksempel regneprøven omtalt med referanser til fagspesifikke matematiske ferdigheter snarere enn til å kunne regne som grunnleggende ferdighet. Dette har betydning for forberedelser og oppfølging, og det peker på behov for å tydeliggjøre formål, forventninger og sammenheng mellom prøvene, ferdighetsområdene og relasjon til læreplanverket.

3.3 Støtte til å forstå resultater og å sette disse i sammenheng

NOU 2023:27 beskriver at dagens kvalitetsvurderingssystem i hovedsak gir kvantitative data, og at dette begrenser nytteverdien for skolens undervisnings- og utviklingsarbeid. Utvalget peker på at kvantitativ informasjon fra prøvene tillegges høy verdi i kvalitetsvurderingen, mens kvalitative vurderinger og annen kontekstuell informasjon i mindre grad settes i et system og tas i bruk.

Prøver i fagfornyelsen beskriver at lærere etterspør bedre og mer brukervennlig tilgang til kvalitativ informasjon om elevenes besvarelser, oppgaveanalyser og forklaringer som kan bidra til å forstå elevenes strategier, og misoppfatninger.

Brukerinnsikten samlet inn av prosjektet viser at skolene bruker betydelige ressurser på å bearbeide prøveresultater i egne systemer for å få et mer nyansert bilde av elevenes kompetanse, mestringsområder og utfordringer. Dette indikerer at løsningene i større grad bør legge til rette for at informasjon fra denne typen strukturerte kartleggingsverktøy kan bearbeides og ses i lys av annen informasjon de har om eleven og klassen. Dette kan også bidra til at kartleggingen i større grad følges opp gjennom systematisk arbeid med å utvikle opplæringen over tid.

3.4 Støtte til oppfølging av elevenes progresjon

NOU 2023:27 peker på at lærere, skoleledere og skoleeiere har et behov for å kunne følge med på elevenes progresjon gjennom opplæringsløpet. Prøver i fagfornyelsen (2020) kobler behovet blant annet til at det er en tydeligere progresjon og sammenheng i læreplanverket LK20/LK20S: *«Lærerne savner bedre verktøy for å måle progresjon: Lærere og skoler har få eller ingen verktøy for å følge elevenes faglige progresjon, og etterspør prøver som bygger på hverandre, der resultatene kan sammenlignes og ses opp mot læreplanene».*

I våre samtaler med skolene har vi blant annet spurt om bruk av informasjon fra prøvene på 9.trinn, og fått til svar at mange opplever denne som viktig fordi de her får en tilbakemelding på elevenes utvikling. Dette skrev vi også om i den første leveransen. Denne

prøven gir et utgangspunkt for dem å vurdere om de har lyktes med å løfte de svake leserne slik at de har markert færre på laveste nivå.

Skolene framhever generelt at de har behov for gode kunnskapskilder for å vurdere om arbeidet på skolen gir forventede resultater for elevens læring. Det vil alltid være usikkerhet knyttet til måling av utvikling over tid, særlig på individnivå. Derfor er det viktig å forstå at resultatene fra denne typen prøver har begrensninger, og at informasjon om endringer over tid må fortolkes med forsiktighet.

3.5 Tilrettelegging for en sammensatt elevgruppe

NOU 2023:27 påpeker at samfunnet har utviklet og endret seg siden dagens system ble etablert, og at elevgruppen i norsk skole er mer sammensatt og mangfoldig. Med dette som utgangspunkt blir det interessant å undersøke hvordan standardiserte verktøy og prøver også kan være fleksible nok til å være relevante for en stor andel av elevene.

Prøver i fagfornyelsen (Udir, 2020) peker på at ordvalg, begrepsbruk, tekstmengde og oppgaveutforming kan bidra til at elevenes språk- og leseferdigheter, eller deres kulturelle bakgrunn, får større betydning for prestasjonene enn de faglige ferdighetene prøvene skal gi informasjon om.

I brukerinnsikten til prosjektet trekkes særlig regneprøven frem som utfordrende, fordi omfattende tekstmengder og komplekse formuleringer kan gjøre det vanskelig å forstå elevenes faktiske kompetanse og på hvilke områder det er behov for å gi mer støtte. I samtaler trekker også noen skoler frem et behov for at leseprøvene i større grad bør gi nyansert informasjon om aspekter som påvirker elevenes lesekompetanse, slik som språkferdigheter og språkkompetanse, avkodingsferdigheter og leseflyt. Her vil det være nødvendig å vurdere om det i tillegg til prøvene også er behov å utvikle andre verktøy.

3.6 Resultatvisninger og støtteressurser

Det er behov for løsninger der prøver, resultater og støtteressurser i større grad henger sammen, og der støtten oppleves som lett tilgjengelig og anvendbar. Lærerne trenger i større grad enn skolelederne og skoleeierne å se hva enkeltelever mestrer for å utvikle egen undervisningspraksis. Slik informasjon finnes i dag, men det er tidkrevende å hente den ut og analysere den. Prøvene gir et presist bilde av faglig nivå uttrykt som skalapoeng og mestringsnivåer, men det er krevende for mange lærere å forstå hvordan dette henger sammen med deres undervisningspraksis og hvordan de skal følge opp resultatene.

Både NOU 2023:27 og Prøver i fagfornyelsen (Udir, 2020) peker på at lærere og skoleledere etterlyser mer konkrete, tilgjengelige og praksisnære støtteressurser knyttet til prøvene. Dette gjelder både bedre veiledning i hvordan de skal tolke resultater og forslag til videre undervisning og oppfølging.

Brukerinnsikten viser at støtteressursene som i dag finnes på udir.no og i prøveadministrasjonssystemet, oppleves som faglig relevante, men ofte for omfattende, lite synlige eller krevende å ta i bruk i en travelt skolehverdag. Flere skoler beskriver at de i liten grad bruker eksisterende støttemateriell, og at behovet for tydeligere sammenheng mellom prøveresultater og konkrete pedagogiske grep er stort.

3.7 Behov i den samiske skolen

Et nytt system for kvalitetsutvikling også skal gi informasjon om elever i den samiske skolen. Når det utvikles nye prøver og verktøy skal disse være tilpasset samiske elever og behovene i den samiske skolen. I tillegg kan det være behov for andre verktøy spesielt tilpasset kvalitetsutvikling i samisk skole. Sametinget har etterspurt bedre informasjon om språkutviklingen til samiske elever som utgangspunkt for at de skal kunne ivareta sitt ansvar. Arbeidsgruppen som har utredet forslag til nye leseprøver gir sine vurderinger av hvordan dette kan løses i kapittel 4. I kapittel 8 drøfter Udir hvordan nye prøver kan svare til behovet for informasjon om læringsutbytte hos elever i samisk skole.

4. Forslag fra arbeidsgruppene

De nye prøvene skal ha et tydelig læringsstøttende formål og være gode pedagogiske verktøy for lærerne i arbeidet med å støtte elevenes læring. For å lykkes med dette vil Udir utvikle de nye prøvene i tett samarbeid med lærere, og vi ønsker å teste ut ulike ideer for å se om de fungerer som intendert. Dette skrev vi om i den første delleveransen til KD i desember.

Våren 2026 har Udir engasjert 13 lærere som har fått i oppdrag å utrede mulige konsepter for nye læringsstøttende prøver. Lærerne har vært organisert i tre faggrupper, matematikk, engelsk og lesing. De representerer ulike deler av landet og skoler med ulike kjennetegn. Gruppenes mandat har vært å gi råd om utvikling av læringsstøttende prøver, med særlig vekt på å synliggjøre brukerbehov og praktiske rammer i skolehverdagen. Gruppene har utviklet konsepter for hvert sitt fagområde, men har også samarbeidet med hverandre og drøftet hverandres ideer.

I sine anbefalinger beskriver gruppene hvordan de mener nye prøver bør utformes, hva de bør kartlegge, og hvilket støttemateriell som bør følge dem. De legger til grunn at prøvene skal være læringsstøttende og at de skal få informasjon som er relevant, forståelig og anvendbar i deres oppfølging av elevenes læring. De har tatt utgangspunkt i føringene i oppdragsbrevet om at prøvene skal ha en tett kobling til læreplanen, at de skal støtte god undervisningskvalitet og at de skal gi et mer nyansert bilde av elevenes kompetanse enn det dagens prøver gir.

Arbeidsgruppene peker på noen overordnede krav til de nye prøvene. De nye prøvene må være faglig kvalitetssikrede, og de må ha et tydelig formål og brukes riktig. Det bør tilbys god fagdidaktisk støtte i tilknytning til prøvene. Prøvene må gi bedre informasjon om elevenes faglige utvikling og gi god nytte i forhold til tiden de bruker på å planlegge, gjennomføre og følge dem opp. Vi bør tilby løsninger som er tilpasset ulike brukernivå.

Før vi presenterer forslagene fra hver av gruppene, gir vi under en sammenstilling av hva som er felles i de tre gruppens forslag.

Forklaring av begreper som brukes i presentasjonen av arbeidsgruppenes forslag

Lineære prøver

I lineære prøver får alle elever tildelt de samme oppgavene i samme rekkefølge. Dagens nasjonale prøver er lineære prøver.

Adaptive prøver

I adaptive prøver tildeles elevene oppgaver med ulik vanskegrad basert på hvilke oppgaver de har fått til tidligere i prøven (AERA et al., 2014). Elevene kan enten tildeles én og én oppgave (oppgave-for-oppgave-adaptivitet) eller de kan tildeles grupper av oppgaver (blokkadaptivitet). Elevene svarer derfor på ulike oppgaver i adaptive prøver.

Blokkadaptive prøver

I blokkadaptive prøver får elevene tildelt grupper eller blokker med oppgaver med ulik vanskegrad (Yan et al., 2014). Første blokk har middels vanskegrad og tas av alle elevene. Deretter tildeles de blokker som passer deres ferdighetsnivå. Antall ganger elevene blir tildelt en blokk (steg), og antall blokker per steg, kan variere, og kompleksiteten i den blokkadaptive prøven påvirker hvor presis informasjon prøven gir om elevene.

Standardisering

Standardiserte prøver har identisk eller nesten identisk innhold for alle som gjennomfører prøven, og forutbestemte regler for gjennomføring, skåring og tolkning av prøvene (AERA et al., 2014). Et av formålene med en standardisert prøve er å skape sammenlignbare resultater.

Kriteriebasert

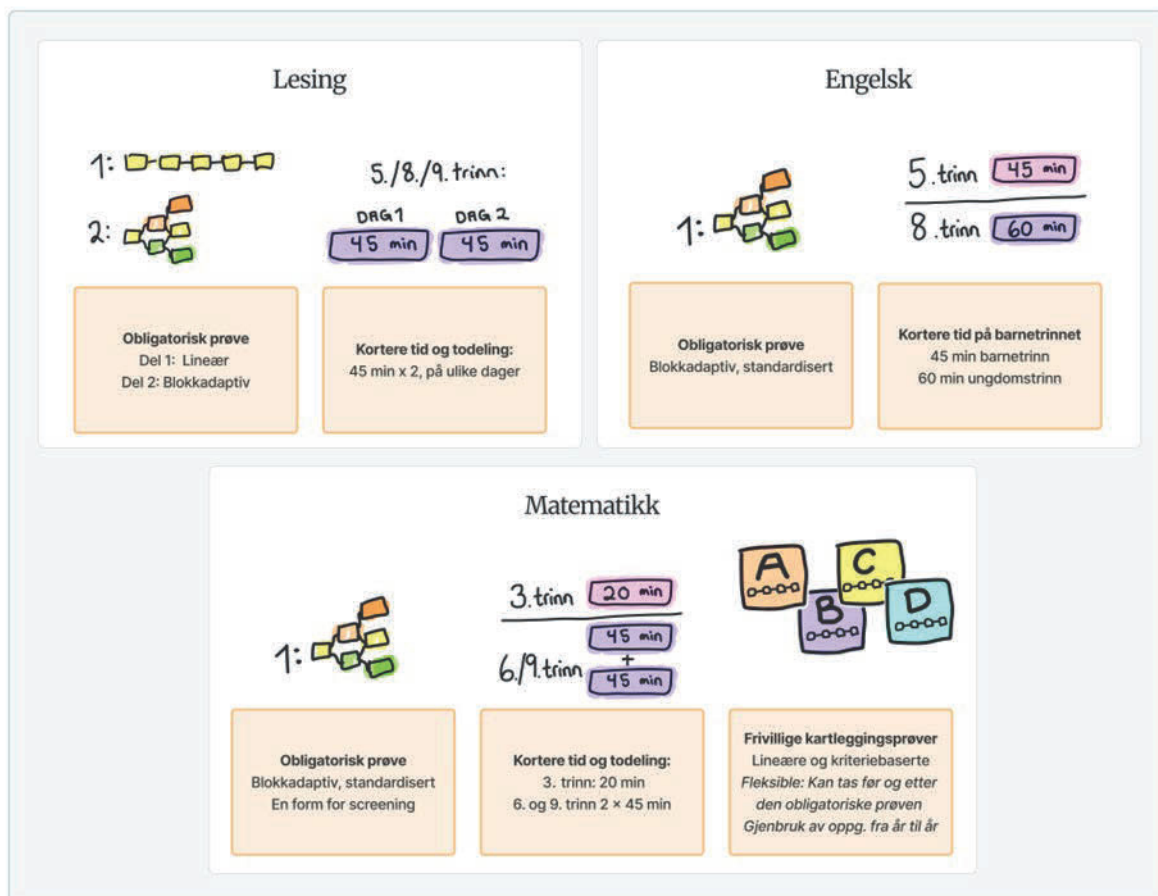
Kriteriebaserte resultater er resultater som blir presentert relativt til satte kriterier, istedenfor for eksempel relativt til et gjennomsnitt (AERA et al., 2014). Et kriteriebasert resultat kan for eksempel være at en viss poengsum reflekterer god måloppnåelse.

Retest

Med retest mener vi i denne sammenhengen en ny prøve som gjennomføres en viss tid etter gjennomføringen av en prøve. Formålet med en retest kan være å få informasjon om utviklingen av elevenes ferdigheter over tid.

4.1 Fellestrekk i konseptene for engelsk, matematikk og lesing

Figur 4 viser en enkel fremstilling av de tre prøvekonseptene. De ulike prøvekonseptene har til felles at elevene skal ha en god opplevelse av prøven, at prøvene må inngå som del av en helhetlig prosess og at prøvene kan gi informasjon om elevenes utvikling over tid. De har også forslag til hvordan resultatvisningen kan utformes slik at den både blir enklere å bruke og samtidig gir et rikere bilde av elevenes faglige ferdigheter. Dette skal samlet sett bidra til at skolene får bedre støtte til å følge opp resultatene. Når vi skal følge opp konseptforslagene videre, må vi avklare i hvilken grad det bør være et felles konsept for prøvene i engelsk, matematikk og lesing, eller om faglige forskjeller tilsier at vi bør åpne for ulike konsepter.



Figur 4: forenklet visuell fremstilling av prøven i de tre konseptene

Prøver som skaper motivasjon og mestring

Alle gruppene er opptatt av at elevene skal ha en positiv opplevelse av å gjennomføre disse prøvene. Hovedforslagene deres handler om lengden på prøven, og at elevene i større grad skal få oppgaver som er tilpasset deres faglige nivå. Slik vil de unngå at en elev for eksempel opplever å mestre svært få oppgaver i løpet av en lang prøvegjennomføring.

Dagens prøver i lesing og regning gjennomføres innenfor 90 minutter. Engelskprøvene varer inntil 60 minutter. Arbeidsgruppene mener prøvene bør bli kortere. De ønsker også prøver som er helt eller delvis blokkadaptive. Dette er for gruppene primært begrunnet i elevenes prøveopplevelse. Med blokkadaptive prøver får elevene oppgaver som i større grad er tilpasset deres nivå.

Arbeidsgruppene i lesing og matematikk anbefaler å dele opp prøven ved at elevene har mulighet til å ta en prøve over flere dager eller ulike delprøver. Begrunnelsen for forslaget er at dette gir mindre belastning for elevene, samtidig som det gir mer informasjon om både hver elev og klassen/gruppen.

Arbeidsgruppene anbefaler at prøvene bør legge til rette for motivasjon og mestring gjennom et mer engasjerende design og innhold. Elevperspektivet for gruppene handler

både om en positiv prøveopplevelse for eleven, og at prøvene skal gi god og sikker informasjon om alle elevene, uavhengig av ferdighetsnivå.

Prøvene som prosess

Gruppene ønsker at de nye prøvene og tilhørende støtteressurser skal bli en integrert del av undervisningen. Arbeidsgruppene har konkrete forslag til modeller som inneholder flere elementer enn hva dagens prøver har. Formålet er å gi god støtte til læringsarbeidet som skjer i klasserommet, og anerkjenne at dette er prosesser som tar tid. Arbeidsgruppene anbefaler også at det utvikles egne ressurser som utgangspunkt for skoleledere og for arbeid i profesjonsfellesskapet.

Bedre støtte til å følge opp elevenes utvikling over tid

Arbeidsgruppene ønsker at nytt konsept skal gjøre det mulig å følge elevenes faglige utvikling over tid. En måte å gjøre dette på, er å kunne gjennomføre den samme eller tilsvarende prøve etter en viss tid. Videre i dokumentet omtaler vi denne type tester som *retest*. Slike retester kan i prinsippet være både frivillige og obligatoriske. Arbeidsgruppene foreslår også flere frivillige prøver eller moduler som kan gjennomføres etter behov. Formålet er å se om skolens arbeid bidrar til elevenes læring, og om elevene har den nødvendige progresjonen. Gruppene mener at resultater bør følge eleven ved lærer- eller skolebytte slik at sentral informasjon ikke forsvinner, men at den kan brukes i læringsarbeidet også etter ulike typer overganger.

Rikere resultatvisninger og bedre informasjon til ulike målgrupper

I tillegg til at resultatene bør følge elevene over tid, anbefaler arbeidsgruppene at det utvikles rike resultatvisninger som viser elever og klasser/gruppers styrker og utviklingsområder. Resultatvisningene bør i større grad enn nå peke fremover. Løsningene skal være tilpasset ulike målgrupper, og skal gi forbedret støtte til å formidle resultatene til elever og foresatte for eksempel med egne visninger for disse.

Alle tre konseptene omtaler en form for flerdimensjonalt resultat, men med litt forskjellig ordbruk: delferdigheter, områder, etc. Vi diskuterer dette her som del av en ny resultatvisning. Dette vil imidlertid også ha konsekvenser for utforming av prøvene, og må da ses i sammenheng med hvor omfattende prøve eller prøver elevene skal gjennomføre. Dette drøftet vi i kapittel 1.3.

Mer treffsikre oppfølgingsressurser

Arbeidsgruppene mener dagens støtteressurser er for generelle. De anbefaler ressurser som er mer konkrete og som er tettere koblet til resultatene. Dette kan for eksempel være forslag til undervisningsopplegg, sjekklister og digitale læringsstier. Omfanget av støtteressurser bør være langt større enn i dag, men støtteressursene skal være frivillige å bruke, og skolene skal kunne bruke dem når de har behov for det i løpet av skoleåret. Arbeidsgruppene anbefaler også at det utvikles ressurser som er rettet direkte mot foresatte.

4.2 Engelskgruppens forslag og anbefalinger

Arbeidsgruppen for engelskprøven (heretter kalt *engelskgruppen* eller *gruppen*) har bestått av fem lærere med lang erfaring som engelsklærere i grunnskolen, både på barne- og ungdomstrinnet. Lærerne jobber ved skoler i ulike deler av landet. Gruppen har forankret konseptet sitt i forskning på vurdering for læring, formativ vurdering, kommunikativ og interkulturell kompetanse og motivasjonsteori, samt i læreplanverket LK20. De har også støttet seg til intervjuene med de strategiske samarbeidsskolene som ble gjennomført høsten 2025, og drøftet med utgangspunkt i egen profesjonserfaring.

Faglig utgangspunkt for forslaget

Teoretisk støtter gruppen seg på forskning om vurdering for læring, språkdidaktiske teorier og rammeverk for kommunikativ og interkulturell kompetanse (inkl. CEFR), blant annet hos Brevik & Helness (2018), Canale & Swain (1980) og Speitz & Myklevold (2022). I tillegg viser gruppen til forskning knyttet til narrativ ramme, motivasjon og validitet. Dette danner bakteppet for gruppens anbefaling om å utvikle et nytt prøvesystem i engelsk som i større grad fungerer som et praksisnært fagdidaktisk verktøy for lærere og skoler.

Gruppen tar utgangspunkt i at engelskfaget både er et språkfag og et dannelsesfag, der interkulturell forståelse og kommunikasjon står sentralt. Prøven skal derfor gi et så rikt bilde som mulig av elevenes faktiske språkkompetanse, og omfatter både reseptive ferdigheter og sentrale sider ved kommunikativ kompetanse. Gruppen foreslår derfor en prøve som måler *Reading & Listening* og *Use of English*. Dette innebærer en utvidelse av konstruktet for prøven sammenlignet med dagens prøve i engelsk. Vi presenterer noen faglige vurderinger av dette i slutten av dette delkapittelet.

Engelskgruppen ønsker at de nye prøvene i større grad enn i dag skal gi et nyansert bilde av elevenes ferdigheter. Samtidig ønsker de at det skal bli lettere for lærere å finne de sentrale læringspunktene gjennom en tydelig resultatvisning og tilgjengelige pedagogiske ressurser som legger til rette for målrettet oppfølging. Gruppen har vært opptatt av at prøven tilpasses elever med ulike tilretteleggingsbehov og kompetansenivåer, og gruppen har brukt mye tid på å diskutere hvordan en slik prøve kan bli en positiv opplevelse for elevene. Hensikten med konseptforslaget er å utvikle et godt pedagogisk verktøy, forankret i en moderne språkdidaktisk tankegang, med verdi for lærerens undervisning og elevenes læring.

Konseptet er utviklet for å møte det gruppen vurderer som de viktigste behovene i skolen. Gruppen viser samtidig til at engelsklærere har ulik formell kompetanse, og at skolars tilgang til tilpassede ressurser og digitale læreverk varierer. Det er derfor et mål å utvikle gode felles verktøy som kan bidra til å redusere forskjeller og gi mer likeverdig støtte til undervisning og oppfølging. Konseptet inkluderer flere frivillige støtteressurser som skal støtte mer tilpasset opplæring og hjelpe lærerne i arbeidet med å tilpasse opplæringen, samtidig som eleven i større grad kobles på egen læring og progresjon.

Konseptbeskrivelse

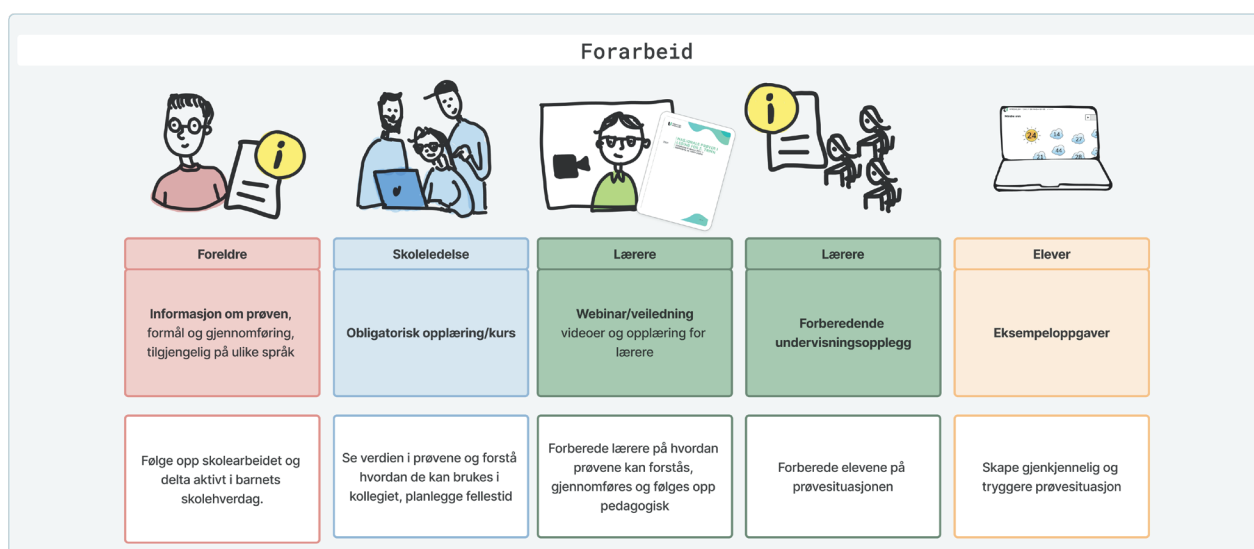
Engelskgruppen ønsker én obligatorisk prøve som er standardisert og blokkadaptiv. Tidsrammen er maks 45 minutter på barnetrinnet og 60 minutter på ungdomstrinnet. Gruppen anbefaler også en narrativ ramme som skal støtte motivasjon og utholdenhet.

Med narrativ ramme mener arbeidsgruppen en rammefortelling eller bruk av en karakter eller et mål som driver handlingen (altså prøvegjennomføringen) videre. Den narrative rammen skal ikke påvirke det som måles i prøven.

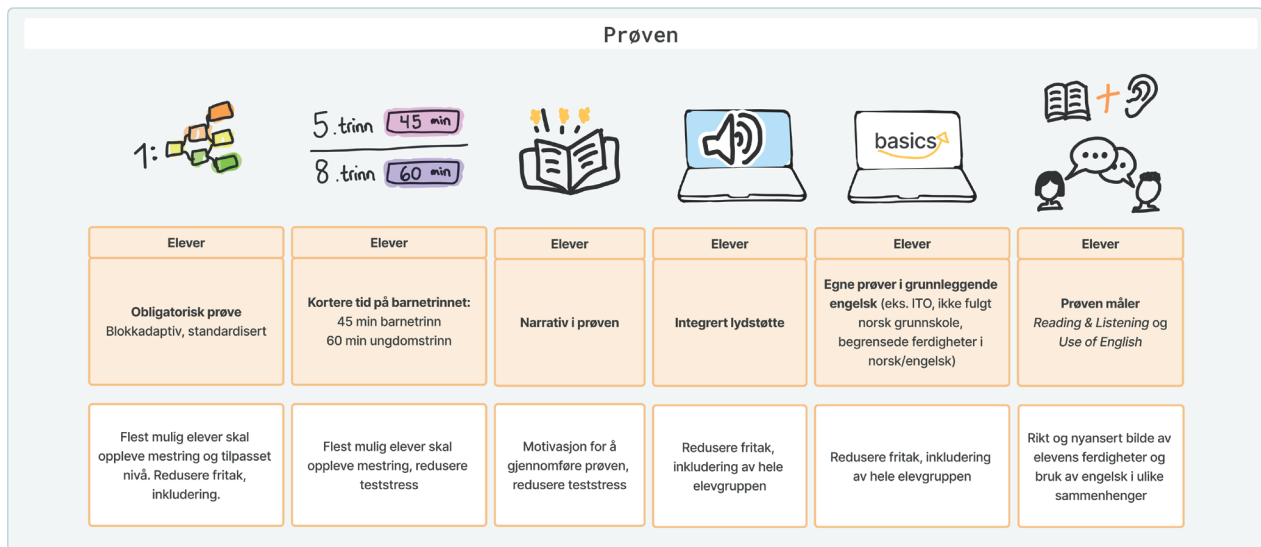
Gruppen ønsker at resultatene ikke bare skal gi informasjon om elevens nivå, men også synliggjøre hva elevene mestrer og hva de trenger å arbeide videre med.

Gruppen foreslår en prøve der *Reading & Listening* videreføres fra dagens nasjonale prøver, og måler forståelse av hovedinnhold, detaljer og implisitt informasjon i tekster. I tillegg introduseres *Use of English* som en ny del, for å gi et mer nyansert bilde av elevenes bruk av engelsk i ulike sammenhenger, med vekt på blant annet sosiolingvistisk kompetanse og interkulturell kommunikasjon.

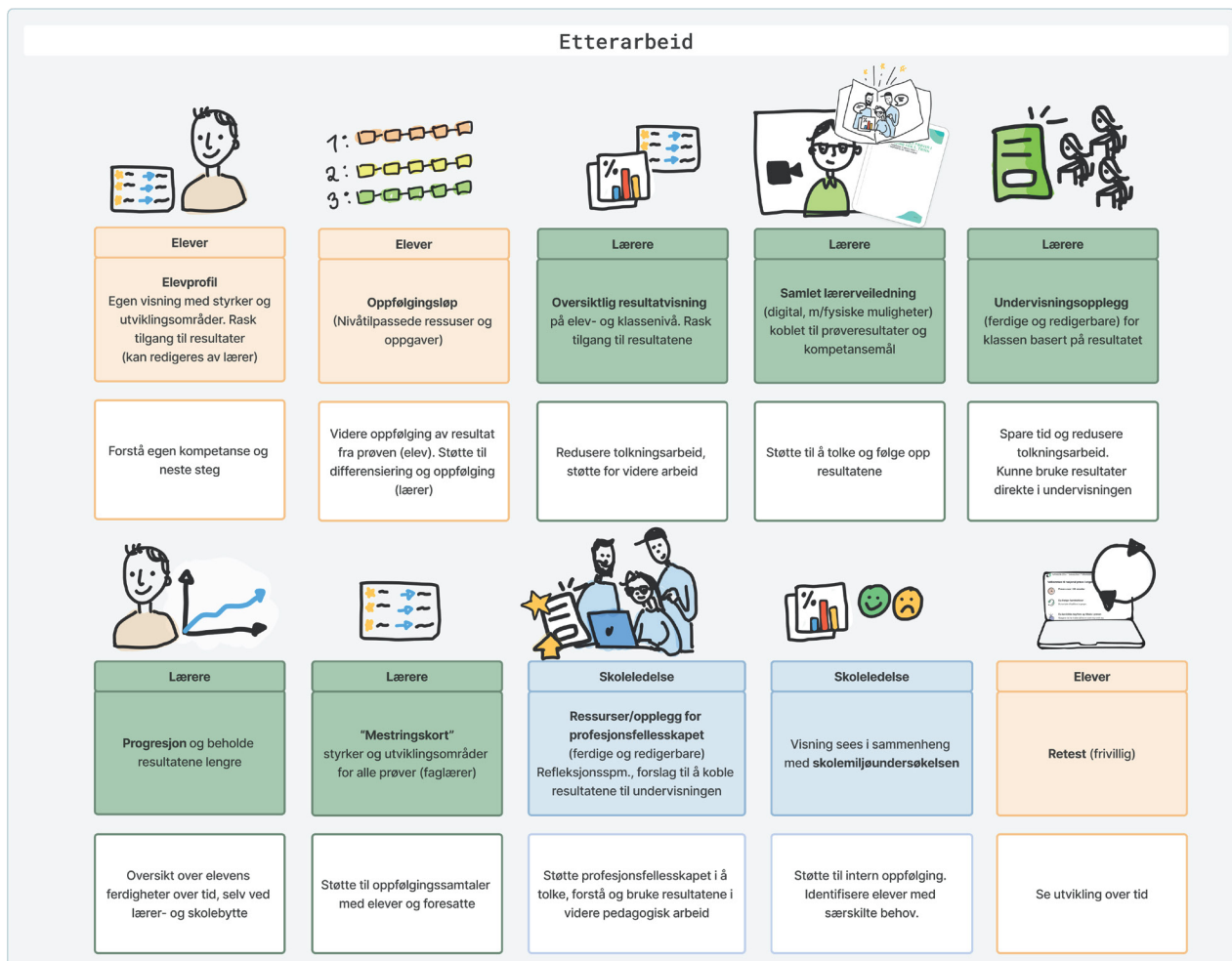
Konseptet skal gi et startpunkt for videre læring, ikke et slutt punkt. Læringsstøtten ligger i helheten av tiltak før, under og etter prøven. Sammenlignet med dagens prøver utvikles særlig arbeidet før og etter prøven. Blant annet skal resultatvisning og støtteressurser kobles bedre sammen. Gruppen ønsker rolletilpassede dashboards som gjør det enklere å omsette resultater til pedagogisk handling. For eleven skal resultatvisningen bidra til at eleven blir en selvstendig aktør i egen læringsprosess ved å synliggjøre noen få styrker og prioriterte utviklingsområder. Det skal også være mulig å se forslag til oppfølgingsløp og utvikling over tid. Oppfølgingsløp defineres av engelskgruppen som strukturerte, nivåtilpassede ressurser elevene kan bruke i videre læring. Løpet tar utgangspunkt i deres prøveresultater, og viser en tydelig vei videre i læringsarbeidet. For læreren skal løsningen være praktisk og tidsbesparende gjennom rask oversikt, og gi mulighet til å følge elevenes aktivitet i oppfølgingsløpene. Det skal være tett kobling mellom klasseresultater og konkrete undervisningsopplegg. Konseptet inkluderer også en frivillig, men anbefalt, retest for å synliggjøre elevenes utvikling og styrke læringsstøttende oppfølging. Under presenterer vi en visuell fremstilling av prøvekonseptet og utdyper forslagene fra gruppen.



Figur 5: visuell oversikt over forarbeid i engelskkonsept



Figur 6: visuell oversikt over prøven i engelskkonseptet



Figur 7: visuell oversikt over etterarbeid i engelskkonseptet

Obligatorisk fellesprøve og frivillig retest

Gruppen ønsker at den obligatoriske blokkadaptive prøven skal gi et felles og konstrukt utgangspunkt for videre arbeid. Prøven skal danne grunnlag for umiddelbar tilbakemelding og tilpasset oppfølging gjennom oppfølgingsløp for elevene.

Retest er frivillig for skole og lærer, men anbefalt. Oppgavene er ikke de samme som i den obligatoriske prøven, men har nytt innhold som måler de samme ferdighetsområdene. Retesten er ment å møte behovet for å synliggjøre elevenes utvikling over tid og gi et bedre grunnlag for oppfølging enn en enkeltmåling alene. Den sees i sammenheng med oppfølgingsløpene, slik at prøve, oppfølgingsløp og retest samlet støtter elevenes forståelse av sammenhengen mellom innsats, læring og utvikling, i tråd med prinsippene for undervisvurdering.

Tilrettelegging og egne prøver i grunnleggende engelsk

Gruppen er opptatt av inkludering, tilrettelegging og av å redusere fritak. De anbefaler derfor at løsningen, i tillegg til lovpålagt universell utforming, inkluderer en integrert opplesingsfunksjon uten mekanisk lydbilde. Dette vil særlig støtte elever som trenger ekstra funksjoner for å få vist kompetansen sin likeverdig, for eksempel ved synsvansker eller dysleksi. Gruppens erfaringer viser at eksterne verktøy (som Lingdys/Lingright/talesyntese) ofte er sårbare for lokale forskjeller i teknikk, oppsett og kompetanse, og at noen elever velger bort lesestøtte fordi den oppleves som treg eller lite hensiktsmessig. En integrert funksjon kan dermed gi mer likeverdig tilgang, redusere praktiske hindringer og være tilpasset ulike elevgrupper slik KD ber om i oppdragsbrevet.

Gruppen påpeker at adaptivitet og støttefunksjoner ikke vil være tilstrekkelig for hele elevgruppen, og ser behov for en kartlegging også for elever på et grunnleggende nivå i engelsk, uavhengig av alder og trinn. Det kan for eksempel gjelde elever med svært begrensede ferdigheter i norsk og/eller engelsk, elever som ikke har fulgt norsk grunnskole gjennom hele opplæringsløpet, elever med individuelt tilrettelagt opplæringsplan (ITO) i engelsk, eller andre elever som av ulike grunner ikke får vist kompetansen sin i den ordinære prøven. De peker på at det finnes kartleggingsprøver i lesing og regning på småtrinnet, men at det mangler et tilsvarende verktøy i engelsk. Gruppen anbefaler derfor en enkel kartlegging i grunnleggende engelsk som alternativ når elever fritas fra ordinær prøve.

Støtteressurser

Engelskgruppens konsept inkluderer et bredt tilfang av støtteressurser. Støtteressurser og andre pedagogiske ressurser skal være direkte koblet til prøven og prøveresultatet. Dette inkluderer altså alt som brukes både før og etter gjennomføring for å støtte videre læring. Gruppen mener at det bør utvikles støtteressurser som gjør prøvene forståelige og pedagogisk relevante for elever, lærere, foresatte og skoleledere. Gruppen ønsker at det skal bli lettere å bruke resultater direkte i undervisningen og redusere tiden lærere bruker til å fortolke og bearbeide resultatene. Ressursene skal altså være tett koblet til skolens ordinære arbeid, og de skal utformes slik at de oppleves som arbeidsbesparende.

Før prøven ønsker gruppen eksempeloppgaver og en lærerveiledning med forslag til undervisningsopplegg (redigerbart) som lærer kan bruke direkte i klasserommet for å

forberede elevene på prøven. For foresatte ønsker gruppen et klart informasjonsskriv om prøven, tilgjengelig på flere språk, for å støtte skole-hjem-samarbeid. For rektor ønsker gruppen tilsvarende redigerbare opplegg til bruk i skolens utviklingsarbeid. Gruppen løfter frem opplæring som en sentral forutsetning for at prøvene skal forstås, tas i bruk og følges opp som del av skolens pedagogiske arbeid, blant annet gjennom kurs/webinarer og lett tilgjengelig støttemateriell. Skoleeier har ansvar for å legge til rette for at dette faktisk gjennomføres.

Gruppen ønsker en samlet digital lærerveiledning koblet til prøveresultater og kompetansemål. Lærerveiledningen skal inneholde forslag til redigerbare minileksjoner, veiledninger per ferdighetsområde, nivådelte tiltak og eksempeloppgaver knyttet til *Reading & Listening* og *Use of English*. Ressursene bør kunne brukes både digitalt og i fysisk form alt ettersom hva som er mest hensiktsmessig. Målet er at ressursene skal gi læreren støtte i å vurdere hva klassen som helhet trenger å arbeide videre med, samtidig som de kan gi læreren et bedre grunnlag for å tilpasse oppfølgingen av enkeltelever – uavhengig av hvor de befinner seg faglig. For skoleledelsen foreslås redigerbare opplegg til profesjonsfellesskapet også i etterarbeidet, og for skoleeier pekes det på tilrettelegging for møteplasser der skoler kan dele erfaringer og “best practice”.

Resultatvisning

En sentral forutsetning i konseptet er at lærere får tilgang til pedagogisk relevante og ferdig tolkede resultater kort tid etter gjennomføring. Gruppen mener lang ventetid svekker både læreres og elevers eierskap til resultatene, og at dette reduserer sannsynligheten for at resultatene brukes aktivt i læringsarbeidet. Engelskgruppen foreslår derfor en sterkere vektlegging av kriteriebasert tolkning, slik at mer av den læringsrelevante tilbakemeldingen kan genereres automatisk uten å være avhengig av senere normering. Dette kan redusere ventetiden betydelig og styrke sammenhengen mellom gjennomføring, oppfølging og videre undervisning.

Engelskgruppen foreslår en resultatvisning som er koblet til støtteressursene. Her vil ulike brukerroller ha egne rolletilpassede dashboard. Dette inkluderer elevdashboard, lærerdashboard, en visning for skoleledelsen og en rapport for hver enkelt elev til kontaktlærer.

Elevdashboard skal gjøre eleven til en selvstendig aktør i sin egen læringsprosess, med en forenklet fremstilling av noen få styrker (3–4) og utvalgte utviklingsområder. Gruppen skisserer også at læreren ideelt sett kan redigere hvilke utviklingsområder eleven ser, for å tilpasse fokus og omfang. Gruppen er opptatt av at lærere kan få støtte til å arbeide med elevenes læring over tid. Dette mener gruppen vil øke motivasjon, eierskap til læringen og legge til rette for kontinuerlig utvikling.

Lærervisningen skal gjøre det enkelt og tidsbesparende å følge opp enkeltelever og grupper, og læreren skal kunne se elevenes aktivitet og faglige progresjon i oppfølgingsløpene. På klassenivå skisseres det en visuell profil for oversikt over ferdighetsområder, med en tett kobling til tiltak i klasserommet i form av ferdige, redigerbare undervisningsopplegg knyttet til delferdighetene som kartlegges. Rapport til kontaktlærer kaller gruppen for “mestringskort”. Mestringskortet skal bestå av styrker og utviklingsområder fra de ulike

prøvene, og støtte mer konkrete og læringsstøttende tilbakemeldinger enn dagens løsning. Det skal være enkelt å lese og skrive ut til bruk i utviklingssamtaler med foresatte og elever.

Gruppen peker også på behov for en resultatvisning som kan støtte skoleledelsens arbeid, blant annet ved å se prøveresultater i sammenheng med andre datakilder, som skolemiljøundersøkelsen. Dette vil gi et mer helhetlig bilde og identifisere elever med særlige behov. Videre ønsker gruppen at resultater følger eleven og lagres over tid. Slik vil utvikling kunne ses i sammenheng også ved lærer- eller skolebytte. En slik endring vil kreve personvernmessige vurderinger.

4.3 Matematikkgruppens forslag og anbefalinger

Arbeidsgruppen for matematikkprøven (heretter kalt matematikkgruppen eller gruppen) har bestått av fire lærere fra ulike deler av landet. De har lang erfaring som matematikklærere: én fra barnetrinnet, to fra mellomtrinnet og én fra ungdomstrinnet. Alle har utdanning i matematikk og matematikkdiraktikk.

Matematikkgruppens arbeid skiller seg fra de andre gruppene ved at de ikke kun har utredet et nytt konsept for en læringsstøttende prøve, men også har måttet vurdere konsekvenser av å endre fra nasjonale prøver i regning til nye læringsstøttende prøver i matematikk. I kapittel 1.4 drøftet vi mulige faglige implikasjoner av denne endringen som del av oppdragsforståelsen.

I gruppens konseptbeskrivelse legges det til grunn at de nye prøvene skal gi innsikt i både elevenes regneferdigheter og deres matematikkompetanse. Regneferdigheter, som tellekompetanse, tallforståelse og det å kunne regne med tall, beskrives som en nødvendig del av grunnlaget for videre utvikling av matematikkompetanse. Gruppen anbefaler derfor at prøvene formidles og utvikles med utgangspunkt i disse to begrepene: regneferdigheter og matematikkompetanse.

Teoretisk støtter gruppen seg særlig på forskning om vurdering for læring og læringsfremmende prøver, blant annet hos Hodgen og Wiliam (2014), Popham (2009) og de Lange (1995). I tillegg viser gruppen til forskning og erfaringer som peker på at lærere trenger faglig kvalitetssikrede prøver, forståelse for prøvens formål, bedre støtte til tolkning og oppfølging av resultater, og et prøvesystem som oppleves som nyttig og håndterlig i en travel skolehverdag. Dette danner bakteppet for gruppens anbefaling om å utvikle et nytt prøvesystem i matematikk som i større grad fungerer som et praksisnært og fagdidaktisk verktøy for lærere og skoler.

Matematikkgruppen ønsker at de nye prøvene skal støtte elevenes læring og gi lærerne et mer nyansert og handlingsrelevant bilde av elevenes kompetanse i matematikk. Det etterlyses tydeligere sammenheng mellom prøver, læreplan, undervisning og oppfølging enn i dag.

Gruppen ønsker et prøvesystem som tilbyr et praksisnært og læringsstøttende matematikdidaktisk verktøy for lærere og skoler. Prøvene skal ikke bare gi informasjon om elevens nivå, men støtte læreren i å forstå elevenes mestring, utfordringer, misoppfatninger og utviklingsbehov. På denne måten skal prøvene i større grad bidra til læringsfremmende underveisvurdering og bedre tilpasset opplæring.

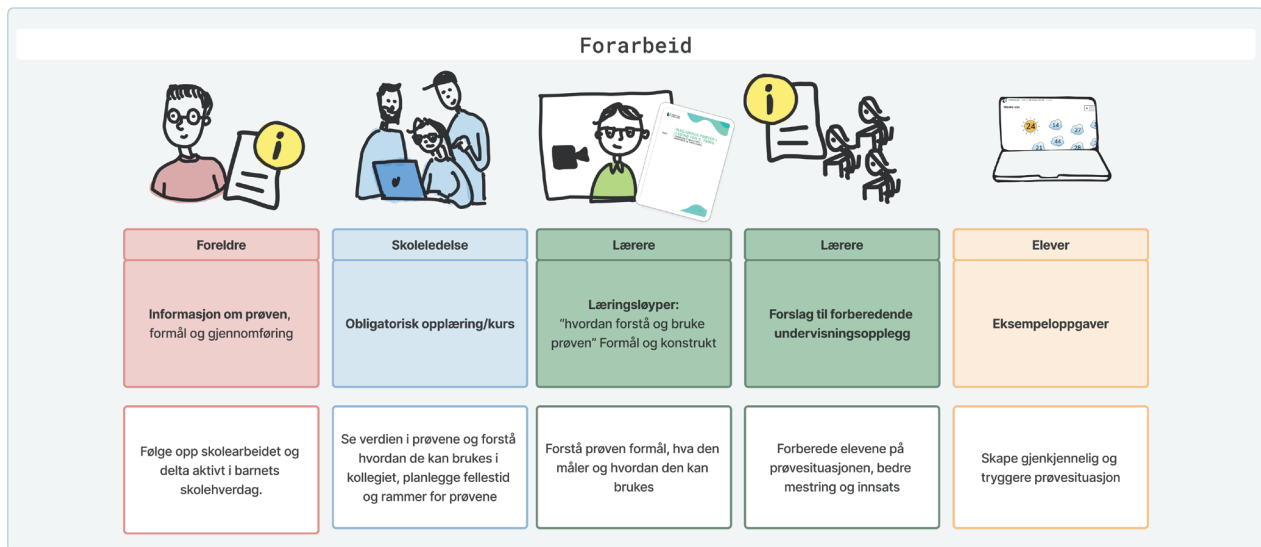
Gruppen legger til grunn at verdien av prøvene ligger i nytteverdien for lærere og elever. Dersom prøveresultatene bidrar til bedre planlegging av undervisning og mer målrettet støtte til elevene, vurderer gruppen at prøvene har høy verdi som læringsstøttende verktøy. Konseptet er derfor utviklet som et forsøk på å balansere de viktigste behovene i skolen, forskning på feltet og det gruppen vurderer som avgjørende suksessfaktorer for å lykkes med læringsstøttende prøver i matematikk.

Konseptbeskrivelse

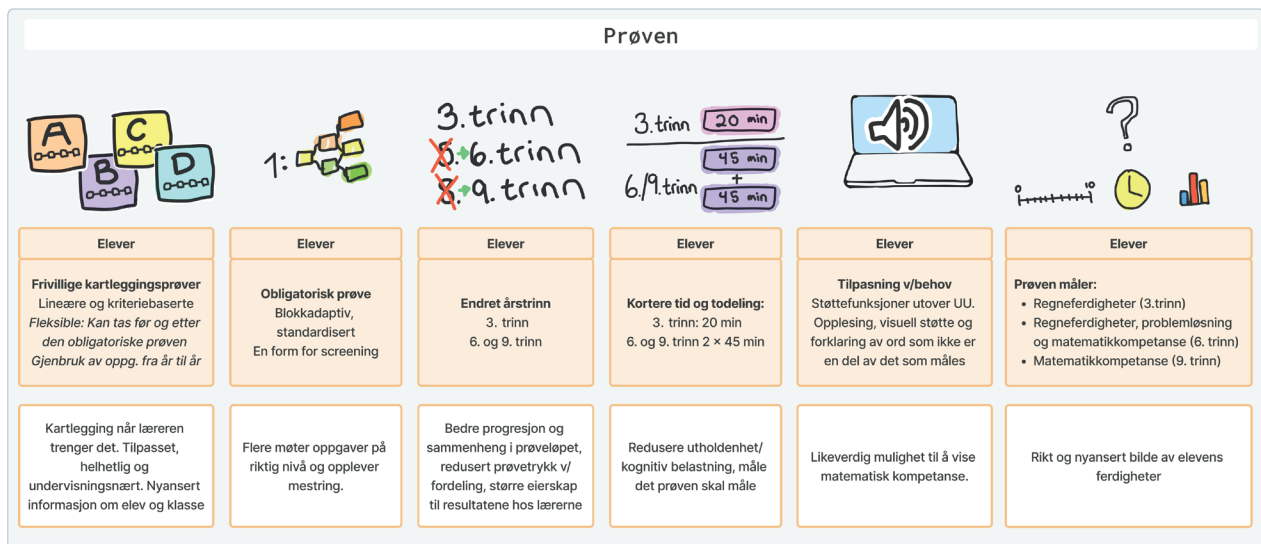
I konseptforslaget anbefales det at nye læringsstøttende prøver i matematikk organiseres som et prøvesystem bestående av flere frivillige kartleggingsprøver og én obligatorisk fellesprøve. De frivillige prøvene skal være lineære og kunne brukes fleksibelt av læreren gjennom året, mens fellesprøven skal være standardisert og blokkadaptiv. Hensikten med denne organiseringen er å kombinere behovet for et felles sammenligningsgrunnlag med behovet for mer detaljert og handlingsrelevant informasjon om elevenes matematiske kompetanse.

Konseptet innebærer også at prøvesystemet knyttes tett til læreplanen i matematikk og til sentrale deler av matematikkompetansen. Gruppen anbefaler at prøvene måler både regneferdigheter og matematikkompetanse. De ønsker at fagmiljøene som utvikler prøvene, har særlig oppmerksomhet mot fem områder: relevante og aldersadekvate kjerneelementer for det aktuelle trinnet, tallforståelse, regneferdigheter, matematisk språk og engasjement for matematikk. På denne måten skal prøvene gi et rikere bilde av elevenes kompetanse enn dagens prøver.

Videre foreslår gruppen at prøvene organiseres på 3., 6. og 9. trinn, og at dagens kartleggingsprøve på 3. trinn inngår som en del av det nye systemet. Dette begrunnes med behovet for bedre progresjon og sammenheng i prøveløpet, redusert prøvetrykk ved å fordele prøver utover ulike år og større eierskap til resultatene hos lærerne. Gruppen mener at prøver på disse trinnene i større grad vil gi lærere relevant informasjon på et tidspunkt der de kjenner elevene godt nok til å kunne følge opp resultatene pedagogisk. Læreplanen i matematikk har kompetansemål på alle trinn unntatt 1. trinn, derfor vil det være fullt mulig å velge andre gjennomføringstrinn for denne prøven.



Figur 8: visuell oversikt over forarbeid i matematikkonseptet



Figur 9: visuell oversikt over prøven i matematikkonseptet



Figur 10: visuell oversikt over etterarbeid i matematikkonseptet

Innhold og struktur i konseptet

Gruppen anbefaler altså et prøvesystem som består av frivillige kartleggingsprøver og én obligatorisk fellesprøve på høsten på 3., 6. og 9. trinn.

De frivillige kartleggingsprøvene skal kunne gjennomføres fleksibelt, både når det gjelder tidspunkt og antall ganger, ut fra lærers behov. Prøvene bør rettes automatisk og i stor grad kunne gjenbrukes, slik at lærere over tid blir godt kjent med oppgavene og kartleggingene. Hensikten er å gi lærere god innsikt i elevenes matematikkompetanse, med et tydelig fokus på positiv testing (deLange, 1995), der elevenes mestring, styrker og eventuelle misoppfatninger synliggjøres. Slik skal prøveresultatene gi et godt grunnlag for tilpasset opplæring og undervisningsvurdering.

Den obligatoriske fellesprøven skal være en standardisert, blokkadaptiv prøve som gir grunnlag for videre oppfølging av elevenes læring. Gruppen foreslår at også fellesprøven i hovedsak bør gjenbrukes fra år til år, slik det allerede gjøres i dagens kartleggingsprøver, eventuelt med enkelte justeringer eller utskifting av oppgaver. Prøven bør være universelt utformet og tilpasses ulike elevgrupper, og det bør åpnes for at den kan tas på nytt i form av en retest året etter. Oppgavene skal tilpasses elevens mestringsnivå og slik bidra til at

flere elever opplever mestring samtidig som læreren får mer presis informasjon om kompetansen i elevgruppen. Gruppen vurderer at en blokkadaptiv løsning vil være mindre kompleks å utvikle enn en fulladaptiv prøve, samtidig som den er lettere for lærere å få oversikt over og å bruke i videre læringsarbeid enn hva en fulladaptiv prøve vil være.

På 3. trinn anbefaler matematikkgruppen prøver som særlig måler elevenes regneferdigheter, herunder telling, tallforståelse og det å regne med tall. De frivillige kartleggingsprøvene foreslås organisert tematisk, mens den obligatoriske fellesprøven skal gi læreren innsikt i klassens regneferdigheter sammenlignet med et nasjonalt gjennomsnitt. Den obligatoriske fellesprøven foreslås med en tidsramme på 20 minutter. Fellesprøven skal gi klasseprofil for hele gruppen, mens mestringsprofil på individnivå i hovedsak anbefales for elever under oppfølgingsgrensen.

På 6. trinn anbefales prøver som både måler regneferdigheter og matematikkompetanse. Her foreslås frivillige prøver knyttet til blant annet kompetansemål i lys av kjerneelementene anvendelse, resonnering og representasjoner, matematisk språk og begreper, tall og regning og matematiske kunnskapsområder. Den obligatoriske fellesprøven foreslås delt i to deler på inntil 45 minutter hver, for å redusere betydningen av utholdenhet og gi et bedre grunnlag for å fange bredden i elevenes kompetanse. Arbeidsgruppen anbefaler at fellesprøven så langt det er mulig gir informasjon både på individnivå og klassenivå, men understreker samtidig at fagmiljøene må vurdere hvor presis slik informasjon kan bli.

På 9. trinn anbefaler gruppen prøver som i større grad enn på 6. trinn måler elevenes matematikkompetanse, særlig i lys av kjerneelementene representasjoner, resonnering og generalisering. Også her foreslås en kombinasjon av frivillige kartleggingsprøver og en todelt blokkadaptiv fellesprøve. Den obligatoriske prøven foreslås med samme tidsinndeling som på 6. trinn. Arbeidsgruppen legger til grunn at fellesprøven først og fremst skal gi læreren indikasjoner på klassens kompetanse, og at fagmiljøene må vurdere hvor langt det er mulig å utvikle presise mestringsprofiler på individnivå.

Støtteressurser

Forslaget fra matematikkgruppen inneholder et bredt tilfang av støtteressurser. Gruppen legger til grunn at prøvesystemet bare vil fungere læringsstøttende dersom lærere og skoleledere får god og variert støtte både før og etter gjennomføringen av prøvene. Støtteressursene skal derfor ikke være et tillegg til prøvene, men en integrert del av konseptet.

Det foreslås blant annet læringsløyper, kompetansepakker, oppgavespesifikke ressurser, undervisningsforslag, lettleste fagtekster og eksempler fra praksis. Ressursene skal støtte lærere i å forstå prøvenes formål, analysere og tolke resultater og bruke disse i videre undervisning. I tillegg foreslås det støtte rettet mot ulike målgrupper, herunder elever, foresatte, lærere, skoleledelse, profesjonsfellesskap og skoleeier.

Gruppen anbefaler også at støtteressursene gjøres lett tilgjengelige gjennom digitale løsninger, og at de kobles tett til resultatvisningen, slik at læreren raskt kan finne fram til relevante tiltak og oppfølgingsressurser. Det understrekes at støtteressursene bør være

tilgjengelige for alle lærere og faglig ansatte gjennom innlogging på Udirs nettsider, og ikke være avhengige av innlogging i selve prøveportalen. På denne måten skal avstanden mellom prøveresultat og pedagogisk handling bli kortere enn i dag.

Resultatvisning

Matematikkgruppen foreslår en resultatvisning som støtter læreren på veien fra data til handling. I konseptet beskrives dette som et dynamisk lærerpanel (dashboard) som raskt skal gi læreren oversikt over formålet med prøvene og tilgang til ulike resultatvisninger. Resultatene skal presenteres gjennom klasseprofiler og elevprofiler som viser hva elevene mestrer, hvilke områder de strever med, og hvilke misoppfatninger eller utviklingsbehov som kan identifiseres på bakgrunn av prøven. Resultatvisningen skal være enkel å navigere i og gi læreren rask tilgang til relevant informasjon på både klasse- og elevnivå. Gruppen ser en god resultatvisning som et sentralt suksesskriterium for å lykkes med de læringsstøttende prøvene.

Et sentralt poeng i forslaget er at resultatvisningen må kobles tett til støtteressurser. Læreren skal kunne se hva oppgavene måler, få beskrivelser av hvilke ferdigheter eller kompetanser de er knyttet til, og ledes direkte videre til relevante ressurser for oppfølging. Gruppen mener at dagens resultatvisninger i for liten grad støtter det pedagogiske arbeidet, og anbefaler derfor at nye løsninger utvikles i tett samarbeid med brukerne og prøves ut trinnvis.

Det foreslås også at resultatene skal kunne vises i ulike formater avhengig av formål og målgruppe, blant annet som faglige visninger for lærer og mer tilgjengelige versjoner for elev og foresatte. Slik skal resultatene kunne brukes både i lærerens analysearbeid, i elevsamtaler og i kommunikasjon med hjemmet. Samtidig er det i den oppdaterte konseptbeskrivelsen tydeligere at resultatvisningen må være enkel å navigere i, ha kort vei til støtteressurser og presentere informasjon på en måte som gjør det mulig for læreren å gå raskt fra resultat til pedagogisk handling.

4.4 Lesegruppens forslag og anbefalinger

Arbeidsgruppen for leseprøven (heretter kalt lesegruppen eller gruppen) har bestått av fire lærere fra grunnskolen og en samisk representant. Tre arbeider på ungdomstrinnet, én på barnetrinnet og én arbeider ved Nasjonalt senter ved samisk i opplæringa. Sammen representerer gruppen ulike deler av landet og de har variert erfaring og ansiennitet.

Faglig utgangspunkt for forslaget

Arbeidsgruppen har forankret sitt konseptforslag i internasjonal forskning på lese teori, motivasjonsteori og lesevurdering, samt nasjonale føringer for vurdering og grunnleggende ferdigheter slik de er beskrevet i læreplanverket (LK20/LK20S). I sitt arbeid har de også støttet seg til intervjuene med de strategiske samarbeidsskolene, i tillegg til egen profesjonserfaring og annet kunnskapsgrunnlag fra praksisfeltet.

I konseptbeskrivelsen legges det til grunn at lesing er en sammensatt ferdighet som handler om å forstå, bruke, reflektere over, vurdere og engasjere seg i tekster. Dette

innebærer at nye prøver i lesing ikke bare bør gi et øyeblikksbilde av elevenes nivå, men også gi innsikt i hva elevene mestrer, har utfordringer med og hva de trenger å arbeide videre med. Gruppen har valgt å holde seg tett opp til den etablerte forståelsen av lesing i læreplanverket og i dagens leseprøver, fordi denne er faglig godt forankret og gir kontinuitet fra dagens nasjonale prøver.

Gruppen anbefaler derfor at prøvene utvikles med utgangspunkt definisjonen PIRLS- og PISA-undersøkelsene bruker på lesing: "Reading literacy is understanding, using, evaluating, reflecting and engaging with texts in order to achieve one's goals, to develop one's potential and participate in society". Denne definisjonen viser at lesing er mer enn bare «avkoding x forståelse», slik den er beskrevet i «the simple view of reading». Lesing er også å forstå og bruke innholdet i en tekst, vurdere og reflektere rundt det man leser og at man engasjerer seg i tekstene. Tilsvarende defineres grunnleggende ferdighet som å kunne forstå, bruke, reflektere over og engasjere seg i innholdet i tekster. Lesing forstås ikke bare som det som er skrevet med ord, men også illustrasjoner, symboler eller andre uttrykksmåter (LK20/LK20S).

Teoretisk støtter gruppen seg særlig på forskning om vurdering for læring, god leseopplæring, lesevurdering og motivasjon, blant annet hos Black & William (1998), McLaughlin (2012) og Afflerback (2016).

For samiske leseprøver vil den samiske språksituasjonen påvirke hva som er hensiktsmessige løsninger for denne prøven. Konseptforslag for de samiske leseprøvene presenteres i et eget delkapittel.

Lesegruppen mener, som de andre gruppene, at dagens prøver gir for lite informasjon om hva elevene mestrer og hva de kan gjøre for å utvikle ferdigheten videre. Nasjonale og internasjonale målinger av norske barns leseferdigheter viser en bekymringsfull utvikling. Derfor er det viktig at de nye prøvene kan bidra positivt når det gjelder elevenes leseferdigheter.

Hensikten med konseptforslaget er å styrke læreren i klasserommet og å gjøre leseprøven til en mer integrert del av opplæringen. En tydeligere kobling mellom forarbeid, gjennomføring og etterarbeid kan bidra til at prøven går fra å være en enkeltstående hendelse til å bli en del av elevenes leseutvikling, der elevene i større grad blir påkoblet egen læring. På denne måten er det læringsstøttende. Prøven skal ikke bare gi informasjon om nivå, men også støtte læreren i å følge opp elevenes læring videre.

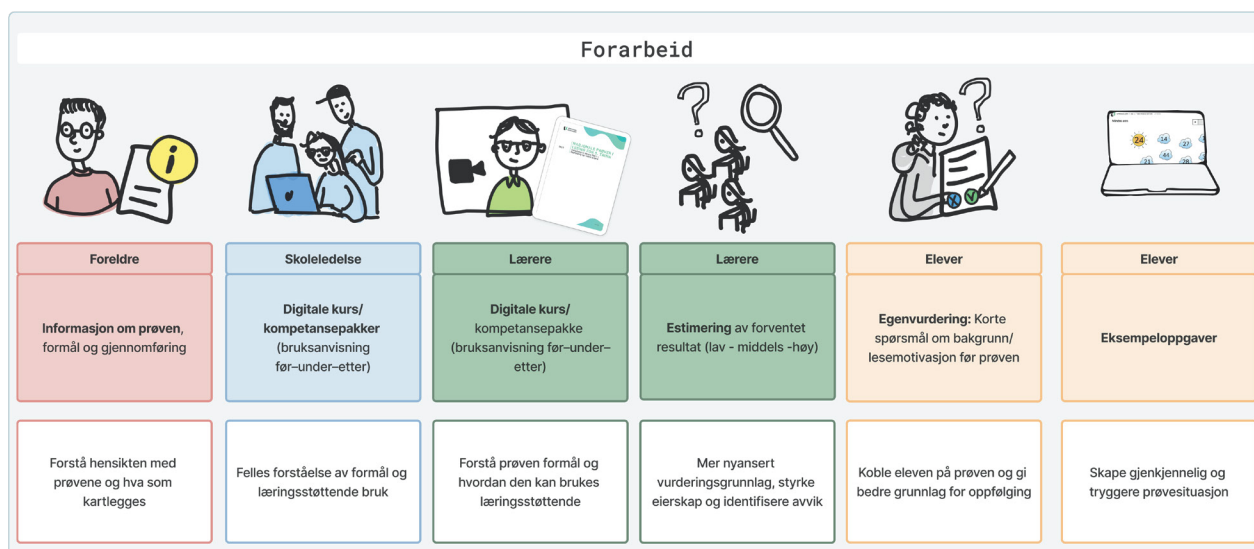
Konseptbeskrivelse

I dette konseptforslaget anbefales det at de nye læringsstøttende prøvene i lesing organiseres som et prøvesystem der selve prøven er todelt, med en lineær del og en blokkadaptiv del. Hver prøvedel er ment å ha en øvre tidsramme på 45 minutter, og delene skal fortrinnsvis gjennomføres på to ulike dager. Hensikten med dette er å redusere betydningen av elevenes testutholdenhet, slik at prøven i større grad måler elevenes reelle lesekompetanse.

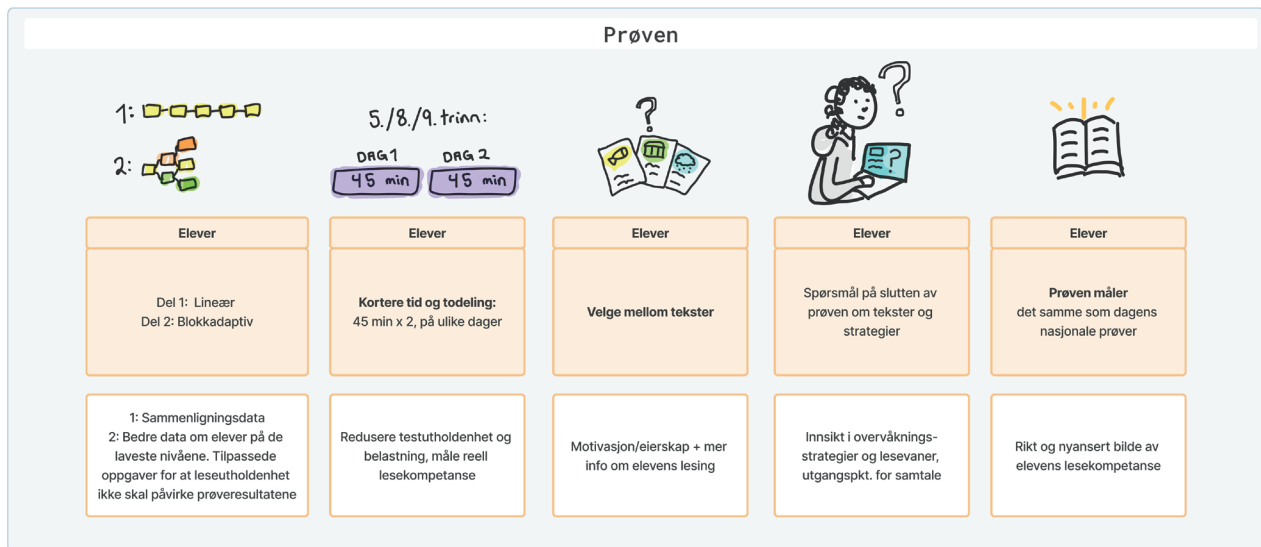
Den lineære delen av prøven skal gi et felles sammenligningsgrunnlag, mens den blokkadaptive delen tilpasser oppgaver mer til elevens nivå og gir mer nyansert informasjon av hva eleven mestrer. Målet er at prøven skal støtte lærerens arbeid med elevenes læring. Dette innebærer at resultatene ikke bare skal gi informasjon om elevens nivå, men også bidra til å synliggjøre hva elevene mestrer og hva de trenger å arbeide videre med. Slik kan prøven bli et bedre verktøy for videre læring, i tråd med intensjonen for nye prøver med læringsstøttende formål.

Konseptet innebærer også at prøven inngår i en helhetlig prosess med forarbeid, gjennomføring og etterarbeid. Overordnet sett handler lesekonseptet om å bruke leseprøven som et verktøy for å utvikle elevenes leseferdigheter. Det læringsstøttende arbeidet skjer både før, under og etter prøven, og vil ha særlig verdi i oppfølgingen av resultatene i etterkant.

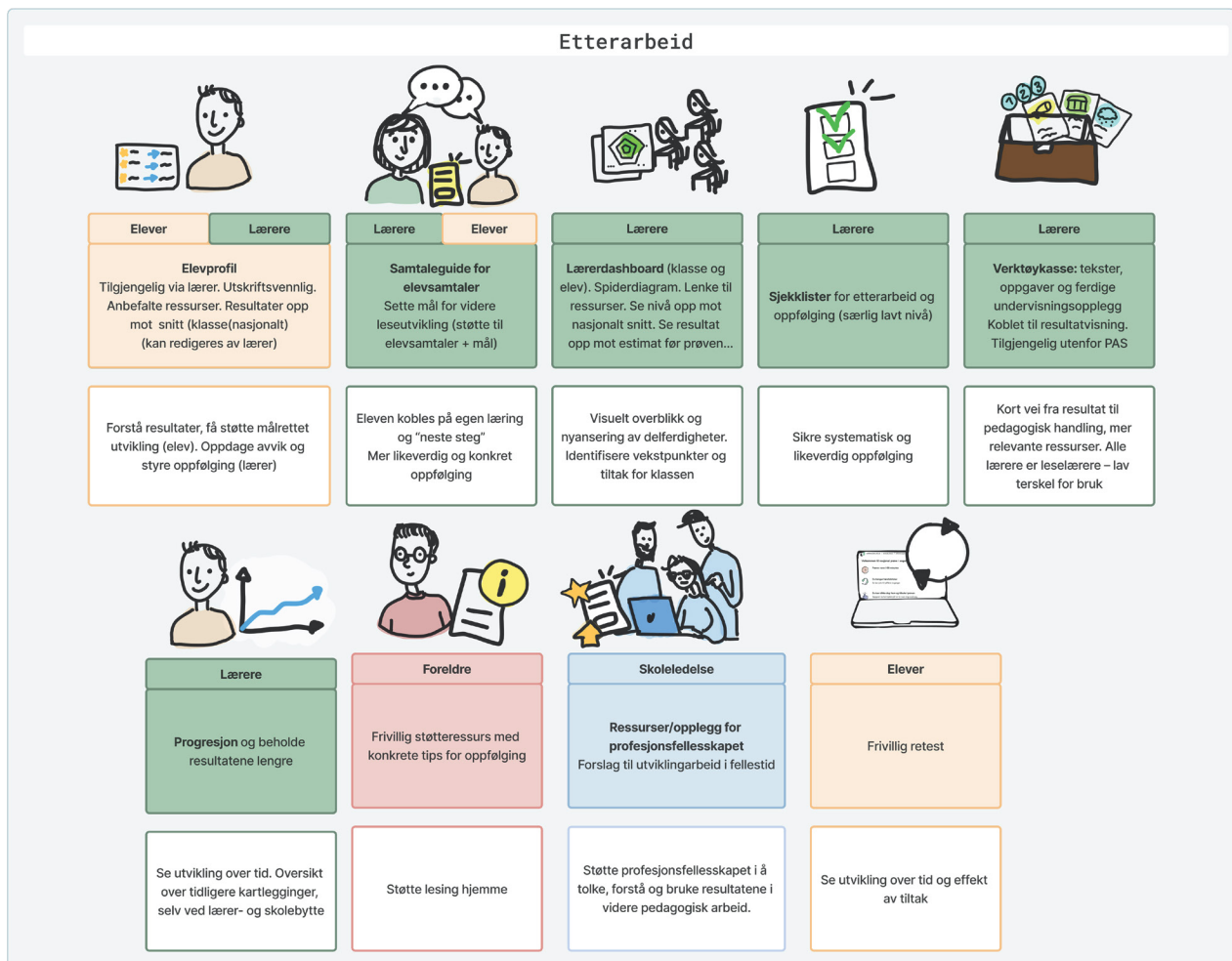
Lesekonseptet bygger videre på styrkene til dagens nasjonale prøver, men utvikler særlig arbeidet både i forkant og etterkant av selve prøvegjennomføringen. Gjennom kombinasjon av prøveresultater, lærervurdering og elevens egen målsetting skal læreren få et bedre grunnlag for å følge opp elevene. I tillegg skal elevene i større grad bidra i sin egen læringsprosess, blant annet gjennom å sette seg egne mål og følge opp disse i etterkant av prøven. Det er planlagt et tydeligere samspill mellom prøvenes resultatvisning og støtteressursene som hører til, og det er særlig i forarbeidet for lærere og elever og etterarbeidet med resultatvisningen og støtteressursene at konseptet blir læringsstøttende. Innholdet i selve prøven, støtteressurser og resultatvisning utdypes i egne avsnitt under.



Figur 11: visuell oversikt over forarbeid i lesekonseptet



Figur 12: visuell oversikt over prøven i lesekonseptet



Figur 13: visuell oversikt over etterarbeid i lesekonseptet

Kort oversikt over innhold og struktur i konseptet

Arbeidsgruppen anbefaler en prøve i lesing som bygger på en tydelig tredeling mellom forarbeid, prøve og etterarbeid. I forarbeidet inngår introduksjon av prøvene til skolene, inkludert informasjon om støtteressurser og resultatvisning til bruk i profesjonsfellesskapet. Det skal også være tilgang til eksempeloppgaver som forbereder elevene på oppgaveformater og prøvesituasjon. I tillegg foreslås det at læreren estimerer elevenes forventede resultat, og at elevene svarer på noen korte spørsmål om bakgrunnsvariabler og lesemotivasjon. Hensikten med lærerens forventede estimat er å sette prøveresultatet inn i en bredere vurderingssammenheng, slik at et eventuelt avvik mellom lærerens inntrykk og elevens resultat kan identifiseres og følges opp.

Den anbefalte prøven består av to deler. Del 1 er lineær og har som formål å gi et felles sammenligningsgrunnlag, der resultater på egen skole kan ses opp mot et nasjonalt snitt. Det er kun statlige prøver som kan gi denne typen informasjon, som er viktig for å overvåke elevenes leseutvikling. Resultatene kan også brukes til forskning og som grunnlag for lokalt skoleutviklingsarbeid. Del 2 er blokkadaptiv, som vil si at oppgavene i større grad tilpasses levens nivå. Denne delen skal gi mer nyansert informasjon om hva den enkelte elev mestrer. Målet er særlig å få bedre innsikt i prestasjonene til elever i begge ender av ferdighetsskalaen, slik at informasjonen i større grad kan støtte lærerens arbeid med elevenes leseutvikling.

Det er også et ønske om at prøven åpner for elevvalg av tekst, for eksempel ved at elevene kan velge mellom to ulike tekster innenfor samme teksttype eller sjanger. Dette skal bidra til økt motivasjon og samtidig gi mer informasjon om elevenes leseferdigheter. Etterarbeidet skal omfatte gjennomgang av resultatvisningen og vurdering av om det er samsvar mellom lærerens forhåndsvurdering og elevenes faktiske resultater. Det skal også gjennomføres elevsamtaler der elev og lærer setter mål for videre leseutvikling. Videre skal det være tilgjengelige støtteressurser for arbeid med god leseopplæring i alle fag etter prøven, støtte til arbeid i profesjonsfellesskapet og retest for å gi innsikt i effekten av iverksatte tiltak og vurdere elevenes progresjon.

Støtteressurser

Forslaget fra arbeidsgruppen omfatter et variert spekter av støtteressurser, som ses som en integrert del av prøvesystemet. Gruppen legger til grunn at prøvene kun blir læringsstøttende dersom lærere og skoleledere får god støtte før, under og etter gjennomføring. Det bør derfor følge med digitale kurs som fungerer som en bruksanvisning til prøvene, og som gir innsikt i hva prøven måler, hvordan resultater skal tolkes og hvordan ressursene kan brukes i oppfølgingen. Tilgang til tidligere prøver og eksempeloppgaver bør videreføres for å forberede elevene og gjøre dem trygge på prøveformen.

I oppfølgingen bør det gis anbefalinger, gjerne gjennom kurs, om effektiv bruk av resultater og ressurser. Det bør også utvikles en verktøykasse med tekster, oppgaver og konkrete undervisningsopplegg, tett koblet til resultatvisningen. Ressursene bør inkludere råd om videre kartlegging og oppfølging av de elevene som har lave leseferdigheter eller som ikke har tilstrekkelig utbytte av den ordinære opplæringen.

Støtten bør også omfatte veiledning til elevsamtaler, inkludert forslag til spørsmål, tilbakemeldinger og arbeid med mål for videre leseutvikling, dokumentert i en elevprofil knyttet til resultatene. Ressursene bør være tilgjengelige også uten direkte tilgang til resultatvisningen, slik at flere lærere kan bruke dem. Verktøykassen bør dessuten inneholde generelle ressurser for arbeid med tekst og begreper. Foresatte bør få tilgang til enkle, frivillige ressurser med konkrete råd for oppfølging av lesing hjemme.

Resultatvisning

Lesegruppen foreslår en resultatvisning der prøveresultatene skal fremstilles i to ulike varianter; et lærerdashbord og en elevprofil.

Lærerdashbordet skal gi de aktuelle lærerne relevant informasjon om elevenes leseferdigheter på klasse- og elevnivå. Alle lærere som har fag der lesing er relevant, skal ha tilgang til resultatene til sine elever. Som nevnt tidligere skal lærere som en del av forarbeidet til prøven estimere resultatet til klassen. For lærerne som har gjort dette, vil estimatet være synlig for dem ved siden av visningen av elevenes faktiske resultater. Dette skal bidra til å styrke lærerens eierskap til prøven og gjøre det lettere å identifisere avvik mellom egen vurdering og elevens resultat.

På klassenivå bør resultatene framstillet på to måter; klassens nivå sammenlignet med nasjonalt snitt og klassens delferdigheter innenfor lesing i et spider-diagram. Spider-diagrammet vil utdype og nyansere klassens resultater ved å vise klassens delferdigheter innenfor lesing. Dette vil tydeliggjøre vekstpunkter for klassen som helhet. Spider-diagrammet bør henge tett sammen med støtteressursene, for eksempel ved at man kan trykke på delferdigheter i diagrammet og deretter få opp relevante støtteressurser som kan brukes for å styrke den spesifikke delferdigheten på klassenivå.

Når læreren ser på lærerdashbordet på elevnivå, bør enkeltelevens kartleggingsresultater fremstilles sammen med anbefalte ressurser. Her bør elevens resultater vises opp mot nasjonalt snitt og klassens snitt, slik at læreren enkelt kan gjenkjenne avvikende data. Resultatene skal være så tydelige at læreren kan, ved å lese resultatene, se hvilke tekst- og oppgavetyper eleven har mestret og ikke mestret. Dette vil gi grunnlag til hva slags tekster og oppgaver læreren burde jobbe mer med sammen med eleven.

Elevprofilen skal synliggjøre resultatene for elev og foresatte. Formålet til elevprofilen er å informere elev og foresatte om relevante deler av kartleggingsresultatet. Det er lærer som distribuerer tilgang til elevprofilen. Resultatene skal følge eleven slik at lærerne til eleven kan se nåværende og tidligere kartleggingsresultater i sammenheng, og følge utvikling over tid. Ved at resultatene følger eleven så vil det også være enklere for lærere å se kartleggingshistorikk hos nye elever som av ulike årsaker har byttet skole eller fått nye lærere. I elevprofilen skal det være mulig for elevene å skrive inn mål i samråd med lærer i etterkant av prøven. Dette vil bevisstgjøre elevene i egen læring og leseutvikling.

Særskilte vurderinger for samiske leseprøver

De nye læringsstøttende prøvene i lesing på samisk vil lages parallelt på de tre samiske språkene som er definert som opplæringspråk i Opplæringsloven – nordsamisk, lulesamisk og sørsamisk. Elevene skal ta prøven på det språket de har som førstespråk.

De samiske leseprøvene bør ligge tett opp til leseprøvene på norsk, men med nødvendige tilpasninger og justeringer. Prøvene må utvikles med utgangspunkt i samisk språk, kultur og samfunnsliv, og i tråd med de til enhver tid gjeldende læreplanene for fagene i Kunnskapsløftet Samisk. I kapittel 7 vil Udir drøfte prøvekonseptet i lys av behovene i samisk skole.

Gruppen anbefaler at de nye læringsstøttende prøvene i lesing på samisk, i likhet med dagens prøver, utvikles parallelt på de tre samiske språkene: nordsamisk, lulesamisk og sørsamisk. Elevene skal gjennomføre prøven på det språket de har som førstespråk.

Prøvene er også en læringsarena for samiske barn. For de samiske prøvene må det derfor tas hensyn til at prøvene inngår i en språklig og kulturell sammenheng der elevenes erfaringer med samisk kan variere betydelig. Det finnes ikke læremidler på alle de tre samiske språkene i alle fag, og den samiske læringsstøttende leseprøven bør derfor ikke bare handle om lesing og leseforståelse, men også om elevenes erfaringer med samisk språk, ulike teksttyper og faglig terminologi på samisk. Elever som ikke har tilgang til læremidler på sitt førstespråk, vil ofte måtte lære fagterminologi på norsk. Dette gjelder også elever som har samisk som førstespråk, men som i hovedsak bare får opplæring på samisk i samiskfaget og ikke i andre fag. Dette vil kunne påvirke elevenes forutsetninger for å forstå tekster på samisk. I praksis innebærer dette at ikke alle elever som gjennomfører leseprøven på samisk, har hatt like muligheter til å utvikle lesekompetanse i tråd med læreplanen i LK20S. De kan ha ulik erfaring med å lese ulike sjangre på samisk og med å bruke samisk fagspråk. Dette kan føre til at tekster som i utgangspunktet er tilpasset aldersgruppen, likevel oppleves som mer komplekse for elever med begrenset erfaring med lesing på samisk. Ved valg av tekster må det derfor legges særlig vekt på språk, ordtilfang, kompleksitet og innhold. Målet bør være å bruke tekster som er tilpasset det enkelte samiske språket og elevgruppens varierte bakgrunn.

Gruppen anbefaler at det i forarbeidet til prøven eller i prøven inkluderes spørsmål om elevenes språklige bakgrunn, leseinteresse, språksituasjonen ved skolen og i hvilken grad samisk er integrert i elevenes hverdag. Dette kan gi viktig informasjon om forhold ved språksituasjonen som kan påvirke lesingen. Samiske elevers erfaringer med samisk i skolen varierer i stor grad. For noen er samisk først og fremst et skolefag, noe som innebærer at språket i hovedsak brukes i samisktimene. Andre elever har et samisk språkmiljø også utenfor undervisningen, som gir mulighet til å bruke språket aktivt. Det finnes også elever som går på samiske skoler der samisk er undervisningsspråk i alle fag, mens andre får opplæring gjennom fjernundervisning, der eksponeringen for språket i større grad er digital og mer isolert fra resten av skolehverdagen.

Når det gjelder selve prøven, vil det være lite hensiktsmessig å legge opp til gjennomføring over to eller flere dager. Organiseringen av samiskopplæringen varierer betydelig mellom skolene. Noen elever har all samiskundervisning samlet på én dag, mens andre har den

fordelt utover uken. Det må også tas hensyn til at mange samiske lærere er ambulerende, noe som gjør det mer praktisk å gjennomføre prøven på én dag ved de ulike skolene.

Det vil samtidig være krevende å utvikle en adaptiv prøve på samisk. For det første vil det være vanskelig å finne et tilstrekkelig antall egnede tekster på alle de tre samiske språkene. Det samme gjelder en prøveordning der elevene skal kunne velge mellom flere tekster. For det andre forutsetter utviklingen av adaptive prøver storskalautprøvinger, noe som vil være vanskelig å gjennomføre i en så liten kohort. Den begrensede størrelsen på elevgruppen vil også påvirke hva som er mulig når det gjelder å hente ut, analysere og presentere resultater.

Mangelen på fagtekster og fagbøker på samisk i mange fag understreker behovet for å gjøre flere faglige tekster tilgjengelige. Dette vil kunne støtte lærernes arbeid, gi elevene tilgang til flere fagtekster og styrke muligheten til å lære på eget førstespråk. Det arbeidet som Nasjonalt senter for samisk i utdanningen har startet med Giellabálgát (samiske språkløyper) i tilknytning til dagens nasjonale prøver, kan være et godt utgangspunkt for videre utvikling av støtteressurser. Disse ressursene bør også omfatte kompetanseheving for lærere og skoleledere. giellabalgat.no er en relativt ny nettside, driftet av miljøet som til nå har utarbeidet nasjonale prøver i lesing på samisk. En del av støtteressursene til dagens nasjonale prøver er allerede tilgjengelige der, og dette arbeidet vil kunne videreføres.

Det er et uttalt behov at de samiske leseprøvene må gi mer informasjon til kvalitetsutvikling enn hva dagens prøver gjør. Dette omtaler vi nærmere i kapittel 7.

4.5 Vurdering av endringer i prøvenes konstrukt

Når prøvene skal endres forutsetter dette også en gjennomgang av prøvenes konstrukt. Her drøfter vi kort forslagene fra arbeidsgruppene, og hva dette kan bety for faglig innhold og begrunnelser. Vi skal lage nye prøver i matematikk i stedet for regning som i dag. Det vil altså være en større endring i faglig innhold for disse prøvene, enn for prøvene i lesing og engelsk. Vi drøfter mulige konsekvenser av denne endringen mer utfyllende.

Engelsk

Engelskgruppen foreslår at prøven utvides med en del for *Use of English*, for å gi et mer nyansert bilde av elevenes engelskkompetanse. Denne delen skal fange sider ved elevenes bruk av engelsk i ulike sammenhenger, blant annet sosiolingvistisk kompetanse og interkulturell kommunikasjon, og sees som en viktig utvidelse av prøvens konstrukt sammenlignet med dagens prøve. Vi har svært begrenset erfaring med å måle produktive ferdigheter, men ser med interesse fram til å undersøke om dette kan fungere i praksis.

Matematikk

Overgangen fra regning til matematikk innebærer at prøvene skal endres fra å måle en grunnleggende ferdighet til å måle kompetanse i et fag. Det vil føre til et skifte i hva prøvene skal vurdere og hvordan resultatene kan brukes i opplæringen.

Dagens prøver bygger på rammeverket for grunnleggende ferdigheter. Rammeverket er et sentralt grunnlagsdokument i utviklingen av både LK06, LK20 og de nasjonale prøvene. For nasjonale prøver i lesing er rammeverket operasjonalisert på en måte som ligger tett på både konstruktet i PISA og læreres etablerte forståelse av hva lesing innebærer. For regning er situasjonen annerledes.

Nasjonale prøver i regning skal måle i hvilken grad elevene kan bruke regning til å løse problemer i en virkelighetsnær kontekst. Rammeverket har problemløsning som utgangspunkt, og oppgavene er knyttet til ulike deler av problemløsningsprosessen. I dagens prøver kommer dette til uttrykk gjennom de tre ferdighetsområdene «gjenkjenne og beskrive», «bruke og bearbeide» og «reflektere og vurdere». Det er disse områdene lærerne får tilbakemeldinger om. Brukerinnsikten vår viser imidlertid at disse kategoriene i begrenset grad er gjenkjennelige for lærere i deres vurderingspraksis. De oppleves ikke alltid som tydelig koblet til kompetansemålene lærerne vurderer etter, og gir derfor ikke nødvendigvis den støtten i oppfølgingen som intensjonen tilsier.

Dette kan tyde på en mer grunnleggende utfordring: Sammenliknet med lesing framstår grunnleggende ferdigheter i regning som mindre tydelig forankret i lærernes forståelse av læreplanverket. I dagligtalen forbindes regning ofte med beregning eller aritmetikk, og begrepet kan derfor få en snevrere betydning enn det rammeverket legger til grunn. Samtidig har bruken av nærliggende uttrykk som «regneferdigheter», «grunnleggende regneferdigheter» og «grunnleggende ferdigheter i regning» vært lite konsekvent og ofte uten tydelig avklaring av innholdet. Når sammenhengen mellom grunnleggende ferdigheter i regning og fagene heller ikke oppleves som tydelig nok, kan dette påvirke hvordan innholdet i prøvene forstås, og i hvilken grad resultatene oppleves som nyttige for lærerne. Informasjon om ferdighetsområder kan derfor være faglig relevant, men likevel framstå kompleks og lite tydelig.

Overgangen fra nasjonale prøver i regning til prøver i matematikk kan gi flere faglige og pedagogiske gevinster, men innebærer også noen prinsipielle og praktiske utfordringer. En sentral fordel er at matematikk som fag er mer etablert og gjenkjennelig enn regning som grunnleggende ferdighet. En slik endring kan derfor bidra til større begrepsmessig klarhet og gjøre det lettere for lærere og skoleledere å forstå hva prøvene måler, og hvordan resultatene kan brukes i undervisningen. Fra tilbakemeldinger vet vi at prøven i dag ofte forstås som en prøve i matematikk og ikke ferdigheten regning. Videre kan en sterkere forankring i matematikkfaget gi prøvene økt relevans for lærernes praksis, særlig dersom tilbakemeldingene knyttes til kompetanseområder som oppleves som tydelige og gjenkjennelige i faget. Dette kan også styrke prøvenes legitimitet og øke sannsynligheten for at de får en tydeligere læringsstøttende funksjon.

Samtidig reiser overgangen viktige spørsmål om hva som går tapt når prøvene beveger seg bort fra regning som grunnleggende ferdighet og over mot matematikk som fagkompetanse. Regneprøvene har hatt et særskilt formål ved å måle elevenes evne til å bruke regning i ulike faglige, praktiske og samfunnsmessige sammenhenger. Dersom de

nye prøvene i større grad orienteres mot matematikkfaget, er det en risiko for at det anvendte og tverrfaglige perspektivet svekkes. En annen utfordring er at et kompetanseorientert matematikkbegrep er bredere og mindre entydig avgrenset enn rammeverket for grunnleggende ferdigheters beskrivelse av regning. Dette kan gjøre det mer krevende å definere prøvenes konstrukt, utvikle presise tilbakemeldingskategorier og sikre at resultatene blir forståelige og nyttige for lærerne. Et annet aspekt er at dersom lærerne forventer at vi skal kunne måle bredden i faget, kan konseptet bli veldig stort og omfattende.

Samlet sett framstår overgangen fra regning til matematikk som et grep som kan styrke prøvenes tydelighet, faglige forankring og pedagogiske nytte, men bare dersom konstruktet får en god avgrensning og kommuniseres på en måte som både er faglig presis og nyttig for lærere i deres arbeid. Det vil være avgjørende å finne en balanse mellom nærhet til læreplanen, robusthet over tid og ivaretagelse av fagets funksjonelle og anvendte sider.

Lesing

Lesegruppe har i sitt arbeid valgt å ta utgangspunkt i den etablerte forståelsen av lesing i læreplanverket, og i dagens leseprøver. Gruppen anbefaler at de nye prøvene fremdeles utvikles med grunnlag i den samme definisjonen på lesing som PIRLS- og PISA-undersøkelsene bruker. Dette er faglig godt forankret, og vil også bidra til kontinuitet. Kontinuitet i leseprøvene er særlig viktig for å kunne følge utviklingen i de synkende leseresultatene.

5. Sentrale utredningsspørsmål og veivalg

I dette kapittelet diskuteres fellestrekkene i forslagene fra arbeidsgruppene. Vi drøfter først forslag som skal bidra til at elevene får en bedre opplevelse av prøven. Vi går så igjennom gruppenes forslag til hvordan informasjon fra prøvene kan bli mer nyttige for lærerne. Dette handler om når lærerne får informasjon, hva de får informasjon om og hvorvidt de har mulighet til å få referansepunkter som hjelper dem å forstå styrker og forbedringsområder for elever og klasser. Vi drøfter også utforming av gode resultatvisninger og god støtte til bruk av prøvene. Til slutt gjør vi noen vurderinger av hva vi mener det er realistisk å få til kort sikt, og hva som krever mer arbeid.

Forklaring av begreper som brukes i drøfting av arbeidsgruppenes forslag

Skalapoeng

Prøveresultatene fra dagens nasjonale prøver omregnes til skalapoeng. Skalapoengene skal reflektere elevenes ferdighetsnivå. Skalaen ble fastsatt i 2022, og 50 skalapoeng representerer gjennomsnittlig poengsum for nasjonale prøver i 2022. Skalaen er i teorien uendelig, men de fleste elever får et sted mellom 20 og 80 skalapoeng.

Mestringsnivåer

Elevene deles i mestringsnivåer som gjenspeiler ferdighetsnivået deres på bakgrunn av skalapoeng. Prøvene på 5. trinn har tre mestringsnivåer, mens prøvene på 8. og 9. trinn har fem mestringsnivåer. Oppgavene i prøvene tildeles et mestringsnivå som gjenspeiler vanskegraden til oppgavene.

Anker og ankeroppgaver

Et anker brukes for å koble sammen prøver for at resultatene fra prøvene skal kunne sammenlignes (de Ayala, 2009). Prøvene kobles sammen ved at noen oppgaver, kalt ankeroppgaver, er med i de ulike prøvene flere år på rad. Resultatene på ankeroppgavene gjør det mulig å plassere prøvene på samme skala, selv om de ellers inneholder ulike oppgaver.

Lenking

Lenking handler om å plassere prøver som måler samme konstrukt på samme skala, slik at resultatene blir sammenlignbare (Kolen & Brennan, 2014). Dagens nasjonale prøver lenkes over tid ved hjelp av ankeroppgaver. Lenkingen gjør at resultatene fra prøvene fra år til år kan sammenlignes. Vertikal lenking vil si at prøver som er utviklet for forskjellige aldersgrupper, men som måler det samme, blir satt på samme skala (Kolen & Brennan, 2014).

Trend

Det nasjonale gjennomsnittet for resultatene på nasjonale prøver sammenlignes over år for å vise utviklingen, eller trenden, i resultatene over tid. Denne sammenligningen mellom tidligere års resultater er mulig fordi prøvene lenkes ved hjelp av ankeroppgaver.

IRT-analyser

I dag brukes Item Response Theory (IRT) og ankerdesign for å regne ut resultatene og å koble prøvene sammen over tid. IRT er en samling med modeller som har til formål å beskrive sammenhengen mellom oppgaver i en prøve og det underliggende konstruktet som prøven skal måle (de Ayala, 2009). IRT medfører høyere presisjon av resultatene enn poengsum.

5.1 Kortere prøver

Gruppene ønsker at de nye prøvene skal gi mer informasjon om den faglige ferdigheten, samtidig som elevene har en bedre prøveopplevelse. De mener dagens prøver er for lange. Som svar på dette foreslår arbeidsgruppene blokkadaptive prøver og prøvekonsepter som består av kortere prøver og ulike delprøver.

I kapittel 1.3 pekte vi på ulike hensyn som veies mot hverandre når denne typen verktøy skal designes. Vi omtaler det som et kompromiss mellom tidsbruk, relevans og presisjon. Lengden på en prøve, både antall oppgaver og tidsbruk, vil ha konsekvenser for hvor bredt konstrukt som kan kartlegges og hvor presise resultatene vil være. For å måle hele konstruktet må prøven inneholde tilstrekkelig mange og relevante oppgaver. Dersom oppgavene i en prøve ikke representerer ferdigheten i konstruktet godt nok kan det gå utover validiteten til resultatene.

Krav til presisjon handler om hvor sikre vi ønsker å være når vi presenterer resultatene fra en prøve. Antall oppgaver i en prøve har konsekvenser for hvor sikre vi kan være (de Ayala, 2009). Hver oppgave gir oss informasjon om en elevs ferdigheter, i større eller mindre grad.

5.2 Blokkadaptive prøver

De tre arbeidsgruppene ønsker blokkadaptive prøver som en del av sine prøvekonsepter. Hensikten er altså både å få mer nyansert informasjon fra prøvene, samtidig som det kan redusere prøvebelastningen for elevene. Arbeidsgruppene er opptatt av at prøvene skal være motiverende for elevene, og at alle elever skal oppleve mestring.

I en lineær prøve får elevene de samme oppgavene i samme rekkefølge. I en adaptiv prøve får elevene nye, tilpassede oppgaver basert på hvordan de har prestert på de foregående oppgavene. Fordelingen av oppgaver kan enten gjøres oppgave for oppgave, eller ved at oppgaver grupperes sammen i blokker. Blokkadaptive prøver er adaptive prøver der tilpasningen av oppgavenes vanskegrad skjer på blokk-nivå, og ikke på enkeltoppgavenivå. Blokkene består av grupper av oppgaver, og blokkene har ulik vanskegrad. I en blokkadaptiv prøve tar alle deltagerne de samme oppgavene i første blokk. Etter gjennomføring av første blokk sendes deltagerne videre til neste steg, og tildeles en oppgaveblokk som er tilpasset deres prestasjonsnivå basert på resultatene i foregående blokk. Dette gjør at elevene får en stadig mer presis tildeling av oppgaver.

Prøver gir bedre informasjon om elevenes ferdigheter når de er tilpasset elevenes ferdighetsnivå, slik de blir i blokkadaptive prøver. En for lett prøve gjør det vanskelig å skille mellom elever med moderate og høye ferdigheter, mens en for vanskelig prøve ikke gir tilstrekkelig informasjon om svake områder hos elever med lave ferdigheter. Blokkadaptive prøver tilpasser vanskegraden slik at hver elev svarer på flere oppgaver som samsvarer med deres ferdighetsnivå, sammenlignet med en tradisjonell, lineær prøve. Samtidig vil eleven måtte svare på færre oppgaver for å oppnå samme presisjonsnivå som ved lineære prøver (Yan et al., 2014). For at blokkadaptiv prøvegjennomføring skal fungere tilfredsstillende, forutsetter dette derimot høy presisjon i estimeringen av oppgaver før gjennomføring, slik at oppgavene kan grupperes i blokker i tråd med testdesignet. Dette stiller krav til et større og mer representativt utprøvningsgrunnlag enn dagens prøver.

Adaptivitet utfordrer tolkning av resultatene

Adaptive prøver kan gjøre det mer krevende å presentere resultater på en forståelig måte for lærerne fordi elevene får ulike oppgaver som også vektes ulikt. Adaptive prøver medfører altså at det blir langt flere oppgaver som læreren må sette seg inn i. Når elevers oppgaveløp kan variere, må det avklares hvilke typer detaljer som kan og bør vises på klasse- og elevnivå, og hvordan visningene skal gi et godt grunnlag for oppfølging uten at læreren må gjøre et omfattende manuelt tolkningsarbeid.

Danmark gjennomførte adaptive nasjonale prøver for grunnskolen over en lang periode (2010 – 2021), men valgte å gå bort fra dette prøvedesignet. Disse prøvene var imidlertid oppgave for oppgave-adaptive, og ikke blokkadaptive. Kritikken mot de danske prøvene gikk i hovedsak ut på at selv om resultatene var veldig presise, var de også svært vanskelig å ta i bruk i klasserommet. Blokkadaptive prøver kan ses som et kompromiss som både kan gi mer informasjon om de høyest og lavest presterende elevene, samtidig som elevene har noen oppgaver felles.

Arbeidsgruppene ønsker at både elever og lærere skal få en kvalitativ tilbakemelding om hva en elev eller klasse mestrer, og hva de må arbeide mer med. Dette forslaget om å presentere noen samlede vurderinger virker å la seg kombinere med blokkadaptive prøver. Gode kvalitative handlingsrettede tilbakemeldinger kan kompensere for at lærere ikke vil se et felles oppgavesett som alle elever har gjennomført, og bruke dette som felles referansepunkt i oppfølging med elevene.

Det finnes allerede løsninger for adaptivitet i dagens system. Kartleggingsprøvene på 1. og 3. trinn er blokkadaptive. Brukerinnsikten bekrefter bekymringen om at det kan være vanskelig for lærere å forstå logikken bak slike prøver, og de kan bli forvirret dersom de ønsker å sammenligne resultater for ulike elever. Det adaptive designet må kommuniseres godt til skolene dersom vi skal gå videre med dette. Endringer i resultatvisning og støtteressurser må også eventuelt testes grundig.

Det er mer tid- og kostnadskrevende å utvikle blokkadaptive prøver, og denne typen prøver vil kreve teknisk utvikling, flere store utprøvinger og mer avanserte analysemetoder. Et annet aspekt er videreutvikling av dagens resultatvisning. Ressursbehovet knyttet til endringer i resultatvisningen som følge av et adaptivt design vil variere med ambisjonsnivået.

5.3 Vurdere resultater mot standard, norm eller gjennomsnitt

Arbeidsgruppene har vurdert behovet for at hver skole skal kunne se hva deres elever mestrer opp mot en standard eller gjennomsnitt. Særlig lese- og matematikkgruppen ser dette som en slags kalibrering og referanse hvor lærer og den enkelte skole kan se hva de mestrer godt eller mindre godt opp mot et nasjonalt snitt eller definerte standarder. Her gjør vi rede for hvilke forutsetninger som må være til stede for å kunne tilby denne typen informasjon. Vi drøfter også noen konsekvenser, men uten å ta stilling til om dette er noe vi bør tilby.

For å kunne presentere standardisert informasjon kreves prøver som har identisk eller tilnærmet identisk innhold for alle som gjennomfører, og forutbestemte regler for

gjennomføring, skåring og tolkning av prøvene (AERA et al., 2014). Standardisering handler om å sikre at prøver gjennomføres, vurderes og analyseres under helt like betingelser hver gang. Dette sikrer at resultatene blir pålitelige, rettfærdige og sammenlignbare. Hvis resultatene fra en prøve blir for tilfeldige, forsvinner også mye av relevansen.

Standardisering gjør det mulig å se prøveresultater i sammenheng på tvers av kontekster. Det er flere måter å tilby skolene en norm de kan vurdere egne elevers ferdigheter mot. En norm kan etableres ved at alle elever innenfor ett kull gjennomfører prøven på et gitt tidspunkt som i dag, eller at normen baseres på et tilstrekkelig stort utvalg. En annen fremgangsmåte vil være å bruke utprøvningsdata for å etablere en norm. Omfanget av utprøving vil være avhengig av designet på prøven.

Arbeidsgruppene foreslår også mer kriteriebaserte tilbakemeldinger. Det vil si kvalitative tilbakemeldinger på hva gruppen eller eleven mestrer, og hva som bør jobbes med videre. Slike tilbakemeldinger krever ikke IRT-analyser av prøveresultatene, og kan være tilgjengelig for lærerne rett etter selve prøvegjennomføringen.

5.4 Følge opp elevenes faglige utvikling gjennom frivillige prøver

Arbeidsgruppene foreslår at vi, som del av det nye prøvekonseptet, utvikler frivillige prøver som kan gi lærerne informasjon om elevenes faglige utvikling (retest).

Retest kan forstås som gjennomføring av en ny prøve som er utformet på tilsvarende måte som hovedprøven, og som er faglig og psykometrisk ekvivalent med tidligere gjennomføring av prøven. Retesten kan også være en ny gjennomføring av samme blokkadaptive prøve, men der eleven får tildelt andre blokker dersom ferdigheten har endret seg.

Det er ikke likegyldig for prøveresultatet hvilket tidspunkt en eventuell retest gjennomføres på. En retest som blir gjennomført kort tid etter første gjennomføring vil i større grad påvirkes av en øvingseffekt enn en prøve som blir gjennomført etter lengre tid. Øvingseffekten betyr at elever ofte presterer bedre på en andre gjennomføring av en prøve (Hausknecht et al., 2007). Arbeidsgruppene har ulike forslag til når en retest bør gjennomføres. Dersom en retest for eksempel skal være tilgjengelig hele tiden, slik at lærerne kan bruke den etter behov, vil det stille store krav til både selve prøven og til informasjonen som blir gjort tilgjengelig i forbindelse med prøven for at resultatet skal bli tolket på riktig måte.

En annen avveining er hvilken informasjon retesten skal gi. Det enkleste vil være å sammenligne poengsum. Det vil derimot forutsette at det er samme prøve som brukes. Den mest hensiktsmessige måten å analysere en retest med en ny prøve på vil være å bruke det samme designet som blir brukt for trend, med ankeroppgaver som går igjen i gjennomføringene. Dette designet vil hjelpe med å sette de to gjennomføringene på samme skala. Gitt at gjennomføringene har blitt gjennomført i like prøvesituasjoner, vil resultatene på de to gjennomføringene da være sammenlignbare. En slik metode vil fungere både for lineære og blokkadaptive prøve. Det er viktig å merke seg at avanserte metoder for analyse også vil medføre økt kompleksitet og kan gjøre prøvene mindre intuitive å tolke for lærerne som skal bruke informasjonen.

Arbeidsgruppene foreslår at en eventuell retest skal være frivillig. Dersom en prøve er frivillig, vil vi ikke kunne legge til grunn at elevene som gjennomfører prøven representerer populasjonen. Et skjevt utvalg vil gjøre det vanskelig å regne ut et troverdig gjennomsnitt, og det betyr at det vil være krevende å benytte data fra gjennomføringer av denne typen prøver som utgangspunkt for en standardisert prøve.

Arbeidsgruppene peker på frivillige moduler eller prøver som kan tas når som helst. Dette står særlig sentralt i matematikkgruppens konsept. Et slikt konsept utgjør en stor forandring fra dagens prosesser og rammer. Det vil kreve mye utvikling å skulle sikre god brukervennlighet i en slik løsning.

5.5 Analyser av oppgavekvalitet, elevferdigheter og trender

Arbeidsgruppene beskriver ulike ønsker og behov knyttet til presentasjon av prøveresultatene. Resultatet av en digitalt gjennomført prøve kan omregnes og formidles på forskjellige måter. Den mest brukte omregningen er poengsum, altså summen av antall poeng en elev har fått på en prøve. Poengsum kan være en rimelig refleksjon av en elevs prestasjon på en prøve. Den store fordelen med denne metoden er at den er velkjent og mer intuitivt forståelig. Ulempen er at mer avanserte metoder, som for eksempel Item Response Theory-analyser, gir mer presise estimer av elevenes ferdighetsnivå. Hvilken analysemetode som egner seg best for det læringsstøttene formålet må utredes videre.

Selv om konseptene arbeidsgruppene presenterer skiller seg fra dagens prøver, kan det fortsatt være behov for utregning av resultater og lenking med etablerte metoder, for å sikre at grunnlaget for tilbakemeldingene er valid og reliabelt. Dette vil være relevant også dersom tilbakemeldingene skal være mer kvalitative, slik som for eksempel kriteriebaserte tilbakemeldinger. Hvordan disse resultatene presenteres til lærerne vil likevel være annerledes enn slik det gjøres i dag. Resultatvisninger drøftes senere i dette kapittelet.

Umiddelbare resultater

Flere av arbeidsgruppene ønsker at resultatene fra de nye prøvene blir tilgjengelig for lærerne umiddelbart. Behovet for raskere resultater går også igjen i brukerinnsikten fra skolene. I dagens nasjonale prøver blir elevsvar og poengsum tilgjengelig umiddelbart etter gjennomføring, mens skalapoeng og mestringsnivå blir tilgjengelig for skolene 2 - 3 uker etter siste gjennomføringsdag. Nasjonalt, fylkes- og kommunale gjennomsnitt som kan sammenlignes på gruppenivå publiseres ca. en måned senere etter siste gjennomføringsdag. Elevferdigheter blir estimert i etterkant av gjennomføringsperioden. Estimering i etterkant gjør at vi unngår at det nasjonale snittet er basert på et skjevt utvalg, og i denne prosessen blir også data fra prøvene kvalitetssikret grundig. Lærerne uttrykker at den lange ventetiden gjør at de ikke får brukt resultatet ordentlig fordi «elevene da har glemt alt». Raskere resultater gjør at tiltak kan settes inn raskere og mer effektivt.

Dersom IRT skal brukes for å estimere elevferdigheter, er det nødvendig å vite hvor vanskelige oppgavene er og hvor godt oppgavene skiller mellom forskjellige ferdighetsnivåer. Disse to aspektene ved oppgavene kalles oppgaveparameterne. Når oppgaveparameterne er kjente, kan elevenes ferdighetsnivå estimeres direkte basert på

elevsvarene. Oppgaveparameterne kan estimeres ved utprøving på et representativt utvalg av elever i forkant av prøven. Dette gjøres allerede i dag, som en del av prøveutviklingen. Den andre forutsetningen for umiddelbare resultater, er et prøvegjennomføringssystem som kan regne ut elevferdigheter etter at hver elev har gjennomført prøven. Prøvegjennomføringssystemet Udir bruker i dag har ikke denne muligheten, og dette må i så fall implementeres.

Flerdimensjonale analyser

Brukerinnsikten peker altså på at det må bli lettere å gå fra funn til handling. Lærere etterlyser klarere kategorier, tydeligere kobling mellom delferdigheter og tiltak, bedre visninger på klasse- og oppgavenivå og mer tilgjengelig visning av prøven i etterkant.

IRT-analysene som brukes i dag antar at det som måles i prøven kun tilhører én underliggende årsak. Innenfor IRT-rammeverket finnes det også modeller som kan ta høyde for at det er flere kategorier som blir målt i den samme prøven. Disse modellen kalles flerdimensjonale modeller. Resultatet av slike modeller vil være informasjon om hver av kategoriene istedenfor bare ett skalapoeng.

Flerdimensjonale modeller forutsetter derimot at det faktisk er flere kategorier som måles i prøvesettet (de Ayala, 2009). Det vil si at prøvene spesifikt må utvikles for å måle flere underkategorier. I tillegg vil det, som spesifisert om kortere prøver, være en balanse mellom ønsket presisjonsnivå og antall oppgaver som blir inkludert. I en prøve som skal måle flere kategorier vil det være færre oppgaver til hver kategori enn det er til konstruktet i en prøve som skal måle ett konstrukt. Det vil kunne bety at hver kategori blir målt med lavere presisjon enn ønsket, eller at prøven/prøvene eventuelt må bli mer omfangsrike.

Informasjon om progresjon

Arbeidsgruppene ønsker prøver som gir informasjon om elevenes progresjon, både i form av retester, men også ved at resultater skal følge eleven over tid. Å følge elevenes resultater over tid vil bety at skolene må ha tilgang til flere prøveresultater og resultater over lengre tid enn i dag. Her vil det være behov for å utrede nærmere både metodiske spørsmål og konsekvenser for personvern og datasikkerhet.

Resultater på prøvene fra både retest og tidligere gjennomføringer kan følge elevene og være tilgjengelig for de forskjellige lærerne denne eleven har. Det kan gjøres uten noe ekstra analysesteg. Man kan både se for seg at lærer og skole får beholde resultatene over tid, eller at resultatene blir tilgjengelig igjen når elevene har gjennomført en ny prøve.

En mer avansert løsning kan være å legge til rette for at man kan følge utvikling i elevenes læring ved at prøver for ulike aldersgrupper kobles via vertikal lenking (se faktaboks i kapittel 5). Dette har vært testet tidligere på dagens nasjonale prøver i regning (Ræder, Andersson & Olsen, 2022).

Alle arbeidsgruppene har løftet fram behovet for å se elevenes utvikling over tid. Når dette skal utforskes videre, må en utredning inkludere de personvernmessige sidene ved å oppbevare og tilgjengeliggjøre informasjon om elever over flere år. Det vil være nødvendig å se på i hvilken grad progresjonsinformasjon faktisk utgjør en viktig rolle i det

læringsstøttende arbeidet. Vi må også vurdere om nivået av datainnsamling fra elever blir for omfattende sett i forhold til formålet til prøvene.

Trend handler om systematiske endringer over tid på aggregert nivå. For store utvalg vil denne informasjonen være pålitelig. For mindre utvalg vil det være mer usikkerhet rundt estimatene. Spørsmålet om trend henger derfor sammen med spørsmålet om hvilken informasjon kommuner og nasjonalt nivå skal få fra de nye prøvene. Å skulle undersøke hvordan elevferdigheter utvikler seg over tid som i dag, forutsetter spesifikke prøvedesign og statistiske modeller for å kunne sammenligne resultater fra ulike kohorter og ulike prøvesett. Enkelt sagt innebærer dette at en del oppgaver gjenbrukes mellom år gjennom et ankerdesign. De oppgavene som gjenbrukes kalles ankeroppgaver og kobler resultater på tvers av år. Det muliggjør etablering av en felles skala mellom årene, slik at vi kan sammenligne elevferdigheter mellom ulike årskull. Dersom det er ønskelig med sammenlignbarhet over tid i de nye læringsstøttende prøvene, vil dette få konsekvenser for både prøveutvikling og analysemetoder. Blant annet vil det kunne redusere fleksibiliteten i verktøyene og gjøre det mer krevende å oppnå andre sentrale målsettinger for prøvene.

5.6 Prosessdata kan gi bedre innsikt

I oppdragsbrevet etterspørres rikere informasjon om elevene. En mulighet vil være å undersøke nærmere *hvordan* elevene arbeider med oppgavene. Når prøver gjennomføres, logger systemet informasjon om hvordan elever interagerer med oppgavene. Denne informasjonen kalles prosessdata. Eksempler på prosessdata er tidsbruk på enkeltoppgaver og på hele prøven, antall ganger en elev klikker seg inn på hver oppgave, detaljerte svar på hver oppgave (for eksempel valg i flervalgsoppgaver), samt om en elev har sett på en oppgave og hoppet over den uten å svare.

Prosessdata kan gi en bedre forståelse av prøveresultater på individnivå. For eksempel kan to elever ha svake prestasjoner på en prøve, men den ene eleven har brukt kun 10 minutter på hele oppgavesettet, mens den andre har brukt 60 minutter. I et slikt tilfelle kan det være nyttig for læreren å vurdere om prestasjonen til den første eleven er mindre valid, fordi eleven har brukt for kort tid til å vise sine ferdigheter. Samtidig kan den andre eleven ha større vansker med å besvare oppgavesettet.

I tillegg er det nyttig å analysere prosessdata på oppgavenivå. For eksempel kan tidsbruk synliggjøre hvilke oppgaver som er tidkrevende, detaljerte svar kan avdekke typiske feil, og informasjon om hvilke oppgaver mange elever hopper over, samt på hvilket trinn i prøven dette skjer, kan gi viktig innsikt i prøvefunksjoneringen.

Prosessdata utgjør altså en verdifull datakilde som går utover tradisjonelle prestasjonsmål, ved å gi innsikt i hvordan en klasse interagerer med og besvarer oppgaver. Dette kan bidra til en mer nyansert tolkning av prøveresultater og være til hjelp i undervisningen, i tråd med arbeidsgruppenes ønsker. Det er imidlertid viktig å vurdere om denne informasjonen blir for omfattende til at lærerne faktisk kan nyttiggjøre seg av den.

I tillegg til prosessdata vil det det være mulig å trekke ut informasjon på gruppenivå. For eksempel kan det være nyttig å vite ikke bare hvor mange som svarte riktig på en oppgave, men også å få informasjon om hvilke elever/hvor mange elever som valgte de ulike gale

svaralternativene (distraktorene). Mønster i feilsvar kan indikere felles misforståelse i en klasse, eller at elever har utfordringer innenfor bestemte ferdighetsområder. Læreren kan bruke denne informasjonen til å justere undervisningen og rette fokus mot sentrale misoppfatninger og ferdighetsområder.

5.7 Tilgjengelige resultater og støtteressurser

For at en prøve skal være læringsstøttende, må resultatene presenteres på en god måte og skolene må få god støtte til å følge opp disse gjennom systematisk arbeid på skolen. I dette kapittelet viser vi hvilke konkrete utfordringer ved dagens løsninger arbeidsgruppene har tatt utgangspunkt i. Dette utdyper situasjonsbeskrivelsen som vi presenterte i gjennomgangen av kunnskapsgrunnlaget i kapittel 3. For arbeidet med nye prøver er det viktig å få tak i konkrete utfordringer og hva skolene ønsker at vi skal gjøre annerledes.

Utfordringer med dagens resultatvisning og støtteressurser

Støtteressursene til dagens prøver består av tre deler: generell støtte for å følge opp de grunnleggende ferdighetene lesing og regning, samt deler av faget engelsk; støtte til å tolke prøveresultatene og støtte til å følge opp årets prøveresultater. Den generelle støtten ligger åpent tilgjengelig, mens øvrig støtte krever innlogging. Brukerinnsikten viser at det er flere utfordringer knyttet til dagens ressurser.

Tilgang til resultater og støtte

Noen av dagens støtteressurser, særlig ressurser knyttet til tolkning av resultater, er kun tilgjengelige i PAS og forutsetter at lærere har tilgang til prøvegrupper som har gjennomført. Selv om tilgangsstyringen forvaltes lokalt, kan dette medføre barrierer for deling og oppfølging på tvers. Tilgangen til ressursene i PAS er også tidsavgrenset. Resultater og tilhørende ressurser er bare tilgjengelige i 10 måneder, noe som kan utfordre kontinuitet i oppfølgingen.

Finne fram til relevant støtte

Dagens støtteressurser oppleves av flere som fragmenterte og lite sammenhengende. Ressursene er fordelt mellom ulike plattformer, noe som kan gjøre det krevende å få oversikt, finne relevant støtte og bevege seg effektivt fra resultater til oppfølging. I tillegg påvirker utforming og format hvordan ressursene brukes. Flere lærere etterspør mer tilgjengelige, praktiske og handlingsrettede formater, blant annet utskriftsvennlige veiledninger og tilgang til oppgaver til bruk i etterarbeid.

Forstå og vurdere støtte

Selv om flere av dagens støtteressurser er utviklet for å støtte tolkning og oppfølging, opplever mange det utfordrende å omsette resultatene i pedagogisk handling. Utfordringene handler blant annet om å forstå skalapoeng og mestringsnivåer, og om å kommunisere resultatene på en meningsfull måte til elever og foreldre.

Omsette i oppfølging

For at lærere skal ta støtteressurser i bruk, må de oppleves som relevante og gi merverdi i praksis. Noen skoler har etablert egne rutiner og oppfølgingsopplegg, og opplever derfor mindre behov for Udirs ressurser. En innvending er at støtteressursene er for generelle, og

at de dermed ikke står i forhold til ambisjonen om tilpasset undervisning. Et genuint dilemma er hvilke lærere disse ressursene tar sikte på å støtte. Det som kan være nyttig for en lærer, kan være elementært for en annen.

5.8 Mer handlingsrettet informasjon fra prøvene

Brukerinnsikten understøtter tilbakemeldingen fra lærergruppene. De største utfordringene ved dagens prøver handler om å tolke visningen og følge opp elevene. Mange beskriver tung navigasjon (mange klikk), og flere skoler lager egne oversikter for å få den helheten de trenger. Flere av skolene etterspør innsyn i oppgaver på en måte som støtter lokalt etterarbeid og et godt skole-hjem-samarbeid. Arbeidsgruppene foreslår resultatvisninger som peker videre mot hvordan skolen bør jobbe med resultatene, og som er tettere knyttet til eleven med egen elevvisning med konkrete fremovermeldinger.

Handlingsrettet visning

Arbeidsgruppene legger stor vekt på at resultatvisningen må være mer handlingsrettet og tidsbesparende, og at den bør kobles tett til relevante støtteressurser for å redusere lokalt tolknings- og utviklingsarbeid. Samtidig peker brukerinnsikten på at utfordringen ofte ikke er å forstå resultatene i seg selv, men at det tar for lang tid å komme fra resultat til konkret oppfølging, blant annet på grunn av mange klikk, fragmentert navigasjon og behov for å lage egne oversikter lokalt. En mer handlingsrettet visning kan derfor være et av de mest virkningsfulle grepene for å øke sannsynligheten for læringsstøttende bruk i praksis.

Samtidig innebærer mer "handlingsrettet" visning en avveining: jo mer resultatvisningen peker mot tiltak, desto viktigere blir det at anbefalingene oppleves som støtte til profesjonelt skjønn, ikke som fasit eller som styring. Dette peker mot at videre arbeid bør prioritere å teste hvilke grep som faktisk sparer tid for lærere, slik som tydeligere kategorier/delområder og raskere vei til relevant støtte, og å validere at dette oppleves nyttig på tvers av skoler med ulik praksis og kapasitet.

Konkrete fremovermeldinger

Et gjennomgående forslag fra arbeidsgruppene er mer konkrete fremovermeldinger som støtter veien videre for elev og lærer, gjerne gjennom resultatvisninger som tydeliggjør styrker og utviklingsområder og kobler disse til relevante oppfølgingsressurser. Dette samsvarer med brukerinnsikten og sitatene fra PAS-undersøkelsen, der flere beskriver dagens mestringsbeskrivelser som for generelle og vanskelig å nyttiggjøre seg av i praksis, og etterlyser mer konkrete signaler om hva elevene bør øve mer på.

Samtidig må det avklares hva prøvene kan gi presise og ansvarlige tilbakemeldinger om? Vi legger til grunn at mye må bero på lærerens vurdering og lokal kontekst. Dette peker mot et behov for klarere kategorier/ferdighetsområder og bedre kobling mellom resultater og *foreslåtte* oppfølgingsgrep. Slik kan lærer få forslag, men uten at prøven framstår som en fasit for undervisningspraksis. Videre arbeid bør derfor teste hvor konkrete fremovermeldingene kan være uten at de oppleves styrende, og hvilke formuleringer som faktisk gir læreren støtte til å handle uten å "overtolke" et resultat. Engelskgruppen foreslår at lærere selv skal kunne legge inn sine vurderinger med utgangspunkt i informasjonen fra

prøvene. Intensjonen med dette er å styrke samspillet mellom et prøveresultat og all annen informasjon læreren allerede har om eleven.

Egne visninger for elev (elevprofil)

Flere av arbeidsgruppene anbefaler egne og tydeligere elevprofiler/elevvisninger, blant annet for å styrke elevens eierskap til læringen. En slik visning kan også gjøre resultatene enklere å bruke i samtaler med elev og foresatte.

Brukerinnsikten viser at språk og begreper i dagens resultatvisning kan være vanskelige å forstå og videreformidle. En ny elevvisning må utformes på en måte som er lett å forstå uten å skape feiltolkning, og at skolens veiledning i samtaler fortsatt vil være viktig.

Et viktig spørsmål er hvordan resultatene bør presenteres for å ivareta motivasjon og mestring, særlig for elever med svært lave resultater. Dette bør prøves ut tidlig med lærere og elever for å sikre at elevvisningen oppleves støttende og ikke demotiverende.

5.9 Behov for støtte til oppfølging

Brukerinnsikten viser at dagens støtteressurser er for fragmenterte og generelle. Derfor vil vi gjennomgå støtteressursene for å gjøre dem mer sammenhengende og lettere tilgjengelige.

De lærerne som rapporterer at det ved deres skole settes av tid til å følge opp resultatene fra dagens prøver, opplever dagens prøver som mer nyttige enn lærere som arbeider ved skoler hvor dette ikke blir prioritert. Denne innsikten er også relevant for nye prøver. Samtidig bør vårt utgangspunkt være at tid er en knapphetsfaktor i skolen, og derfor fjerne unødvendige hindre for god oppfølging av resultatene.

Støtteressurser

Arbeidsgruppene anbefaler flere og mer målrettede ressurser som en integrert del av prøvene. Samtidig ser vi at eksisterende støtte i PAS brukes lite selv om den ligger tett på resultatvisningen. Utfordringene vi ser i dag handler ikke nødvendigvis om å utvikle mer materiell, men om å utvikle ressurser som er lett tilgjengelige, tydelig strukturerte og tett koblet til det læreren faktisk trenger å gjøre etter prøven. NOU 2023:27 peker også på at ressursbruk bør vurderes opp mot faktisk bruk, og at for stort omfang kan gjøre det vanskelig å finne og nyttiggjøre seg relevant støtte (s. 96).

Flere av arbeidsgruppene sine anbefalinger peker på sjekklister, samtaleguider og ferdige undervisningsopplegg som virkemidler for mer lik oppfølging og kortere vei fra resultat til tiltak. Samtidig reiser dette en avveining: jo mer konkret støtten er, desto større er risikoen for at den innskrenker lærernes handlingsrom og oppleves som en form for "myk styring", slik NOU 2023:27 også peker på (s. 96).

En mulig løsning som også foreslås av noen av arbeidsgruppene, er at ressursene gjøres redigerbare og presenteres som støtte til profesjonelt skjønn, snarere enn som fasit. Dette kan bidra til å gi lærere konkrete startpunkter og spare tid, samtidig som det er rom for lokal tilpasning og faglig vurdering.

Tid, arbeidsbelastning og eierskap

Flere anbefalinger forutsetter at skoleledelse og skoleeier legger til rette for tid og strukturer, for eksempel gjennom opplæring, fellestid og planlagt oppfølging i profesjonsfellesskapet. Brukerinnsikten og PAS-tilbakemeldinger viser at manglende tid er en sentral barriere i dag, og at oppfølging for flere er personavhengig dersom den ikke settes i system.

Dette tilsier at støtteressurser alene ikke er nok. Det må også utvikles støtte som gjør det enklere å etablere rutiner lokalt. Samtidig må det drøftes hvor langt nasjonale føringer bør gå, og hvilke former for opplæring som gir effekt uten å oppleves som pålegg som ikke lar seg gjennomføre innenfor skolens rammer.

Arbeidsgruppen i lesing foreslår at lærere i forkant av prøven skal gi en vurdering av elevene. Forhåndsvurdering av elevenes ferdigheter eller kompetanse kan styrke lærerens eierskap til resultatene. En slik vurdering kan gi et ekstra tolkningsgrunnlag ved at læreren kan se egen vurdering i lys av prøveresultater. Samtidig vil forarbeidet også innebære en økt arbeidsbelastning, som må vurderes opp mot den samlede tidsbarrieren som allerede preger oppfølgingen. Dersom et slikt grep skal fungere i praksis, bør det organiseres på en måte som gjør det gjennomførbart, for eksempel i team rundt elevgruppen, slik konseptet antyder. I det videre arbeidet bør derfor vi teste om estimat faktisk gir økt kvalitet i tolkning og oppfølging, og om merbelastningen oppleves som akseptabel sammenlignet med gevinsten.

6. Konsekvenser av forslagene for videre digitalisering

I delleveranse 1 skrev vi om at utvikling av læringsstøttende prøver vil kreve ressurser til oppgaveutvikling, testing, kvalitetssikring og analyser, samt utvikling av IT-systemene våre. Under gjør vi rede for muligheter og begrensninger i våre eksisterende systemer, og hvordan det vil påvirke arbeidet med å realisere de nye prøvene. Vurdering av systemstøtte vil være en vesentlig del av vurderingen knyttet til økonomiske og administrative konsekvenser for Udir. Vurderingene i denne rapporten er foreløpige og må utdypes og kvalitetssikres.

De sentrale systemene vi benytter for å gjennomføre prøvene er Prøve- og administrasjonssystemet (PAS, "administrasjonssystemet") og Eksamens- og prøvesystemet (EPS, "gjennomføringssystemet"). PAS håndterer administrering og resultatvisning, og EPS er systemet hvor oppgaver og prøver utvikles og gjennomføres.

Vi legger til grunn at dagens nasjonale prøver skal fortsette inntil nye læringsstøttende prøver er ferdig innfaset. Dette betyr at utviklingen av løsninger for nye læringsstøttende prøver må gå i parallell med forvaltning og vedlikehold av eksisterende prøver. Dette gjør at ressurser knyttet til forvaltningen av dagens prøver ikke i sin helhet kan benyttes i utviklingen av nye prøver. Det betyr også at eventuelle systemendringer må gjøres på en måte som ikke forstyrrer de eksisterende prøvene. I innfasingsperioden vil det totale ressursbehovet øke.

Vi har igangsatt utredning av nytt teknisk veikart for systemporteføljen for eksamen og prøver, for å modernisere systemene og kunne møte behov fra oppdragene for å videreutvikle eksamen og prøver. Utredningsarbeidet i første fase for systemporteføljen i prøve- og eksamenstjenesten vil pågå ut 2026 med oppfølging våren 2027. Formålet er blant annet å avklare hvordan systemene kan moderniseres for å redusere risiko og kompleksitet i gjennomføringen, og hvordan nye behov for videreutvikling av prøver og eksamen kan ivaretas.

6.1 Prøveadministrasjonssystemet (PAS)

PAS er et egenutviklet IT-system fra 2015 som brukes til å administrere gjennomføringen av prøver og eksamen. Systemet omfatter også dagens resultatrapport hvor lærer får tilgang til resultater for sine elevgrupper. Dette begynner nå å bli et gammelt IT-system som ikke uten videre kan bygges om til å møte helt andre behov.

I dag er systemet tett knyttet til de eksisterende prøveformene og nåværende prosesser. Alle endringer som fraviker dagens prosesser hva angår påmelding, gjennomføring og oppfølging, vil medføre utviklingskostnader. Større endringer vil også kunne endre driftskostnadene, for eksempel økt mengde data som skal lagres over en lengre periode. For å kunne gjøre prøvene mer relevante og møte forslagene fra arbeidsgruppene, kan det være behov for endringer på roller, tilganger, brukerflyter og prosesser. Dette kan bety store endringer som treffer kjernen av dagens systemportefølje. Slike endringer vil være omfattende, kreve mye IT-utvikling, og innebære risiko for uønskede forandringer i dagens prøver og prosesser.

For å møte behovene i læringsstøttende prøver kan det også bli behov for endring eller nyutvikling som påvirker integrasjonen mellom PAS og skoleadministrative systemer (SAS) og krever utvikling hos deres leverandører.

6.2 Eksamen- og Prøvesystemet (EPS)

Eksamen- og prøvesystemet (EPS) er systemet hvor innholdet i oppgaver og prøver utvikles og hvor elevene gjennomfører prøvene. Funksjonaliteten i dette systemet rammer inn hvilke oppgaveformater som kan utforskes og utvikles på kort sikt. Dette er en internasjonal hyllevare for å gjennomføre prøver, og det er ikke rom for videreutvikling av dette systemet innenfor dagens kontrakt. Større endringsbehov vil medføre en ny anskaffelsesprosess.

En av grunnene til at systemet ble valgt, var at det inneholdt et rikt utvalg av ulike oppgaveformater for elevene. Dersom utviklingen av de læringsstøttende prøvene ikke kan gjøres innenfor dette systemet, er det derfor sannsynlig at det vil kreve utvikling av et helt nytt system, fremfor anskaffelse av en eksisterende hyllevare. Både utvikling av et nytt gjennomføringssystem eller anskaffelse av en ny hyllevare vil være et stort IT-prosjekt som vil kreve større utredningsarbeid før vi kan gi en slik anbefaling.

I dag brukes det samme gjennomføringssystemet for Udirs prøver (nasjonale og kartleggingsprøver) og skriftlig eksamen i et flertall av fag. Det betyr at videreutvikling for å møte behovene i de nye læringsstøttende prøvene må ses i sammenheng med plan for videreutvikling av kartleggingsprøver og eksamen, jfr. oppdragsbrev 25-04 – *Utrede en sikrere og forbedret eksamensordning i videregående opplæring* og 25-14 – *Tilrettelegging av eksamen yrkesfag*.

Dagens kontrakt om Eksamens- og prøvesystemet går ut i 2029. Det betyr at vi uansett må avklare eventuelle behov for en ny anskaffelse. Dette er en del av Udirs arbeid med å utrede nytt teknisk veikart. Funnene fra denne kartleggingen vil være sentrale i vurderingen av tid og kostnad for å møte behovene i de nye læringsstøttende prøvene. Det betyr at Udir kan komme tilbake med en mer presis vurdering av konsekvenser av nye prøver for systemutvikling etter at denne anbefalingen foreligger.

6.3 Hva er mulig på kort, mellomlang og lang sikt

Vi vil understreke at vi ikke har utredet konsekvenser av overgangen til læringsstøttende prøver for IT-systemene. Forslagene fra arbeidsgruppene medfører en stor endring, og dette vil mest sannsynlig være kostbart å utvikle. Vurdering av konsekvenser for systemutviklingen inngår i utredningen nytt teknisk veikart for prøver og eksamenstjenesten og vil ikke foreligge før ved årsskiftet. Det er derfor ikke mulig å gi økonomiske estimater for dette arbeidet i delleveranse 2. Vi vil likevel gi noen foreløpige vurderinger av hva vi mener det er realistisk å endre på kort sikt, og hva som vil ta mer tid.

Endringer som er mulig på kort sikt

På kort sikt vil fokuset det kommende året være på utredning og utprøvinger med utgangspunkt i forslagene fra lærergruppene i denne rapporten. Vi legger til grunn at utprøvinger av oppgavetyper og prøver på kort sikt må skje innenfor de tekniske oppgaveformatene som kan brukes i eksisterende gjennomføringssystem (EPS), og at

videreutvikling av systemet ikke er aktuelt før ny anskaffelse eller ny systemutvikling på mellomlang sikt.

Kunnskapsgrunnlaget og konseptene fra arbeidsgruppene peker mer på betydningen av forbedringer i hvordan rapporter og støtte til oppfølging av elevene i klassen utformes for lærer, enn hva slags tekniske formater som finnes for selve oppgavene i prøven.

Vi ser derfor ikke begrensningen i dagens system som en avgjørende hindring for å prøve ut nye oppgaver, men forslag om hvordan ulike deler av en prøve skal kunne settes sammen av lærer og hvordan modellene for adaptivitet i prøvene skal være, kan gjøre at mer fleksible og komplekse deler av et nytt prøvekonsept må utvikles på lengre sikt.

Utprøving av endringer og prototyping av nye løsninger for resultatvisning og støtteressurser bør utforskes så snart som mulig, siden disse løsningene er mest etterspurt blant lærerne. På disse områdene er det sannsynlig at det kan gjøres større forbedringer som gir verdi for lærer, uten å nødvendigvis kreve endringer i prøvenes utforming eller gjennomføring. Utprøvingene kan både brukes til å gjøre forbedringer i dagens prøver, og også vise vei mot gode løsninger for de nye prøvene. Vi ønsker å komme i gang med denne typen utprøvinger relativt raskt etter at vi har fått tilbakemeldinger på denne leveransen fra KD.

Endringer som er mulig på mellomlang sikt

Over de neste 1-3 årene vil Udir måtte jobbe parallelt med faglig utvikling av prøvene og utvikling av nødvendig systemstøtte. Arbeidet forutsetter god brukerinvolvering og partsforankring i alle faser.

Omfanget av endringer som kreves avhenger av valgt prøvekonsept og ambisjonsnivå, men kan bli omfattende. Flere av forslagene fra arbeidsgruppene vil medføre store forandringer. Om endringene gjennomføres i dagens systemportefølje, er det også en risiko for at nyutvikling fører til feil eller utilsiktede endringer i gjennomføringen av de eksisterende prøvene.

Gitt omfanget av endringer kan det være aktuelt å utvikle ny systemstøtte for å realisere noen av forslagene fra lærergruppene. Dette vil også redusere risikoen for feil i gjennomføringer av dagens prøver. Ved å bygge ny systemstøtte, vil man også redusere risiko knyttet til teknisk gjeld fra dagens system. Ulempen er økt utviklingskostnad, og at det vil kunne ta lenger tid før innfasing. Utredningen av dette er pågående, og gjøres i sammenheng med analyse av behov og løsninger for neste anskaffelse av gjennomføringssystem når kontraktperioden utløper.

Vår utgangshypotese er at digitalisering av en ny resultatvisning, en forbedret støtteressurs og et dashboard for lærer, trolig kan gjøres utenfor dagens systemer. Dette vil i så fall måtte valideres gjennom prototyping og deretter utvikles smidig gjennom iterativ og inkrementell systemutvikling av minimumsversjoner som kan videreutvikles ut fra erfaring i bruk. Arbeidet forutsetter god brukerinvolvering.

Det vil være viktig å benytte etablerte samarbeid på tvers av relevante systemer. Dette kan for eksempel gjelde samarbeid mellom Udirs systemer, Feide og leverandørene for skoleadministrative systemer. Integrasjon med eksterne aktører bør ikke oppleves som en begrensende faktor i den kontinuerlige videreutviklingen av de læringsstøttende prøvene.

Endringer som er mulig på lengre sikt

Gitt varighet på eksisterende kontrakter og utviklingstid for nye løsninger eller anskaffelser, er det naturlig å tenke at større endringer ikke vil kunne realiseres før tidligst 3-5 år frem i tid. Innenfor denne tidsperioden legger vi til grunn at vi må forholde oss til de kjernesystemene vi har. Samtidig tar vi sikte på å teste ut nye løsninger relativt raskt, og at vi kan utvikle bedre støtte til skolene også med utgangspunkt i de systemene vi har nå. Når nødvendig systemstøtte for de nye læringsstøttende prøvene er utviklet til en slik grad at de kan innføres i fullskala, vil de eksisterende nasjonale prøvene kunne avvikles. Avhengig av tidligere valg kan det være behov for utvikling relatert til utfasing av dagens prøver.

Etter at de nye prøvene er ferdig innført, er det nødvendig at de blir kontinuerlig videreutviklet. De nye læringsstøttende prøver bør ikke sees på som et prosjekt som ferdigstilles, leveres og vedlikeholdes, men som en levende tjeneste som videreutvikles inkrementelt og iterativt i dialog med lærere og skoleeiere. Udir mener at det er av avgjørende betydning at involveringen med praksisfeltet og fagmiljøer kan fortsette for å holde prøvene og systemet både faglig og teknisk oppdatert. Dette vil bidra til at vi møter forventningen i oppdragsbrevet om at nye prøver skal være *“oppdaterte, relevante og brukervennlige”*.

6.4 Våre forslag til oppfølging og videre arbeid

Det er altså usikkert hvordan drifts- og videreutviklingskostnader vil se ut, både på kort, mellomlang og lang sikt. Dette vil avhenge av ønsket ambisjonsnivå og valgt prøvekonsept.

På kort sikt skal Udir gjennomføre den pågående utredning av dagens systemportefølje og beslutte en plan for veien videre fra et teknisk perspektiv. Her må vi ta spesielt hensyn til den utgående kontrakten for dagens gjennomføringssystem. Parallelt med dette vil læringsstøttende elementer prøves ut og brukertestes for å skaffe erfaringer både til langsiktig utvikling og for forbedringer vi kan implementere på kort sikt. Vi mener det vil være mulig å gå raskt i gang med å prøve ut noen av forslagene fra lærergruppene om forbedret resultatvisning og opprydding i støtteressurser.

På mellomlang sikt utvikles nødvendig systemstøtte for læringsstøttende prøver med bakgrunn i valgt teknisk retning, og fortsatt med kontinuerlig brukerinvolvering. I denne fasen vil utviklingskostnadene være høye. For å estimere kostnadene vil det være nødvendig å slutføre utredningen av nytt teknisk veikart og validere prototyper for ny funksjonalitet.

På lengre sikt vil kostnadsbildet ved nyutvikling og forvaltning av tjenesten være avhengig av anskaffelse av ny gjennomføringsløsning og implementasjonen av denne for læringsstøttende prøver, eller om det alternativt vil utvikles ny systemstøtte for gjennomføring.

På lang sikt bør prøvene, systemet og ressursene videreutvikles som en levende tjeneste, slik at de fortsetter å oppleves som oppdaterte, relevante og brukervennlige.

7. Informasjon om læringsutbytte for elever i den samiske skolen

Både oppdragsbrevet (OB 2025-018) og Meld. St. 34 (2023-2024) er tydelig på at et nytt system for kvalitetsutvikling også skal gi informasjon om læringsutbyttet for elever i den samiske skolen. Når det utvikles nye prøver og verktøy skal disse være tilpasset samiske elever og behovene i den samiske skolen. I tillegg kan det være behov for andre verktøy spesielt tilpasset kvalitetsutvikling i samisk skole.

Det er langt færre elever som gjennomfører prøver på samisk enn på norsk, noe som gjør at det er vanskelig å utarbeide meningsfylt statistikk på samme nivå som for de norske prøvene. Derfor er det et sterkt behov for å forbedre den kvalitative informasjonen som nye prøver kan gi, og for å gi en mer skreddersydd støtte til hvordan resultatene kan følges opp i opplæringen og i profesjonsfellesskapet.

For å kunne utvikle prøver, resultatvisninger og støtteressurser som både gir god informasjon om elevene i samisk skole, og også relevant støtte til oppfølging, er det helt nødvendig å innhente enda dypere innsikt i behovene i samisk skole. Behovene vil for eksempel knytte seg til forbedret kvalitativ informasjon fra prøvene, prøver som er utviklet på samisk kultur og samiske språks premisser, ulikhetene i elevenes opplæring og den store mangelen på læremidler i samisk skole.

7.1 Behov for forbedret kvalitativ informasjon fra samiske prøver

Gjennomføring av prøver i små elevpopulasjoner gir noen utfordringer når disse dataene skal analyseres. Små elevgrupper gjør at vi ikke kan bruke statistiske modeller som baserer seg på storskala populasjon. For eksempel blir det lite hensiktsmessig å bruke ankeroppgaver for å sammenligne utvikling av ferdigheten over tid. Eventuelle blokkadaptive prøver må også konstrueres på en annen måte.

Ettersom hele elevpopulasjonen er langt mindre enn et vanlig piloteringsutvalg, piloterer de samiske prøvene ikke etter samme kravene som de norske prøvene. Konsekvensen av dette er at de samiske prøvene ikke kan gi presis informasjon om elevenes ferdighetsnivå. Samtidig gir gjennomføring i små elevpopulasjoner noen helt klare fordeler, som det vil være mulig å utnytte bedre. Blant annet vil det være enklere å følge opp resultatene på den enkelte skole, og å ha en mer direkte oppfølging av elevgruppene. Dersom dette potensialet utnyttes på riktig måte, vil det være mulig å gi skolene mer informasjon både på elev- og gruppenivå.

En mulig løsning som bør utforskes videre er at resultatvisningen også kan vise mer detaljert informasjon som støtter tolkning og refleksjon, og ikke bare en totalskår. Det vil

for eksempel være mulig å gi rikere informasjon om oppgavenes kobling til prøvekonstruktet sammen med elevresultatene. Ved å bruke prosessdata som for eksempel tidsbruk, antall forsøk og elevens valg av svaralternativer er det også mulig å gi mer kvalitativ informasjon om elevenes ferdigheter.

7.2 Behov for prøver som er utviklet på egne premisser

Både NOU 2023:27 og Meld. St. 34 (2023–2024) peker på behovet for å utvikle egne prøver på samisk. Samiske språk og opplæringssituasjoner har særegne forutsetninger, og oversatte prøver vil ikke nødvendigvis fange opp samiske elevers faktiske ferdigheter. Prøver som utvikles på de samiske språkene og kulturens premisser, vil gi mer relevante resultater og dermed et bedre grunnlag for læringsstøtte og oppfølging av elever i samisk skole. Dette gjelder både fleksible verktøy i begynneropplæringen og de nye prøvene med læringsstøttende formål.

I lesing utvikler vi allerede egne leseprøver for de samiske språkene. Når det gjelder matematikk, følger norsk og samisk skole samme læreplan (MAT01-05). Denne læreplanen ivaretar særskilt samisk innhold. Det kan være behov for å utrede om det skal utvikles egne oppgaver til de nye matematikkprøvene for samiske skoler, for å bedre ivareta samiske perspektiver. Det er også behov for mer kvalitetssikret oversettelse av matematikkprøven, utført av personer med kompetanse både i matematikk og i de ulike samiske språkene. Vi er i dialog med Senter for samisk i opplæringa ved Samisk høyskole om dette.

7.3 Behov for relevante sammenligningsgrunnlag

Samiske elever utgjør en heterogen gruppe, med store variasjoner i forutsetninger, språkbakgrunn og erfaringer. Noen elever går på skoler der all undervisning foregår på samisk, andre elever har undervisning på samisk i enkelte fag eller mottar undervisning i samisk språk som fjernundervisning. Dermed blir det ekstra viktig å møte behovet for et meningsfullt sammenligningsgrunnlag og støtteressurser til videre oppfølging. Støtteressurser som er basert på norsk skole vil ikke treffe denne gruppen godt nok. Det kan også være behov for å spisse ulike ressurser mot ulike grupper av både skoler, lærere og elever.

7.4 Behov for læremidler på samisk

Meld. St. 34 (2023–2024) peker på en kritisk mangel på samiske læremidler, noe som gjør det krevende å undervise og å gi elevene tilstrekkelig støtte i samiske fag. Dette gjør at prøver og tilhørende støtteressurser får en utvidet funksjon. Innholdet i prøvene bidrar med tekster, oppgaver og eksempler som også kan brukes i undervisningen. Dette gjelder særlig i leseopplæringen, men også i regning/matematikk. Det er derfor behov for å se prøveutvikling, resultatvisning og oppfølgingsmateriale i lys av den generelle læremiddelsituasjonen, og sikre at ressursene som utvikles er enkle å ta i bruk og tilgjengelige for lærere i samisk skole.

7.5 Oppsummering

Det er flere behov som må undersøkes grundigere for å kunne utforme prøver, resultatvisninger og støtteressurser som kan bidra til god oppfølging i samisk skole. Udir har et tett samarbeid med Senter for samisk i opplæringa ved Samisk høgskole. Dette senteret er tildelt oppdraget med å utvikle dagens prøver. I kontrakten med senteret er det også angitt at senteret skal bidra i arbeidet med å utvikle prøvene videre, og at dette skal skje i tett samarbeid med samiske skoler, lærere og fagmiljøer. Udir vil i samarbeid med Senter for samisk i opplæringa ved Samisk høgskole sørge for at det innhentes brukerinnsikt og videre analyse av behov. Vi vil også undersøke muligheten for å analysere dataene fra de samiske prøvene på en slik måte at de gir mer informasjon til fagmiljø, skoler og Sametinget som utgangspunkt for videre utvikling.

8. Nye verktøy på 1.trinn

Delleveranse 3 i oppdraget gjelder verktøy som lærere kan bruke i den tidlige lese- og regneopplæringen. På kort sikt bes vi om å øke fleksibiliteten i dagens kartleggingsprøver på 1.trinn. Disse skal etter hvert erstattes av nye fleksible verktøy, og vi bes om å foreslå hvordan disse kan utformes.

I dette kapittelet gjør vi først rede for pågående prosesser som påvirker eller påvirkes av arbeidet med nye verktøy til bruk i begynneropplæringen. Vi orienterer så om hvilke kortsiktige endringer vi har gjort med dagens kartleggingsprøver på 1. trinn, før vi med utgangspunkt i foreliggende kunnskapsgrunnlag peker på behov de nye fleksible verktøyene må møte. Til slutt redegjør vi for hvordan vi vil arbeide videre med denne delen av oppdraget.

Udir har innhentet informasjon om skolenes erfaringer med dagens kartleggingsprøver, og om hvilke behov skolene beskriver i arbeidet med tidlig lese- og regneopplæring. Det er gjennomført dybdeintervjuer med et mindre utvalg samarbeidsskoler, og innhentet tilbakemeldinger ved hjelp av de årlige spørreundersøkelsene som er beskrevet i kap. 2.5. Udir har i tillegg hatt dialog med fagmiljøene som utvikler dagens kartleggingsprøver og Statped for å innhente råd.

Det primære målet i denne fasen har vært å etablere et godt utgangspunkt for arbeidet med denne delen av oppdraget. Vi har lagt vekt på å samle bred innsikt i og kunnskap om hvilke behov dagens kartleggingsprøver dekker, hva som er utfordrende, og hvilke egenskaper nye verktøy bør ha for å fungere i skolens praksis.

8.1 Pågående prosesser og endringer av relevans for arbeidet

Udir har flere oppdrag og pågående prosesser som har relevans for arbeidet med nye verktøy til bruk i begynneropplæringen. Arbeidet bør også sees i sammenheng med kompetanseutvikling og veiledning, slik at lærere får nødvendig støtte både i bruken av verktøyet og i oppfølgingen av elevenes utvikling.

Tidlig innsats, tilpasset opplæring og individuell tilrettelegging

Kartlegging og fleksible verktøy til bruk i begynneropplæringen er nært knyttet til arbeidet med oppdrag 2025-01 - *Kunnskap, status og tiltak for å sikre bedre tidlig innsats, tilpasset opplæring og individuell tilrettelegging*. En helhetlig tilnærming vil bidra til at tiltak for kartlegging, forebygging, tidlig innsats og individuell tilrettelegging vurderes og utredes i sammenheng.

Endringer i struktur og innhold i skolehverdagen på 1.–2. trinn

I oppdrag 2026-010² om *Endringer i struktur og innhold i skolehverdagen på 1.–2. trinn* arbeider vi med modeller som angir overordnet retning for hvordan skolehverdagen på 1.–2. trinn kan organiseres. Oppdraget tar for seg hvordan rammene for opplæringens innhold og timer i skolen skal reguleres for å sikre en god overgang fra barnehage til skole og særlig tett oppfølging av elevenes utvikling av lese- skrive- og regneferdigheter, mer lek og fysisk aktivitet og bedre muligheter for utvikling av sosiale og emosjonelle ferdigheter.

Justering av læreplanen i matematikk

Fra høsten 2026 vil det gjennomføres en justering av læreplanen i matematikk for grunnskolen. Formålet med justeringen er å tydeliggjøre det faglige innholdet, synliggjøre regneferdigheter og justere beskrivelsen av regning som grunnleggende ferdighet slik at progresjonen i ferdigheten blir tydeligere. Som støtte til læreplanen er det lagt inn forklaringstekster med didaktisk støtte til mange kompetansemål i de aktuelle trinnene. Det må utredes i hvilken grad endringen i kompetansemålene kan påvirke vurderingene og arbeidet med mer fleksible verktøy, slik at verktøyene samsvarer med det justerte faglige innholdet.

Digitale ferdigheter og bruk av digitale enheter

Det er besluttet at bruken av digitale enheter for elever på 1.-4. trinn skal være særskilt varsom, og ikke dominere undervisningen. Redusert bruk av digitale enheter for de yngste elevene er forskriftsfestet, og vil gjelde fra 1. august. Udirs råd for god digital praksis og skjermbruk i skolen endres også i tråd med dette. Endringen gjør at vi i mindre grad kan legge til grunn at elevene er kjent med digitale enheter på de laveste årstrinn, og at en digital prøvesituasjon kan være krevende for elever og lærere. Dette kan også gi usikre resultater. En mulig konsekvens av endringen vil være å utvikle analoge verktøy som erstatning eller supplement til digitale kartleggingsprøver. Vi drøfter dette kort i kap. 8.4.

8.2 Endringer for å gjøre dagens prøver mer fleksible

Fra skoleåret 2025–2026 har Udir gjennomført endringer i dagens ordning for kartleggingsprøver på 1. trinn. Udir har nå lagt til rette for en mer fleksibel gjennomføring, men prøvenes innhold og oppsett er ikke endret.

Fra og med i år er kartleggingsprøvene ikke lenger låst til én enkelt periode. Skolene kan velge å gjennomføre denne prøven enten på våren i 1. trinn eller på høsten 2. trinn. Skolene

² Tillegg nr. 6 til tildelingsbrev til Utdanningsdirektoratet for 2025 med oppdrag 2025-010 om Endringer i struktur og innhold i skolehverdagen på 1.-2. trinn

må velge én av disse gjennomføringsperiodene. Vi har også åpnet for at skolene kan melde opp enkeltelever og ikke kun hele klasser som i dag. Det betyr at skolene kan gjennomføre prøvene kun for elever de mener det er nyttig å kartlegge nærmere.

Som følge av endringene er det gjort en justering i merknaden til Opplæringsforskrifta § 19-6, tredje ledd, for å tydeliggjøre at kartleggingsprøvene er frivillige og at skolene kan tilpasse bruken til egne behov. Samtidig angir opplæringsloven §28-1 at kommunen har beslutningsmyndighet når det gjelder bruken av kartleggingsprøvene. Det innebærer at kommuner kan fastsette føringer for gjennomføring i sine skoler. Veiledningen på udir.no er oppdatert for å støtte skolene i den fleksible bruken av kartleggingsprøvene.

8.3 Behovsbeskrivelser i kunnskapsgrunnlaget

Under peker vi på noen sentrale behov med utgangspunkt i kunnskap fra ulike kilder. Dette gir noen holdepunkter for hva det vil være viktig å vektlegge i arbeidet med å utvikle nye fleksible verktøy til bruk i den tidligere lese- og regneopplæringen.

Dagens kartleggingsprøver forstås ulikt

Et tema som går igjen i kunnskapsgrunnlaget, og som også oppdragsbrevet trekker frem, handler om formålet for kartleggingsprøvene. Dagens kartleggingsprøver har som formål å identifisere de elevene som trenger ekstra hjelp i lese- og regneopplæringa. Prøvene gir kun informasjon om disse elevene. Både NOU 2023:27 og Prøver i fagfornyelsen (2020) peker på at formålet med dagens kartleggingsprøver forstås ulikt. Fagmiljøene har pekt på at betegnelsen «kartleggingsprøve» kan skape forventninger om at prøven skal gi bredere informasjon om elevgruppen enn det som samsvarer med formålet. Skolene gir på sin side tilbakemelding om at de ønsker å få informasjon om alle elevene som gjennomfører prøvene, og ikke bare elevene som er i oppfølgingsområdet. På bakgrunn av dette er det viktig å avklare formål og kommunisere dette klart til skolene, slik at lærerne har riktige forventninger til hvilken informasjon verktøyene vil gi.

Skolenes erfaringer med gjennomføringen

Skolene gir tilbakemelding om at enkelte oppgaver eller ferdighetsområder i dagens prøver kan oppleves å ligge foran det elevene har møtt i undervisningen på gjennomføringstidspunktet. Læreplanen har ingen kompetansemål før etter 2. trinn, men lærere oppfatter at kartleggingsprøvene indikerer et forventet nivå for 1. trinn. Dette kan bidra til å forskyve målene for opplæringen. Brukerinnsikten viser også at skolene opplever at gjennomføringen ikke er godt nok tilpasset elevenes alder. Dette gjelder blant annet lengde, omfang og tekstmengde, samt det digitale formatet. I tillegg til dette gir skolene tilbakemelding om at de ønsker å få bedre informasjon om løsningsprosessene til elevene. Dette gjelder både for at lærerne bedre skal kunne forstå resultatene for elever i oppfølgingsområdet, og hvordan informasjonen om elevene kan brukes til å planlegge undervisningen. Innhold i prøver eller verktøy, tidspunkt for gjennomføring og forventninger til progresjon må henge sammen med læreplan, lokal opplæringspraksis og elevenes faktiske opplæring.

Digitale eller analoge verktøy?

I sammenstillingen fra samarbeidsskolene er det gjennomgående at flere skoler ville valgt papir, papir ved behov eller en kombinert løsning fremfor et rent digitalt format. I brukerinnsikten knyttes utfordringer ved digital gjennomføring blant annet til elevenes ulike digitale erfaringer, tekniske og praktiske forhold og spørsmålet om hva digitale prøver faktisk måler. Noen skoler viser til at digitale enheter kun brukes i gjennomføring av prøvene, og ikke ellers i undervisningen. Bruk av digitale prøver for de yngste problematiseres også i NOU 2023: 27.

Temaet knyttet til digital gjennomføring må også ses i lys av nasjonale råd og føringer om skjermbruk. Redusert bruk av digitale enheter for de yngste elevene forskriftsfestes denne våren, og vil gjelde fra 1. august. Det betyr at vi i mindre grad kan legge til grunn at elevene er kjent med digitale enheter, og at en digital prøvesituasjon kan kreve mye oppfølging.

En digital prøvesituasjon kan stille krav som ikke nødvendigvis er en del av det prøven er ment å måle. Samtidig kan digitale løsninger gi mange muligheter, blant annet for adaptivitet og kortere gjennomføring. Samtidig gir flere skoler tilbakemeldinger om at det er fordeler ved det digitale. De trekker for eksempel frem at det er tidsbesparende, og at det er mindre administrativt arbeid ved denne typen prøver. Digitale løsninger vil også kunne bidra til mindre etterarbeid når det gjelder oversikt over og tolkning av prøveresultatene. Dersom de nye verktøyene skal være helt eller delvis papirbaserte, må det gjøres en egen vurdering for om det for eksempel er behov for en plattform for å samle og tolke resultater. Dette vil i så fall bety at prøven er analog for eleven, men digital for lærer.

Lærernes vurderinger og systematikk i oppfølgingen

Skolens arbeid med å følge med på elevenes utvikling er del av den ordinære opplæringen, og skolene har plikt til å gi intensiv opplæring til elever på 1.–4. trinn som står i fare for ikke å ha forventet progresjon i lesing, skriving eller regning jfr. §11-3 i Opplæringslova.

Kvalitetsutviklingsutvalget viser til forskning som peker på at lærere ofte har mye informasjon om elever med utfordringer, men at de ikke nødvendigvis har tid, ressurser eller kompetanse til å følge opp informasjonen. Brukerinnsikten gir eksempler på at lærere allerede gjør mange vurderinger tidlig i skoleløpet. Flere skoler beskriver at de bruker bokstavprøver, observasjoner og andre former for kartlegging av delferdigheter for å få oversikt over informasjonen de anser som nødvendig. Denne informasjonen brukes blant annet til å tilpasse undervisningen. Samtidig tyder innspillene på at det er stor variasjon i hvordan skolene setter informasjon og oppfølgingstiltak i system.

I brukerinnsikten uttrykkes det behov for bedre oversikt over resultater på tvers av for eksempel prøver og vurderinger, og for at informasjon i større grad kan følge eleven over tid. Å støtte skolene i å sette informasjon og iverksatte oppfølgingstiltak i et system kan blant annet være nyttig ved lærerbytte og i samarbeid med skolens pedagogiske støttesystemer. Også skolene vi har intervjuet melder tilbake at de gjerne ønsker bedre støtte til å følge utviklingen i elevenes læring.

Kvalitetsutviklingsutvalget refererer til forskning som viser at lærernes innhenting av

informasjon ofte blir mer systematisk fra mellomtrinnet og videre i opplæringsløpet, og i større grad brukes til å vurdere elevenes faglige utvikling. Gitt betydningen av tidlig innsats er det da relevant å undersøke om det er behov for bedre støtte til systematisk innhenting og oppfølging av informasjon i den tidlige lese- og regneopplæringen.

Behov for bedre verktøy og støtte i vurderingsarbeidet

Som del av oppfølgingen av oppdrag nr. 2025-01 - *Kunnskap, status og tiltak for å sikre bedre tidlig innsats, tilpasset opplæring og individuell tilrettelegging*, har Udir oversendt en kunnskapsoversikt til KD som gjennomgår arbeidet med tidlig innsats, inkludering og tilrettelegging for barn og elever med særskilte behov (Utdanningsdirektoratet, 2026). Her løftes mangelen på kvalitetssikrede kartleggingsverktøy fram som et eget tema. Det pekes på betydningen av å bruke verktøy med dokumentert kvalitet når lærere skal avdekke elever som kan ha behov for videre oppfølging. Samtidig vises det til at det finnes et bredt spekter av verktøy i bruk, hvor flere i begrenset grad er kvalitetssikret eller systematisk undersøkt (Arnesen et al., 2017). Hovedbudskapet er altså at det er behov for kvalitetssikrede verktøy som kan støtte lærerne i å følge elevenes utvikling over tid og vurdere effekten av tiltak som iverksettes.

Fagmiljøene som utvikler dagens kartleggingsprøver påpeker at lærere har behov for støtte til å systematisere egne observasjoner og vurderinger, særlig i starten av 1. trinn. Det skjer mye på kort tid i denne fasen, og det kan være krevende for lærer å få tidlig oversikt over eventuelle tilretteleggingsbehov. Fagmiljøene løfter også frem at vurderingsstøtte kan ha andre former enn en prøve, for eksempel sjekklister, veiledet observasjon eller dynamisk kartlegging. Dette understøttes av intervjuene vi har gjennomført. Disse skolene er positive til dynamisk kartlegging som en mulighet for å få utdypende informasjon om elevene når lærere har behov for det.

Kartleggingsprøver kan fungere som god beslutningsstøtte i tidlig leseopplæring (Walgermo et al., 2025). Kvalitetsutviklingsutvalget refererer til metastudier som viser at lærere generelt er godt rustet til å vurdere elevene, men at det likevel er svært krevende å vurdere elever som trenger ekstra oppfølging. Dette gjelder både elever som presterer lavt og elever med høyt læringspotensial. Det krever særskilt kompetanse og gode verktøy for å avdekke slike behov (Machts et al., 2016, sitert i NOU 2023:27). Nye verktøy bør altså supplere lærerens øvrige kunnskap om elevene og støtte lærernes profesjonsfaglige vurderinger.

Informasjon om alle eller noen?

Kvalitetsutviklingsutvalget mener det er problematisk at dagens kartleggingsprøver kun gir tolkbar informasjon om elever i oppfølgingsområdet, selv om alle elevene må gjennomføre prøven. Intervjuene våre bekrefter dette. Skolene ønsker informasjon om alle elever, utover at de er utenfor oppfølgingsområdet. Skolene understreker også at mer informasjon ikke nødvendigvis er bedre informasjon dersom resultatene blir vanskelige å tolke, eller dersom resultatene brukes til noe de ikke er ment for.

8.4 Mulige endringer på lengre sikt

Det er flere retninger som kan være relevante å utforske i det videre arbeidet med å utvikle nye fleksible verktøy til bruk i den tidlige lese- og regneopplæringen. Et sentralt spørsmål er hvilken rolle fremtidige verktøy skal ha i skolen. Kunnskapsgrunnlaget peker mot behov for å vurdere hvordan verktøyene kan fungere både som støtte i lærernes arbeid med vurdering, tilpasning og oppfølging av elever. Dette kan innebære å undersøke om verktøyene i større grad enn i dag kan bidra som beslutningsstøtte i lærernes vurderinger.

Flere av behovsbeskrivelsene peker mot bedre synliggjøring av elevenes delferdigheter, strategier og utvikling over tid. I dialogen med fagmiljøene er det løftet frem muligheter knyttet til verktøy og støtteressurser som kan gi lærerne mer innsikt i informasjonen de ønsker om blant annet hvordan elevene tenker gjennom enkel feilanalyse eller oppfølgingsstøtte knyttet til elevenes svar.

Behovet for god sammenheng mellom vurdering og oppfølging går igjen i mange beskrivelser og undersøkelser. Dette kan være løsninger som gjør det mulig å følge elevenes læring over tid, og som kombinerer kartlegging og profesjonsfaglige vurderinger. Dette vil kunne bidra til mer helhetlig oppfølging av elevene. Udir må samarbeide tett med brukergruppene om hvordan slik støtte kan utvikles på en måte som gir mening og oppleves nyttig.

Nye verktøy bør utformes med mål om kortere økter, tilpasset oppgaveutforming og gjennomføringsformer som tar hensyn til elevenes alder og føringer i læreplanen. Erfaringene fra skolene og dialogen med fagmiljøene viser at både digitale, papirbaserte og kombinerte løsninger vil ha styrker og begrensninger, og at valg av format bør vurderes i lys av hva verktøyene skal gi informasjon om og hvilke krav de stiller til elevene.

Spørsmålet om hvilken elevgruppe verktøyene skal gi informasjon om, er også et viktig avklaringspunkt i det videre arbeidet. Vi vil undersøke løsninger som kan møte skolens behov for informasjon, samtidig som verktøyene beholder et tydelig formål.

8.5 Veien videre, nye verktøy til bruk i begynneropplæringen

Denne våren har vi satt ned en egen utredningsgruppe dedikert til å jobbe med nye fleksible verktøy. Gruppen er satt sammen av utvalgte lærere fra arbeidsgruppene i lesing, matematikk og engelsk, og vil i slutten av juni levere et retnings- og anbefalingsnotat til Udir. Arbeidet skal ta utgangspunkt i kunnskapsgrunnlaget og beskrivelser og drøftinger i denne delleveransen. Samtidig vil vi sørge for at arbeidet ses i lys av andre pågående prosesser, som for eksempel planlagte endringer i struktur og innhold i skolehverdagen på 1.-2. trinn (OB 2026-010). Vi må vurdere om endringene på 1-2.trinn blir av en slik art at det er hensiktsmessig å vente med å ferdigstille de nye verktøyene til etter at disse endringene er avklart.

9. Anbefalinger om videre arbeid, nye prøver og nye verktøy

Udir vil følge opp forslagene fra arbeidsgruppene på flere måter. Sammen med fagmiljøer som har spisskompetanse på prøveutvikling vil vi utrede hvordan nye prøver kan få høy kvalitet og samtidig møte lærernes behov best mulig.

Forslagene fra arbeidsgruppene vil innebære store endringer. Å lage løsninger for behovene de skisserer vil kreve en omfattende utviklingsprosess. Vi mener det vil være en god løsning å organisere dette som en prosess med smidig utvikling. Da vil vi både kunne redusere risiko, legge til rette for at løsningene som utvikles treffer behovene bedre og redusere risikoen for å utvikle løsninger som blir for kostbare og komplekse.

Mange av ønskene fra arbeidsgruppene krever forbedrede resultatvisninger, og bedre støtteressurser. Vi vil undersøke hvilke forbedringer vi kan få til med utgangspunkt i den informasjonen vi får ut fra prøvene i dag og hvilke muligheter vi har med dagens system som ikke utnyttes i dag, for eksempel prøveformater vi ikke benytter. Vi vil også kartlegge hvilke av behovene vi ikke kan møte med dagens system, og hvordan vi kan utvikle ny systemstøtte for å møte disse behovene. Alle tre spor kan og bør gjennomføres parallelt. Der vi kan utvikle ny systemstøtte på siden av dagens kjernesystemer, vil vi kunne jobbe raskere. Store endringer i administrasjonssystemet eller prøvegjennomføringssystemet, vil ta lenger tid.

9.1 Vurdering av hvordan vi kan følge opp forslagene

Vi drøfter her noen konkrete konsekvenser av forslagene som har kommet fra lærergruppene.

Flere gjennomføringer og frivillige moduler

Å åpne for flere gjennomføringer av prøver vil ikke nødvendigvis kreve teknisk utvikling. I teorien kunne man også innført en enkel form for retest uten større endringer. Dagens resultatvisning støtter imidlertid ikke direkte sammenligning av flere prøver. Her kreves det altså videreutvikling.

Både frivillige moduler og retester fører til at det gjennomføres flere prøver, som betyr at det produseres mer prøve- og resultatdata. Forslaget om å vise resultater over lenger tid fører også til at dataene lagres lenger. Dette kan bety økte driftskostnader relatert til datalagring. Det må utredes om dette også utløser behov for utvikling av nye, mer kostnads- og ytelseseffektive lagringsløsninger. Her vil også konsekvenser for personvern måtte utredes nærmere.

Arbeidsgruppene, særlig matematikkgruppen, ønsker frivillige moduler eller prøver som kan tas når som helst. Dette utgjør en stor forandring fra dagens prosesser og rammer. Det vil kreve mye utvikling å skulle sikre god brukervennlighet i en slik løsning, og vi må unngå at dette skaper uønskede effekter på prosessen til eksisterende prøver. Å skulle sammenligne resultater fra forskjellige prøvegjennomføringer vil kreve utviklingsarbeid.

Matematikkgruppen foreslår også at prøvene heller bør gjennomføres på 6. og 9. trinn. Brukerinnsikten viser at det kan være grunnlag for å utrede valg av årstrinn for gjennomføring videre.

Analyse av resultatene

Vi vil utforske hvordan prosessdata fra prøver kan brukes for å møte ønskene fra lærergruppene. Et eksempel kan være å forstå bedre *hvordan* elevene arbeider. Vi vil også utforske nye oppgaveformater og prøvedesign med formål om å teste om forslagene fra arbeidsgruppene fungerer. Dette må vi utrede sammen med fagmiljøer i UH-sektoren som har kompetanse på prøveutvikling og metodiske spørsmål.

Målet med arbeidet er å etter hvert kunne implementere et nytt prøvedesign, med analyser som tilbyr relevant informasjon til lærere om deres elever i tråd med forslagene i kapittel 4. Vi vil da ha fått undersøkt ulike løsninger, og vil kunne bruke erfaring fra utprøvingene i implementering av et nytt prøvedesign og nye analysestrategier.

Støtteressurser og resultatvisning

Vi ser både strukturelle og prosessuelle utfordringer knyttet til utvikling og forvaltning av dagens støtteressurser. Utfordringene handler om hvordan støtteressurser utvikles og forvaltes, om tekniske begrensninger i systemene, og om hvor tilgjengelige ressursene er for lærerne i praksis.

Tilbakemeldinger fra lærere på dagens løsninger vil være et viktig grunnlag for å vurdere hvilke endringer som gir nytte og verdi for arbeidet med nye prøver. Ved å teste ut tiltak og innhente erfaringer kan vi få bedre kunnskap om hva som faktisk gir verdi, og bruke denne innsikten i det videre utviklingsarbeidet.

Behovsbeskrivelsene synes å peke mot løsninger som kombinerer strukturert informasjon om elevenes læring med støtte til kvalitative kriteriebaserte vurderinger. Slik kan resultater fra prøvene ses i sammenheng med annen informasjon læreren har om eleven.

Gjøre støtteressursene lettere tilgjengelige

Manglende kjennskap til støtteressursene og utfordringer med å finne dem ser ut til å være viktige barrierer for bruk. Tydeligere kommunikasjon kan bidra til økt bruk og bedre sammenheng mellom prøvegjennomføring og pedagogisk oppfølging. Dagens organisering fører til unødvendig lange utviklings- og oppdateringsløp. Ved valg av framtidig løsning vil vi særlig vurdere:

- hvor lett er det for lærere å få tak i ressursene
- hvor raskt kan vi publisere og forbedre innhold
- hvordan innholdet kan forvaltes på en sikker måte

Målet er å gjøre ressursene enklere å finne og bruke, samtidig som vi kan publisere og forbedre dem raskere. Det gir oss også bedre forutsetninger for å lære av arbeidet med dagens prøver, og ta læringen med inn i utviklingen av støtteressurser til de nye prøvene.

Når vi har konkludert på hvilken plattform vi ønsker å benytte for å publisere støtteressurser, må vi også utrede om det kan være behov for å videreutvikle denne for å møte behovene for nye prøver. Samtidig bør muligheten for integrasjon mellom et prøveadministrasjonssystem og en plattform for publisering av støtteressurser utredes. Dette vil kunne føre til behov for utviklingsarbeid i begge systemer.

Konsekvenser av endringer

Alle større endringer vil medføre utvikling. Når det gjelder analyse av prøveresultater, vil nye prøvedesign og analysestrategier kreve mange utprøvinger og tett kontakt med fagmiljøer med kompetanse på prøveutvikling. Adaptive prøver forutsetter systemer som kan regne ut elevferdigheter underveis i gjennomføringen, eller at vi kobler til programmeringsverktøy der slike algoritmer allerede er implementert. Dersom systemet skal presentere helt ny informasjon eller gjør nye analyser eller fortolkninger, vil det kreve større utvikling.

Mengden utvikling som kreves, og dermed kostnaden, vil være avhengig av ambisjon og hvor stor variasjon vi ønsker i resultatvisningene. Én tolkning av "rikt og nyansert" innebærer at informasjonen er tettere knyttet til det spesifikke faget eller ferdigheten på prøven, for eksempel ved å belyse delferdigheter eller elementer av faget. Økt differensiering mellom resultatvisningene til de forskjellige prøvene betyr mer utvikling. Det samme gjelder ønsket om tilpassede visninger for forskjellige brukergrupper. Å arbeide smidig med dette betyr at vi ikke bør legge omfattende ressurser på nye løsninger uten at disse er prøvd ut i mindre skala og at de møter et tydelig behov. Her må det treffes en balanse mellom kostnad, kvalitet, og innfasingstid.

9.2 Kommentarer og tilbakemelding fra organisasjonene i skolesektoren

I kapittel 2.3 skrev vi om hvordan vi har lagt til rette for drøftinger av arbeidet med organisasjonene i skolesektoren så langt, og at vi har vært opptatt av å se arbeidet med nye prøver i lys av målsettingene for et nytt helhetlig system for arbeid med kvalitetsutvikling i skolen.

Organisasjonene fikk tilsendt leveransen som gikk til KD i starten av juni. Denne leveransen ble drøftet i referansegruppen for vurdering 10. juni. I etterkant av dette møtet har organisasjonene sendt sine skriftlige innspill. Disse innspillene er lagt ved denne justerte versjonen av delleveranse 2.

I referansegruppemøtet 10. juni gav organisasjonene støtte til at Udir har valgt å arbeide tett på praksisfeltet i utredningen av mulige konsepter for nye læringsstøttende prøver. Organisasjonene gir uttrykk for at utredningen gir ny innsikt og bidrar til å bringe arbeidet videre. Det kom positive tilbakemeldinger på de konkrete forslagene fra arbeidsgruppene, og på hovedretningen i arbeidet. Gruppen mener den innovative metodikken som er valgt her gir oss et godt utgangspunkt for å utvikle prøver som kan bli nyttige for skolene, samtidig som det er behov for gode drøftinger med organisasjonene.

Organisasjonene etterlyste samtidig tydeligere beskrivelser av Udirs vurdering av forslagene, og hva vi ønsker å gå videre med. Det ble også reist spørsmål om potensialet for å forbedre støtteressurser, med begrunnelse i at det også tidligere er lagt ned ressurser på dette. Udir anerkjenner at valget om å løse oppdraget skrittvis, gir en mer åpen prosess

med både utfordringer og styrker. I oversendelsesbrevet til leveransen løfter vi behovet for å arbeide tettere med organisasjonene, og et behov for å forankre eventuelle konsekvenser for fremdriften i arbeidet. Det krever tid og ressurser å skape enighet om veivalgene.

9.3 Plan for videre arbeid

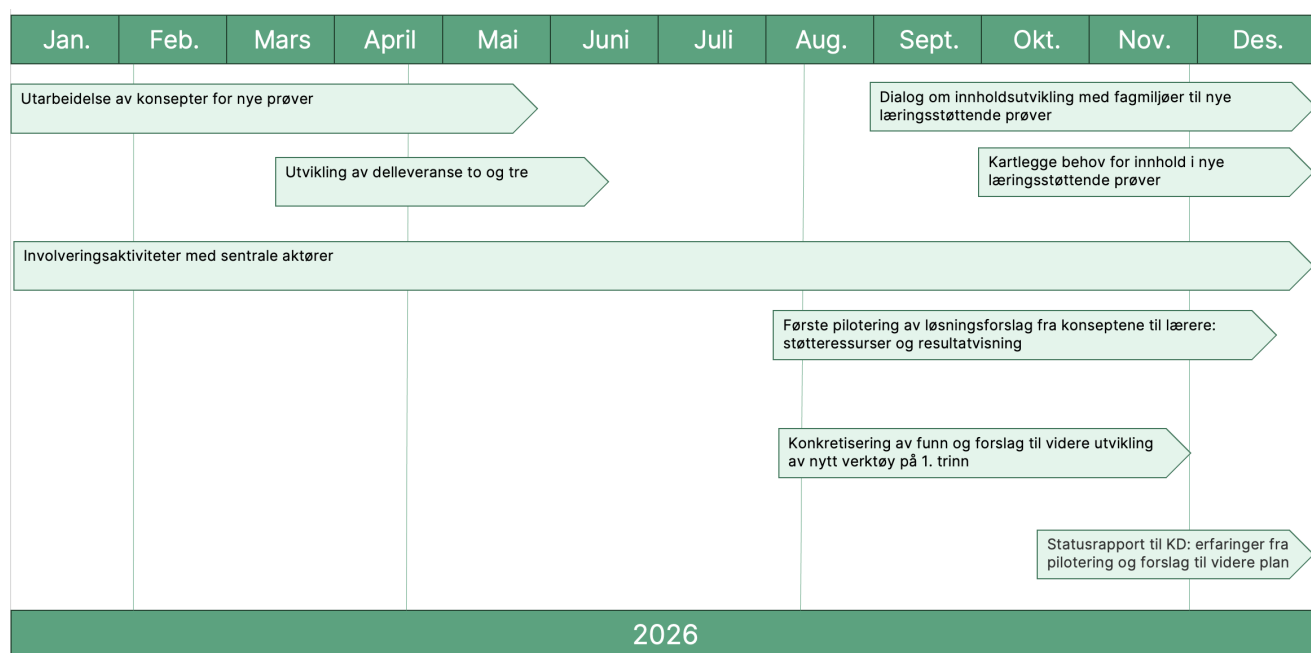
Udir skal sørge for at nye prøver og verktøy kan inngå i et helhetlig og systematisk arbeid med kvalitetsutvikling i skolen. Sammen med andre kilder skal de bidra til å gi et bredere grunnlag for dialog om kvalitet i skolen. 1. september oversender vi forslag til utforming av ny skolemiljøundersøkelse. Vi vil samordne utviklingen av ny skolemiljøundersøkelse og utvikling av nye prøver der det er hensiktsmessig, for å gi skolene bedre støtte til å ivareta viktige sammenhenger mellom trivsel og læring. Et slikt helhetlig arbeid med kvalitetsutvikling vil også inkludere kompetanseutviklingstiltak.

I den videre prosessen ønsker vi å gi mulighet for at skoler og andre interesserte kan gi tilbakemelding på forslagene til utforming av nye prøver. Dette vil styrke forankringen av konseptene, kan gi nye forslag og innspill av betydning for utviklingsarbeidet. Udir vil vurdere tidspunkt for en bredere involveringsprosess, og om det er lettest for skolene å ta stilling til konseptene slik de foreligger nå eller om de bør utvikles noe mer før de deler med et større publikum.

Forslagene fra arbeidsgruppene blir nå bearbeidet. Med utgangspunkt i forslagene, vil vi utarbeide planer for det videre arbeidet med kortsiktige og langsiktige mål. Noen utviklingsspor lar seg teste og pilotere raskt. Andre oppgaver vil kreve ytterligere utredning og innhenting av kunnskap før beslutning om endelig løsning kan tas.

Udir vil involvere UH-miljøer i arbeidet med faglige konstrukter, prøveformer, analyser og støtteressurser. Det er behov for å både utrede valg knyttet til konstrukt, ambisjonen om å kartlegge delferdigheter, og konsekvenser av dette. I disse drøftingene vil vi fortsatt legge vekt på bruksverdien for profesjonsfelleskapet, samtidig som vi skal levere prøver med god faglig kvalitet.

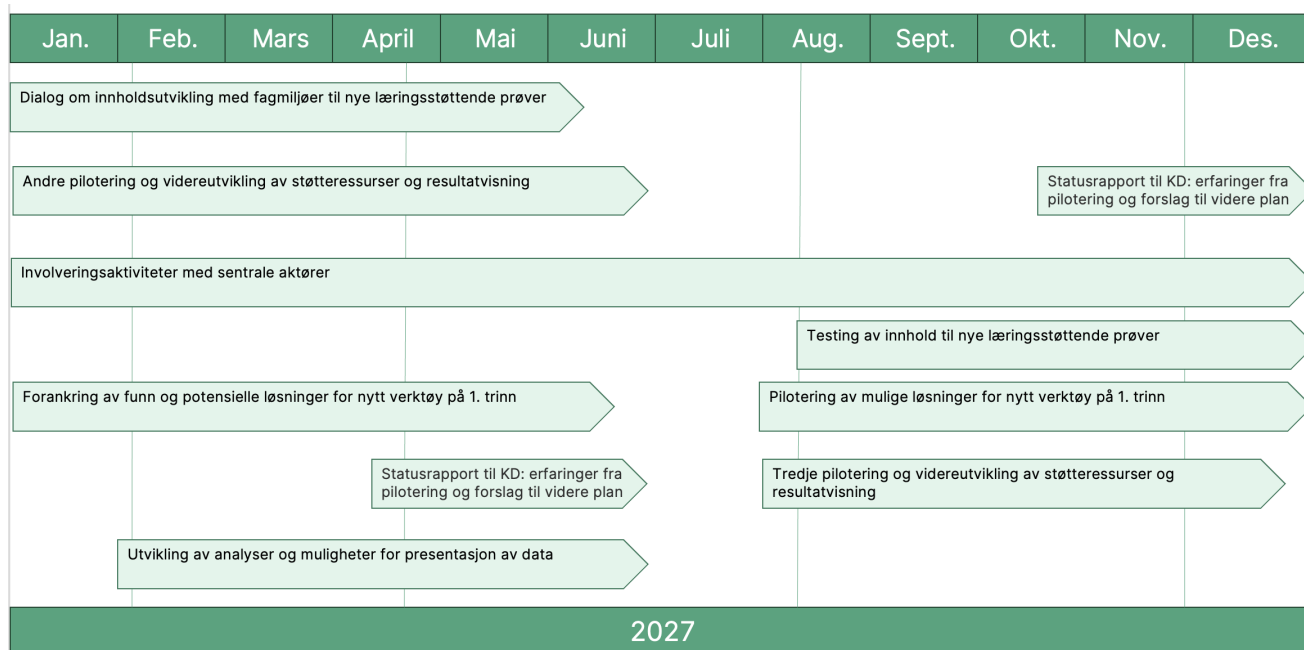
På kort sikt vil Udir starte pilotering av en ny resultatvisning, forbedre støtteressurser og videreutvikle konsepter og innhold i nye prøver for å best ivareta det læringsstøttende formålet. Udir vil ta utgangspunkt i konseptene fra lærergruppenes arbeid, men vil fortsette den smidige tilnærmingen og bearbeide konseptene gjennom brukerinvolvering.



Figur 14: illustrasjon av aktiviteter på kort sikt. Listen er ikke uttømmende.

Det tar to år å lage dagens prøver ut fra dagens kvalitetskrav til prøveutvikling. Hvis vi legger til grunn at fremtidige prøver også vil ta to år å lage og Udir anskaffer aktuelle prøvemiljøer for innholdsutvikling i løpet av 2027, vil første mulighet for en fullskala gjennomføring av nye prøver med læringsstøttende formål være skoleåret 2029/2030. Vi vil understreke at vi mener det vil være aktuelt å prøve ut og skalere opp nye løsninger før dette for å skaffe erfaringer med disse løsningene.

Endelig tidspunkt for fullskalagjennomføring av nye læringsstøttende prøver må avklares i dialog med departementet. Det er flere momenter som vil virke inn på dette, slik som eksempelvis it-systemer, juridiske, organisatorisk m.m.



Figur 15: illustrasjon av aktuelle aktiviteter på lengre sikt. Listen er ikke uttømmende.

Oppfølging av leveransen og innspill fra organisasjonene

Vi har lest organisasjonenes kommentarer med stor interesse, og vi er glade for at de støtter hovedretningen for arbeidet. Organisasjonen peker samtidig på at det i kommende leveranser vil være behov for å undersøke nærmere hva som er skolelederes og skoleeieres behov fra nye læringsstøttende prøver og hvordan disse kan ivaretas i det nye prøvekonseptet. Organisasjonene løfter fram viktige perspektiver i sine innspill, og vi vil følge opp disse i det videre arbeidet. Udir anerkjenner at vi må bearbeide de ulike forslagene mer, og at partene må få konkrete anbefalinger som de kan ta stilling til. Dette vil vi å komme tilbake til i referansegruppen for vurdering høsten 2026.

For å ivareta partenes muligheter til ytterligere medbestemmelse i prosessen med utviklingen av nye prøver har Udir satt opp en egen møteserie med ledelsen i direktoratet og ledelsen for organisasjonene i utdanningssektoren. Totalt har det blitt gjennomført tre møter og ett seminar med denne gruppen. Tilbakemeldingene fra møtene har vært positive. Udir ser samtidig et behov for å involvere organisasjonene i utdanningssektoren tettere fremover, og at vi både legger til rette for overordnede diskusjoner om retningsvalg og drøfting av konkrete forslag. Vi varsler i oversendelsesbrevet at vi ønsker å forankre med KD hvordan vi kan legge til rette for god involvering samtidig som vi har god fremdrift i arbeidet.

Kilder

- Afflerbach, P., Cho, B.-Y., Kim, J.-Y., Crassas, M. E., & Doyle, B. (2013). Reading: What else matters besides strategies and skills? *The Reading Teacher*, 66(6), 440–448.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Arnesen, A., Braeken, J., Ogden, T., & Melby-Lervåg, M. (2019). Assessing children's social functioning and reading proficiency: A systematic review of the quality of educational assessment instruments used in Norwegian elementary schools. *Scandinavian Journal of Educational Research*, 63(3), 465–490.
<https://doi.org/10.1080/00313831.2017.1420685>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
<https://doi.org/10.1080/0969595980050102>
- Brevik, L. M., & Helness, H. L. (2018). Engelsk VG1: Nye læringsstøttende prøver. *Bedre Skole*, 30(2), 48–53.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57–66.
<https://doi.org/10.1080/14703290601081332>
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. The Guilford Press.
- de Lange, J. (1995). Assessment: No change without problems. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment* (pp. 87–172).
- Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., & Moriarty Gerrard, M. O. (2007). Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology*, 92(2), 373–385.
<https://doi.org/10.1037/0021-9010.92.2.373>
- Hodgen, J., & Wiliam, D. (2006). *Mathematics inside the black box: Assessment for learning in the mathematics classroom*. GL Assessment.

- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). Springer. <https://doi.org/10.1007/978-1-4939-0317-7>
- McLaughlin, M. (2012). Reading comprehension: What every teacher needs to know. *The Reading Teacher*, 65(7), 432–440.
- Meld. St. 34 (2023-2024). *En mer praktisk skole. Bedre læring, motivasjon og trivsel på 5. – 10. trinn*. Kunnskapsdepartementet. <https://www.regjeringen.no/contentassets/dcc936e85f4c43bf90ecf0bafabb00a8/no/pdfs/stm202320240034000dddpdfs.pdf>
- Messick, S. (1987). Validity. ETS Research Report Series, 1987(1), i–208. <https://doi.org/10.1002/j.2330-8516.1987.tb00244.x>
- Nordenbo, S. E., Allerup, P., Andersen, H. L., Dolin, J., Korp, H., Larsen, M. S., Olsen, R. V., Svendsen, M. M., Tiftikçi, N., Wendt, R. E., & Østergaard, S. (2009). Pædagogisk brug af test: Et systematisk review. Danmarks Pædagogiske Universitetsforlag.
- NOU 2023: 27. (2023). *Et nytt system for kvalitetsutvikling – for elevens læring og trivsel*. Kunnskapsdepartementet. <https://www.regjeringen.no/contentassets/85cc09d37f604868b31cd0da68aaa200/no/pdfs/nou202320230027000dddpdfs.pdf>
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into practice*, 48(1).
- Ræder, H. G., Andersson, B., & Olsen, R. V. (2022). Numeracy across grades: Vertically scaling the Norwegian national numeracy tests. *Assessment in Education: Principles, Policy & Practice*, 29(6), 653–673. <https://doi.org/10.1080/0969594X.2022.2147483>
- Speitz, H., & Myklevold, G. A. (2022). The subject of English in LK20: A catalyst for multilingualism? *Acta Didactica Norden*, 16(2), Article 5. <https://doi.org/10.5617/adno.9048>
- Utdanningsdirektoratet. (2026). *Del 1: Kunnskapsoppsummering om arbeidet med tidlig innsats, inkludering og tilrettelegging for barn og elever med særskilte behov – fem år etter Meld. St. 6 (2019–2020) Tett på – tidlig innsats og inkluderende fellesskap i barnehage, skole og SFO*. <https://www.udir.no/contentassets/acbd1c159c6c4aa78fc367948a81fe39/delrapp-ort-1-kunnskapsoppsummering-ob-2025-01-del-1.pdf>
- Utdanningsdirektoratet. (2020). *Prøver i fagfornyelsen – hvilke behov skal prøvene dekke?* <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/prover-i-fagfornyelsen/>

Yan, D., Lewis, C., & von Davier, A. A. (2014). Overview of computerized multistage tests. In D. Yan, A. A. von Davier, & C. Lewis (Eds.), *Computerized multistage testing: Theory and applications* (pp. 3–20). CRC Press.

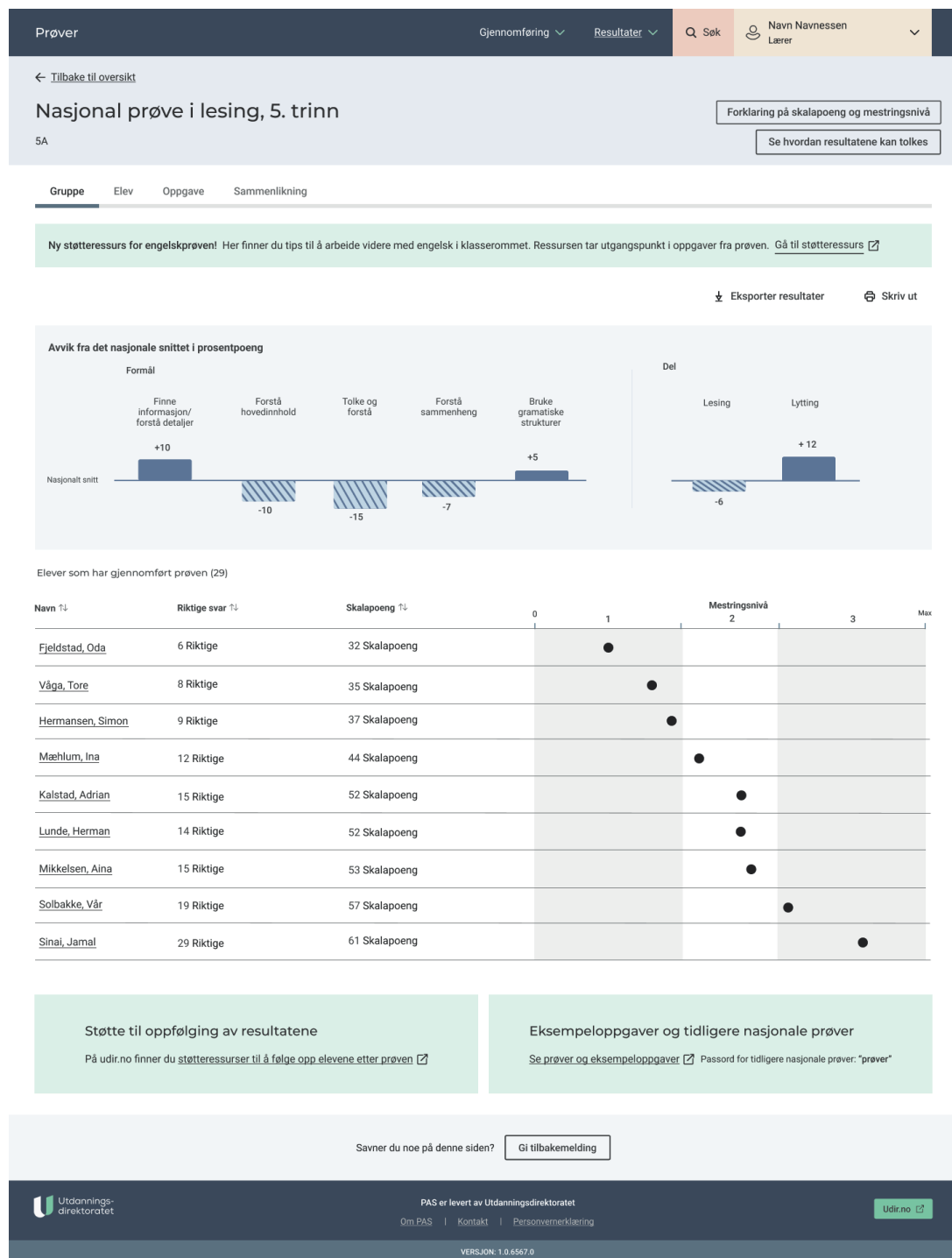
Walgermo, B. R., Uppstad, P. H., & Lundetræ, K. (2025). Den utfordrende, men nødvendige vurderingen av lesing de første årene på skolen. *Spesialpedagogikk*, 64–75.

Vedlegg

Vedlegg 1: Eksempler på hvordan resultater blir presentert i dagens løsning

Gruppevisning

Gruppevisning av prøven (fiktive navn). Viser oversikt over elevenes mestringsnivå, skalapoeng og antall riktige svar, samt gruppens resultater sammenlignet med nasjonalt snitt i prøvens delferdigheter.



Elevvisning

Elevvisning av prøven. Viser hvordan eleven har skåret på ulike oppgaver, hvilket mestringsnivå eleven ligger på, og tilhørende nivåbeskrivelse.

Prøver

Gjennomføring ▼Resultater ▼

Q Søk

Navn Navnessen
Lærer ▼

[← Tilbake til oversikt](#)

Nasjonal prøve i engelsk, 8. trinn

8A

Forklaring på skalapoeng og mestringsnivå

Se hvordan resultatene kan tolkes

GruppeElevOppgaveSammenlikning

Velg elev å se resultater for

Elev Elevsen ▼

[Eksporter resultater](#)

[Skriv ut](#)

Mestringsnivå 3 (43-55 skalapoeng) | Elevens skalapoeng: 52

Tid brukt: 60 min

Generelt tilsier dette nivået at eleven kan:

- forstå en del mindre vanlige ord og uttrykk
- finne / kjenne igjen informasjon i en del komplekse setninger og i tilpassede tekster av ulike lengder
- forstå hovedinnholdet i en tilpasset tekst
- finne / forstå informasjon i skriftlige og muntlige tekster selv når det finnes konkurrerende informasjon
- koble informasjon fra ulike deler av en tekst
- bruke konteksten for å forstå vanskelige deler av en tekst
- tolke / ha en overordnet forståelse av en tekst og finne svaret selv om det ikke er eksplisitt uttrykt i teksten
- plassere informasjon fra en tekst i riktig rekkefølge
- kjenne igjen og bruke grunnleggende grammatiske strukturer og funksjonsord i en kontekst

Oppgaver fordelt på delferdighet og mestringsnivå

Klikk på oppgavene for å se en besvarelse

Delferdighet	Mestringsnivå 1	Mestringsnivå 2	Mestringsnivå 3	Mestringsnivå 4	Mestringsnivå 5
Finne informasjon/forstå detaljer	1	2	15, 16, 27	3, 14, 26	
Tolke og forstå	4, 19	5, 6	8, 25, 28	29	12
Forstå hovedinnhold		17, 18	7, 9, 20	10, 21	11
Forstå sammenheng				33	13
Forstå avansert vokabular			22, 23, 24		
Bruke gramatiske strukturer			30, 32	31	

☒ Oppgave 1 - 14 måler lytting, oppgave 15 - 33 måler lesing

☒ Riktig svar ☒ Delvis riktig ☒ Feil svar ☒ Ikke besvart ☒ Ankeroppgave

Prosentandel riktig og feil svar på lesing og lytting

Lytting

23 %

77 %

Lesing

48 %

52%

☒ Prosentandel riktige svar ☐ Prosentandel feil svar eller ikke besvart

Savner du noe på denne siden?

Gi tilbakemelding

Utdanningsdirektoratet

PAS er levert av Utdanningsdirektoratet

[Om PAS](#) | [Kontakt](#) | [Personvernerklæring](#) | [Tilgjengelighetserklæring](#)

Udir.no

VERSJON: 1.0.6567.0

Oppgaveoversikt

Oppgaveoversikt i prøven. Viser mer informasjon om hver oppgave, inkludert gruppens løsningsprosent og sammenlikning med nasjonalt snitt.

Prøver
Gjennomføring
Resultater
Søk
Navn Navnessen
Lærer

Tilbake til oversikt

Nasjonal prøve i regning, 5. trinn
5A

Forklaring på skalapoeng og mestringsnivå
Se hvordan resultatene kan tolkes

Gruppe
Elev
Oppgave
Sammenlikning

Eksporter resultater
Skrive ut

Oppgaveoversikt for 5A

Oppgave	Gruppens løsningsprosent	Nasjonal løsningsprosent	Avvik	Mestringsnivå	Ferdighetsområde	Fag
1	89%	97%	-8%	1	Bruke og bearbeide	Kro, Mat
2	79%	75%	4%	1	Bruke og bearbeide	Mat
3	32%	64%	-32%	1	Reflektere og vurdere	Mat, Sam
4	89%	64%	25%	1	Gjenkjenne og beskrive	Mat
5	5%	23%	-18%	3	Bruke og bearbeide	Mat
6	68%	73%	-5%	1	Bruke og bearbeide	Mat
7	32%	28%	4%	3	Bruke og bearbeide	Mat

Utdanningsdirektoratet
PAS er levert av Utdanningsdirektoratet
Om PAS | Kontakt | Personvernerklæring | Tilgjengelighetserklæring
Udir.no

VERSION: 1.0.6567.0

Sammenlikning

Sammenlikning av resultater. Viser hvordan gruppen fordeler seg på mestringsnivåene sammenlignet med andre grupper på skolen, samt på kommune-, fylkes- og nasjonalt nivå.

Prøver
Påmelding
Gjennomføring
Resultater
Søk
Navn Navnessen
Skoleadministrator

Tilbake til oversikt

Nasjonal prøve i regning, 8. trinn
8C

Forklaring på skalapoeng og mestringsnivå
Se hvordan resultatene kan tolkes

Gruppe
Elev
Oppgave
Sammenlikning

Elevenes mestringsnivå i prosent

Tabellvisning

Hele landet (52092 elever)	A.13%	B.13%	C.38%	D.19%	E.13%
Fylket (6399 elever)	A.12%	B.20%	C.38%	D.21%	E.9%
Kommunen (631 elever)	A.16%	B.26%	C.38%	D.16%	E.4%
Ungdomsskolen (70 elever)	A.10%	B.28%	C.38%	D.17%	E.7%
8A (24 elever)	A.13%	B.21%	C.46%	D.13%	E.8%
8B (24 elever)	B.38%	C.38%	D.21%	E.4%	
8C (22 elever)	A.19%	B.24%	C.29%	D.19%	E.10%

A Mestringsnivå 1
B Mestringsnivå 2
C Mestringsnivå 3
D Mestringsnivå 4
E Mestringsnivå 5

Utdanningsdirektoratet
PAS er levert av Utdanningsdirektoratet
Om PAS | Kontakt | Personvernerklæring | Tilgjengelighetserklæring
Udir.no

VERSION: 1.0.6567.0

Vedlegg 2: Innspill fra Statped til delleveranse 2

Innspill fra Statped på delleveranse 2 og 3 - oppdrag 2025-018

Innspill på delleveranse 2 og 3 - oppdrag 2025-018

Statped har i dag som oppgave å bistå Udir med tilrettelegging av obligatoriske prøver og eksamen til tegnspråklige elever. I denne oppgaven ligger også tilrettelegging av kartleggingsprøver i lesing og regning.

Vi ser at den tilrettelagte kartleggingsprøve i lesing ikke kartlegger det prøven har som formål å kartlegge. Frem til nå har Statpeds tilrettelegging av kartleggingsprøver i lesing blitt gjort ved at både det muntlige og det skriftlige innholdet, blir presentert på tegnspråk, eleven svarer med å skrive. Når eleven avleser tegnspråk og svarer ved å klikke på skriftlige norske ord/skrive norsk, avkoder eleven to språk; tegnspråk og norsk. Når man skal lese ut resultatene av kartleggingen, vil det da være vanskelig å vite hvilket språk barnet eventuelt strever med. Å kartlegge ett språk ved hjelp av et annet, gir svært usikre resultater. Ett eksempel er en lytteoppgave hvor eleven får et lydspor som sier: «slange – skriv slange». Den tegnspråklige eleven vil da se tegnet for slange med beskjed om å skrive slange. Eleven testes da i om hen oppfatter tegnet for slange og må deretter oversette tegnet til norsk og skrive det norske ordet; slange. Den hørende eleven blir testen i å lytte ut og skrive lyder. Den hørende eleven og den tegnspråklige eleven testes dermed i to ulike ferdigheter og det stilles derfor spørsmål om den tegnspråklige tilretteleggingen er relevant i forhold til kartleggingsprøvens formål.

For de tegnspråklige elevene kan vi se for oss to ulike løsninger:

1. All kartlegging baseres på lesing av skrift, mens oppgaveinstruksene gis på tegnspråk.
2. Tegnspråklige barn med hørselsrest kan gjennomføre kartleggingen med tale, men ha tilgang til tegnspråklige oppgaveinstrukser.

Det er behov for å se nærmere på om disse to løsningene vil svare ut formålet med kartleggingen. På Statpeds læremiddelseminar våren 2023 var kartleggingsprøver satt på agendaen. Statped mottok da tydelige tilbakemeldinger på at kartleggingsprøven ikke ble brukt blant de tegnspråklige elevene og at det var støtte for Statpeds syn på utfordringer som ligger i dagens tilrettelegging. Det ansees derfor som nødvendig å involvere praksisfeltet for å finne gode løsninger for de nye læringsstøttende prøvene for tegnspråklige i lesing for å sikre kartlegging av de ferdigheter det er ment å kartlegge. Det samme vil gjelde de nye læringsstøttende prøvene i engelsk da den prøves per i dag baserer seg på lyd, noe de nye prøvene også kanskje vil gjøre?

Det er videre et ønske om å utvikle læringsstøttende prøver for avlesing av tegnspråk. I delleveranse 3 sier Lesegruppen at konseptet er forankret i internasjonal forskning på lese teori, motivasjonsteori og lesevurdering. Det sies også at konseptforslag for de

samiske prøvene presenteres i eget kapittel. For å utvikle leseprøver, engelskprøver og avlesingsprøver for tegnspråklige, bør tilsvarende konseptforslag utarbeides på bakgrunn av forskning på tegnspråklige elevers læring av skriftlig språk som ikke baseres på auditiv tilnærming, og på teori om tilegnelse og avlesing av norsk tegnspråk. Her bør aktuelle fagmiljø som Statped, OsloMet og Språkrådet kobles på i samarbeid med lærere med kompetanse og erfaring i undervisning av tegnspråklige elever i de aktuelle fagene. Statpeds øvrige brukergrupper bør også involveres i testfase av de nye prøvene.

Det bør også tas hensyn til tidsbegrensning i de nye prøvene, da dette kan påvirke resultater om tid er en faktor og elevene kan komme ut med dårlig skår pga tid og ikke faglig nivå. Hvis tidsfaktoren er utslagsgivende for hvor elevene havner i de blokkadaptive prøvene, så bør dette også hensyntas. Blinde og sterkt svaksynte elever vil ofte bruke mer tid på å komme gjennom tekster ettersom de er avhengige av å bruke hjelpemiddelteknologi og bruker mer tid på lesing generelt. Lang tidsbruk betyr ikke nødvendigvis faglig usikkerhet. Statped har tidligere meldt behov for en egen kartleggingsprøve i lesing av punktskrift, og det kan være hensiktsmessig å vurdere også i denne sammenhengen. Å lære å lese taktilt er metodisk annerledes og krever spesialkompetanse og tett oppfølging. I dag savnes en standardisert kartlegging i lesing av punktskrift og gode målinger av progresjon og støtte til videre oppfølging.

Når det gjelder kartlegging av matematikk så nevnes universell utforming. Her er ønsket Statped å påpeke at det utover universell utforming også bør sjekkes hvorvidt oppgavene er pedagogisk relevante for de forskjellige elevgruppene. Selv om en oppgave er universelt utformet, så er det ingen garanti for at den er pedagogisk relevant for en blind elev. Her kan Statped bistå med videre kompetanse og testing underveis.