

## **Vedlegg 1:**

### **Design på effektstudie om videreutdanning av lærere: problemstillinger som drøftes i dialogen med tilbyderne**

Oppdragsgiver legger i utgangspunktet opp til at følgende problemstillinger skal drøftes i dialogen med tilbyderne. Tilbyderne oppfordres til å drøfte også andre problemstillinger dersom dette er nødvendig for å gi oppdragsgiver et godt beslutningsgrunnlag for en eventuell effektevaluering.

#### **1. UTFALLSVARIABLER/RESULTATMÅL**

Departementet som oppdragsgiver legger som premiss at målene for videreutdanning er å bidra til økt læringsutbytte for elevene i grunnskolen, eventuelt også at lærere forblir lengre i yrket som følge av tilbud om videreutdanning. Tilbyderne kan også foreslå andre relevante utfallsvariabler/resultatmål.

Med det som utgangspunkt må tilbyderne drøfte i hvilken grad eksisterende datakilder om læringsutbytte er tilstrekkelig for å få kunnskap om videreutdanningen har effekt på læringsutbyttet til elevene. Det er sannsynligvis mange ulike problemstillinger innenfor dette området, men som minimum bør følgende problemstillingene drøftes og avklares i dialogen med tilbyderne:

- Nasjonale prøver måler grunnleggende ferdigheter, mens standpunkt- og eksamenskarakterer måler kompetanse i undervisningsfag. I hvilken grad er nasjonale prøver egnet som datakilde for prosjektet gitt at videreutdanningen er innrettet mot undervisningsfag?
- Standpunkt- og eksamenskarakterer på 10. trinn består av en relativt grovkornet skala, også på individnivå. I hvilken grad er karakterstatistikk egnet til å fange opp effekter, særlig dersom effektene er kvantitativt små i størrelse?
- Standpunkt karakterer settes lokalt og kan dermed påvirkes av deltakerne i videreutdanningen. I hvilken grad er det sannsynlig at dette vil påvirke beregninger av eventuelle effekter av deltakelse videreutdanning?
- Selv om eksisterende datakilder er på individnivå kan det tenkes at elevene både har hatt lærere i det aktuelle undervisningsfaget som har søkt om videreutdanning (og enten deltatt i videreutdanning eller ikke) og lærere som ikke har søkt om videreutdanning. Tilbyderne bør drøfte i hvilken grad det er et metodisk problem at

utfallsvariabelen/resultatmålet kan tenkes å være påvirket av begge disse gruppene av lærere, og eventuelt hvordan dette er tenkt løst i designet.

- Tilbyderne må drøfte når effekten skal måles (trinn i opplæringsløpet), gitt tilgangen på eksisterende datakilder og deltakelsen av lærere i videreutdanningen. Her bør tilbyderne diskutere kortsiktige vs. langsiktige effekter. I den sammenheng bør også varigheten av måleperioden diskuteres, herunder hvor lenge eventuelle kontrollgrupper av lærere som ikke deltar i videreutdanning må bestå for at studien skal kunne avdekke både kortsiktige og langsiktige effekter, dvs. hvor lenge i tid må disse lærerne nektes tilbud om videreutdanning for at designet skal la seg gjennomføre.
- Tilbudet bør inneholde en redegjørelse av hvordan eventuelle registeropplysninger om lærerne kan koples til effektmålingen.

På bakgrunn av drøftingen av disse problemstillingen og eventuelt andre tilbyderne mener er relevante for saken, må tilbyderne drøfte hvorvidt det er nødvendig med andre datakilder som kan gi informasjon om måloppnåelse, for eksempel spesialkonstruerte elevtester. Dersom det anbefales utvikling av spesialkonstruerte tester eller bruk av tester utviklet for andre formål må drøftingen inneholde en vurdering av kostnader og praktiske utfordringer og eventuelle negative konsekvenser knyttet til implementering i skolene som omfattes av studien.

## **2. FORKLARINGSVARIABLE/INTERVENSJON**

Departementet legger til grunn en relativt vid forståelse av intervensjonen eller forklaringsvariabelen, og lar det være opp til tilbyderne å drøfte hvordan tiltaket bør avgrenses metodisk for at effektevalueringen skal la seg gjennomføre og danne grunnlaget for relevante analyser for politikktutviklingen.

Med det som utgangspunkt bør tilbyderne drøfte hvilke typer sammenligninger gir mest relevant kunnskap for politikktutviklingen, og hva som er sannsynlige effektstørrelser av de ulike avgrensningene ut fra eksisterende forskning. En ikke-uttømmende liste som bør/kan drøftes er:

- Tilbyderne bør drøfte hvordan og i hvilken grad den valgte avgrensningen av tiltaket skaper en tydelig kontrast til dagens situasjon:
  - Er det mest relevant å forstå tiltaket som en pedagogisk intervensjon som består av konkrete videreutdanningstilbud, for eksempel konkrete didaktiske ferdigheter eller innholdselementer i undervisningsfagene?
  - Er det mest relevant å forstå tiltaket som en type systemintervensjon som på sikt skal endre lærernes formelle fagspesifikke kompetanse (antall studiepoeng eller tilsvarende i undervisningsfagene), som enten oppnås gjennom grunnutdanningen eller senere videreutdanning?
- Dersom tiltaket forstås som en type systemintervensjon bør tilbyderne drøfte i hvilken grad effekter estimert ut fra deltakelse i ulike videreutdanningstilbud avspeiler en tilsvarende eller sammenlignbar intervensjon, dvs. i hvor stor grad er det variasjonen mellom de ulike videreutdanningstilbudene, hvorvidt dette er et problem for

effektestimater og hvordan dette eventuelt kan løses i forskningsdesignet. Det er mange universiteter og høyskoler som tilbyr videreutdanning og det finnes flere tilbud på samme institusjon. Mange av tilbudene har eksistert over lang tid.

- Tilsvarende bør tilbyderne drøfte hvordan implementeringen av intervensjonen kan påvirke utfallsvariabelen, særlig dersom det legges til grunn en forståelse av tiltaket som systemintervensjon. Overføringseffekter til elevenes læringsutbytte kan være kompliserte og forutsetter sannsynligvis stor intervensjonsstyrke. Dette innebærer en lang kausalkjede fra departementet stiller ressurser til rådighet for videreutdanning til elevene oppnår bedre læringsutbytte enn de ellers ville gjort, slik som om utdanningsinstitusjonene øker kapasiteten til videreutdanning, tilbudene har høy kvalitet, skolene frigjør tilstrekkelig med tid og lærerne er i stand til å ta til seg ny kunnskap og integrere den i egen undervisningspraksis. Tilbyderne bør drøfte i hvilken grad evalueringen kan eller bør gi kunnskap om prosessfaktorer, herunder om det er mulig å lage effektestimater for deltakelse i videreutdanningstilbud finansiert gjennom stipendmodell og vikarmodell.
- Uavhengig av innretning bør tilbyderne estimere effekter for spesifikke undervisningsfag. Tilbyderne bør dermed drøfte i hvilken grad det kan etableres effektestimater for alle undervisningsfag i grunnskolen, gitt tilbudene og deltakelsen i videreutdanningen. Dersom det må prioriteres mellom undervisningsfagene vil oppdragsgiver at matematikk og naturfag prioriteres høyest, deretter norsk og engelsk.
- I forskningslitteraturen er det fremholdt at effekten av fagspesifikk kompetanse for læringsutbytte er avtagende over en viss terskelverdi. Tilbyderne bør dermed drøfte hvordan en eventuell slik effekt kan la seg dokumentere på norske data, og om det er mulig å beregne omtrent hvor en slik terskelverdi befinner seg, for eksempel definert ut fra antall studiepoeng i undervisningsfag, og om den varierer etter trinn i grunnskolen, for eksempel om den er forskjellig på barne- og ungdomstrinnet.
- Tilbyderne bør også kunne si noe om konsekvenser av frafall av deltakerne i tiltaksgruppen (at lærerne ikke fullfører utdanningstilbudet)
- Tilbyderne bør drøfte behovet for ”stabilitet i bakgrunnsbetingelsene”, for eksempel at lærere i kontrollgrupper eller de øvrige lærerne på skolene som enten tilhører eksperiment- eller kontrollgruppen ikke tilbys andre relevante tiltak, slik som videreutdanning finansiert av skoleeier selv eller privat. Dersom dette er et problem bør tilbyderne drøfte hvordan dette eventuelt kan løses, enten praktisk eller metodisk, jf. også punkt om varigheten på målingene av utfallsvariabler.

### **3. METODISK DESIGN**

Departementet som oppdragsgiver legger som premiss at det metodiske designet kan identifisere kausale effekter på det høyest mulige evidensnivået, at gjeldende måltall for deltakelse i videreutdanning ikke må reduseres og at enkeltkommuner ikke må bli tvunget til å vente med kompetanseheving på grunn av forskningsdesignet.

Med det som utgangspunkt bør tilbyderne drøfte og foreslå en konkret design som troverdig kan identifisere den kausale effekten av intervensjonen (forklaringsvariabelen) for elevenes læringsutbytte, eventuelt også lærernes tilbøyelighet til å forbli i yrket eller andre relevante utfallsvariabler.

Dersom forskningsdesignet forutsetter reell randomisering av utvalget som skal få tildelt studieplass til videreutdanningstilbud bør tilbyderne drøfte hvilke implikasjoner dette har for dagens system for tildeling av studieplasser. Dersom designet er basert på et premiss om tilfeldig tildeling (dvs. et slags lotteri) av studieplasser ved oversøking må tilbyderne drøfte og beskrive tydelig hvordan tildeling av studieplasser skal foregå, hvor stor oversøking i hvert undervisningsfag designet forutsetter og konsekvenser for evalueringen dersom det ett eller flere år ikke er oversøking til videreutdanning. Tilbyderne som legger til grunn reell randomisering må også drøfte disse problemstillingene:

- I dagens system for tildeling av studieplasser til videreutdanning, søker først lærere til skoleeier, deretter godkjenner skoleeiere et utvalg av disse søknadene og tilslutt fordeler Utdanningsdirektoratet de godkjente søkerne til eksisterende studieplasser etter ett sett med kriterier, jf. vedlagte bakgrunnsdokumenter. Tilbyderne bør drøfte hvordan man kan oppnå best mulig randomisering, og på hvilket nivå randomiseringen bør/kan foregå (lærer, skole, kommunenivå).
- Tilbyderne bør drøfte om randomiseringen må gjøres for alle søkerne eller om det kan etableres en type underutvalg mens resten av søkerne kan tilbys videreutdanning etter dagens system for tildeling av studieplasser, for eksempel om det kan gjennomføres randomisering innenfor en avgrenset region/utvalg av kommuner.
- Tilbyderne bør drøfte om tildelingen av studieplasser delvis kan styres av andre kriterier enn tilfeldighet, for eksempel ved at skoler/lærere med visse kjennetegn gis høyere sannsynlighet for deltakelse i eksperimentgruppen i randomiseringen. Tilbyderne bør drøfte hvor store eksperiment- og kontrollgruppene må være for avdekke presise kausale effekter, gitt at effektene som minimum skal estimeres for spesifikke undervisningsfag, og hvorvidt dette er realistisk gitt deltakelsen på de ulike tilbudene om videreutdanning.

Tilbyderne oppfordres til å drøfte eventuelle metodiske design som ikke forutsetter reell randomisering av tildelingen av studieplasser. I så tilfelle bør tilbyderne drøfte i hvilken grad slike metodiske design øker risikoen for mer upresise og usikre effektestimater, og hvordan dette eventuelt kan reduseres metodisk eller på andre måter.

#### **4. ANALYSER**

Tilbyderne må drøfte og foreslå aktuelle rammeverk for effektstørrelse som effektestimaterne kan fortolkes innenfor. For oppdragsgiver er det viktig å få en forståelse for hvor substansielt betydningsfulle effektene som eventuelt dokumenteres er for elevenes læringsutbytte, og hva som er et rimelig sammenligningsgrunnlag for effektstørrelse, dvs. sammenlignbare tiltaks effektstørrelser.

Tilbyderne skal også drøfte hvorvidt og eventuelt hvordan effektestimatene kan danne grunnlaget for en beregning av samfunnsøkonomisk nytte av investeringer i videreutdanning.

Tilbyderne bør også drøfte hvorvidt det er mulig å analysere eventuelle ringvirkninger på skolenivå av deltakelse i videreutdanning for enkelte lærere, for eksempel om effekter for elever som ikke deltar i studien eller andre typer ringvirkninger.

## **5. SAMARBEID OG ORGANISERING**

For å gjennomføre en effektevaluering er samarbeid med lærere, skoleledere, skoleeiere og partene (Utdanningsforbundet, KS osv) avgjørende. Tilbyderne må drøfte og foreslå løsninger for hvordan dette samarbeidet bør organiseres for å sikre at effektevalueringen lar seg gjennomføre, særlig gjelder det for etableringen av kontrollgrupper.