

Forskningsstöd för kompetensutvecklingsmodellens effekter på undervisning och lärande

Nils Kirsten och Jannika Lindvall



Sammanfattning

Vi ger i denna rapport en översikt till forskning om effekter av lärares kompetensutveckling på undervisning och elevers lärande, samt värderar några kompetensutvecklingsmodeller och deras möjliga omsättning i nationell kompetensutveckling i Sverige.

Rapportens slutsatser och rekommendationer är sammanfattningsvis:

1. Det finns gott forskningsstöd för att högkvalitativ kompetensutveckling kan förbättra undervisning och elevers lärande, medan den kompetensutveckling som lärare vanligtvis deltar i sannolikt inte leder till sådana effekter.
2. Kompetensutveckling som förbättrar elevers lärande är med stor sannolikhet samhällsekonomiskt lönsam.
3. Vi lämnar tre huvudsakliga rekommendationer, oavsett vilken modell för lärares kompetensutveckling som väljs:
 - a. Lärare bör få stöd genom flera olika mekanismer eftersom det ökar sannolikheten att lärare på ett kvalitativt och varaktigt sätt omsätter nyvunna kunskaper till undervisningen.
 - b. Modellen bör bestå av fristående delar med relativt kort varaktighet eftersom insatser inte behöver ha lång varaktighet för att uppvisa effekter på undervisning och elevers lärande och insatser med kortare varaktighet är mer sparsamma med lärares tid och skolans resurser.
 - c. Modellen bör rekommendera¹ undervisningssätt med gott stöd i tillförlitlig forskning om vilka undervisningssätt som förbättrar elevers lärande (med hänsyn till faktorer såsom skolämne, årskurs och kontext) eftersom det annars är osannolikt att elevers lärande förbättras.
4. För att öka sannolikheten att nationell kompetensutveckling förbättrar undervisning och elevers lärande bör Skolverket:
 - a. Utveckla strukturer för att värdera forskningsstöd för kompetensutvecklingens innehåll, t ex vilka undervisningssätt som gynnar elevers lärande, med hänsyn till faktorer såsom skolämne, årskurs och kontext.
 - b. Använda effektvärdering inom ramen för försöksverksamhet som underlag för utvecklingsarbete innan insatser skalas upp nationellt, samt värdera effekter av uppskalade insatser.

Rapporten är författad av Nils Kirsten, universitetslektor i pedagogik vid Uppsala universitet, och Jannika Lindvall, universitetslektor i didaktik vid Mälardalens universitet.

Vi vill tacka ett flertal personer som har bidragit med förslag, kommentarer och frågor i olika skeden av arbetet med denna rapport: Sam Sims, Jennifer Gore, Andrew Miller, Andreas Ryve, Ola Helenius, Malin Tväråna, Pontus Bäckström, Iben Christiansen, Corrado Matta och Per Nilsson. Rapportens slutsatser och rekommendationer är dock helt våra egna.

¹ Här avses att en kompetensutvecklingsinsats ger rekommendationer, snarare än rekommendationer i författningsteknisk bemärkelse.

Innehåll

1. Inledning.....	4
2. Vad är en kompetensutvecklingsmodell?	4
3. Vad krävs för att kompetensutveckling ska främja elevers lärande?.....	5
4. Kriterier för tillförlitlighet i forskningen om kompetensutvecklingens effekter.....	6
5. Litteratursökning	8
6. Resultat i forskning om kompetensutvecklingens effekter på undervisning och lärande i grundskola och gymnasium.....	8
6.1 Effekter av kompetensutvecklingsprogram.....	8
6.2 Egenskaper med kausal påverkan på kompetensutvecklingens effekter.....	10
6.3 Analyser av moderatorer	11
7. Resultat i forskning om kompetensutvecklingens effekter på undervisning och elevutfall i förskola.....	13
7.1 Effekter av kompetensutvecklingsprogram.....	13
7.2 Analyser av moderatorer	14
8. Möjliga modeller för lärares kompetensutveckling	15
8.1 Videodemonstrationer	17
8.2 Kollegaobservationer med stöd av observationsprotokoll	17
8.3 Undervisningsmaterial	18
7.4 Coachning.....	18
9. Överväganden vid omsättning till nationell kompetensutveckling.....	19
9.1 Kostnadseffektivitet.....	19
9.2 Motsvara behov i många skolämnen och årskurser	22
9.3 Uppskalning	23
9.4 Individuellt deltagande eller samverkan mellan lärare	24
9.5 Organisatoriska förutsättningar som påverkar kompetensutvecklingens utfall	24
9.6 Insatsers varaktighet.....	25
9.7 Kompetensutvecklingsmodellens relation till professionsprogrammet	26
9.8 Strukturer för att värdera forskningsstöd för kompetensutvecklingsinsatsers innehåll	27
9.9 Skillnader mellan den föreslagna modellen och befintliga modeller för kompetensutveckling.....	28
10. Effektvärdering av kompetensutvecklingsinsatser.....	28
Referenser.....	31
Bilaga 1: Inkluderade forskningsöversikter avseende grundskola och gymnasium	39
Bilaga 2: Inkluderade forskningsöversikter avseende förskola.....	40

1. Inledning

Denna rapport är författad på uppdrag av Skolverket, med syftet att ta fram ett forskningsbaserat textunderlag om kompetensutvecklingsmodeller för lärare. Skolverkets arbete i frågan utgår i sin tur från ett regeringsuppdrag att ta fram en kompetensutvecklingsmodell för ämnesdidaktisk fördjupning (Utbildningsdepartementet, 2024).

Rapportens huvudfokus är att beskriva forskningsstödet för effekten av kompetensutveckling på undervisning och elevers lärande, inklusive mekanismer och förutsättningar som kan förklara och påverka sådana effekter. En begränsning i rapporten är därmed att kompetensutvecklingens effekter på andra typer av utfall i liten utsträckning berörs, vilket beror på att forskningen om effekter på andra typer av utfall är betydligt mer begränsad. Detta är förstås olyckligt eftersom effekter på exempelvis lärarsjukskrivningar, läraromsättning, samt elevers närvaro i skolan, hälsa, långsiktiga arbetsmarknadsutfall och politiska deltagande också är mycket betydelsefulla.² Detta faktum medför att det är viktigt att försöksverksamhet följer upp effekter bredare än på undervisning och elevers lärande.

I rapportens sista del diskuteras även överväganden kring hur slutsatser från tidigare forskning kan omsättas till nationell kompetensutveckling i Sverige.

För rapporten har avsatts 128 h arbete fördelat mellan två personer, vilket förstås begränsar möjligheten för detaljfördjupning. Det behövs således fortsatt utredning om hur rekommendationerna kan omsättas till praktik på ett sätt som förbättrar undervisning och elevers lärande.

2. Vad är en kompetensutvecklingsmodell?

Tidigare forskning har använt en rad olika begrepp för att kategorisera kompetensutveckling och dess egenskaper, såsom *modell* (A. Kennedy, 2005), *program*, *form* (Sims, Fletcher-Wood, O'Mara-Eves, m.fl., 2023), *angreppssätt* (Borko m.fl., 2010), *kännetecken* (Desimone, 2009), *strategier/metoder* (M. M. Kennedy, 2016), *komponenter* (Markussen-Brown m.fl., 2017), *innehållsfokus* (Lindvall m.fl., 2018) och *mekanismer* (Sims, Fletcher-Wood, O'Mara-Eves, m.fl., 2023). Det råder tyvärr inte konsensus kring innebörden av sådana begrepp inom kompetensutvecklingsforskning. För att åstadkomma tydlighet i denna rapport använder vi nedanstående definitioner.

- *Förutsättning* avser egenskaper som inte i sig orsakar förändring, men kan underlätta/försvåra påverkan, såsom insatsens tidpunkt, lokal, antal deltagare och medium.
- *Innehållsfokus* är de aspekter av undervisning och lärande som insatsen behandlar, vilket både kan omfatta ämnesinnehåll (såsom algebra eller vattnets kretslopp), ämnesövergripande innehåll (såsom motivation eller minne) och hur det ska omsättas i undervisningen.
- *Mekanism* avser i denna rapport en kausalt aktiv komponent i kompetensutvecklingsprogrammet, med andra ord en komponent som gör att programmet orsakar en effekt (Sims, Fletcher-Wood, O'Mara-Eves, m.fl., 2023). Mekanismer både orsakar en kausal effekt och förklarar hur den kausala effekten uppstod. Exempelvis är

² Det finns dock enstaka studier som indikerar positiva effekter av kompetensutveckling på läraromsättning och liknande utfall (Fletcher-Wood & Zuccollo, 2020). Som diskuteras i slutet av avsnitt 9.1 finner även många studier att positiva effekter på elevresultat medför positiva långsiktiga effekter avseende exempelvis arbetsmarknadsutfall, hälsa och politiskt deltagande.

aktiviteten modellering en mekanism om man finner att aktiviteten hjälper lärare att utveckla undervisningstekniker vilket i sin tur påverkar undervisningens kvalitet.

- *Program* avser en specifik kompetensutvecklingsinsats, som består av en viss sammansättning av mekanismer, förutsättningar och innehållsfokus, såsom matematiklyftet eller läsllyftet.
- *Modell* liknar begreppet program eftersom det avser en specifik sammansättning av mekanismer och förutsättningar, men är mer abstrakt då flera program som delvis liknar varandra kan kategoriseras i samma modell, exempelvis att både matematiklyftet och läsllyftet kan sägas utgå ifrån ”kollegialt lärande-modellen” (Kirsten & Carlbaum, 2020). Andra exempel på modeller är coachning och kollegaobservationer (se avsnitt 8). Eftersom program inte är identiska i alla aspekter behöver kategorisering i modeller utgå ifrån likheter i vissa egenskaper och bortse från olikheter i andra egenskaper.

3. Vad krävs för att kompetensutveckling ska främja elevers lärande?

En grundläggande utgångspunkt är att om kompetensutveckling ska förbättra elevers lärande behöver minst två aspekter hålla tillräckligt hög kvalitet: insatsens innehållsfokus och dess mekanismer och förutsättningar (metoder för att hjälpa lärare att förstå och omsätta innehållet till undervisningen). Dessa två aspekter har i kompetensutvecklingsforskning benämnts kompetensutvecklingens ”theory of instruction” (*undervisningsteori*) och ”theory of teacher change” (*lärarutvecklingsteori*) (Wayne m.fl., 2008). Man kan beskriva relationen mellan undervisningsteori och lärarutvecklingsteori som en ekvation:

$$\boxed{\text{Förbättrat lärande hos elever}} = \boxed{\text{Undervisningsteori}} \times \boxed{\text{Lärarutvecklingsteori}}$$

Figur 1. Samspelet mellan undervisnings- och lärarutvecklingsteori.

Enligt ekvationen måste både undervisningsteorin och lärarutvecklingsteorin vara korrekta för att elevers lärande ska förbättras och det räcker att en av teorierna är felaktig för att effekten på elevers lärande ska bli noll. Om lärare får utomordentligt stöd i att utveckla undervisningen i riktning mot ineffektiva undervisningssätt förbättras ju inte elevers lärande, liksom bristfälligt stöd i att omsätta effektiva undervisningssätt till praktiken inte leder till förbättrat lärande hos elever. Modellen är på flera sätt en förenkling. Exempelvis framgår inte att olika undervisningsteorier och lärarutvecklingsteorier kan ha olika gott förklaringsvärde i olika skolämnen, årskurser, skolkontexter, samt för nya och erfarna lärare. Det är också tänkbart att insatser påverkar elevers lärande genom andra mekanismer än förändrad undervisning, såsom att de minskar läraromsättning och sjukfrånvaro, vilket leder till en mer stabil skolsituation.

I denna rapport berör vi i första hand lärarutvecklingsteori (egenskaper hos kompetensutveckling som främjar förändrad undervisning). Det är dock högst centralt att även insatsernas innehållsfokus har starkt forskningsstöd. Vid utvecklingen av insatsernas innehåll bör exempelvis slutsatser om vilka undervisningssätt som förbättrar elevers lärande dras utifrån systematiska sammanställningar av forskning som tillförlitligt kan avgränsa kausala effekter av olika undervisningssätt på elevers lärande. Vi återkommer till detta under punkt 9.8.

4. Kriterier för tillförlitlighet i forskningen om kompetensutvecklingens effekter

Det huvudsakliga forskningsunderlaget för denna rapport utgörs av experimentella och kvasi-experimentella studier om effekterna av lärares kompetensutveckling på undervisning och elevers lärande. Anledningen till att forskningsunderlaget avgränsas till sådana studier är att de använder en design som förmår avgränsa kausala effekter, med andra ord att dessa kan visa huruvida kompetensutvecklingen påverkar undervisning och elevers lärande. Studier som enbart presenterar korrelationer mellan deltagande och utfallsmått eller baseras på deltagares uppfattningar om effekter exkluderas eftersom den typen av resultat inte kan användas för att dra slutsatser om kompetensutvecklingens påverkan på undervisning och elevers lärande. Det är dock viktigt att understryka att icke-experimentella studier fyller andra viktiga funktioner än att isolera kausala samband. De kan exempelvis visa utveckling över tid, kvalitativa skillnader mellan skolor, lärare och elever, samt användas som grund för att utveckla hypoteser om möjliga kausala samband som i sin tur kan testas i studier med en design som kan isolera kausala samband.

Den främsta metoden för att dra kausala slutsatser i studier av kompetensutveckling är det randomiserade kontrollerade försöket, där potentiella deltagare randomiseras till interventions- och kontrollgrupp. Därigenom undviker man att skillnader mellan deltagarna före insatsen påverkar utfallsmåtten på ett systematiskt sätt. Skillnaden i utfall mellan interventions- och kontrollgrupperna kan därför tolkas som en kombination av kausala effekter av insatsen och en statistisk osäkerhet, där omfattningen av den senare kan modelleras. I kvasi-experimentella studier saknas randomisering till interventions- och kontrollgrupp före insatsen och sådana studier delar istället in deltagarna i interventions- och kontrollgrupper med alternativa metoder som söker undvika att eventuella systematiska skillnader mellan grupperna före insatsen påverkar resultaten (se t ex kap 3 i Hägglund & Johansson, 2021, för en lättläst översikt till metoder i experimentella och kvasi-experimentella studier).

Eftersom även högkvalitativa studier med kausal design kännetecknas av osäkerhet behöver man sammanställa ett större antal enskilda studier för att möjliggöra säkrare slutsatser. Det kan ju exempelvis hända att randomiseringen av slumpen råkade fördela de skickligaste lärarna till försöksgruppen eller att något gick snett under studiens genomförande. För att inte sammanställningen av resultat från enskilda studier ska bli godtycklig används så kallade systematiska forskningsöversikter, där de söktermer, inkluderingskriterier och analysmetoder som används för att sammanställa enskilda studier specificeras. I systematiska forskningsöversikter som inkluderar meta-analys beräknas en genomsnittlig effekt utifrån de ingående studierna på ett sätt som viktat resultat från studier med bättre precision högre. Meta-analyser skiljer sig därmed från forskningsöversikter som saknar formella analysmetoder för att sammanställa resultaten och därför drar slutsatser på sätt som lämnar större utrymme för subjektiva tolkningar, såsom de narrativa översikterna av Timperley m fl (2007), Desimone (2009), Wei m fl (2009), Darling-Hammond m fl (2017), samt Hill och Papay (2022). Även om avsaknaden av formella analysmetoder i dessa översikter leder till att det blir svårare att värdera slutsatsernas forskningsstöd kan de ha andra styrkor, såsom att ägna större fokus åt att generera teori som kan testas i fortsatt forskning.

Även i gedigna systematiska forskningsöversikter med formella analysmetoder finns dock risk för snedvridna resultat. Ett skäl till detta är att incitamentstrukturer inom akademien missgynnar studier som visar nolleffekter. Studier som uppvisar nolleffekt är svårare att publicera i akademiska tidskrifter, vilket skapar drivkrafter för forskare att främst skriva artiklar utifrån försök som uppvisar tydliga effekter och att specificera analyser så att effekter

framträder. Det finns dock flera metoder för att motverka detta problem. Det är mycket vanligt att forskningsöversikter genomför analyser av så kallad ”publication bias”, där man gör skattningar av eventuella tendenser till ”överpublicering” av studier med tydliga effekter. Forskare uppmanas även att förregistrera studieupplägg för att motverka att analysmetod väljs utifrån analysresultaten. Vissa forskningsfinansiärer kräver också att studieresultat presenteras i enlighet med förregistrerade studieupplägg oavsett utfallet. Eftersom detta bör leda till att snedvridning på grund av incitamentsstrukturer minimeras kan man anta att publikationer baserade på sådana krav visar mer tillförlitliga resultat. Översikter indikerar också att studier med förregistrerad studiedesign visar mindre effekter på elevresultat, både inom utbildningsvetenskap i allmänhet (Kraft, 2020) och i studier av lärares kompetensutveckling (Sims, Fletcher-Wood, O’Mara-Eves, m.fl., 2023). Den senare översikten indikerade dock att skillnaden mellan förregistrerade och icke-förregistrerade studier delvis kan bero på att de förregistrerade studierna i högre utsträckning använde utfallsmått som har svårt att uppfatta små effekter (såsom nationella prov). Man kan utifrån ovanstående anta att även systematiska forskningsöversikter av hög kvalitet överskattar verkliga effekter något.

Utöver att resultaten i forskningsöversikter kännetecknas av viss osäkerhet i sig tillkommer osäkerhet kring i vilken utsträckning resultaten kan generaliseras till andra sammanhang än de undersökta. Exempelvis har en stor andel av de experimentella och kvasi-experimentella studierna inom utbildningsvetenskap genomförts i USA (se avsnitt 6.1) och det är inte säkert att insatserna fungerar på samma sätt i Sverige.

Här kan även nämnas att en enskild experimentell eller kvasi-experimentell studie vanligen endast testar en avgränsad hypotes, såsom huruvida ett visst kompetensutvecklingsprogram förbättrar elevers lärande, eller huruvida coaching förstärker effekten av en insats. Ofta är dock mer komplexa förklaringsmodeller mer intressanta och praktiskt värdefulla, exempelvis modeller som inbegriper mekanismer som förklarar varför undervisning och lärande förbättras och förutspår hur detta påverkas av kontextuella faktorer. En så komplex förklaringsmodell kan tyvärr inte testas i en enskild studie. Däremot kan jämförelser av resultaten i ett flertal experimentella och kvasi-experimentella studier sammantaget belysa modellens förklaringsvärde. Av denna anledning kan också sägas att resultaten av experimentella och kvasi-experimentella studier blir mer värdefulla ifall de utformas för att bidra till att undersöka förklaringsvärdet av praktiskt relevanta förklaringsmodeller (jfr Sims, Anders, m.fl., 2023; Sweller m.fl., 2024). Som vi återkommer till i avsnitt 6.2 ägnar vi därför särskilt fokus åt studier som söker förklara vilka egenskaper hos kompetensutveckling som är viktiga för dess effekter på undervisning och lärande.

Denna rapport utgår alltså i första hand ifrån systematiska forskningsöversikter med meta-analyser över experimentella och kvasi-experimentella studier, eftersom detta utgör den bästa existerande grunden för att dra kausala slutsatser om kompetensutvecklingens effekter på undervisning och elevers lärande. Eftersom det även i sådana översikter kvarstår viss osäkerhet och det inte är givet att slutsatserna kan generaliseras till en svensk kontext bör slutsatserna i första hand betraktas som de bästa gissningarna om sannolika utfall utifrån tillgänglig forskning. För att veta i vilken mån en insats fungerar i Sverige behöver det testas genom försöksverksamhet (se avsnitt 10).

5. Litteratursökning

Valet av litteratur som huvudsaklig grund för denna rapport baserades alltså på följande inklusionskriterier:

1. Studien är en meta-analys
2. Meta-analysen inkluderar enbart studier av experimentell eller kvasi-experimentell karaktär
3. Meta-analysen inkluderar studier av lärares kompetensutveckling i förskolan, grundskolan och/eller gymnasiet
4. Utfallsvariablerna är undervisningspraktik och/eller elevutfall

Sökningen av litteratur som motsvarade inklusionskriterierna skedde i tre steg. I steg ett sökte vi i Google Scholar efter kombinationer av tre begrepp: *meta-analysis*, *teacher*, *professional development* eller *professional learning*, och gick igenom de första 100 resultaten per sökning. I steg två sökte vi efter meta-analyser avseende lärares kompetensutveckling i några av de mest inflytelserika tidskrifterna inom utbildningsvetenskap och som fokuserar antingen på meta-analyser eller lärares kompetensutveckling: *Educational Research Review*, *Journal of Research on Educational Effectiveness*, *Journal of Teacher Education*, *Review of Educational Research*, *School Effectiveness and School Improvement* och *Teaching and Teacher Education*. I steg tre inkluderade vi studier som vi tidigare hade kännedom baserat på våra forskningsintressen om lärares kompetensutveckling, samt tog kontakt med kollegor inom fältet med liknande forskningsintresse. Vi begränsade sökningen till forskningsöversikter publicerade mellan 2014-2024 av flera anledningar: a) Senare publicerad forskning visar ett mer aktuellt forskningsläge. b) Det är större sannolikhet att senare studier har genomförts i skolkontexter som liknar dagslägets. c) Meta-analyser som publicerades före 2014 (t ex Scher & O'Reilly, 2009) baserades på få originalstudier, vilket beror på att antalet experimentella och kvasi-experimentella inom utbildningsvetenskap har ökat exponentiellt sedan år 2000.

Sökningarna enligt steg 1-3 genererade sammantaget 13 forskningsöversikter avseende grundskola/gymnasium, varav fem berörde effekter på undervisning och elva berörde effekter på elevresultat (tre översikter berörde effekter inom både områden), se bilaga 1. Sökningarna genererade också åtta forskningsöversikter avseende förskola, varav fyra berörde effekter på undervisning och fem berörde effekter på elevresultat (två översikter berörde effekter inom både områden), se bilaga 2.

Vi redovisar resultaten i forskningsöversikterna avseende grundskola/gymnasium och förskola separat, i avsnitt 6 respektive 7.

6. Resultat i forskning om kompetensutvecklingens effekter på undervisning och lärande i grundskola och gymnasium

6.1 Effekter av kompetensutvecklingsprogram

Vi definierade i avsnitt 2 följande begrepp för att beskriva och kategorisera olika aspekter av kompetensutveckling: *förutsättning*, *innehållsfokus*, *mekanism*, *program* och *modell*. I relation till dessa begrepp är forskningsläget avseende kompetensutvecklingens kausala effekter på undervisning och lärande starkast avseende program, det vill säga i vilken mån deltagande i specifika kompetensutvecklingsprogram påverkar undervisning och elevers lärande. Anledningen till detta är att de flesta studier med kausal design har randomiserat deltagande till kompetensutvecklingsprogram respektive kontrollgrupp, snarare än till olika

förutsättningar, innehållsfokus eller mekanismer. Vi inleder därför med att beskriva effekter av deltagande i program.

Genomsnittseffekten av deltagande i kompetensutvecklingsprogram är enligt översikterna:

- Mellan 0,42-1,11 standardavvikelse på undervisning enligt fem meta-analyser (Garrett m.fl., 2019; Gonzalez m.fl., 2022; Kahmann m.fl., 2022; Kraft m.fl., 2018; Rice m.fl., 2024).
- Mellan 0,01-0,52 standardavvikelse på elevresultat enligt elva meta-analyser (Balta m.fl., 2023; Basma & Savage, 2018, 2023; Didion m.fl., 2020; Fletcher-Wood & Zuccollo, 2020; Kahmann m.fl., 2022; Kraft m.fl., 2018; Lynch m.fl., 2019; Pellegrini m.fl., 2021; Rice m.fl., 2024; Sims, Fletcher-Wood, O'Mara-Eves, m.fl., 2023).

De studerade kompetensutvecklingsprogrammen har till störst del behandlat läsförståelse, matematik och naturvetenskap, samt till viss del även ämnesöverskridande områden såsom ledarskap i klassrummet och formativ bedömning. De fem forskningsöversikterna om effekter på undervisning baserades på mellan 13 och 43 originalstudier, med ett genomsnitt om 27 originalstudier. De 11 forskningsöversikterna om effekter på elevresultat baserades på mellan 7 och 104 originalstudier, med ett genomsnitt om 32 originalstudier. För en översikt över alla meta-analyser, se bilaga 1.

De originalstudier som forskningsöversikterna baseras på har i huvudsak genomförts i USA. Ett exempel på detta är att av de 104 studier som inkluderades i forskningsöversikten av Sims m.fl. (2023) genomfördes 73 i USA, 25 i Storbritannien och sex i andra länder (Australien, Italien, Kanada och Nederländerna). Resultaten i översikterna är alltså inte nödvändigtvis generaliserbara till Sverige.

Det kan behövas viss vägledning för att tolka måttet på effekter som anges ovan (standardavvikelse). Detta mått jämför förändringen som orsakas av insatsen med den genomsnittliga spridningen i den undersökta gruppen. En standardavvikelse motsvarar exempelvis skillnaden mellan genomsnittet i en elevgrupp och prestationen för en elev som presterar bättre än 84 procent av gruppen. Den främsta fördelen med detta mått är att det möjliggör jämförelser mellan studier även när dessa baseras på olika resultatmått, exempelvis poäng i standardiserade prov i olika årskurser, länder och skolämnen. Vid tolkningen av standardavvikelse är det dock viktigt att ta hänsyn till vilket utfallsmått standardavvikelsen beräknades utifrån. En förbättring om 0,1 standardavvikelse i ett brett mått såsom nationella provet är betydligt mer imponerande än motsvarande förbättring i ett smalt mått såsom för- och eftertest av 50 glosor som klassen har tränat på. För att bättre förstå vad måttet standardavvikelse innebär kan det konkretiseras på olika sätt. I årskurs 5 motsvarar exempelvis ett års lärande i genomsnitt 0,4 standardavvikelse i ett standardiserat prov (såsom nationella provet), varav skolan orsakar 40 procent av förbättringen (Kraft, 2020). En förbättring om 0,05 standardavvikelse i ett standardiserat prov innebär därmed omkring en månads lärande.

Forskningsöversikterna visar också att kompetensutvecklingen har större påverkan på undervisningen än på elevers lärande, vilket inte är förvånande eftersom insatser i första hand påverkar lärare och endast indirekt påverkar elever. Ett par forskningsöversikter uppskattar att det krävs 1 standardavvikelse förbättrad undervisning för att uppnå 0,21 respektive 0,27 standardavvikelse förbättrade elevresultat (Gonzalez m.fl., 2022; Kraft m.fl., 2018). Detta beror dock i högsta grad på hur undervisningen förändras: ifall insatsen förordar undervisningssätt som inte är effektiva kan man inte förvänta sig förbättrade elevresultat.

Det är även viktigt att nämna att ovanstående forskningsöversikter analyserar effekten av högkvalitativa kompetensutvecklingsprogram. Eftersom det är kostsamt och arbetskrävande att genomföra experimentella och kvasi-experimentella studier genomförs sådana i första hand avseende ovanligt genomarbetade kompetensutvecklingsprogram, där innehåll och upplägg är väl motiverade av forskning. Det betyder att effekterna som beskrivs i ovan nämnda forskningsöversikter inte kan antas gälla för den kompetensutveckling som lärare vanligtvis deltar i. Studier av mer typisk kompetensutveckling indikerar snarare att dess effekt på elevresultat är nära noll (D. N. Harris & Sass, 2011; Jacob & McGovern, 2015; Kirsten m.fl., 2023).

6.2 Egenskaper med kausal påverkan på kompetensutvecklingens effekter

Som nämnt avser existerande forskningsöversikter i första hand effekten av kompetensutvecklingsprogram. Forskningsgrunden för att dra kausala slutsatser om vad som gör dessa program effektiva (med andra ord deras förutsättningar, innehållsfokus och mekanismer) är betydligt svagare. Det är synd eftersom mycket få skolor har tillgång till just de kompetensutvecklingsprogram som har studerats. Lokala beslutsfattare skulle således ha större nytta av kunskaper om vilka egenskaper hos kompetensutveckling som är viktiga för dess effekter på undervisning och elevers lärande, eftersom sådana kunskaper kan användas oavsett tillgången på kompetensutvecklingsprogram i närområdet.

De flesta forskningsöversikter gör dock ansatser till analyser av vad som gör insatser effektiva, där aspekter såsom förutsättningar, innehållsfokus och mekanismer betraktas som *moderatorer*. I sådan analys undersöks i vilken mån program som har en viss egenskap uppvisar högre effekter än program utan egenskapen. Nackdelen med moderatoranalyser är att de visar korrelation snarare än kausal effekt. Det kan exempelvis vara så att de program som har en viss egenskap (exempelvis ämnesspecifikt fokus eller lärarsamarbete) också skiljer sig på andra sätt mot de insatser som saknar egenskapen. Det är därför inte möjligt att avgöra om skillnaden i effekt mellan insatser med/utan egenskapen beror på denna egenskap eller på någon annan skillnad mellan insatserna.

Sims m fl (2023) föreslår mot bakgrund av detta problem två metoder för att dra kausala slutsatser om effekten av egenskaper hos kompetensutveckling. Den första och mest tillförlitliga metoden är att randomisera deltagare inom samma program till olika egenskaper, såsom med/utan ämnesfokus, med/utan lärarsamarbete, med/utan individuell handledning, kort/lång insats, osv. Sådana studier kallas ibland A/B-test eller att ett randomiserat kontrollerat försök har flera "arms". Tyvärr är antalet studier med randomisering till olika interventionsegenskaper litet eftersom projektfinansieringen sällan räcker till mer än en interventions- och kontrollgrupp. Sims m fl (2023) föreslår och utprövar därför ett angreppssätt som är uppdelat i två steg. I det första steget identifieras mekanismer vars kausala effekter har visats i randomiserade studier av A/B-typ inom andra områden (i detta fall kognitionsvetenskap, beteendevetenskap och läkarutbildning). I det andra steget testas en förregistrerad hypotes att de identifierade mekanismerna förklarar skillnaden i effekt mellan olika kompetensutvecklingsprogram för lärare. Analysen i detta andra steg är visserligen fortfarande av korrelationskaraktär, men en kausal tolkning blir mer sannolik eftersom teorin baseras på identifierade kausala mekanismer inom andra fält och hypotesen är förregistrerad.

Resultaten i Sims m fl (2023) indikerar att kompetensutvecklingsprogram med fler mekanismer uppvisar större effekter på elevers lärande, där program som

innehåller mekanismer inom fyra olika områden (teoretiska insikter, motivera undervisningssättet, utveckla tekniker, samt omsättning till praktik, se tabell 1) uppvisar tre gånger så stora effekter på elevresultat som program med mekanismer inom färre områden (0,15 respektive 0,05 standardavvikelser). Forskningsöversikten hittar däremot inte några tydliga mönster avseende att vissa mekanismer eller mekanismkombinationer är särskilt viktiga.

Område	Mekanism
Teoretiska insikter	Hantera kognitiv belastning Repetera tidigare kunskaper
Motivera undervisningssättet	Sätta mål Trovärdig källa Beröm/förstärkning
Utveckla tekniker	Instruktion Praktiskt socialt stöd Modellering Återkoppling Övning
Omsätta till praktik	Påminnelser Självvärdering Planera handling Kontext-specifik repetition

Tabell 1. Mekanismer inom olika områden, översatt från Sims m fl (2023).

Resultaten i studien av Sims m fl (2023) kan tolkas som att lärare behöver flera olika typer av stöd för att på ett kvalitativt och varaktigt sätt omsätta nyvunna kunskaper till undervisningen. Det finns även enstaka studier med randomisering inom samma kompetensutvecklingsprogram (A/B-test) som stöder en sådan slutsats. Två sådana studier indikerade att videodemonstration av undervisningssätt ökar effekten på deltagande lärares undervisning jämfört med att deltagarna lär sig om undervisningssättet genom att läsa och diskutera forskningssammanställningar med konkreta exempel på omsättning till klassrummet (Mancenido m.fl., 2023; Sims, Fletcher-Wood, Godfrey-Faussett, m.fl., 2023). Flera andra studier använde randomisering inom programmet för att undersöka om individuell coaching i tillägg till kompetensutveckling ökade insatsens effekter jämfört med enbart kompetensutveckling, vilket i dessa studier visade sig vara fallet (Cohen m.fl., 2020, 2021; M. H. Davis m.fl., 2018; Mitchell, 2008). I relation till tabell 1 kan modellen videodemonstration kategoriseras som ett exempel på mekanismen modellering, medan modellen coaching utgörs av flera mekanismer, såsom praktiskt socialt stöd och återkoppling (ibland även ytterligare mekanismer såsom påminnelser och planera handling).

6.3 Analyser av moderatorer

Som nämnt genomför flera av forskningsöversikterna analyser av moderatorer, med andra ord jämförelser av effekter i program med/utan en egenskap. Nedan redovisas resultaten för alla moderatorer som har undersökts i minst två av ovan nämnda forskningsöversikter. Som nämnt bör resultaten av dessa analyser betraktas som korrelation snarare än kausalitet. Dessutom baseras moderatoranalyserna oftast på färre underliggande studier än forskningsöversikternas generella resultat, då moderatoranalyserna avgränsas till de av de inkluderade studierna som har rapporterat respektive moderator. Detta leder till sämre precision i analyserna och ofta icke statistiskt signifikanta resultat även när effekterna är relativt stora. Vi redogör i

nedanstående sammanställning för resultaten oavsett statistisk signifikans, men uppmanar alltså till försiktighet i tolkningen.

Kompetensutvecklingens förutsättningar:

- *Varaktighet*: Sju forskningsöversikter indikerar att insatser med kort och lång varaktighet ger lika stora effekter på undervisning och lärande (Basma & Savage, 2018; Didion m.fl., 2020; Garrett m.fl., 2019; Gonzalez m.fl., 2022; Kahmann m.fl., 2022; Kraft m.fl., 2018; Rice m.fl., 2024). Omfattningen av de insatser som jämfördes i dessa forskningsöversikter varierade mellan under 10 h till över 100 h och nolleffekten av varaktighet gällde oavsett om varaktighet mättes som totalt antal timmar, antal veckor som insatsen pågick eller antal träffar per vecka/månad. Två översikter indikerade till och med att insatser med kortare varaktighet gav större effekter på elevresultat: de insatser vars totala omfattning var under 30 h (Basma & Savage, 2018) eller som fördelades över färre än 26 veckor (Garrett m.fl., 2019). Å andra sidan indikerade forskningsöversikterna av Basma m fl (2023) samt Balta m fl (2023) en positiv effekt av varaktighet. Det sammantagna mönstret är alltså att det är svårt att se något tydlig effekt av varaktighet, vilket kan tolkas som att insatsers kvalitet är viktigare än deras omfattning. Vi återkommer till denna fråga i avsnitt 7.2 och 9.6.
- *Skala*: Fem forskningsöversikter indikerar att mer storskaliga insatser uppvisar mindre effekter (Balta m.fl., 2023; Garrett m.fl., 2019; Gonzalez m.fl., 2022; Kahmann m.fl., 2022; Kraft m.fl., 2018). Översikten av Lynch m fl (2019) visar dock ingen skillnad beroende på insatsens skala. En förklaring till att storskaliga insatser ofta visar mindre effekter kan vara att det kan vara svårt att upprätthålla kvalitet och målgruppsanpassning i större skala. En annan förklaring kan vara att storskaliga insatser i större utsträckning använder standardiserade tester som utfallsmått, vilka är mindre känsliga för effekter än studiespecifika mått.
- *Lärarsamarbete*: En forskningsöversikt indikerar att lärarsamarbete och att flera lärare på samma skola deltar inte påverkar effekten av en insats (Kahmann m.fl., 2022), medan detta enligt en annan översikt ökar effekten (Lynch m.fl., 2019). Enligt en översikt med lärares kunskaper, attityder och undervisning som utfallsmått var effekten av lärarsamarbete och att flera lärare på samma skola deltar negativt (Kowalski m.fl., 2020).³ Den senare översikten spekulerar i om den negativa effekten kan bero på att en del deltagare i program med samarbete och flera deltagare från samma skola deltar motvilligt. I så fall skulle en positiv effekt av lärarsamarbete kunna döljas av en negativ effekt av motvillighet.

Kompetensutvecklingens innehållsfokus:

- Två forskningsöversikter indikerar att fokus på ämnesinnehåll är positivt (Gonzalez m.fl., 2022; Kraft m.fl., 2018), medan en översikt indikerar lika stora effekter utan sådant fokus (Didion m.fl., 2020).
- Två forskningsöversikter indikerar att insatser som fokuserar på både ämnesinnehåll och pedagogik uppvisar större effekter än insatser utan sådant kombinerat fokus (Gonzalez m.fl., 2022; Lynch m.fl., 2019).

³ Vi nämner denna översikt här eftersom den är intressant i sammanhanget, trots att den exkluderades från vår litteratursökning eftersom den inte avgränsar effekten på undervisning från effekter på lärares kunskaper och attityder.

Modeller för kompetensutveckling:

- *Coaching*: Fyra forskningsöversikter indikerar att insatser där kompetensutveckling kombineras med individuell coaching uppvisar större effekter (Garrett m.fl., 2019; Kraft m.fl., 2018; Lynch m.fl., 2019; Rice m.fl., 2024).
- *Undervisningsmaterial*: Två forskningsöversikter indikerar att insatser där kompetensutveckling kombineras med undervisningsmaterial uppvisar större effekter (Kraft m.fl., 2018; Lynch m.fl., 2019), medan lika många översikter indikerar att insatser utan undervisningsmaterial uppvisar lika stora effekter (Gonzalez m.fl., 2022; Rice m.fl., 2024).

Sammantaget indikerar ovanstående moderatoranalyser:

1. Insatser behöver inte ha lång omfattning för att uppvisa effekter på undervisning och lärande.
2. Det är svårare att uppnå stora effekter i storskaliga program.
3. Det saknas tydliga belegg för att lärarsamarbete och att flera lärare på samma skola deltar i sig ökar effekterna på undervisning och lärande.
4. Det finns indikationer på att insatser som berör både ämnesinnehåll och pedagogik uppvisar större effekt, men det finns även många insatser med ämnesöverskridande innehåll som uppvisar positiva effekter.
5. Insatser som innehåller individuell coaching i tillägg till kompetensutveckling tenderar att uppvisa större effekt.
6. I vissa fall uppvisar insatser som kombinerar undervisningsmaterial med kompetensutveckling större effekter.

7. Resultat i forskning om kompetensutvecklingens effekter på undervisning och elevutfall i förskola

7.1 Effekter av kompetensutvecklingsprogram

Analysen av effekter av kompetensutveckling i förskolan baseras på åtta forskningsöversikter med meta-analyser över effekter på undervisningen och/eller elevutfall. Två av forskningsöversikterna undersöker effekter på både undervisning och elevutfall och två av forskningsöversikterna inkluderar mer än en meta-analys, avseende olika typer av kompetensutvecklingsinsatser och/eller elevutfall.

Fyra av forskningsöversikterna använde något vidare inkluderingskriterier än enbart experimentella och kvasi-experimentella studier. Brunsek m fl (2020) samt Jensen och Ramussen (2019) inkluderade i ett första steg studier med olika metodologi, men meta-analyserna baserades enbart på experimentella/kvasi-experimentella studier. Av de 27 studierna som inkluderades i analysen av effekter på undervisning i Egert m fl (2018) hade åtta en icke-experimentell design med för- och eftertest för endast en grupp. I denna översikt var dock genomsnittseffekten densamma i studier med experimentell och icke-experimentell design. Avseende Lee och Sung (2023) valde vi att redovisa meta-analysens resultat avgränsat till experimentella och kvasi-experimentella studier, då resultatet avseende de icke-experimentella studierna skilde sig betydligt från resultatet för experimentella och kvasi-experimentella studier.

Genomsnittseffekten av deltagande i kompetensutvecklingsprogram är enligt översikterna:

- Mellan 0,39-0,68 standardavvikelse på undervisning enligt fyra meta-analyser (Egert m.fl., 2018, 2020; J. Y. Lee & Sung, 2023; Markussen-Brown m.fl., 2017).
- Mellan 0,06-0,54 standardavvikelse på elevutfall enligt 19 meta-analyser som rapporteras i fem forskningsöversikter (Brunsek m.fl., 2020; Egert m.fl., 2018; Jensen & Rasmussen, 2019; Joo m.fl., 2020; Markussen-Brown m.fl., 2017).

Effekterna på undervisning avser aspekter såsom lärare-barn-interaktioner, vilket mäts av externa observatörer genom observationsprotokoll såsom *Classroom Assessment Scoring System*. Effekter på elevutfall avser i huvudsak språkliga förmågor (såsom ordförråd, fonetisk medvetenhet och bokstavskunskap) och socioemotionellt beteende (såsom interaktioner med lärare och andra barn, uppmärksamhetsförmåga och problembeteende).

De tre meta-analyserna om effekter på undervisning baserades på mellan 15 och 36 originalstudier, med ett genomsnitt om 29 originalstudier. De 19 meta-analyserna om effekter på elevresultat baserades på mellan 3 och 11 originalstudier, med ett genomsnitt om 6 originalstudier. Underlaget för slutsatserna om effekter på elevutfall är alltså litet och betydligt mindre än i forskningsöversikterna avseende grundskola/gymnasium, vilket innebär att slutsatserna avseende elevutfall i förskolan är mer osäkra. För en översikt över alla meta-analyser om kompetensutveckling i förskolan, se bilaga 2.

I likhet med forskningsöversikterna om kompetensutveckling för grundskola/gymnasium är genomsnittseffekterna på undervisning större än effekterna på elevutfall. Två översikter undersökte korrelationen mellan lärar- och elevutfall i studier med båda utfallsmått. Medan den ena studien fann att lärarutfall förklarade 53 procent av variationen i elevutfall (Egert m.fl., 2018), fann den andra mer oklara och icke statistiskt signifikanta korrelationer från -0,1 till 0,62 (Markussen-Brown m.fl., 2017). Båda dessa analyser baserades dock på ett litet antal originalstudier (mellan fem och elva).

7.2 Analyser av moderatörer

Tre av forskningsöversikterna avseende förskola analyserade egenskaper i de undersökta programmen som kunde förklara skillnader i effekter. Dessa analyser jämför effekter i program med/utan en viss egenskap och resultaten bör tolkas som deskriptiva snarare än som kausala effekter av respektive egenskap. Nedan redovisas resultaten för alla moderatörer som har undersökts i minst två av ovan nämnda forskningsöversikter.

Kompetensutvecklingens förutsättningar:

- *Varaktighet*: Två forskningsöversikter fann att insatser med kort/lång varaktighet hade mycket snarlika effekter (Egert m.fl., 2018, 2020) och resultaten var liknande oavsett om varaktighet mättes i antal timmar, frekvens eller över hur lång tidsperiod insatsen pågick. En av dessa översikter indikerade dock att medellånga program (45-60 h) hade störst effekt (Egert m.fl., 2018). I en annan översikt (Markussen-Brown m.fl., 2017) indikerade vissa mått på varaktighet en positiv association med effekter på undervisning, vilken dock försvann när även antal mekanismer/modeller (såsom kurser och coaching) lades till som förklarande variabel. Detta tyder på att de större effekterna i program med längre varaktighet som översikten fann berodde på att dessa program hade fler mekanismer snarare än på varaktigheten i sig. En av forskningsöversikterna som fann att insatser med kort/lång varaktighet genererade lika stora effekter inkluderade enbart studier som använde observationsverktyget *Classroom Assessment Scoring System* som utfallsmått

(Egert m.fl., 2020), vilket utesluter att de likartade effekterna i denna översikt drivs av att korta program använder snävare utfallsmått.

- *Skala*: En översikt fann att program i stor/liten skala uppvisade lika stora effekter (Egert m.fl., 2018), medan en annan översikt fann att storskaliga insatser uppvisade mindre effekter (Markussen-Brown m.fl., 2017).

Modeller för kompetensutveckling:

- *Coachning*: Två översikter fann att insatser som innehöll coachning uppvisade större effekter, särskilt när coachningen kombinerades med andra modeller och/eller mekanismer (Egert m.fl., 2018; Markussen-Brown m.fl., 2017). En översikt visade ingen statistiskt signifikant effekt av coaching (Egert m.fl., 2020).
- *Undervisningsmaterial*: En översikt fann att program som inkluderade lektionsplaneringar/undervisningsmaterial uppvisade större effekter (Markussen-Brown m.fl., 2017), medan en annan översikt fann att program med fokus vid att implementera sådant material uppvisade mindre effekter (Egert m.fl., 2018).
- *Antal mekanismer/modeller*: Två forskningsöversikter visade att program som innehöll fler *mekanismer/modeller* (såsom coachning, lektionsplaneringar/undervisningsmaterial, återkoppling och videodemonstration) uppvisade större effekter på undervisning och denna association har större förklaringsvärde än respektive *mekanism/modell* i sig själv (Egert m.fl., 2020; Markussen-Brown m.fl., 2017). Dessa forskningsöversikter drar alltså samma slutsats om betydelsen av antal mekanismer som översikten av Sims m fl (2023), se avsnitt 6.2.

Sammantaget liknar resultaten från dessa moderatoranalyser i stor utsträckning de avseende grundskola/gymnasium, med vissa nyansskillnader:

1. Insatser behöver inte ha lång omfattning för att uppvisa effekter på undervisning och lärande, men resultaten indikerar samtidigt att viss varaktighet underlättar för program att kunna erbjuda lärare flera typer av stöd.
2. Det finns tecken på mindre effekter i storskaliga program, men mindre tydligt än i översikterna om grundskola/gymnasium.
3. Det finns indikationer på att insatser som innehåller individuell coachning uppvisar högre effekt.
4. Oklara resultat avseende program som inkluderar lektionsplaneringar/undervisningsmaterial.
5. Program som inkluderar fler mekanismer/modeller uppvisar större effekter på undervisning och denna tendens är tydligare än effekten av respektive mekanism/modell.

8. Möjliga modeller för lärares kompetensutveckling

Den forskning som vi har beskrivit ovan delar inte på ett tydligt sätt in kompetensutvecklingsinsatser i modeller, utan drar snarare slutsatser om effekten av kompetensutvecklingsprogram och olika egenskaper hos kompetensutveckling. För att konkretisera vår diskussion tar vi dock utgångspunkt i fyra modeller för lärares kompetensutveckling som har stöd i forskningslitteraturen: 1) videodemonstrationer, 2) kollegaobservationer med stöd av observationsprotokoll, 3) undervisningsmaterial, och 4) coachning. De mekanismer och förutsättningar som ingår i dessa modeller kan dock kombineras på olika sätt och det finns inte forskningsstöd för att generellt rekommendera en av modellerna framför en annan.

Oavsett modell för lärares kompetensutveckling rekommenderar vi att:

1. Lärare bör få stöd genom flera olika mekanismer eftersom det ökar sannolikheten att lärare på ett kvalitativt och varaktigt sätt omsätter nyvunna kunskaper till undervisningen (se avsnitt 6.2 och 7.2).
2. Modellen bör bestå av fristående delar med relativt kort varaktighet eftersom insatser inte behöver ha lång varaktighet för att uppvisa effekter på undervisning och elevers lärande (se avsnitt 6.3 och 7.2) och insatser med kortare varaktighet är mer sparsamma med lärares tid och skolans resurser.
3. Modellen bör rekommendera undervisningssätt med gott stöd i tillförlitlig forskning om vilka undervisningssätt som förbättrar elevers lärande (med hänsyn till faktorer såsom skolämne, årskurs och kontext) eftersom det annars är osannolikt att elevers lärande förbättras (se avsnitt 3 samt 9.8).

När vi i fortsättningen av detta avsnitt beskriver de fyra modellerna fokuserar vi i första hand på den första av ovanstående punkter, då modellerna främst skiljer sig avseende vilka mekanismer och/eller kombinationer av mekanismer som ingår. Exempelvis kan modellen att tillhandahålla lärare undervisningsmaterial ses som en kombination av mekanismerna att ge instruktion om tekniker och att ge möjlighet att planera handling.

Notera att resultaten från flera av meta-analyserna indikerar att kompetensutvecklingsprogram som inbegriper kombinationer av flera olika mekanismer och/eller modeller visar större effekter på undervisning och/eller elevresultat (Egert m.fl., 2020; Markussen-Brown m.fl., 2017; Sims, Fletcher-Wood, O'Mara-Eves, m.fl., 2023). Dock finns inte stöd för att beskriva vissa mekanismkombinationer som särskilt effektiva. När Sims m fl (2021) grupperar program i särskilda mekanismkombinationer ("instructional coaching", "lesson study" och "strong teacher learning community") uppvisar dessa lika stora effekter på elevresultat och analysen indikerar dessutom att effekten av respektive modell drivs av att programmen inbegriper ytterligare mekanismer än de som definierar respektive modell. Det finns även moderatoranalyser i andra forskningsöversikter som ger indikationer i samma riktning, då program med flera modeller/mekanismer (såsom coaching, lektionsplaneringar/undervisningsmaterial och videodemonstration) uppvisar större effekter på undervisning, samt att denna effekt inte kan hänföras till någon av de enskilda modellerna (Egert m.fl., 2020; Markussen-Brown m.fl., 2017). Av denna anledning utgör redogörelserna för olika modeller för kompetensutveckling under 8.1-4 exempel på möjliga sätt att kombinera mekanismer på snarare än rekommendationer om särskilt effektiva modeller. Notera även, vilket vi också ger exempel på under 8.1-4, att det är möjligt att inom program kombinera olika modeller eller mekanismer från flera modeller.

De exempel på modeller av kompetensutvecklingsprogram vi ger grundas i fyra typer av vetenskapligt stöd:

1. Modellerna inkluderar mekanismer som enligt forskningsöversikten av Sims m fl (2023) förbättrar undervisning och elevresultat.
2. Modellerna har stöd i A/B-tester.
3. Modellerna har stöd i moderatoranalyser.
4. Det finns högkvalitativa effektstudier av kompetensutvecklingsprogram av respektive modell som visar positiva effekter.

8.1 Videodemonstrationer

I relation till tabell 1 kan videodemonstrationer (e.g., filmer av lektioner eller specifika undervisningsmoment) sägas motsvara mekanismen modellering. Det finns flera exempel på randomiserade kontrollerade studier av kompetensutvecklingsprogram innehållande videodemonstrationer. Ett exempel är kompetensutvecklingsprogrammet *My Teaching Partner*, som kombinerar workshops baserade på ett bibliotek med videodemonstrationer med digital coachning där lärare filmar sina lektioner och får feedback av en coach med stöd av ett observationsprotokoll. Detta program har i flera randomiserade kontrollerade studier i både förskola och skola visat på positiva effekter på undervisning och/eller elevernas resultat (Allen m.fl., 2011, 2015; Hamre m.fl., 2012; Pianta m.fl., 2008).⁴

Det finns även flera exempel på randomiserade kontrollerade studier av videodemonstrationer som har genomförts med lärarstudenter. Två sådana studier testade om videodemonstration av undervisningssätt ökar effekten på deltagande lärarstudenters undervisning jämfört med att deltagarna lär sig om undervisningssättet genom att läsa och diskutera forskningssammanställningar med konkreta exempel på omsättning till klassrummet (Mancenido m.fl., 2023; Sims, Fletcher-Wood, Godfrey-Faussett, m.fl., 2023) och såg positiva effekter för lärarstudenter som fick se videodemonstrationer.

8.2 Kollegaobservationer med stöd av observationsprotokoll

Kollegaobservationer motsvarar i tabell 1 mekanismerna att lärare får koppla teoretiska insikter till tekniker genom instruktion och modellering. I den mån lärare möts för reflektion efter observationen omfattas även mekanismerna återkoppling och praktiskt socialt stöd. Om det sker flera observationsomgångar möjliggörs även mekanismerna övning och kontextspecifik repetition.

På senare år har flera randomiserade kontrollerade experiment där kollegaobservationer används som grund för kompetensutveckling genomförts. Dessa studier tyder på att sannolikheten att kollegaobservationer främjar undervisning och lärande är större om observationerna utgår ifrån etablerade observationsprotokoll. Exempelvis visar en studie från England att ett kompetensutvecklingsprogram som bygger på kollegaobservationer i en ”lesson study”-modell inte visar några effekter på elevernas resultat (Murphy m.fl., 2021). I detta program används inte ett observationsprotokoll som grund för observationerna, utan utgångspunkterna för observationerna bestäms av lärarna i förhållande till lektionen som man har planerat tillsammans. Flera studier av kollegaobservationer med etablerade observationsprotokoll visar istället positiva effekter (Burgess m.fl., 2021; Gore m.fl., 2017, 2021). Det mest kända exemplet är kanske det australiensiska kompetensutvecklingsprogram *Quality Teaching Rounds* där lärare i grupper om fyra observerar varandra och sedan, med stöd av ett observationsprotokoll, diskuterar lektionerna tillsammans. Detta program har i flera randomiserade kontrollerade experiment visat på positiva resultat på både undervisningspraktiken och elevernas resultat (Gore m.fl., 2017, 2021), vilket inte setts för kontrollgrupper eller lärargrupper som observerat varandra utan observationsprotokollet (Gore m.fl., 2021).

De observationsprotokoll som användes i ovan nämnda studier (*Framework for Teaching* respektive *Quality Teaching Model*) är ämnesöverskridande. Det finns dock även flera väletablerade ämnesspecifika observationsprotokoll såsom *Mathematical Quality of*

⁴ Insatsen som studeras av Hamre m fl (2012) innehåller inte coachning, utan insatsen består av en kurs med videodemonstrationer samt till viss del kollegaobservationer.

Instruction och Teaching for Robust Understanding för matematik, *Protocol for Language Arts Teaching Observations* för läs- och skrivundervisning, *Authentic Intellectual Work* för samhällsorienterade ämnen, samt *UTeach Teacher Observation Protocol* för matematik och naturvetenskap. Det finns även protokoll som är anpassade för förskolan, såsom *Classroom Assessment Scoring System* och *Early Language and Literacy Classroom Observation*. Det finns skäl att misstänka att användning av årskurs- och ämnesspecifika observationsprotokoll skulle kunna generera tydligare effekter på undervisning och lärande eftersom studier indikerar att innebörden av undervisningsskicklighet delvis skiljer sig åt mellan skolämnen (Burgess m.fl., 2022). Vi känner dock inte till några experimentella eller kvasi-experimentella effektstudier som har undersökt kollegaobservationer med ämnesspecifika observationsprotokoll. I jämförelser mellan hur väl olika observationsprotokoll predicerar elevers lärande visar det sig också att ämnesöverskridande observationsprotokoll uppvisar minst lika högt förklaringsvärde som ämnesspecifika (Kane & Staiger, 2012).

8.3 Undervisningsmaterial

Att kombinera kompetensutveckling med tillhandahållande av nytt undervisningsmaterial (såsom lektionsplaneringar och läromedel) motsvarar i första hand mekanismerna att ge instruktion om tekniker och att ge möjlighet att planera handling. Flera av meta-analyserna, både för skola (t ex Garrett m.fl., 2019; Lynch m.fl., 2019) och förskola (t ex Markussen-Brown m.fl., 2017) visar på att kompetensutvecklingsprogram tillsammans med undervisningsmaterial har en större effekt jämfört med enbart kompetensutveckling även om det inte alltid är statistiskt signifikant. Det finns även exempel på A/B-test som indikerar att kompetensutveckling kombinerat med undervisningsmaterial är mer effektivt än att enbart erbjuda kompetensutveckling (C. J. Harris m.fl., 2015). I sammanhanget intressant är att erbjudandet av undervisningsmaterial, utan kompetensutveckling, också kan vara tillräckligt för att erhålla positiva resultat på elevernas prestationer. Ett A/B-test visar exempelvis att både tillhandahållandet av undervisningsmaterial, samt tillhandahållande av undervisningsmaterial kombinerat med kompetensutveckling hade positiva effekter på elevernas matematikresultat, även om effekter för undervisningsmaterial och kompetensutveckling var större (Jackson & Makarin, 2018).

Ett exempel på modellen undervisningsmaterial, som testats i effektstudier, finns även i Sverige inom matematikämnet. Materialet *Tänka, Räkna och Resonera*, vilket har utvecklats för förskoleklass, har i ett randomiserat kontrollerat experiment med enbart undervisningsmaterialet (utan kompetensutveckling för lärarna) visat på positiva effekter på elevresultat (Sterner m.fl., 2020). Liknande effekter uppvisas även när programmet har skalats upp (Sterner m.fl., 2023).

Kvaliteten på de undervisningsmaterial som erbjuds är sannolikt betydelsefullt för om positiva effekter ska uppnås, exempelvis att designen av lärarhandledningen bygger på designprinciper för *educative curriculum materials*, det vill säga hur materialet kan stödja lärares lärande (se t ex Arias m.fl., 2017; E. A. Davis m.fl., 2017; Roblin m.fl., 2018).

7.4 Coaching

Coaching, det vill säga att lärare efter observation av en eller flera lektioner får återkoppling på sin undervisning av en coach, kan motsvara ett flertal av mekanismerna i tabell 1, då coachen kan ge många typer av både teoretiskt och praktiskt stöd. Coaching är också den modell som har lyfts fram som effektiv av flest meta-analyser. Exempelvis fokuserar en meta-analys specifikt på kompetensutveckling med en coaching-komponent och finner statistiskt

signifikant resultat på både undervisning och elevresultat (Kraft m.fl., 2018). Vidare indikerar flertalet av övriga meta-analyser (t ex Garrett m.fl., 2019; J. Y. Lee & Sung, 2023; Markussen-Brown m.fl., 2017; Rice m.fl., 2024) att kompetensutveckling med en coaching-komponent har mer positiv effekt på undervisningen och/eller elevresultaten jämfört med kompetensutveckling utan coaching, även om resultaten inte alltid är statistiskt signifikanta. Det finns även A/B-test av kompetensutvecklingsprogram där man specifikt jämfört kompetensutveckling + coaching med enbart kompetensutveckling och funnit att kompetensutveckling med coaching har signifikant större effekt (Mitchell, 2008). Ett annat exempel är Neuman och Wright (2010) som visade att förskolelärare som fick coaching visade på signifikant bättre undervisningskvalitet jämfört med den grupp som fick kompetensutveckling i form av en kurs samt kontrollgruppen.

Avseende coaching är det viktigt att beakta på vilka grunder och hur coachingen ges. Exempelvis visar meta-analyser från både skola (Kraft m.fl., 2018) och förskola (J. Y. Lee & Sung, 2023; Markussen-Brown m.fl., 2017), att coaching har större effekt när det kombineras med andra aktiviteter, såsom workshops, gruppsessioner med andra lärare och/eller användande av undervisningsmaterial. Ett annat exempel är det tidigare nämnda programmet My Teaching Partner som kombinerar coaching med workshops baserade på videodemonstrationer. Dessa resultat antyder att lärare kan dra nytta av att först få bygga en baskunskap om undervisningsteorin innan de engagerar sig direkt med en coach.

9. Överväganden vid omsättning till nationell kompetensutveckling

Avsnitten ovan ger en bild av bästa tillgängliga forskning om vilken typ av kompetensutveckling som visar effekt på undervisning och elevers lärande. Det finns dock många ytterligare frågor som behöver besvaras vid omsättning till nationell kompetensutveckling, såsom kostnadseffektivitet, anpassning till olika målgrupper och försöksverksamhet. Vi diskuterar sådana aspekter nedan.

9.1 Kostnadseffektivitet

Det är oftast tillräckligt med relativt små effekter på elevers lärande och utveckling för att insatser i skolan ska vara samhällsekonomiskt lönsamma, eftersom sådana effekter kan leda till högre lön, bättre hälsa och lägre kriminalitet, vilket i ett livsperspektiv motsvarar omfattande ekonomiska resurser. Två forskningsöversikter indikerar exempelvis att om skolutgifter ökar motsvarande \$1000 per elev förbättras elevresultat med i genomsnitt 0,04 standardavvikelse och deltagande i högskol utbildning med 3,3 procent (Jackson & Persico, 2023). Med utgångspunkt i tidigare studier som indikerar att en percentil förbättrade skolresultat motsvarar 0,5 procent högre lön och att 10 procent högre högskoledeltagande motsvarar fem procent högre lön uppskattar Jackson och Persico (2023) att netto-kostnadskvoten för ökade skolutgifter motsvarande \$1000 per elev är minst 2,47, med andra ord att en investerad krona ger minst 2,47 kronor i utdelning. Eftersom denna beräkning inte tar hänsyn till andra samhällsekonomiska vinster av utbildning, såsom förbättrad hälsa och lägre kriminalitet, underskattar beräkningen sannolikt de samhällsekonomiska vinsterna. Beräkningen tar inte heller hänsyn till samhällsekonomiska vinster av effekter på lärare såsom bättre hälsa och minskad sjukfrånvaro och läraromsättning. Forskare i England har exempelvis uppskattat de direkta kostnaderna för att utbilda en lärare till minst £19 000 och kostnaderna för produktivitetsförlusten när en lärare byter skola till minst £12 500 (Fletcher-Wood & Zuccollo, 2020).

En så dyr insats som \$1000 per elev med så små effekter på testresultat som 0,04 standardavvikelser är alltså sannolikt ändå samhällsekonomiskt lönsam. Eftersom kompetensutvecklingsinsatser vanligtvis är betydligt billigare än så och kan generera minst lika stora effekter är de sannolikt ännu mer kostnadseffektiva. Ett exempel kan hämtas från Grönqvist m fl (2021) som uppskattade att matematiklyftet kostade omkring €80 per elev och förbättrade testresultat med 0,03 standardavvikelser. Utifrån detta beräknar de nytto-kostnadskvoten till 27, med andra ord att en investerad krona ger 27 kronor i utdelning.⁵ Eftersom de forskningsöversikter som vi utgår ifrån i denna rapport indikerar att högkvalitativa kompetensutvecklingsprogram i genomsnitt förbättrar elevresultat mellan 0,01-0,52 standardavvikelser bör det vara möjligt att nå betydligt högre effekter än de som matematiklyftet uppvisade. Eftersom forskningsöversikterna dessutom indikerar att insatser med kort/lång varaktighet leder till lika stora effekter bör dessutom kostnaderna kunna sänkas betydligt jämfört med matematiklyftet, där deltagande lärare avsatte cirka 80 h under ett år, motsvarande omkring fem procent av deras årsarbetstid (Kirsten, 2020). Det verkar alltså realistiskt att uppnå en betydligt högre nytto-kostnadskvot än 27.

Även vid billiga insatser krävs förstås att de uppvisar positiva effekter på betydelsefulla utfall för att vara samhällsekonomiskt lönsamma. Eftersom studier indikerar att den kompetensutveckling som lärare vanligtvis deltar i inte förbättrar elevresultat (D. N. Harris & Sass, 2011; Jacob & McGovern, 2015; Kirsten m.fl., 2023) och studier från USA och England uppskattar kostnaderna för den typiska kompetensutvecklingen till omkring tre procent av skolans kostnader (Killeen m.fl., 2002; Miles m.fl., 2004; Van den Brande & Zuccollo, 2021)⁶ skulle förbättrad kvalitet av kompetensutvecklingen kunna leda till betydande samhällsekonomiska vinster.

Ur ett kostnadseffektivitetsperspektiv bör insatser generera så stora effekter som möjligt till så låg kostnad som möjligt. Tabell 2 nedan, baserad på Kraft (2020), åskådliggör denna princip i relation till typiska effektstorlekar och kostnader för insatser i skolan. Färgerna i tabellen ger en ungefärlig illustration av kostnads-nyttokalkylen för olika insatser, där insatser i gröna rutor är klart samhällsekonomiskt lönsamma, insatser i rosa rutor är på gränsen till samhällsekonomiskt lönsamma, medan insatser i den röda rutan sannolikt inte är samhällsekonomiskt lönsamma.

Effektstorlek	Kostnad per elev		
	Låg (< 5000 kr)	Medel (5000 - 40 000 kr)	Hög (> 40 000 kr)
Liten (< 0,05 SD)			
Medel (0,05 - 0,2 SD)			
Stor (> 0,2 SD)			

Tabell 2. Insatsers kostnader och effekter, baserat på Kraft (2020).

I relation till tabell 2 är kompetensutvecklingsprogram utifrån existerande skattningar av kostnader och effekter en billig typ av skolinsats som uppvisar små eller medelstora effekter.

⁵ För läsflyftet beräknades ingen nytto-kostnadskvot, men effekten på det huvudsakliga elevresultatmålet (nationella prov-resultat i svenska) var mindre – omkring 0,01 standardavvikelser – och även kostnaderna per lärare var antagligen något lägre eftersom skolorna inte fick ersättning för deltagande lärare, till skillnad från upplägget i matematiklyftet (Holmlund m.fl., 2024).

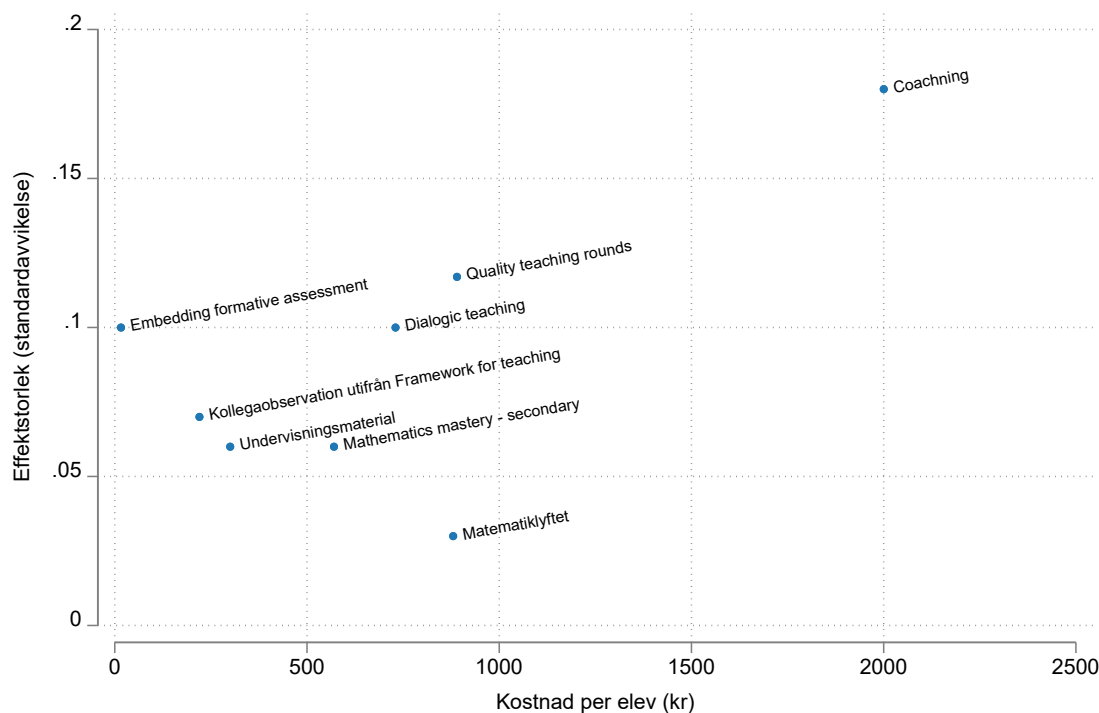
⁶ Eftersom svenska lärare deltar mindre i kompetensutveckling än lärare i USA och England är kostnaderna för kompetensutveckling i Sverige antagligen något lägre. Före kommunaliseringen motsvarande exempelvis lärares kompetensutveckling omkring 1,4 procent av skolans kostnader (Kirsten, 2020).

Nedan ges några exempel:

- *Kollegaobservationer*: Burgess m fl (2021) undersökte en insats där lärare observerade delar av sex eller tolv lektioner utifrån observationsprotokollet ”Framework for teaching”. Forskarna uppskattade kostnaden för insatsen till omkring £450 per lärare, med effekter på elevresultat om 0,07 standardavvikelser och där elevresultaten förbättrades för både observerade och observerande lärare. Om varje lärare undervisar minst 20 elever blir kostnaden per elev högst £22. Kollegaobservationsmodellen *Quality teaching rounds* uppskattas kosta \$127 per elev och generera en effekt om 0,117 standardavvikelser på elevers testresultat i matematik (Deloitte, 2020). Nyttokostnadskvoten för *Quality teaching rounds* beräknades till mellan 40 och 150, med andra ord att en investerad krona ger mellan 40 och 150 kronor i utdelning.
- *Lektionsplaneringar/undervisningsmaterial*: Att ge lärare tillgång till högkvalitativa lektionsplaneringar och undervisningsmaterial är ett exempel på en mycket billig insats. En studie av en sådan insats uppskattade kostnaden per lärare till \$431 (Jackson & Makarin, 2018), med andra ord en tiondel av kostnaden för coachning. För att kunna jämföra dessa siffror med övriga insatser slår vi ut siffran över 20 elever, så att kostnaden per elev blir cirka \$22 (egentligen undervisade de deltagande lärarna omkring 90 elever under ett läsår enligt studien). Effekten på elevresultat uppskattades till 0,06 standardavvikelser, med särskilt stora effekter för svaga lärare och större effekter för lärare som fick stöd i tillägg till lektionsplaneringarna. Nyttokostnadskvoten beräknades till över 900, med andra ord att en investerad krona ger 900 kronor i utdelning.
- *Kompetensutvecklingsprogram med olika typer av stöd för omsättning till praktik*: Fletcher-Wood och Zuccollo (2020) ger exempel på två kompetensutvecklingsinsatser i England (”Embedding formative assessment” och ”Dialogic teaching”) som kostade £1,2 respektive £54 per elev, med effekter mellan 0,1 och 0,15 standardavvikelser på elevresultat. En kompetensutvecklingsinsats riktad mot ”mastery learning” inom matematik visade en effekt om 0,06 standardavvikelser på elevresultat och kostade £42 per elev (Education Endowment Foundation, 2016). Matematiklyftet uppskattades vara något dyrare och generera mindre effekter: cirka €80 per elev och 0,03 standardavvikelser effekt på elevresultat (Grönqvist m.fl., 2021).
- *Coachning*: Enligt översikten av Kraft m fl (2018) förbättrar coachning i genomsnitt elevresultat med 0,18 standardavvikelser, vilket är en medelstor effekt. Samtidigt är coachning en relativt dyr form av kompetensutveckling, där kostnaden har uppskattats till \$3300-5200 per lärare (Kraft m.fl., 2018), även om detta fortfarande är billigt jämfört med att exempelvis öka lärartätheten. Om varje lärare undervisar minst 20 elever blir kostnaden per elev alltså omkring \$200. Kraft m fl (2018) föreslår att kostnaderna för coachning möjligen kan minskas genom exempelvis digital coachning, samt att det kan vara lämpligt att rikta coachning till utvalda lärare och skolor där det kan göra särskilt stor skillnad. Det finns också exempel på initiativ för att förbättra coachningens skalbarhet med bibehållen kvalitet, såsom i *My teaching partner* och *Steplab*, där coachning kombineras med webbaserade resursbanker med forskningsbaserade rekommendationer som kan användas utifrån den enskilda lärarens behov. Ett antal studier av *My teaching partner* visar positiva effekter på elevresultat (Allen m.fl., 2015; Gregory m.fl., 2017).

Som nämnt kategoriseras alla dessa insatser som billiga och med liten eller medelstor effekt enligt ramverket i tabell 2. Kostnadseffektiviteten skiljer sig dock ändå en del mellan insatserna, vilket illustreras i figur 3. Illustrationen bör betraktas som högst ungefärlig eftersom siffrorna över kostnader baseras på enskilda insatser och beräkningarna kan ha gjorts på olika sätt i olika insatser och därmed inte är jämförbara.

I figuren skulle en idealisk insats (billig och med stor effekt) hamna högt uppe till vänster, medan det är mindre önskvärt att hamna långt nere till höger (dyrt och liten effekt). Det framgår exempelvis i figuren att ”Embedding formative assessment” verkar vara mer kostnadseffektiv än matematiklyftet. Coachning skiljer sig från övriga insatser genom att vara både dyrare och effektivare. Det är inte helt givet om det är mest önskvärt med en dyr och effektiv eller billig och något mindre effektiv insats.



Figur 3. Några exempel på kompetensutvecklingsinsatserns kostnader och effekter.

Som Kraft (2020) noterar beror givetvis värderingar av insatserns lämplighet inte enbart på kostnads-nyttokalkyler, utan även på frågor såsom insatserns genomförbarhet och skalbarhet, lokala prioriteringar, tillgång till kompetens, samt vilket stöd eller motstånd olika typer av insatser väcker i olika intressegrupper. Även när en insats har gott forskningsstöd kvarstår också betydande osäkerhet om hur väl den kommer fungera i en ny kontext. Jämförelserna ovan utgår dessutom endast ifrån effekter på elevresultat. Det kan hända att bilden skulle se något annorlunda ut med andra utfallsmått då testresultat missar en del av lärarens påverkan på elevers långsiktiga utveckling (Jackson, 2018). Det finns exempelvis studier som tyder på att lärarens påverkan på aspekter såsom skolkännsro och betyg har större betydelse för långsiktiga utfall för svaga elever, medan lärarens påverkan på elevers testresultat har större betydelse för starka elever (Backes m.fl., 2024). Många studier indikerar dock att förbättrade elevresultat leder till långsiktiga effekter inom områden såsom kriminalitet, högskoledeltagande, hälsa och politiskt deltagande och intresse (Björkenstam m.fl., 2011; Bömmel & Heineck, 2023; Jackson & Persico, 2023; Sörberg Wallin m.fl., 2018).

9.2 Motsvara behov i många skolämnen och årskurser

I vårt uppdrag ingår att värdera hur kompetensutveckling kan riktas mot många skolämnen och årskurser. En bakgrund till detta är att många lärare har saknat kompetensutveckling som är riktad mot just deras undervisning. Som nämnt i avsnitt 6.3 finns det också en del forskning som tyder på att ämnesinriktad kompetensutveckling uppvisar större effekter på elevresultat,

även om också ämnesövergripande kompetensutveckling kan förbättra elevresultat och bäst effekt verkar uppnås med ett kombinerat fokus på ämnesinnehåll och pedagogik.

Om kompetensutvecklingsinsatser ska inriktas mot enskilda ämnesområden och årskurser innebär det givetvis att utvecklingsarbetet inför kompetensutvecklingen utökas väsentligt. Om kollegaobservationer utgår från ett allmändidaktiskt observationsprotokoll krävs exempelvis inte lika mycket resurser för att utveckla/översätta och utprova ett protokoll. Det finns även belägg för att allmändidaktiska protokoll kan vara effektiva för olika ämnen och årskurser. I denna rapport har vi gett exempel på flera sådana (t ex Burgess m.fl., 2021; Gore m.fl., 2021; Povey m.fl., 2023). Ämnesdidaktisk fördjupning skulle istället kunna uppnås genom att observationerna sker inom lärargrupper där lärarna undervisar samma ämnen, eller att en resursbank med exempel på hur olika faktorer i protokollet kan realiseras i undervisningen konstrueras med avseende på skolämne och årskurs.

Anpassning till olika skolämnen och årskurser är en än större utmaning för modellen undervisningsmaterial, då material behöver utvecklas för alla ämnen, ämnesområden och årskurser. Även om detta är kostsamt kan det samtidigt ge långsiktiga vinster om det görs med hög kvalitet. Exempelvis är det i nuläget mycket sällsynt med konkreta och utprovade rekommendationer, lärarhandledningar och undervisningsmaterial i olika skolämnen och årskurser som baseras på tillförlitlig forskning om effekter på elevers lärande och utveckling. Det är också svårt för lärare och läromedelsförlag att överblicka forskning på ett tillräckligt högkvalitativt sätt för att utveckla sådant material. I synnerhet bör denna typ av material kunna utgöra ett viktigt stöd för nya lärare.

Om utvecklingen av årskurs- och ämnesspecifikt material är en kärnfråga behöver man göra olika typer av prioriteringar för att arbetet inte ska bli ohanterligt. Några exempel på utgångspunkter för detta är:

1. Förbättring i dessa ämnesområden och/eller årskurser skulle ha stor betydelse för elevernas fortsatta skolgång.
2. Det finns tillräcklig forskningsgrund för att ge rekommendationer om undervisningssätt i ämnesområdet.
3. Uppskattade effekter av rekommenderade undervisningssätt är stora.
4. Det saknas i nuläget tillräckligt bra vägledning för lärare i området.
5. Det finns tecken på att lärares arbetssätt kan förbättras i området.
6. Undervisningssätten är enkla att lära sig och lyckas med och kräver inte stora resurser i termer av tid eller pengar.

9.3 Uppskalning

En utmaning avseende lärares kompetensutveckling är att bibehålla effekterna även när programmen skalas upp till att inkludera många lärare. Fem forskningsöversikter indikerar att mer storskaliga insatser uppvisar mindre effekter (Balta m.fl., 2023; Garrett m.fl., 2019; Gonzalez m.fl., 2022; Kahmann m.fl., 2022; Kraft m.fl., 2018), medan en översikt visar på ingen skillnad (Lynch m.fl., 2019).

En utmaning berör mängden stöd som behövs från en extern person (t ex en coach). Exempelvis visar resultaten från Kraft m fl (2018) att de genomsnittliga effekterna av större coachingprogram (>100 lärare) är en bråkdel av de effekter som hittas för mindre program (< 100 lärare). Vad detta beror på kan studien inte svara på, men en möjlig orsak kan vara svårigheterna i att behålla kvaliteten på coachingen när många coacher ska utbildas. Detta problem är inte lika stort när det gäller kollegaobservationer med observationsprotokoll,

videodemonstrationer och undervisningsmaterial, då behovet av stöd från en extern person är mer begränsad i dessa modeller och ofta avgränsas till en inledande workshop (ibland ännu mindre).

Med tanke på att coachning är en dyr insats som ställer höga krav på hög kompetens hos coacherna kan man tänka sig att en sådan insats i första hand bör riktas mot grupper där behovet av stöd är särskilt stort, såsom nya lärare och lärare på skolor med svåra förutsättningar.

9.4 Individuellt deltagande eller samverkan mellan lärare

Som nämnt i avsnitt 6.3 kan man utifrån forskningsöversikter inte finna något tydligt samband mellan lärarsamverkan och effekter på undervisning och elevresultat. Det finns även mer generellt (det vill säga utanför kompetensutvecklingsforskning) få tillförlitliga studier om effekter av lärarsamverkan, då de flesta studier i området bygger på observationer, intervjuer eller korrelationer (såsom huvuddelen av studierna i översikterna av Stoll m.fl., 2006; Vangrieken m.fl., 2015, 2017). De få existerande studier som har sökt avgränsa kausala effekter av lärarsamverkan finner blandade resultat (Banerjee m.fl., 2017; Kraft m.fl., 2016; M. Lee & Louis, 2019; Ronfeldt m.fl., 2015). En möjlig tolkning av dessa resultat är att lärarsamverkan i sig inte främjar undervisning och lärande, men att vissa typer av lärarsamverkan som riktas mot vissa typer av undervisningsförändring kan göra det. I relation till ramverket i tabell 1 kan exempelvis lärarsamverkan användas för att uppnå vissa mekanismer, såsom praktiskt socialt stöd, återkoppling och modellering. Dessa mekanismer kan dock även uppnås på andra sätt. Även ifall lärarsamverkan används kan det ske med andra än lärarna på den egna skolan, exempelvis andra lärare som deltar i en kompetensutvecklingskurs med ämnesspecifik inriktning.

Alla fyra modellerna som beskrivs i avsnitt 8.1-4 är möjliga att genomföra enskilt i bemärkelsen att det inte behövs fler än en deltagande lärare per skola. Videodemonstrationer och undervisningsmaterial är exempelvis enkelt att ta del av digitalt. Dock ökar sannolikheten för effekter på undervisning och elevers lärande om läraren får stöd genom exempelvis kursträffar som inriktas mot mekanismer såsom de i tabell 1. Eftersom coachning och kollegaobservation förutsätter interaktion är de inte möjliga att genomföra helt enskilt, men eftersom interaktionen kan ske digitalt med utgångspunkt i filmad undervisning krävs inte samverkan med andra lärare på samma skola.

9.5 Organisatoriska förutsättningar som påverkar kompetensutvecklingens utfall

Forskningsgrunden är otillräcklig för att dra tydliga slutsatser om vilka organisatoriska förutsättningar som krävs för att öka sannolikheten att kompetensutvecklingsinsatser främjar undervisning och elevresultat. Detta beror på att få studier har undersökt dessa frågor med metoder som förmår avgränsa kausala samband. En forskningsöversikt sammanställde dock slutsatserna i 21 implementeringsutvärderingar ("Implementation and Process Evaluations") av kompetensutvecklingsinsatser (Sims m.fl., 2021), med andra ord utvärderingar av vad som påverkade kvaliteten i insatsens genomförande. Utvärderingarna baserades i huvudsak på enkäter, intervjuer och observationer. Översikten över dessa utvärderingar fann följande deskriptiva mönster:

- Alla insatser hade implementeringsproblem, där vissa deltagande lärare eller skolor av olika anledningar inte följde insatsens riktlinjer.
- Stödsystemet, insatsens upplägg och skolkontexten påverkar implementeringen. För att uppnå högkvalitativ implementering behöver insatsen hålla hög kvalitet och lärare behöver få stöd att omsätta idéerna till praktik. Om lärare fick se exempel på hur god

implementering ser ut och fick tydliga riktlinjer var det lättare att lyckas. Rektorer behöver även stödja insatsen för att engagemanget ska bibehållas och insatsen behöver överensstämma med skolans prioriteringar. Lärare behöver också tillräckligt med tid för att genomföra insatsen.

- Ju bättre programmet är anpassat för lärares behov, desto större chans att det implementeras väl. Program bör kräva små tidsinvesteringar av lärare eftersom lärares scheman redan är fulla.

Sammanfattningsvis tyder sammanfattningen av utvärderingarna på att det lättare att lyckas med högkvalitativ omsättning till praktiken i insatser som kräver lite tid, rekommenderar undervisningsätt som är enkla att översätta till praktiken, ger konkret stöd för omsättningen, samt har stöd av rektor.

9.6 Insatsers varaktighet

Som nämnt i avsnitt 6.3 och 7.2 tyder ett flertal forskningsöversikter på att insatser med kort/lång varaktighet uppvisar lika stora effekter på undervisning och lärande. Det kan tyckas vara förvånande. Det bör exempelvis vara enklare att ge flera typer av stöd till lärare i en längre insats. En forskningsöversikt fann också en viss sådan effekt, då längre insatser i denna översikt uppvisade större effekter, ett samband som dock försvann när antal mekanismer/former lades till som ytterligare förklarande variabel (Markussen-Brown m.fl., 2017). Med andra ord tyder analysen på att varaktighet inte i sig ökade insatsernas effekter, men att viss varaktighet kan öka sannolikheten att lärare får flera typer av stöd. Att de flesta forskningsöversikter trots detta inte finner en positiv effekt av varaktighet skulle i så fall kunna bero på att längre insatser inte utnyttjar möjligheten att ge flera typer av stöd. En ytterligare förklaring som forskare har nämnt är att lärare i första hand tar till sig förslag på undervisningsförändringar som är enkla att omsätta i undervisningen (Desimone & Stuckey, 2014). Det är förstås även så att insatser som tar mycket tid i anspråk minskar den tid som lärare kan använda för annat som kan förbättra elevers lärande, såsom att förbereda lektioner.

Eftersom lärares tid och skolans resurser är begränsade föreslår vi utifrån ovanstående att kompetensutveckling erbjuds i form av fristående delar med relativt kort varaktighet. Varje insats behöver samtidigt ha en tillräcklig omfattning för att rymma flera typer av stöd till lärare i att inhämta nya kunskaper och omsätta dessa till praktik på ett kvalitativt och hållbart sätt. Hur mycket tid som krävs för detta kan dock variera. Nedan ges några exempel utifrån de modeller av kompetensutveckling som vi beskrev i avsnitt 8.1-4. I tillägg till nedanstående kan man tänka sig att lärare även ägnar extra tid åt planering och reflektion i samband med deltagande i kompetensutveckling, för att planera hur nya lärdomar kan omsättas i undervisningen.

- *Videodemonstration*: Två mycket korta insatser som gav stöd genom teoretiska texter, seminariediskussioner och videodemonstrationer hade en varaktighet om tre timmar eller mindre under samma dag (Mancenido m.fl., 2023; Sims, Fletcher-Wood, Godfrey-Faussett, m.fl., 2023). Dessa insatser undersökte dock enbart effekter på undervisning – inte på elevers lärande – och hade ett mycket begränsat innehåll. Det kan därför vara lämpligt att utöka både innehåll och antal mekanismer och då genomföra insatsen under exempelvis ett par timmar per gång under några tillfällen.
- *Kollegaobservation med observationsprotokoll*: Ett exempel på ett kompetensutvecklingsprogram utifrån modellen kollegaobservation som redovisas i Burgess m fl (2021), där deltagande lärare först får en introduktion till observationsprotokollet och därefter observerar 15-20 minuter undervisning under minst sex tillfällen, gärna med efterföljande återkoppling. Man kan tänka sig att detta medför en

total omfattning om tio timmar fördelat över sju dagar. I ett annat exempel på ett kompetensutvecklingsprogram med modellen kollegaobservation, *Quality teaching rounds* (Gore m.fl., 2017, 2021), omfattar programmet fyra heldagar för observationer och möten i observationsgrupperna, samt två ytterligare dagars inledande introduktion för de två lärare per skola som leder arbetet.

- *Lektionsplaneringar/undervisningsmaterial*: En insats som gav lärare tillgång till lektionsplaneringar och undervisningsmaterial samt stöd i form av webinarier och samarbetsgrupper (Jackson & Makarin, 2018) krävde ingen minsta tid av lärare, då användning och deltagande var frivilligt. Trots att användning av lektionsplaneringarna visade sig förbättra elevernas lärande och effekten var störst för de lärare vars elever tidigare hade haft svag kunskapsutveckling visade det sig att lärare i genomsnitt använde få lektionsplaneringar och särskilt få bland de lärare som inte fick stöd i tillägg till undervisningsmaterialet. Forskarna gissar att detta kan bero på att det kräver en inledande tidsinvestering för att granska förslag till lektionsplaneringar och att planera hur man själva ska iscensätta dessa. Å andra sidan kan en färdig lektionsplanering spara tid jämfört med att planera lektionen från grunden. Det verkar utifrån detta vara lämpligt att kombinera lektionsplaneringar och undervisningsmaterial med ytterligare stöd, såsom en introduktion till materialets teoretiska grund och betydelse, presentation och demonstration av tekniker, samt orientering i materialet och värdering av vad som kan vara lämpligt att använda. Även uppföljningsmöten och liknande kan sannolikt förbättra effekten av materialet.
- *Coaching*: I ett exempel på en coachningsintervention genomfördes sex ”coachningscykler” under ett år (Allen m.fl., 2015). Varje sådan cykel började med att läraren filmar en lektion och skickar filmen till coachen. Därefter valde coachen ut en kort filmsnutt och bad läraren reflektera kring denna i relation till vissa aspekter i ett observationsprotokoll. Slutligen möttes läraren och coachen i ett 20-30 minuter telefonsamtal. Sammantaget kräver varje cykel alltså ett par timmars arbete för läraren, motsvarande drygt 10 timmar över ett år.

9.7 Kompetensutvecklingsmodellens relation till professionsprogrammet

I skolverksuppdraget som denna rapport utgår ifrån (se avsnitt 1) ingår även att lämna förslag på hur kompetensutvecklingsmodellerna kan fungera inom ramen för professionsprogrammet. Det finns i nuläget endast en föreslagen struktur för professionsprogrammet att ta ställning till (Skolverket, 2024). I den föreslagna strukturen ingår kompetensutveckling i form av högskolekurser samt formella krav för meriteringsnivåer, vilka är tänkta att implementeras stegvis från 1 januari 2025. Det finns olika möjliga sätt att relatera den form av kompetensutveckling som vi föreslår till professionsprogrammet.

1. Den form av kompetensutveckling som vi föreslår kan utgöra ett komplement till professionsprogrammet genom att täcka in ett bredare behov av kompetensutveckling än professionsprogrammet kan erbjuda, exempelvis genom att mindre tidskrävande insatser gör det möjligt för en större andel av lärarna att delta än vad som är möjligt med typiska högskolekurser.
2. Den form av kompetensutveckling som vi föreslår kan användas istället för traditionella högskolekurser inom professionsprogrammet. En motivering för att göra så är att forskningsstödet för att den typ av kompetensutveckling som vi föreslår förbättrar elevers lärande och utveckling är betydligt starkare än stödet för att traditionella högskolekurser gör det, då ett flertal studier indikerar att kunskapsutvecklingen inte är bättre för elever som undervisas av lärare med lärarexamen (Taylor, 2023). Denna generella bild kan dock nyanseras, då studier indikerar att vissa aspekter eller former av lärarutbildning är gynnsamt för elevresultat, såsom ämnesspecialisering (Boyd m.fl., 2007; M. M. Kennedy m.fl., 2008), effektiva mentorer under praktikperioder (Goldhaber m.fl., 2020) och

handledning utifrån observerad undervisning (Cohen m.fl., 2020). En tolkning skulle kunna vara att högskolekurser med lika god undervisnings- och lärarutvecklingsteori som de kompetensutvecklingsprogram som har undersökts i meta-analyser (se avsnitt 3, 6 och 7) kan ge lika goda effekter på undervisning och lärande, men att så sällan är fallet. Även om högskolekursernas kvalitet skulle höjas kan man dock fortfarande ifrågasätta om så tidskrävande kurser som 15 hp för meriteringsnivå 1 och 60 hp för meriteringsnivå 2 (enligt nuvarande förslag till professionsprogram) är motiverat, med tanke på insatser med kort/lång varaktighet verkar ge lika stora effekter på undervisning och elevresultat. Det faktum att professionsprogrammet även ska utgöra grund för värdera meriteringsnivåer utgör dock en komplikation ifall traditionella högskolekurser ersätts av kortare kurser. För att möjliggöra formell meritering med den typ av kompetensutveckling som vi föreslår behövs i så fall någon form av examination i tillägg till kompetensutvecklingsmodellens komponenter.

9.8 Strukturer för att värdera forskningsstöd för kompetensutvecklingsinsatsers innehåll

Modellen i avsnitt 3 anger att det endast finns förutsättningar att förbättra elevresultat om de undervisningssätt som insatsen förordar främjar elevers lärande (med andra ord: om insatsens undervisningsteori stämmer). Vi har i denna rapport inte utrymme att värdera forskningsstödet för olika undervisningssätt, men vi vill betona vikten av att detta görs oavsett kompetensutvecklingsmodell. I nuläget sker tyvärr sällan sådan värdering av kompetensutvecklingsinsatsers innehåll på ett systematiskt sätt i Sverige. Inom insatserna matematiklyftet, läslyftet och specialpedagogik för lärande ställdes exempelvis inte krav på att de undervisningssätt som förordades skulle stödjas av forskning som tillförlitligt kunde identifiera kausala effekter av undervisningssätt på elevers lärande och utveckling (jfr Boesen m.fl., 2015).

Om målet är att kompetensutvecklingen ska förbättra elevers lärande och utveckling bör man alltså inför utvecklingen av kursinnehåll kräva att rekommendationer om undervisningssätt baseras på studier som kan identifiera kausala effekter på utfall såsom elevers lärande, i första hand högkvalitativa experimentella och kvasi-experimentella studier. Om möjligt bör också de nationella kontexter där studierna har genomförts beaktas, där studier som genomförts i Sverige (eller i nationell utbildningskontext som liknar Sverige) ges större vikt.

För att värdera forskningsgrunden för rekommendationer om undervisningssätt krävs kompetens att bedöma tillförlitligheten i forskningsresultat, vilket exempelvis kan uppfyllas genom samverkan mellan universitet/högskolor och Skolforskningsinstitutet. Ett exempel på hur forskningsgrundade rekommendationer kan utvecklas finns i ramverken för de engelska nationella kurserna ”National Professional Qualifications” (NPQs) (se Department for Education, 2024), vilka utvecklas av expertgrupper bestående av forskare och företrädare för praktiktäna organisationer, samt med stöd av *Education Endowment Foundation*, en organisation som, likt Skolforskningsinstitutet, sammanfattar och sprider forskning i skolfrågor. I England finns även ytterligare organisationer som i samverkan mellan forskare och lärare sammanställer och omsätter forskning till undervisningspraktik, samt tar initiativ till forskning utifrån praktikens behov, såsom *Chartered College of Teaching* och *Ambition Institute*.

9.9 Skillnader mellan den föreslagna modellen och befintliga modeller för kompetensutveckling

Den typ av kompetensutveckling som vi föreslår skiljer sig på ett flertal sätt från kollegialt lärande-modellen som används inom bland annat matematik- och läslyftet samt specialpedagogik för lärande.

- Kompetensutvecklingen består av fristående insatser med kortare varaktighet (se avsnitt 9.6), snarare än en termin eller ett år som inom kollegialt lärande-modellen.
- Kompetensutvecklingen ger tydliga rekommendationer om undervisningsätt, till skillnad från kollegialt lärande-modellen som var försiktig med att ge tydliga rekommendationer eftersom det ansågs kunna undergräva lärares handlingsfrihet.
- Kompetensutvecklingen rekommendationer om undervisningsätt grundas i forskning som kan dra kausala slutsatser om effekter av undervisningsätt på elevers lärande och utveckling.
- Lärare får mer stöd i att omsätta forskningsbaserade rekommendationer till undervisningspraktik, genom exempelvis videodemonstrationer eller kollegaobservationer med observationsprotokoll. Inom kollegialt lärande-modellen förekommer visserligen videor i vissa moduler, men i dessa utgör videodemonstrationer endast en liten del, där intervjuer med lärare, forskare och elever utgör en större del av filmerna. De videodemonstrationer som ändå förekommer i kollegialt lärande-modellens filmer utgör ofta små glimtar och inte sällan i ovanliga undervisningskontexter såsom att läraren undervisar små elevgrupper.

10. Effektvärdering av kompetensutvecklingsinsatser

Som nämnt kan man inte vara säker på att en insats kommer ge positiva effekter på undervisning och elevers lärande i Sverige, även om omfattande och tillförlitlig forskning indikerar att insatsen uppvisar goda effekter. Det beror både på att det även i tillförlitliga studier kvarstår potentiella felkällor och på att det är svårt att avgöra hur väl en insats kommer fungera i ett nytt sammanhang. Detta innebär att insatsers effekter behöver undersökas i Sverige, även vid ett gott internationellt forskningsläge.

Med effekter avses i vilken mån en insats uppfyller sitt syfte, exempelvis att förbättra undervisning och elevers lärande. Utvärderingar som baseras på deltagares åsikter ger tyvärr mycket lite information om sådana effekter. Ett exempel kan hämtas från utvärderingen av läslyftet, där enkätundersökningar visade att svensklärare och lågstadielärare var mest nöjda med insatsen, samt att 27 procent av de deltagande lärarna uppskattade att läslyftet i stor utsträckning har bidragit till förbättrad läsförståelse (Carlbaum m.fl., 2016, 2020). En senare utvärdering av effekter på elevers kunskapsutveckling visade dock nästan obefintlig effekt i svenska och lågstadiet, samt störst effekt i NO och SO på högstadiet (Holmlund m.fl., 2024), alltså omvänt resultat jämfört med vad enkätsvaren indikerade. För att avgöra effekter på undervisning och elevers lärande krävs alltså utvärderingar som direkt undersöker detta, så kallade effektvärderingar.

Den enklaste metoden för att möjliggöra slutsatser om kausala effekter av en insats är att potentiella deltagare (exempelvis lärare eller skolor som uttrycker intresse för att delta i en kompetensutvecklingsinsats) randomiseras till insats- och kontrollgrupp. Randomiseringen medför att grupperna inte skiljer sig åt på ett systematiskt sätt före insatsen och att eventuella skillnader som uppstår under insatsens gång kan hänföras till insatsen. Randomiseringen medför i sig inte någon extra kostnad, men möjliggör att utvärderingar kan dra slutsatser om effekter, utöver slutsatser om deltagande, processer och uppfattningar. Det är dock viktigt att

tidigt i planeringen av en insats även förbereda utvärderingen och att avsätta medel för detta. I detta ingår att planera vilka utfallsmått som man är intresserad av: vad hoppas man att insatsen ska förbättra eller åtminstone inte förvärra och vilka mått kan användas för att mäta detta? Vissa utfallsmått finns i redan tillgängliga data, medan ytterligare empiriinsamling kan behövas i andra fall.

I synnerhet vid försöksomgångar är det dessutom lämpligt att undersöka flera varianter av samma insats, exempelvis om lika stora effekter kan uppnås i en billigare insats, eller om effekterna på elevernas lärande blir större om lärarna exempelvis får tillgång till a) undervisningsmaterial, b) videodemonstrationer, eller c) individuell coaching. För att undersöka detta behöver deltagare randomiseras till de olika förutsättningarna.

Randomisering bör helst ske på skolnivå, det vill säga att skolor randomiseras till interventions- och kontrollgrupp, eftersom det annars finns en risk att deltagande lärare på en skola påverkar icke-deltagande lärare på samma skola, så att insatsen indirekt även inbegriper kontrollgruppen. Ifall mer än en lärare per skola ändå deltar behöver analyser dessutom ta hänsyn till klustring, med andra ord att lärare inom en skola är mer lika varandra än lärare som arbetar på olika skolor. Detta medför att den reella urvalsstorleken blir mindre (och därmed möjligheten att dra statistiskt signifikanta slutsatser) jämfört med att lika många lärare deltar med endast en lärare per skola (Torgerson & Torgerson, 2013).

För att förbättra möjligheterna att dra statistiskt signifikanta slutsatser bör studier använda förtest i tillägg till eftertest, eftersom det gör effektskattningarna mer precisa (Torgerson & Torgerson, 2013). Förtestet behöver inte vara detsamma som eftertestet, så länge resultaten i testerna korrelerar med varandra, med andra ord att de åtminstone delvis mäter samma förmågor.

Möjligheten att dra statistiskt signifikanta slutsatser ökar även om så kallade proximala utfallsmått används, det vill säga mått som fångar just de utfall som insatsen avser att påverka. Om insatsen exempelvis avser att påverka elevers taluppfattning kan en eventuell effekt lättare upptäckas i ett taluppfattningstest än i ett bredare test såsom ett nationellt prov (vilket är ett så kallat distalt mått). Studier indikerar att uppskattade effekter ofta är 2-4 gånger större i proximala mått än i distala mått (Kraft, 2020). Som tidigare nämnt är uppmätta effekter dessutom ofta betydligt större på undervisningen än på elevresultat (Gonzalez m.fl., 2022; Kraft m.fl., 2018), vilket innebär att effekter på undervisningen är lättare att identifiera. Dessutom ger information om effekter på undervisning viktig information om möjliga orsaker till att man ser eller inte ser effekter på elevresultat.

Om man använder proximala mått och för- och eftertester kan exempelvis 50 interventions- och kontrollskolor vara tillräckligt för att identifiera statistiskt signifikanta effekter. Ett annat möjligt tillvägagångssätt är att genomföra ett stegvis ("sequential") försök. Då kan man i en första omgång rekrytera 30 interventions- och kontrollskolor och sedan genomföra ytterligare omgångar till dess att statistiskt signifikanta effekter eller precisa nolleffekter kan identifieras (jfr Sims, Anders, m.fl., 2023).

Det är givetvis även viktigt att använda tillförlitliga mått. Exempel på detta är välbeprövade tester inom områden såsom avkodning, läsförståelse och taluppfattning, som bedöms maskinellt eller av externa bedömare enligt bedömningsanvisningar med litet tolkningsutrymme. Exempelvis innehåller LegiLexi ett flertal tester av den karaktären avseende viktiga förmågor för läsutveckling i årskurs 1-6, vilka både omfattar smalare mått av

bland annat avkodningsförmåga och bredare mått avseende läsförståelse. Ett alternativ är istället att använda betyg och nationella prov. En utmaning med dessa typer av tester är dock att de kan bedömas på ett olikvärdigt sätt, vilket kan leda till missvisande resultat, även om viss olikvärdighet i bedömningen kan kompenseras för i analyser ifall mönstren som orsakar olikvärdighet är kända, såsom att vissa skolor konsekvent sätter mer generösa betyg (jfr Holmlund m.fl., 2024). Om nationella prov blir externt rättade minskar sannolikt problemen med olikvärdighet betydligt.

En vanlig kritik mot standardiserade tester är att vissa förmågor är svåra att mäta i en sådan form. Ett exempel på detta är skrivförmåga, där det har visat sig vara svårt att uppnå hög överensstämmelse mellan olika lärares bedömningar. Det är samtidigt, trots svårigheterna, viktigt att mäta skrivförmåga eftersom det är ett viktigt mål i skolan. Även utfallsmått med mindre grad av överensstämmelse mellan bedömare kan användas i effektvärdering, så länge bedömningen inte systematiskt skiljer sig mellan interventions- och kontrollgrupp. Låg överensstämmelse mellan bedömare leder dock till lägre precision i effektmåtten, ett problem som dock kan minskas genom att kombinera flera mått på samma utfall.

Oavsett utfallsmått skulle analyser av effekter få större precision och bli mer tillförlitliga om skolornas uppgifter om elev-lärarkopplingar samlas in, det vill säga vilken lärare som undervisar en viss klass.

Även om försöksverksamhet i första hand bör använda proximala mått kan studier även dra nytta av existerande utfallsmått av mer distal karaktär i offentliga data, exempelvis betyg och nationella prov-resultat, avklarade poäng i högskolan, samt lön vid 25 eller 30 års ålder. Skolnärvaro kommer framöver bli ytterligare ett viktigt mått i offentlig statistik. I offentlig statistik finns även mått som kan användas för att mäta effekter på lärare, såsom läraromsättning och sjukskrivningar. För att dra statistiskt signifikanta slutsatser om effekter på distala mått krävs dock oftast stora urvalsstorlekar, vilket innebär att sådana analyser mer sannolikt blir informativa när insatser har skalats upp. Exempelvis kunde IFAU:s utvärderingar av matematik- och läslyftet dra slutsatser om effekter på distalt uppmätta elevresultat (nationella prov) med hög precision utifrån data från omkring 2000 interventionsskolor och mer än 1000 kontrollskolor (Grönqvist m.fl., 2021; Holmlund m.fl., 2024).

Referenser

- Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing Secondary School Instruction and Student Achievement: Replication and Extension of the My Teaching Partner-Secondary Intervention. *Journal of Research on Educational Effectiveness*, 8(4), 475–489. <https://doi.org/10.1080/19345747.2015.1017680>
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement. *Science*, 333(6045), 1034–1037. <https://doi.org/10.1126/science.1207998>
- Arias, A. M., Smith, P. S., Davis, E. A., Marino, J.-C., & Palincsar, A. S. (2017). Justifying Predictions: Connecting Use of Educative Curriculum Materials to Students' Engagement in Science Argumentation. *Journal of Science Teacher Education*, 28(1), 11–35. <https://doi.org/10.1080/1046560X.2016.1277597>
- Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2024). How to Measure a Teacher: The Influence of Test and Nontest Value-Added on Long-Run Student Outcomes. *Journal of Human Resources*. <https://doi.org/10.3368/jhr.1023-13180R2>
- Balta, N., Fukkink, R., & Amendum, S. J. (2023). The Effect of Job-embedded Professional Development on Teacher and Student Outcomes: A Multi-Level Meta-Analysis. *International Educational Review*, 1(1), 1–23. <https://doi.org/10.58693/ier.111>
- Banerjee, N., Stearns, E., Moller, S., & Mickelson, R. A. (2017). Teacher Job Satisfaction and Student Achievement: The Roles of Teacher Professional Community and Teacher Collaboration in Schools. *American Journal of Education*, 123(2), 203–241.
- Basma, B., & Savage, R. (2018). Teacher Professional Development and Student Literacy Growth: a Systematic Review and Meta-analysis. *Educational Psychology Review*, 30(2), 457–481.
- Basma, B., & Savage, R. (2023). Teacher Professional Development and Student Reading in Middle and High School: A Systematic Review and Meta-Analysis. *Journal of Teacher Education*, 74(3), 214–228.
- Björkenstam, C., Weitoft, G. R., Hjern, A., Nordström, P., Hallqvist, J., & Ljung, R. (2011). School grades, parental education and suicide--a national register-based cohort study. *Journal of Epidemiology and Community Health*, 65(11), 993–998.
- Boesen, J., Helenius, O., & Johansson, B. (2015). National-scale professional development in Sweden: theory, policy, practice. *ZDM*, 47(1), 129–141.
- Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. I P. Peterson, E. Baker, & B. McGaw (Red.), *International encyclopedia of education* (Vol. 1–7, s. 548–556). Elsevier.
- Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The Effect of Certification and Preparation on Teacher Quality. *The Future of Children*, 17(1), 45–68.
- Brunsek, A., Perlman, M., McMullen, E., Falenchuk, O., Fletcher, B., Nocita, G., Kamkar, N., & Shah, P. S. (2020). A meta-analysis and systematic review of the associations between professional development of early childhood educators and children's outcomes. *Early Childhood Research Quarterly*, 53, 217–248. <https://doi.org/10.1016/j.ecresq.2020.03.003>

- Burgess, S., Rawal, S., & Taylor, E. S. (2021). Teacher Peer Observation and Student Test Scores: Evidence from a Field Experiment in English Secondary Schools. *Journal of Labor Economics*, 39(4), 1155–1186. <https://doi.org/10.1086/712997>
- Burgess, S., Rawal, S., & Taylor, E. S. (2022). *Characterising Effective Teaching*. University of Bristol; Nuffield Foundation. <https://www.nuffieldfoundation.org/wp-content/uploads/2022/05/Burgess-Characterising-Effective-Teaching-Full-Report-April-2022.pdf>
- Bömmel, N., & Heineck, G. (2023). Revisiting the causal effect of education on political participation and interest. *Education Economics*, 31(6), 664–682. <https://doi.org/10.1080/09645292.2022.2141199>
- Carlbaum, S., Andersson, E., & Hanberger, A. (2016). *Utvärdering av Läslyftet. Delrapport 2: Erfarenheter av Läslyftet läsåret 2015/16*. Umeå Centre for Evaluation Research, Umeå universitet. https://www.skolverket.se/download/18.4fc05a3f164131a74183db2/1537192996436/Laslyftet_utvardering_delrapport2.pdf
- Carlbaum, S., Hanberger, A., Andersson, E., Roe, A., Tengberg, M., & Kärnebro, K. (2020). *Utvärdering av Läslyftet. Slutrapport från den nationella utvärderingen av Läslyftets genomförande och effekter i olika skolformer*. Umeå Centre for Evaluation Research, Umeå universitet. https://www.skolverket.se/download/18.a35553a171bb8ce903cf4/1591805722854/Laslyftet_utvardering_Slutrapport_2020_UCER.pdf
- Cohen, J., Krishnamachari, A., & Wong, V. C. (2021). Experimental Evidence on the Robustness of Coaching Supports in Teacher Education. I *EdWorkingPapers.com*. Annenberg Institute at Brown University. <https://edworkingpapers.com/ai21-468>
- Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher Coaching in a Simulated Environment. *Educational Evaluation and Policy Analysis*, 42(2), 208–231. <https://doi.org/10.3102/0162373720906217>
- Darling-Hammond, L., Hylar, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/Effective_Teacher_Professional_Development_REPORT.pdf
- Davis, E. A., Palincsar, A. S., Smith, P. S., Arias, A. M., & Kademian, S. M. (2017). Educative Curriculum Materials: Uptake, Impact, and Implications for Research and Design. *Educational Researcher*, 46(6), 293–304. <https://doi.org/10.3102/0013189X17727502>
- Davis, M. H., McPartland, J. M., Pryseski, C., & Kim, E. (2018). The effects of coaching on English teachers' reading instruction practices and adolescent students' reading comprehension. *Literacy Research and Instruction*, 57(3), 255–275. <https://doi.org/10.1080/19388071.2018.1453897>
- Deloitte. (2020). *Quality Teaching Rounds – Cost Benefit Analysis*. <https://qtacademy.edu.au/wp-content/uploads/2020/10/Deloitte-Access-Economics-QTR-Cost-Benefit-Analysis-Final-report-9-Sept.pdf>
- Department for Education. (2024, februari 23). *National professional qualifications frameworks*. <https://www.gov.uk/government/publications/national-professional-qualifications-frameworks-from-september-2021>

- Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, 38(3), 181–199.
- Desimone, L. M., & Stuckey, D. (2014). Sustaining professional development. In L. E. Martin, S. Kragler, D. J. Quatroche, & K. L. Bauserman (Eds.), *Handbook of professional development in education: Successful models and practices, PreK-12* (s. 467–482). Guilford Press.
- Didion, L., Toste, J. R., & Filderman, M. J. (2020). Teacher Professional Development and Student Reading Achievement: A Meta-Analytic Review of the Effects. *Journal of Research on Educational Effectiveness*, 13(1), 29–66.
<https://doi.org/10.1080/19345747.2019.1670884>
- Education Endowment Foundation. (2016). *Ark Mathematics Mastery: Secondary*.
<https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/mathematics-mastery-secondary/>
- Egert, F., Dederer, V., & Fukkink, R. G. (2020). The impact of in-service professional development on the quality of teacher-child interactions in early education and care: A meta-analysis. *Educational Research Review*, 29, 100309.
<https://doi.org/10.1016/j.edurev.2019.100309>
- Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of In-Service Professional Development Programs for Early Childhood Teachers on Quality Ratings and Child Outcomes: A Meta-Analysis. *Review of Educational Research*, 88(3), 401–433.
<https://doi.org/10.3102/0034654317751918>
- Fletcher-Wood, H., & Zuccollo, J. (2020). *The effects of high-quality professional development on teachers and students*. Education Policy Institute, Ambition Institute, and Wellcome Trust. <https://epi.org.uk/publications-and-research/effects-high-quality-professional-development/>
- Garrett, R., Citkowicz, M., & Williams, R. (2019). How Responsive Is a Teacher's Classroom Practice to Intervention? A Meta-Analysis of Randomized Field Studies: *Review of Research in Education*, 43(1), 106–137.
- Goldhaber, D., Krieg, J., & Theobald, R. (2020). Effective like me? Does having a more productive mentor improve the productivity of mentees? *Labour Economics*, 63, 101792. <https://doi.org/10.1016/j.labeco.2019.101792>
- Gonzalez, K. E., Lynch, K., & Hill, H. C. (2022). *A Meta-Analysis of the Experimental Evidence Linking STEM Classroom Interventions to Teacher Knowledge, Classroom Instruction, and Student Achievement* (EdWorkingPaper No 22-515). Annenberg Institute at Brown University. <https://doi.org/10.26300/d9kc-4264>
- Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., & Lubans, D. (2017). Effects of professional development on the quality of teaching: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education*, 68, 99–113.
- Gore, J., Miller, A., Fray, L., Harris, J., & Prieto, E. (2021). Improving student achievement through professional development: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education*, 101, 1–12.

- Gregory, A., Ruzek, E., Hafen, C. A., Mikami, A. Y., Allen, J. P., & Pianta, R. C. (2017). My Teaching Partner-Secondary: A Video-Based Coaching Model. *Theory Into Practice*, 56(1), 38–45. <https://doi.org/10.1080/00405841.2016.1260402>
- Grönqvist, E., Öckert, B., & Rosenqvist, O. (2021). *Does the 'Boost for Mathematics' Boost Mathematics?* IFAU. <https://www.ifau.se/Forskning/Publikationer/Working-papers/2021/does-the-boost-for-mathematics-boost-mathematics/>
- Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., Howes, C., LaParo, K., & Scott-Little, C. (2012). A Course on Effective Teacher-Child Interactions: Effects on Teacher Beliefs, Knowledge, and Observed Practice. *American Educational Research Journal*, 49(1), 88–123. <https://doi.org/10.3102/0002831211434596>
- Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., Cheng, B. H., & Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching*, 52(10), 1362–1385. <https://doi.org/10.1002/tea.21263>
- Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, 95(7), 798–812.
- Hill, H. C., & Papay, J. P. (2022). *Building Better PL: How to Strengthen Teacher Learning*. RPPL. <https://annenbergbrown.edu/sites/default/files/rppl-building-better-pl.pdf>
- Holmlund, H., Häggblom, J., & Lindahl, E. (2024). *The Boost for reading*. IFAU. <https://www.ifau.se/globalassets/pdf/se/2024/wp-2024-6-the-boost-for-reading.pdf>
- Hägglund, P., & Johansson, P. (2021). *Upp till bevis: en ESO-rapport om experiment som underlag för politik*. Regeringskansliet.
- Jackson, K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy*, 126(5), 2072–2107.
- Jackson, K., & Makarin, A. (2018). Can Online Off-the-Shelf Lessons Improve Student Outcomes? Evidence from a Field Experiment. *American Economic Journal: Economic Policy*, 10(3), 226–254. <https://doi.org/10.1257/pol.20170211>
- Jackson, K., & Persico, C. (2023). Point column on school spending: Money matters. *Journal of Policy Analysis and Management*, 42(4), 1118–1124. <https://doi.org/10.1002/pam.22520>
- Jacob, A., & McGovern, K. (2015). *The Mirage: Confronting the Hard Truth about Our Quest for Teacher Development*. TNTP. <https://eric.ed.gov/?id=ED558206>
- Jensen, P., & Rasmussen, A. W. (2019). Professional Development and Its Impact on Children in Early Childhood Education and Care: A Meta-Analysis Based on European Studies. *Scandinavian Journal of Educational Research*, 63(6), 935–950. <https://doi.org/10.1080/00313831.2018.1466359>
- Joo, Y. S., Magnuson, K., Duncan, G. J., Schindler, H. S., Yoshikawa, H., & Ziol-Guest, K. M. (2020). What Works in Early Childhood Education Programs?: A Meta-Analysis of Preschool Enhancement Programs. *Early Education and Development*, 31(1), 1–26. <https://doi.org/10.1080/10409289.2019.1624146>

- Kahmann, R., Droop, M., & Lazonder, A. W. (2022). Meta-analysis of professional development programs in differentiated instruction. *International Journal of Educational Research*, 116, 102072. <https://doi.org/10.1016/j.ijer.2022.102072>
- Kane, T. J., & Staiger, D. O. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*. Bill & Melinda Gates Foundation.
- Kennedy, A. (2005). Models of continuing professional development: a framework for analysis. *Journal of In-Service Education*, 31(2), 235–250.
- Kennedy, M. M. (2016). How Does Professional Development Improve Teaching? *Review of Educational Research*, 86(4), 945–980.
- Kennedy, M. M., Ahn, S., & Choi, J. (2008). The value added by teacher education. I *Handbook of research on teacher education* (s. 1247–1271). Routledge. <https://api.taylorfrancis.com/content/chapters/edit/download?identifierName=doi&identifierValue=10.4324/9780203938690-134&type=chapterpdf>
- Killeen, K. M., Monk, D. H., & Plecki, M. L. (2002). School district spending on professional development: Insights available from national data (1992-1998). *Journal of Education Finance*, 28(1), 25–49.
- Kirsten, N. (2020). *Kompetensutveckling som styrning. Om statliga insatser för lärares kompetensutveckling och hur de förhandlas i lärares lokala praktik* [Acta Universitatis Upsaliensis]. <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-406594>
- Kirsten, N., & Carlbaum, S. (2020). Kompetensutveckling för professionella lärare? Introduktionen av kollegialt lärande i svensk skola. *Pedagogisk forskning i Sverige*, 25(1), 7–34.
- Kirsten, N., Lindvall, J., Ryve, A., & Gustafsson, J.-E. (2023). How effective is the professional development in which teachers typically participate? Quasi-experimental analyses of effects on student achievement based on TIMSS 2003–2019. *Teaching and Teacher Education*, 132, 104242. <https://doi.org/10.1016/j.tate.2023.104242>
- Kowalski, S. M., Taylor, J. A., Askinas, K. M., Wang, Q., Zhang, Q., Maddix, W. P., & Tipton, E. (2020). Examining Factors Contributing to Variation in Effect Size Estimates of Teacher Outcomes from Studies of Science Teacher Professional Development. *Journal of Research on Educational Effectiveness*, 13(3), 430–458. <https://doi.org/10.1080/19345747.2020.1726538>
- Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*, 88(4), 547–588.
- Kraft, M. A., Marinell, W. H., & Shen-Wei Yee, D. (2016). School Organizational Contexts, Teacher Turnover, and Student Achievement: Evidence From Panel Data. *American Educational Research Journal*, 53(5), 1411–1449.
- Lee, J. Y., & Sung, J. (2023). Effects of in-service programs on childcare teachers' interaction quality: A meta-analysis. *Teaching and Teacher Education*, 124, 104017. <https://doi.org/10.1016/j.tate.2023.104017>

- Lee, M., & Louis, K. S. (2019). Mapping a strong school culture and linking it to sustainable school improvement. *Teaching and Teacher Education, 81*, 84–96.
- Lindvall, J., Helenius, O., & Wiberg, M. (2018). Critical features of professional development programs: Comparing content focus and impact of two large-scale programs. *Teaching and Teacher Education, 70*, 121–131. <https://doi.org/10.1016/j.tate.2017.11.013>
- Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the Research Base That Informs STEM Instructional Improvement Efforts: A Meta-Analysis. *Educational Evaluation & Policy Analysis, 41*(3), 260–293.
- Mancenido, Z., Hill, H. C., Coppersmith, J. G., Carter, H., Pollard, C., & Monschauer, C. (2023). Practice-Based Teacher Education Pedagogies Improve Responsiveness: Evidence from a Lab Experiment. I *EdWorkingPapers.com*. Annenberg Institute at Brown University. <https://edworkingpapers.com/ai23-873>
- Markussen-Brown, J., Juhl, C. B., Piasta, S. B., Bleses, D., Højen, A., & Justice, L. M. (2017). The effects of language- and literacy-focused professional development on early educators and children: A best-evidence meta-analysis. *Early Childhood Research Quarterly, 38*, 97–115. <https://doi.org/10.1016/j.ecresq.2016.07.002>
- Miles, K. H., Odden, A., Fermanich, M., & Archibald, S. (2004). Inside the Black Box of School District Spending on Professional Development: Lessons from Five Urban Districts. *Journal of Education Finance, 30*(1), 1–26.
- Mitchell, J. F. (2008). *The effects of literacy coaching on teacher strategy use and student reading comprehension* [PhD Thesis, Western Connecticut State University]. <https://search.proquest.com/openview/e31c6e555c5ed56895f3e274f0844860/1?pq-origsite=gscholar&cbl=18750>
- Murphy, R., Weinhardt, F., & Wyness, G. (2021). Who teaches the teachers? A RCT of peer-to-peer observation and feedback in 181 schools. *Economics of Education Review, 82*, 102091. <https://doi.org/10.1016/j.econedurev.2021.102091>
- Neuman, S. B., & Wright, T. S. (2010). Promoting Language and Literacy Development for Early Childhood Educators: A Mixed-Methods Study of Coursework and Coaching. *The Elementary School Journal, 111*(1), 63–86. <https://doi.org/10.1086/653470>
- Pellegrini, M., Lake, C., Neitzel, A., & Slavin, R. E. (2021). Effective Programs in Elementary Mathematics: A Meta-Analysis. *AERA Open, 7*, 2332858420986211. <https://doi.org/10.1177/2332858420986211>
- Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 23*(4), 431–451. <https://doi.org/10.1016/j.ecresq.2008.02.001>
- Povey, J., San Martin Porter, M., Kennedy, E., Hussain Potia, A., Bellotti, M., Cook, S., & Austerberry, S. (2023). *Building capacity for quality teaching in Australian schools: Queensland replication study - final report*. <https://nova.newcastle.edu.au/vital/access/services/Download/uon:48736/ATTACHMENT02>

- Rice, M., Lambright, K., & Wijekumar, K. (Kay). (2024). Professional Development in Reading Comprehension: A Meta-analysis of the Effects on Teachers and Students. *Reading Research Quarterly*, 59(3), 424–447. <https://doi.org/10.1002/rrq.546>
- Roblin, N. P., Schunn, C., & McKenney, S. (2018). What are critical features of science curriculum materials that impact student and teacher outcomes? *Science Education*, 102(2), 260–282. <https://doi.org/10.1002/sce.21328>
- Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher Collaboration in Instructional Teams and Student Achievement. *American Educational Research Journal*, 52(3), 475–514.
- Scher, L., & O'Reilly, F. (2009). Professional Development for K–12 Math and Science Teachers: What Do We Really Know? *Journal of Research on Educational Effectiveness*, 2(3), 209–249. <https://doi.org/10.1080/19345740802641527>
- Sims, S., Anders, J., Inglis, M., Lortie-Forgues, H., Styles, B., & Weidmann, B. (2023). *Experimental education research: rethinking why, how and when to use random assignment* (CEPEO Working Paper Series Nr 23–07). UCL Centre for Education Policy and Equalising Opportunities. <https://econpapers.repec.org/paper/uclcepeow/23-07.htm>
- Sims, S., Fletcher-Wood, H., Godfrey-Faussett, T., Mccrea, P., & Meliss, S. (2023). *Modelling evidence-based practice in initial teacher training: causal effects on teachers' skills, knowledge and self-efficacy*. UCL Centre for Education Policy and Equalising Opportunities. <https://econpapers.repec.org/RePEc:ucl:cepeow:23-09>
- Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., Van Herwegen, J., & Anders, J. (2023). Effective Teacher Professional Development: New Theory and a Meta-Analytic Test. *Review of Educational Research*. <https://doi.org/10.3102/00346543231217480>
- Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Van Herwegen, J., & Anders, J. (2021). *What Are the Characteristics of Effective Teacher Professional Development? A Systematic Review and Meta-Analysis*. Education Endowment Foundation. <https://eric.ed.gov/?id=ED615914>
- Skolverket. (2024). *Redovisning av uppdrag om ett nationellt professionsprogram. Dnr U2022/023199, U2023/01870*. <https://www.skolverket.se/publikationer?id=12833>
- Sterner, G., Nagy, C., & Nyström, P. (2023). A scaled-up mathematics intervention in preschool classes. *Scandinavian Journal of Educational Research*, 0(0), 1–11. <https://doi.org/10.1080/00313831.2023.2250352>
- Sterner, G., Wolff, U., & Helenius, O. (2020). Reasoning about Representations: Effects of an Early Math Intervention. *Scandinavian Journal of Educational Research*, 64(5), 782–800. <https://doi.org/10.1080/00313831.2019.1600579>
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional Learning Communities: A Review of the Literature. *Journal of Educational Change*, 7(4), 221–258.
- Sweller, J., Zhang, L., Ashman, G., Cobern, W., & Kirschner, P. A. (2024). Response to De Jong et al.'s (2023) paper “Let’s talk evidence – The case for combining inquiry-based and direct instruction”. *Educational Research Review*, 42, 100584. <https://doi.org/10.1016/j.edurev.2023.100584>

- Sörberg Wallin, A., Zeebari, Z., Lager, A., Gunnell, D., Allebeck, P., & Falkstedt, D. (2018). Suicide attempt predicted by academic performance and childhood IQ: a cohort study of 26 000 children. *Acta Psychiatrica Scandinavica*, *137*(4), 277–286.
- Taylor, E. S. (2023). Chapter 2 - Teacher evaluation and training. I E. A. Hanushek, S. Machin, & L. Woessmann (Red.), *Handbook of the Economics of Education* (Vol. 7, s. 61–141). Elsevier. <https://doi.org/10.1016/bs.hesedu.2023.03.002>
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. New Zealand Ministry of Education.
- Torgerson, D., & Torgerson, C. (2013). *Randomised trials in education: An introductory handbook*. https://www.researchgate.net/publication/273421357_Randomised_trials_in_education_An_introduutory_handbook
- Utbildningsdepartementet. (2024). *Regleringsbrev för budgetåret 2024 avseende Statens skolverk. Dnr U2023/01741, U2024/01246*. <https://www.esv.se/statsliggaren/regleringsbrev/Index?rbId=24519>
- Van den Brande, J., & Zucollo, J. (2021). *The cost of high-quality professional development for teachers in England*. Wellcome Trust & Education Policy Institute. https://epi.org.uk/wp-content/uploads/2021/07/2021-Cost-of-quality-teacher-cpd_EPI.pdf
- Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, *15*, 17–40.
- Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education*, *61*, 47–59. <https://doi.org/10.1016/j.tate.2016.10.001>
- von Hippel, P. T. (2024). Multiply by 37 (or Divide by 0.027): A Surprisingly Accurate Rule of Thumb for Converting Effect Sizes From Standard Deviations to Percentile Points. *Educational Evaluation and Policy Analysis*, *01623737241239677*. <https://doi.org/10.3102/01623737241239677>
- Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting With Teacher Professional Development: Motives and Methods. *Educational Researcher*, *37*(8), 469–479.
- Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession*. National Staff Development Council. <https://www.academia.edu/download/32304170/NSDCstudytechnicalreport2009.pdf>

Bilaga 1: Inkluderade forskningsöversikter avseende grundskola och gymnasium

Effekt på undervisning	Årtal	Publikationstyp	Analyserade studier	Effektstorlek	Fokus	Referens	
1	Garrett et al 2019	2019	Tidskriftsartikel	40	0.42	Vilket skolämne som helst	https://doi.org/10.3102/0091732X19830634
2	Gonzalez et al 2022	2022	Working paper	26	0.57	STEM (NO, teknik, ingenjörskonst och matematik)	https://doi.org/10.26300/d9kc-4264
3	Kahmann et al 2022	2022	Tidskriftsartikel	14	0.63	Differentierad undervisning	https://doi.org/10.1016/j.ijer.2022.102072
4	Kraft et al 2018	2018	Tidskriftsartikel	43	0.49	Coaching	https://doi.org/10.3102/0034654318759268
5	Rice et al 2024	2024	Tidskriftsartikel	13	1.11	Läsförståelse	https://doi.org/10.1002/rrq.546
<i>Genomsnitt</i>		<i>2021</i>		<i>27.2</i>	<i>0.64</i>		
<i>Min</i>		<i>2018</i>		<i>13</i>	<i>0.42</i>		
<i>Max</i>		<i>2024</i>		<i>43</i>	<i>1.11</i>		

Effekt på elevresultat	Årtal	Publikationstyp	Analyserade studier	Effektstorlek	Fokus	Referens	
1	Balta et al 2023	2023	Tidskriftsartikel	19	0.523	Vilket skolämne som helst	https://doi.org/10.58693/ier.111
2	Basma & Savage 2018	2018	Tidskriftsartikel	17	0.225	Läsning	https://doi.org/10.1007/s10648-017-9416-4
3	Basma & Savage 2023	2023	Tidskriftsartikel	14	0.062	Läsning, middle & high school	https://doi.org/10.1177/00224871231153084
4	Didion et al 2020	2020	Tidskriftsartikel	28	0.18	Läsning (elementary & middle school)	https://doi.org/10.1080/19345747.2019.1670884
5	Fletcher-Wood & Zuccollo 2020	2020	Rapport	42	0.09	Vilket skolämne som helst	https://epi.org.uk/publications-and-research/effects-high-quality-professional-development/
6	Kahmann et al 2022	2022	Tidskriftsartikel	17	0.08	Differentierad undervisning	https://doi.org/10.1016/j.ijer.2022.102072
7	Kraft et al 2018	2018	Tidskriftsartikel	31	0.18	Coaching	https://doi.org/10.3102/0034654318759268
8	Lynch et al 2019	2019	Tidskriftsartikel	95	0.21	STEM (NO, teknik, ingenjörskonst och matematik)	https://doi.org/10.3102/0162373719849044
9	Pellegrini et al 2021	2021	Tidskriftsartikel	10	0.03	Matematikinhåll och -pedagogik	https://doi.org/10.1177/2332858420986211
10	Pellegrini et al 2021	2021	Tidskriftsartikel	7	0.19	Pedagogik (t ex kooperativt lärande)	https://doi.org/10.1177/2332858420986211
11	Pellegrini et al 2021	2021	Tidskriftsartikel	12	0.01	Implementera undervisningsmaterial	https://doi.org/10.1177/2332858420986211
12	Rice et al 2024	2024	Tidskriftsartikel	24	0.193	Läsförståelse	https://doi.org/10.1002/rrq.546
13	Sims et al 2023	2023	Tidskriftsartikel	104	0.05		https://doi.org/10.3102/00346543231217480
<i>Genomsnitt</i>		<i>2021</i>		<i>32.3</i>	<i>0.16</i>		
<i>Min</i>		<i>2018</i>		<i>7</i>	<i>0.01</i>		
<i>Max</i>		<i>2024</i>		<i>104</i>	<i>0.523</i>		

Bilaga 2: Inkluderade forskningsöversikter avseende förskola

	Effekt på undervisning	Årtal	Publikationstyp	Analyserade studier	Effektstorlek	Fokus	Referens
1	Lee & Sung 2023	2023	Tidskriftsartikel	35	0.62	Interaktionskvalitet	https://doi.org/10.1016/j.tate.2023.104017
2	Egert et al 2020	2020	Tidskriftsartikel	15	0.39	Interaktionskvalitet/CLASS	https://doi.org/10.1016/j.edurev.2019.100309
3	Egert et al 2018	2018	Tidskriftsartikel	36	0.68	Undervisningskvalitet	https://doi.org/10.3102/0034654317751918
4	Markussen-Brown et al 2017	2017	Tidskriftsartikel	30	0.59	Undervisningsprocesser	https://doi.org/10.1016/j.ecresq.2016.07.002
	<i>Genomsnitt</i>	2020		29.0	0.57		
	<i>Min</i>	2017		15	0.39		
	<i>Max</i>	2023		36	0.68		

	Effekt på elevutfall	Årtal	Publikationstyp	Analyserade studier	Effektstorlek	Fokus	Referens
1	Brunsek et al 2020	2020	Tidskriftsartikel	3	0.14	Expressiv vokabulär	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	3	0.47	Närhet/socioemotionell	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	3	0.12	Uppmärksamhet/hyperaktivitet	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	4	0.43	Internatisera problem	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	6	0.52	Problembeteende	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	4	0.24	Sociala färdigheter	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	4	0.54	Expressiv vokabulär	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	8	0.3	Bokstavskunskap	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	3	0.24	Textmedvetenhet	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	7	0.06	Receptivt språk	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	3	0.18	Rimmedvetenhet	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	5	0.37	Fonemisk medvetenhet	https://doi.org/10.1016/j.ecresq.2020.03.003
	Brunsek et al 2020	2020	Tidskriftsartikel	4	0.28	Skrivförmåga	https://doi.org/10.1016/j.ecresq.2020.03.003
2	Egert et al 2018	2018	Tidskriftsartikel	9	0.14	Språk, beteende, etc	https://doi.org/10.3102/0034654317751918
3	Markussen-Brown et al 2017	2017	Tidskriftsartikel	5	0.21	Receptiv vokabulär	https://doi.org/10.1016/j.ecresq.2016.07.002
	Markussen-Brown et al 2017	2017	Tidskriftsartikel	9	0.3	Fonologisk medvetenhet	https://doi.org/10.1016/j.ecresq.2016.07.002
	Markussen-Brown et al 2017	2017	Tidskriftsartikel	11	0.12	Bokstavskunskap	https://doi.org/10.1016/j.ecresq.2016.07.002
4	Jensen & Rasmussen 2019	2019	Tidskriftsartikel	8	0.35	Språk, beteende, etc	https://doi.org/10.1080/00313831.2018.1466359
5	Joo et al 2020	2020	Tidskriftsartikel	9	0.26	Språk, beteende, etc	https://doi.org/10.1080/10409289.2019.1624146
	<i>Genomsnitt</i>	2019		5.7	0.28		
	<i>Min</i>	2017		3	0.06		
	<i>Max</i>	2020		11	0.54		