

Rolf Vegar Olsen

Universitetet i Oslo

Sverre Tveit

Universitetet i Agder

Julius Kristjan Björnsson

Universitetet i Oslo

DOI: <http://dx.doi.org/10.5617/adno.6647>

Redaksjonelt

Temanummer om nasjonale prøver og eksamener i norsk og svensk grunnopplæring

Elever i de fleste land har regelmessig prøver som lærerne bruker til å få informasjon om elevenes læring. Dette er gjerne prøver lærerne selv lager, og som de retter og tolker resultatene fra. Slike prøver gir lærerne informasjon som brukes i oppfølgingen av elevene, i arbeidet med evaluering og planlegging av undervisningen og som grunnlag for sluttvurdering i fag. Dette temanummeret handler ikke om prøver som lærerne selv har ansvar for, men er i sin helhet viet prøver som i sin utforming og gjennomføring har en nasjonal innramming. Flere av artiklene har likevel stor relevans for lærere som ønsker mer innsikt i utfordringer knyttet til det å prøve, måle og beskrive elevenes læring.

I dette temanummeret retter vi altså oppmerksomheten mot prøver som kan betraktes som eksterne, sentralgitte og/eller standardiserte. Siden årtusenskiftet har slike prøver økt i både antall og betydning i norsk og svensk grunnopplæring. Temanummerets artikler som belyser den norske konteksten, inkluderer *eksamensprøver, nasjonale prøver og kartleggingsprøver. Karakterstøttende prøver og læringsstøttende prøver* omfattes også av denne vide forståelsen av prøver, men er ikke i fokus i dette temanummeret. I Sverige omfatter en tilsvarende vid forståelse av begrepet «prov» framfor alt det vi kjenner som «*nationella prov*», og temanummeret belyser flere av disse prøvene og også «det *nationella provsystemet*» som helhet.

Selv om det er store forskjeller i disse prøvenes formål, innhold og utforming, er det store likhetstrekk mellom profesjonsfellesskapene som har vokst fram i begge land i arbeidet med å utvikle og kvalitetssikre prøvene. *Prøveutviklere* har gjerne ansvar for et bestemt skolefaglig område, mange har lærerbakgrunn eller har på annen måte fagdidaktisk kompetanse, og de tilhører gjerne fagmiljøer knyttet til akademiske institusjoner. Noen av disse fagmiljøene har lange tradisjoner og en sterk vitenskapelig forankring i universitetsmiljøer. Andre har vokst fram i takt med nasjonale myndigheters innføring av nye eller utvidelse av eksisterende prøvetyper de siste to tiårene. Sammenslåinger av

universitets- og høyskolemiljøer og etableringen av ulike *sentre* har også ført til endringer i prøveutvikleres formelle og vitenskapelige tilknytning.

Særlig prøveutviklernes fagspesialisering innebærer at de representerer litt ulike prøvetradisjoner. Dette kan være en styrke ettersom de enkelte skolefagene kan ha behov for spesifikke tilnærminger til hvordan kunnskaper og ferdigheter prøves, men innebærer samtidig at prøveutviklere mangler fellesarenaer hvor de kan møtes og dele erfaringer og perspektiver. Den typiske prøveutviklerens lærerbakgrunn gjør at mange av fagmiljøene har en sterk tilhørighet til lærerprofesjonen. Samtidig møtes de med helt andre faglige forventninger og normer som preger akademiske institusjoner og vitenskapelig praksiser. Dette spennet i prøveutviklernes profesjonelle forankring innebærer at de kan ses på som en profesjon som fortsatt er i etableringsfasen.

Den faglige spesialiseringen, kombinert med en rask vekst i antallet prøver, kan lede til idiosynkratiske og lukkede praksiser dersom det ikke parallelt etableres nye fellesarenaer for faglig utvikling. Det er en slik fellesarena vi ønsket å bidra til da vi tok initiativ til dette temanummeret. Nå som vi slutfører det redaksjonelle arbeidet med temanummeret, sitter vi igjen med et klart inntrykk av at en sterkere profesjonsidentitet er i ferd med å ta form på tvers av de fagspesifikke prøvetradisjonene og de akademiske institusjonene. Den gryende felles profesjonsidentiteten består blant annet i framveksten av felles normer og standarder knyttet til hva som kan sies å kjennetegne «gode» prøver. Profesjonen har i stor grad en felles verktøykasse som inneholder en rekke metoder som benyttes for å lage «gode» prøver. Prøveutvikleren besitter på denne måten en ganske kompleks kompetanse som i mangel av mer presise begreper kan sies å representere «the art, craft and science of test development». I og med at miljøene har tilhørighet til akademiske institusjoner, følger det med forventninger om at profesjonen skal være forskningsbasert, altså at «the science of...» skal være vektlagt i praksisen, noe vi håper dette temanummeret skal bidra til.

I både Sverige og Norge ser vi en slik sterkere forskningsorientert identitet i prøvemiljøene. I tillegg til å være drevet fram av prøveutviklernes akademiske tilhørighet, er denne dimensjonen av praksisen vektlagt av de tydelige og generiske rammeverkene som i dag finnes for de nasjonale prøvene i begge land. I disse rammeverkene vektlegges blant annet ideer og kunnskap fra testteori (psykometri), med begreper som reliabilitet og validitet i kjernen. Dette gir felles terminologi som er svært betydningsfull for dannelsen av en profesjon. Artiklene i dette temanummeret representerer gode eksempler på prøveutvikling som en vitenskapelig aktivitet, samtidig som de bidrar til å gi en mer helhetlig oversikt over rammene for utviklingen av prøvene.

I den norske konteksten innebærer framveksten av prøveutviklerprofesjonen at de tradisjonsrike eksamensprøvene kommer i et nytt lys og blir møtt med andre forventninger enn tidligere. Dette er prøver som til forskjell fra de nasjonale prøvene i Norge ikke utarbeides av prøveutviklere med tilknytning til viten-

skapelige fagmiljøer med testteoretisk bakgrunn. Eksamensprøvenes kvaliteter og legitimitet kjennetegnes ved at den forvaltes av lærerprofesjonen selv i et tett samspill med nasjonale myndigheter – Utdanningsdirektoratet – hvor mange sentrale prøveutviklere selv har lærerbakgrunn. Prøvene har også en lang tradisjon: Politikere, arbeidsgivere, foreldre (og besteforeldre) kjenner eksamen fra sin egen skoletid, og selv om prøvene har endret seg i både form og innhold, gir dette likevel enn «historisk kapital» og tyngde som bidrar til å gi prøvene tillit i samfunnet. Eksamensprøvene gjennomgår altså helt andre prosedyrer for kvalitetssikring, forvaltet av representanter for en lærerprofesjon som har faglig autoritet i kraft av undervisningserfaring og fagdidaktisk ekspertise. Parallelt med at arbeidet med dette temanummeret har vært i sluttfasen, har flere av artikkelforfatterne bidratt i et utvalg som skal gi råd om videre utvikling av eksamensprøvene og eksamenssystemet i arbeidet med *fagfornyelsen*. Flere av artiklene gir innsikter som kan være en ressurs i dette arbeidet, blant annet de mange artiklene som viser nytteverdien i å evaluere prøvers kvalitet fra et testteoretisk perspektiv.

Den svenske konteksten skiller seg fra den norske ved at «nationella prov» er fagprøver som brukes til karaktersetting, og dermed minner de mer om de norske eksamensprøvene enn «nasjonale prøver». Samtidig har de noen felles trekk med de nasjonale prøvene i Norge ved å være basert på testteoretiske prinsipper og ved å ha en viktig funksjon i styringen av utdanningssystemet. Med unntak for prøvene i det tredje årstrinnet inngår de svenske «nationella prov» i grunnlaget for lærernes karaktersetting. Samtidig skal prøvene gi kunnskap som grunnlag for kvalitetsutvikling både nasjonalt og lokalt, og det har vært et mål at prøvene skal være til hjelp for lærerne i den faglige oppfølgingen av elevene.

Denne bredden av formål som har kjennetegnet «nationella prov», representerer en stor utfordring for prøveutviklerne. De mange formålene har blant annet sin bakgrunn i at prøvene som tradisjonelt skulle støtte lærerne i karaktersettingen, etter hvert fikk en viktig funksjon i å kvalitetssikre skolenes og lærernes praksiser for karaktersetting – spesielt i kjølvannet av den sterke desentraliseringen av utdanningssystemet og framveksten av friskoler. Med de siste årenes reformer ble utvidelsen av prøvesystemet begrunnet både med behovet for ytterligere kvalitetssikring og med at prøvene kunne være en ressurs for lærernes formative vurderinger, i tråd med det økende fokuset på *vurdering for læring* («bedömning för lärande»). Samlet har dette ført til sprikende formål i prøvene. Dette har blitt kritisert av flere fagmiljøer, og har resultert i et omfattende pågående arbeid med å utvikle et nytt rammeverk for «nationella prov».

Bakgrunnen for temanummeret er de store forskjellene mellom prøvene i de to landene. Samtidig har prøveutviklerne og prøvesystemene også mange felles utfordringer. Det ligger dermed til rette for erfaringsdeling på tvers av landene, som kan komme til nytte for både prøveutviklere og nasjonale myndigheter med ansvar for prøvesystemene. De tre redaktørene har på ulike måter vært i kontakt

med prøveutviklingsmiljøer i Norge og Sverige i løpet av de siste årene, samt i arbeidet med eksamen i norsk grunnopplæring. Invitasjoner ble sendt ut til dette nettverket av kollegaer. Flere av artiklene er også et resultat av at temanummeret ble annonsert med en åpen invitasjon om å sende inn bidrag. Ambisjonen med temanummeret har vært at prøvemiljøene selv skulle få en anledning til å definere problemstillinger som de anser som viktige knyttet til prøver i sitt skolefaglige område. Til sammen gir derfor artiklene god innsikt i teoretiske perspektiver, historiske og politiske betingelser og empiriske beskrivelser av sentrale aspekter for flere av de prøvene som i dag er så viktige i norsk og svensk skole.

I det følgende presenterer vi kort noen av de perspektivene som flere av bidragene hver for seg og samlet belyser, men uten at vi refererer til de enkelte artiklene:

- *De overordnede rammene for nasjonale prøver eller eksamen i Norge og Sverige:* Flere av bidragene presenterer historiske, utdanningspolitiske og overordnede testteoretiske refleksjoner for prøvesystemene i våre to land. Selv om de fleste av prøvene er innført for å bidra til elevenes læring og/eller skolens og systemets kvalitetsutvikling, så har de bakgrunn i spesifikke politiske begrunnelser og behov. I dette ligger det også implisitte spenninger og teoretiske perspektiver for hvordan man kan sertifisere, styre og støtte læring og undervisning ved hjelp av prøver.
- *Likeverdige og rettferdige prøver:* Dette er et stort tema som spenner fra etiske refleksjoner til mer konkrete og spesifikke psykometriske analyser av oppgaver i prøver. Spesielt i engelskspråklig litteratur har «fairness» blitt løftet fram som et sentralt kjennetegn ved gode prøver. I korte trekk omhandler dette at prøvene skal fungere likt for alle prøvetakere: Forskjeller i skår skal kun reflektere at elever har ulik dyktighet – ingenting annet.
- *Pålitelige (reliable) prøveresultater:* Dette er et tema som berøres i mange av de presenterte bidragene. Spesielt er det mange bidrag som presenterer analyser av sensorreliabilitet for prøver hvor elevenes tekstlige produksjon skal vurderes. Det er godt kjent at den største feilkilden ved denne typen prøver er nettopp den (av nødvendighet) subjektive vurderingen som sensorer må foreta når større elevarbeider skal vurderes. Av den grunn blir det viktig å ha design for sensurering som åpner opp for at denne feilkilden blir synliggjort.
- *Standardsetting og lenking:* Det er økende oppmerksomhet knyttet til at prøveresultater presentert i form av et tall som reflekterer andel riktige svar (eller tilsvarende) ikke i seg selv er meningsbærende. Man trenger noe mer, en ytre sammenlikning eller standard som tallverdien kan knyttes til. På hver sine måter er både såkalt standardsetting og lenking av prøver tilnærminger som søker å gjøre nettopp dette.

- *Prøver med ulike formål:* Bidragene dekker prøver som hver for seg har ulike formål. De nasjonale prøvene i Sverige skal eksempelvis være et verktøy som skal hjelpe lærerne når de skal gi elevene karakterer i fag, kartleggingsprøvene i Norge skal hjelpe lærerne til tidlig å identifisere elever som kan ha behov for mer støtte og spesifikk hjelp, eksamen i norsk skole skal primært angi elevenes sluttkompetanse i fag, og flere av prøvene skal fungere som monitorering og støtte for skolens kvalitetsutvikling. Flere av prøvene har ofte også flere formål som skal tilfredsstilles på samme tid, noe som ikke er helt enkelt å få til.
- *Prøver i spesifikke faglige områder:* Til sammen dekker bidragene prøver i en rekke ulike skolefaglige områder (skrivning, lesing, regning, geografi, samfunnskunnskap og engelsk).

Til sammen gir dermed dette temanummeret en god oversikt over nasjonale prøver og eksamener i norsk og svensk skole. Det er vår oppfatning at det fra før av er et svært begrenset utvalg av gode og forskningsbaserte tekster som beskriver og drøfter temaer knyttet til prøver i skolen. Testteoretiske vitenskapsmiljøer og utdanningsforskning generelt er dessuten stadig mer internasjonalt orientert, og faglig utvikling foregår dermed i stor grad i engelsk språkdrakt. Dette kan komme i veien for god faglige kommunikasjon på morsmålet, både mellom prøveutviklere og mellom den framvoksende prøveutviklerprofesjonen og lærerprofesjonen.

En av ambisjonene vi har hatt med dette temanummeret, er at artiklene også skal kunne inngå som pensum i lærerutdanninger. Ettersom vurderingsterminologi kan være kompleks nok i seg selv, håper vi de norske og svenske artiklene kan bidra til at kommende lærere i de to landene får nye og mer tilgjengelige kilder til innsikt i dette fagfeltet. Når lærerstudenter kommer ut i yrket, vil det forventes at de skal kunne forholde seg kritisk, reflektert og konstruktivt til den informasjonen som de får (eller ikke får) om elevene sine når resultater fra prøvene blir rapportert.

De enkelte artiklene er likevel ikke utformet eksplisitt med dette formålet. De har undergått ordinær fagfellevurdering hvor vitenskapelig kvalitet har blitt vurdert og kommentert av (for forfatterne) anonyme personer. Vi retter derfor en spesiell takk til alle kollegaer som har gitt svært gode og kritiske tilbakemeldinger. Som redaktører har det vært svært tilfredsstillende å se hvordan fagfellenes tilbakemeldinger har bidratt til å forbedre alle artiklene betydelig.

Om redaktørene for dette temanummeret

Rolf Vegar Olsen er professor ved Universitet i Oslo. Hans forskningsinteresser omfatter blant annet hvordan vi kan måle elevenes kunnskaper, ferdigheter og holdninger i nasjonale og internasjonale storskalastudier.

Institusjonstilknytning: Centre for Educational Measurement, Universitetet i Oslo, Postboks 1161 Blindern, 0318 Oslo.

E-post: r.v.olsen@cemo.uio.no

Sverre Tveit er universitetslektor ved Universitet i Agder. Hans forskningsinteresser omfatter blant annet policyutvikling og legitimering, vurdering og evaluering i både høyere utdanning og grunnopplæringen.

Institusjonstilknytning: Institutt for pedagogikk, Universitetet i Agder, Postboks 422, 4604 Kristiansand.

E-post: sverre.tveit@uia.no

Julius Kristjan Björnsson er forsker ved Universitet i Oslo og leder av EKVA – Enhet for kvantitative utdanningsanalyser. Hans forskningsinteresser omfatter blant annet psykometri og prøveutvikling, storskala utdanningsanalyser og internasjonale komparative studier.

Institusjonstilknytning: Institutt for lærerutdanning og skoleforskning, Universitetet i Oslo, Postboks 1099 Blindern, 0317 Oslo.

E-post: j.k.bjornsson@ils.uio.no