

Evidensbasert Undervisning - Eit Filosofisk Forsvar

Steinar Bøyum

Abstract

In the debate about evidence-based teaching one often encounters philosophical arguments. Such arguments are characterized by being principled and general, in the sense of not resting on whether meta-analyses or randomized controlled experiments are well or poorly executed. In this article I look at several such philosophical arguments against evidence-based teaching and the relevance of RCT's for education: instrumentalism, normativity, judgment, contextuality, complexity, and identification. I conclude that none of these arguments are valid. This does not mean that evidence-based education is unproblematic, but it does mean that the best arguments are practical rather than philosophical.

Key words: Evidence-based teaching, RCT, instrumentalism, contextuality, philosophy of science

Evidensbasert undervisning

- eit filosofisk forsvar

Debatten om evidensbasert praksis har rast i fleire år. Han verken er eller bør vere over. Tematikken rører ei rekke sentrale spørsmål om profesjonar, politikk, kunnskap og vitskap. Internasjonalt er det også mykje litteratur som diskuterer evidensbasert *undervisning*. Her verkar det som om pedagogar, i alle fall i Noreg, stort sett har vore skeptiske, kritiske eller negative til evidensbasering. Så medan evidensbasert praksis har blitt ein dominerande diskurs innan til dømes medisin, så er det snarare opposisjonen som dominerer innan pedagogikkfaget, trass eit visst press ovanfrå (Grossman & Leach, 2008).

På mange måtar har evidensbasering vore offisiell politikk for utdanningsforskning i Noreg i fleire år allereie. Her har Noreg gått i spora til OECD, som i lengre tid har vore opptatt av at utdanningsforskning skal bidra til evidensbasert praksis (“EBP”) (OECD, 2003).¹ Sant nok skjuler regjeringa seg litt bak OECD, men ut frå konteksten er det tydeleg at regjeringa i sin strategi for utdanningsforskning går inn for “evidence-based policy and practice, der ein gjennom systematiske kunnskapsoversikter får overblikk over til dømes kva som verkar og ikkje verkar på utdanningsfeltet” (Kunnskapsdepartementet, 2008, s. 192). Vidare vil regjeringa prioritere “effektforskning, forskning basert på longitudinelle studier og kvantitativ forskning” (s. 195), samt “vidareføre samarbeidet med Campbell Collaboration” (s. 196). Etableringa av eit nasjonalt kunnskapssenter for utdanning føyer seg inn i dette mønsteret.

¹ I USA har dette, ikkje uventa, gått endå lengre (Walters, Lareau, & Ranis, 2009).

Sjølvsagt om det er uklart i kva grad det vil prege senteret sin aktivitet, så er det rimeleg å anta at etableringa i det minste er inspirert av evidensbevegelsen.

I denne artikkelen skal eg sjå på evidensdebatten i eit filosofisk lys. Meir spesifikt skal eg ta for meg ei rekke filosofiske argument som har vore retta mot evidensbasert undervisning (“EBU”). Det primære målet er å vise at ideen om EBU ikkje er spesielt problematisk filosofisk sett. Det betyr ikkje at EBU er uproblematisk generelt sett, men det betyr at kritikk må kome frå ein annan stad enn filosofien, til dømes frå praktiske eller politiske argument. Indirekte håpar eg også å bidra til ein meir sakleg debatt enn det me har hatt til no. Evidensdebatten har uheldigvis vore prega av rask retorikk snarare enn grundig analyse. Særleg er det utstrakt bruk av stråmannsargumentasjon (Mayer, 2001). I denne artikkelen vil eg derfor prøve å følgje eit hermeneutisk barmhjertighetsprinsipp, og drøfte EBU på sitt mest edruelege, ikkje mest ekstreme.

Med “filosofiske” argument meiner eg ikkje noko mystisk. Her betyr det berre at det er snakk om *prinsipielle* argument, og ikkje argument som kviler på om systematiske kunnskapsoversikter eller randomiserte kontrollerte eksperiment (“RCT”) blir godt eller dårleg utført. Eit filosofisk argument seier altså at sjølv i sin mest vellykka form har ideen om EBU grunnleggande svakheiter som gjer han pedagogisk inadekvat eller irrelevant. Dette betyr at eg ikkje skal gå detaljert inn på spesifikt metodologiske spørsmål, til dømes om RCT.² Alle metodar har sine sterke og svake sider, men slike ting kan eg ikkje gå inn på her. Ein må gjerne kritisere dårlege RCT-design eller peike på utfordringar med å gjere ein RCT, men dette aleine utgjer ingen filosofisk eller prinsipiell kritikk av EBU eller RCT.

² Ein liknande strukturert artikkel som går nærare inn på metodiske ting, er Cook & Payne (2002).

Sidan evidensdebatten involverer så mange ulike spørsmål, er det viktig å avgrense seg. Eg skal i det følgjande ikkje gå inn på historia til evidensbevegelsen (Daly, 2005). Eg skal heller ikkje skal gå inn på dei politiske sidene, til dømes den rolla som dedikerte organisasjonar spelar (Campbell, EPPI, Dansk Clearinghouse mfl.). Vidare skal eg nøye meg med å sjå på EBP i eit pedagogisk lys, altså EBU, om enn det kan vere at nokre av poenga mine også kan gjelde evidensbasert praksis generelt. Eg skal heller ikkje ta for meg meir omfattande spørsmål om kausalitet i samfunnsvitskapane, som særleg krinsar rundt statistisk analyse (Bhrolcháin, 2001; Humphreys, 1986; Russo, 2008). Sidan evidensdebatten er av allmenn interesse ser eg det som ein eigenverdi å unngå for mange teknikalitetar. Og til sist skal eg ikkje diskutere dei som kritiserer evidensbevegelsen for det den *ikkje* er oppteken av, til dømes at den berre ser på kva som verkar og ikkje kvifor. Bortsett frå at slik kritikk ofte berre rammar stråmenn, blir det i dagens spesialiserte forskingsverd stadig meir problematisk å kritisere nokon for det dei ikkje seier noko om. Ingen kan seie alt, og at ein går inn for X (t.d. RCT) betyr ikkje at ein går imot Y (t.d. intervjustudiar).

Kva betyr evidensbasering?

For å få ein fornuftig diskusjon om EBP, må ein først avklare kva ein legg i omgrepet. Ein snever definisjon er ikkje hensiktsmessig her. Dersom ein definerer “evidensbasert” til å bety “basert på akkurat den typen systematiske kunnskapsoversikter som Cochrane-samarbeidet lagar”, så vil det skyve fokus over på karakteren av desse oversiktene i seg sjølve i staden for på kva dei er oversikter *over*, altså karakteren til primærstudiane som inngår i meta-analysen. Det er ikkje primært oversiktene som skapar kontroversane, men kva forskning som får vere med i desse. Det er med andre ord det underliggende synet på kva som er gangbar forskning det står strid om.

Ein vid definisjon er også lite hensiktsmessig i denne konteksten. Dersom ein definerer “evidensbasert” til å berre bety “basert på kunnskap” eller “influert av forskning”, tilslører ein dei kontroversane som EBP er omgitt av. I andre kontekstar kan det vere verdifullt å definere omgrepet vidt, til dømes for å minne om at evidenskritikarar ikkje er imot at profesjonell praksis skal vere kunnskapsbasert og at evidensbevegelsen sitt kunnskapsomgrep ikkje er det einaste mulige.³ Men i ein artikkel som denne, som skal diskutere det filosofisk kontroversielle med EBP, er ein slik vid definisjon ufruktbar, for det som gjer EBP kontroversielt, er ikkje at praksis skal vere basert på forskning eller kunnskap, men på ein bestemt *type* forskning og kunnskap.

Medan det er ueinigheit om kva som eventuelt er nytt i og med evidensbevegelsen, er det liten tvil om kva som er mest kontroversielt, nemleg det såkalla evidenshierarkiet, der randomiserte kontrollerte feltforsøk blir rekna som den metoden som gjer mest solid kunnskap om kva tiltak som verkar best (Grimen, 2009; Lipworth, Carter, & Kerridge, 2008). Debatten om EBU har derfor i stor grad blitt til ein debatt om relevansen av RCT for pedagogisk forskning og i kva grad resultat frå slike skal leggest til grunn for pedagogisk praksis. “Evidensbasert” (og “vitskapeleg”) fungerer gjerne som “code words” for RCT, og i det følgjande skal eg ofte bruke nemninga “EBU/RCT” for å framheve dette (Berliner, 2002, s.18).⁴ Dersom ein lausriver EBU frå RCT, forsvinn mykje av sprengkrafta i tematikken, noko som ikkje er hensiktsmessig her, sjølv om formålet mitt er å demontere snarare enn å detonere dette sprengstoffet. “Evidensbasert undervisning” skal derfor i denne artikkelen bety undervisning

³ Ein vid definisjon kan også vere fagpolitisk klokt. Dette ser me i definisjonen av EBP som til dømes Norsk Psykologforening (og APA) har vedtatt: “Evidensbasert psykologisk praksis (EBPP) er integrering av den beste tilgjengelige forskningen med klinisk ekspertise sett i sammenheng med pasientens egenskaper, kulturelle bakgrunn og ønskemål” (Norsk Psykologforening, 2007)

⁴ Den nære samanhengen mellom kritikk av EBU og kritikk av relevansen av RCT for undervisning, er også sentral i to viktige bøker om temaet (Smeyers, 2009; Walters et al., 2009). Samtidig er fleire av argumenta me skal sjå på av ein slik art at dei også vil ramme eit prosjekt som John Hatties, sjølv om han ikkje legg like stor vekt på RCT som mange andre i evidensbevegelsen (2009).

som (i større grad enn no) baserer val av metodar og program på resultat frå forskning, primært på RCTs (og liknande design) og systematiske oversikter over slike.

Evidens og positivisme

Evidensdebatten kan ofte minne om den gamle positivismestriden (Cook & Payne, 2002, s. 154-156). Mange av dei same argumenta dukkar opp igjen, og vil også gjere det her. Likevel skal eg ikkje seie så mykje om positivisme. Det er fleire grunnar til det. For det første blir “positivisme” ofte brukt i ei så vag tyding at det er vanskeleg å vite korleis ein skal respondere. Merkelappen har då i realiteten lite meir enn ein retorisk funksjon, nærare bestemt som skremmebilde, og forvirrar meir enn han klargjer (Oakley, 2001). Det vil derfor gagne debatten om ein alltid presiserer kva type positivisme ein snakkar om, eller endå betre, refererer direkte til den epistemologiske eller vitskapsteoretiske påstanden ein er ute etter. Aller best ville det vore om termen vart pensjonert.⁵

For det andre er det heller ikkje opplagt at EBU/RCT “føreset” eller “kviler på” positivisme. Ein av dei skarpaste evidenskritikarane i Noreg er Tor-Johan Ekeland, og influert av Hans Skjervheim reknar han det som utvilsamt at evidensbevegelsen kviler på eit positivistisk kunnskapssyn.

I 1977 skreiv Hans Skjervheim essayet “Det instrumentalistiske mistaket” der han kritiserte pedagogikken for å sjå danningprosessar som noko ein kan måle og kontrollere på same måte som tekniske prosessar i industrien. Det var ein pedagogikk som på den tida var på god veg inn i ei evidensbasert blindgate. Dagens situasjon er langt verre. Forkledd som “evidensbasert kunnskap” blomstrar nypositivismen. (2007, s. 10)

⁵ Michael Friedman viser kor stor avstanden er mellom det ein i akademia kallar “positivisme” og det positivistane faktisk stod for (1999).

I likskap med Harald Grimen er eg skeptisk til om denne samankoplinga er fruktbar (Grimen, 2009, s. 194-198). Eit positivistisk vitenskapssyn, kva enn det no er, treng ikkje implisere støtte til EBU, og *vice versa*. Dei er logisk uavhengige av kvarandre. Det er heller ikkje nokon opplagt samanheng mellom RCT og positivisme.⁶ På same måte som tilhøvet mellom teori og praksis er tilhøvet mellom vitenskapsfilosofi og metodologi logisk sett altfor komplisert til at ein kan fange det inn med uttrykk som “føreset” eller “basert på”. Det er riktig at ein metodologi, som alt anna, har filosofiske føresetnader, men det er ikkje det same som å seie at ein metodologi føreset ein *bestemt* filosofi, for filosofien er ikkje så fattig at ikkje motstridande filosofiar kan ha ting felles. Sagt med andre ord, ein kan vere tilhengjar av EBU og vere noko heilt anna enn ein positivist, og ein kan vere positivist og vere noko heilt anna enn ein tilhengjar av EBU.

Eg vil påstå at EBU er meir epistemologisk beskjeden enn motstandarane går ut frå. Det utelukkar ikkje at enkeltforskarar kan gi ein meir epistemologisk lada utlegning av EBU, men i seg sjølv treng ikkje EBU kome med særleg tung filosofisk bagasje. Kjernen i EBU er for det første at det innan utdanning ofte er nyttig å finne ut kva som verkar best og for det andre at RCT (og systematiske oversikter over slike) er den beste måten å finne det ut på. Verre er det ikkje, og dette treng ikkje tyde på positivisme, men kan tvert imot grunngjevast ut frå ei rekke forskjellige vitenskapsfilosofiske retningar, noko sentrale figurar som Nancy Cartwright (Cartwright & Stegenga, 2008), David H. Hargreaves (1999) og Donald Campbell (1987) er bevis på.

⁶ Unntaket er naturlegvis dersom ein stipulativt definerer “positivisme” på ein måte som impliserer støtte til RCT, men i så fall blir også påstanden om ein samanheng tautologisk.

Instrumentalismeargumentet

Typisk for Skjervheim og positivismestriden, og som altså for nokon lever vidare i evidensdebatten, er dei moralske undertonane. Positivisme er i Skjervheim sine skrifter ikkje berre ontologisk, men *moralsk* suspekt. Innvendinga om at EBU er “instrumentalistisk” går i samme toneart (Biesta, 2007; Elliott, 2001). Dessverre er det ikkje lett å formulere denne innvendinga presist.

Kva betyr “instrumentalisme”? Kjernen i omgrepet synest vere ein tilstand der ein har klart definerte mål for ein gitt aktivitet, og at desse måla måla er eksterne i høve til aktiviteten.

Forskjellen mellom interne og eksterne mål kan ein illustrere med eit eksempel: Å score er eit internt mål i fotball, medan å tene pengar er eit eksternt mål. Interne mål er altså konstitutive for ein viss praksis og kan ikkje lausrivast frå denne, slik eksterne mål kan. Instrumentalisme er når eksterne mål dominerer, slik at interne mål står i fare for å bli til middel til dei eksterne måla, og har altså å gjere med kva grunnar ein har for å utføre ein viss aktivitet eller delta i ein viss praksis.⁷

Allereie her ser me likevel eit problem med å bruke omgrepet instrumentalisme kritisk. Når ein tenkjer seg om, er svært mange av livets gjeremål *både* instrumentelle og ikkje-instrumentelle, med unntak av ting som å pusse tennene, som for dei fleste er reint instrumentelt. Det er heller ikkje slik at bestemte aktivitetar *er* instrumentelle eller ikkje-instrumentelle. Det som for nokon er reint instrumentelt, til dømes å trene eller å gå på jobb, er for andre autotelisk, dvs. mål i seg sjølv. Og det som på eit tidspunkt er autotelisk, kan på

⁷ I nyare utdanningspolitiske samanhengar betyr ofte instrumentalisme at ein er for opptatt av økonomiske omsyn, til dømes at utdanning skal bidra til nasjonens konkurransevne i den globale økonomien. Men denne betydningen er ikkje så relevant i denne filosofiske konteksten, sjølv om det kan godt vere at fokuset på kvalitet og “what works” verkeleg er instrumentalistisk på denne måten.

eit anna bli instrumentelt, og omvendt. Slik forstått er ikkje instrumentalisme automatisk moralsk problematisk.

Instrumentalisme blir moralsk problematisk først når visse *sosiale* aktivitetar blir instrumentelle. Då behandlar ein andre menneske som middel, og innanfor eit kantiansk moralsyn er dette galt fordi ein då behandlar dei som om dei var ting. Dette er kjernen i instrumentalismeargumentet: EBU involverer ei *objektivering* av elevane (Ekeland, 2009, s. 148). Med dette meiner ein ikkje primært at ein objektiverer elevane i sjølve forskinga, men at undervisninga som skal bygge på forskingsresultata inneber ei slik objektivering. Tanken synest å vere at dersom ein har klart definerte mål for undervisninga, samt førehandbestemte handlingsreglar for korleis ein skal nå desse måla (dvs. undervisningsmetodar som med god sannsynlegheit verkar), så gjer ein elevane om til ting. Men her er det ein fundamental ambivalens, som det ikkje er lett å kvitte seg med, og som har sine kjelder i Skjervheim sine egne tekstar.

I følge Skjervheim er det ein tendens til at ein misforstår pedagogisk praksis som pragmatiske eller tekniske handlingar (Skjervheim, 1996). Dette er handlingar der ein har eit mål, og i tillegg eit sett tekniske reglar, fortrinnsvis basert på vitenskapleg kunnskap, som gjev retningslinjer for korleis ein kan nå målet. Slike tekniske handlingar er vellukka når ein på denne måten når det målet ein har sett seg. Motstykket til dette er praktiske handlingar, der ein handlar, ikkje ut frå tekniske reglar, men moralske normer. Ein handlar ovanfor andre på den måten som er riktig, ikkje på den måten som er effektiv. Å forstå (eller misforstå) pedagogikk som ein teknisk vitenskap, er då å tenkje seg at den pedagogiske vitenskapen skal finne fram til tekniske reglar for undervisning. Dette er å gjere eit instrumentalistisk mistak. Ein behandlar ein praksis som om den var ein teknologi.

Men kvifor tenkjer Skjervheim seg at “teknisk” og “praktisk” er motsetningar? Kvifor skulle ikkje dei tekniske reglane ein handlar etter vere i samsvar med moralske normer? Tanken om at dei er motsetningar kjem frå følgjande føresetnad: Når ein handlar teknisk i høve til nokon, behandlar ein dei som middel, og dette er moralsk galt. Den som ein handlar i relasjon til, blir redusert til middel i eins eigne mål/middel-kalkylar, i staden for autonome personar som sjølve har formål med det dei gjer. Den tekniske forma for pedagogisk forskning, som undersøker slike mål/middel-relasjonar, lir då under den moralske bresten at den behandlar personar som objekt.

Det er ikkje vanskeleg å sjå at Skjervheim sitt argument er relevant for spørsmålet om EBU. Uheldigvis er det sentrale uttrykket, “å behandle nokon som middel”, ambivalent, noko som fører Skjervheim, og evidenskritikarar som tenker på same måten, opp i eit alvorleg dilemma.

Den eine betydningen er at det er galt å behandle nokon som middel *überhaupt*. Det blir då galt å ha bestemte mål for ein aktivitet ein utøver saman med andre, til dømes undervisning, og indirekte blir då også pedagogisk mål/middel-forskning moralsk suspekt. Dette standpunktet er verkeleg radikalt, men også heilt urimeleg. Foreldre har til dømes fleire mål med måten dei oppdrar og omgår barna sine på: At dei skal ta omsyn til andre, lære om verda rundt seg, bli glade og oppfinnsame i leik, kunne lese og rekne, knyte skoa sine sjølve osv. Det er vanskeleg å sjå kva som kan vere galt i at ein prøver å finne gode middel til å nå desse måla (gitt, naturlegvis, at ein følgjer vanlege moralske reglar). Det same gjeld skulen. I læreplandokument finn ein uttrykt ei rekke mål for opplæringa: Lære å lese, kunne gangetabellen, lære å symje, tenkje kritisk, ikkje mobbe osv. Kva metodar og tiltak er mest effektive for å nå desse måla? Det er vanskeleg å tru at Skjervheim meinte at det er moralsk

illegitimt å forske på slike spørsmål, men han uttrykte seg ofte som om han meinte det, og han blir ofte gjengitt som om det var det han meinte.

På dette punktet pleier gjerne forsvararar av Skjervheim å trekke seg tilbake. Det var naturlegvis ikkje dette han meinte, seier dei. Og då kjem den andre tydinga inn: Det er galt å behandle nokon *berre* som middel. Dette leier til eit langt meir moderat standpunkt, som Skjervheim indikerer når han seier: “Den pragmatiske, instrumentelle fornuft har sin rimelege plass, også i pedagogikken. Men den har og sine grenser ...” (1996, s. 248).⁸ Ein annan stad skriv han at det fins “teknikk-liknande spørsmål innan pedagogikken”, men at den tekniske modellen ikkje kan vere “grunnmodellen” (1992, s. 216). I denne tolkinga er det altså *reindyrka* instrumentalisme som er moralsk galt: Å drive mål/middel-forskning utan å bry seg om kvaliteten på måla og legitimiteten av midla *i det heile*, kan vere moralsk problematisk. Dette synet er adekvat nok, men også heilt ukontroversielt, og den kritiske brodden blir då borte.

Problemet er dermed at Skjervheim sin retorikk spelar opp til eit radikalt (men urimeleg) standpunkt, men at poenget hans ofte forfell til eit moderat (men ukontroversielt) standpunkt. Det me ikkje kan få til med dette argumentet, er ein kritikk som berre treff den typen pedagogikk ein ikkje likar, til dømes EBU/RCT. Enten meiner evidenskritikaren at det å behandle nokon som middel *überhaupt er* galt, men då er det svært mykje av det me gjer galt, eller så meiner han at det å behandle nokon *berre* som middel er galt, og i så fall treng det heller ikkje vere noko galt med verken EBU eller positivistisk pedagogikk.⁹ Det er nemleg

⁸ Som John L. Austin sa, “There’s the bit where you say it and the bit where you take it back” (1962, s. 2).

⁹ Føresetninga er naturlegvis at ein følgjer alminnelege moralske reglar, men dette gjeld også “ikkje-evidensbasert” undervisning og er heilt uavhengig av kva type forskingsmetode som ligg under (dessuten kan evidensforskning sjølvsagt ende opp med å tilrå metodar som allereie er i bruk).

ingenting ved EBU som tilseier at ein skal behandle elevar utelukkande som middel, like lite som refleksjon over mål og middel i opplæringa av eins egne barn tilseier det. Å setje seg mål for det ein gjer, og så leite etter best mulige måtar å nå det målet på, er daglegdagse aktivitetar som ikkje er filosofisk lada på nokon bestemt måte, og dette gjeld både “interne” og “eksterne” mål. Filosofisk sett er EBU ikkje anna enn ei ekstrapolering og systematisering av slik kvardagsinstrumentalisme.

Normativitetsargumentet

Det neste argumentet me skal sjå på, nyttar også eit moralsk poeng, men på ein litt annan måte: “Om noe virker eller ikke, om noe har “god effekt” eller ikke, er normative spørsmål” (Ekeland, 2009, s. 147). Den mest djuptgripande bruken av dette argumentet finn me hjå David Carr.¹⁰

Carr (2001) set fleire argument i spel mot EBU, men det mest grunnleggande poenget hans er at undervisning er normativt eller verdilada. Kva syn ein har på utdanning og undervisning involverer moralske oppfatningar om kva eit godt liv er, kva menneskeleg vekst er, kva eit godt samfunn er osv. Og slike spørsmål er det ueinigheit om, både i og mellom samfunn. Ein kan derfor ikkje svare på dei med empirisk forskning. Forskjellen mellom eit tradisjonelt og eit progressivt syn på utdanning er til dømes ikkje primært ei ueinigheit om undervisningsmetodar sin kausale effektivitet, men ei normativ ueinigheit med både moralske og politiske aspekt. Slike konflikhtar kan ein ikkje avgjere “ein gong for alle” med hjelp av vitskapelege metodar, til dømes kontrollerte forsøk. Problemet med EBU, seier Carr, er at ein behandlar som teknisk og empirisk noko som i grunnen er moralsk og normativt.

¹⁰ Andre evidenskritikarar, til dømes Biesta (2007, s. 9-10), kviler også på Carrs argumentasjon.

Dette normativitetsargumentet involverer to feilslutningar, ein om ueinigheit og ein om normativitet. For det første kan me vere einige om meir underordna mål sjølv når me er ueinige om meir overordna mål (og *vice versa*). Me kan til dømes vere ueinige om kva det overordna målet for utdanning bør vere, men einige om at det inkluderer å kunne ein del historie. For det andre betyr ikkje det at eit område er strukturert rundt normative omgrep at det same området ikkje kan involvere mange spørsmål som sjølv ikkje er normative. Som Peters trakk fram i si klassiske analyse er sjølve utdanningsomgrepet normativt lada, men gitt at utdanning inkluderer det å kunne lese, kan det vere eit faktisk spørsmål om nokon kan lese og korleis dei (best) lærer seg det (Peters, 1967, s. 6-8). Oppsummert betyr dette at ein kan vere einige om visse *konstituentar* i utdanning og så forske i kva som er dei beste midla for å realisere desse.

Det er forunderleg at Carr ikkje vurderer dette, all den tid han i ein annan kontekst gir oss eit eksempel på nettopp ein normativ praksis der ein kan ha vitskapeleg belegg for val av metodar, nemleg odontologi (2001, s. 471). Både medisin og odontologi er normativt strukturerte felt, fordi sentrale omgrep som helse, sunnhet og sjukdom er verdilada. Men dersom det fins slike felt som er normative og der forskning, til og med i form av kontrollerte forsøk, kan spele ei avgjerande rolle, kvifor ikkje utdanning? Carr tenkjer seg kanskje at det er *vanskelegare* å kome fram til slike evidensbaserte metodar innan utdanningsfeltet, og det kan godt vere, for det er mange forskjellar mellom elev og pasient og mellom skule og sjukehus, men for det første er det vanskeleg å vite det før me har prøvd, og for det andre undergrev Carr då sitt eigentlege argument, nemleg at det er *prinsipielt* galt med ei teknisk-vitskapeleg forståing av eit normativt lada område.

I likskap med Skjervheim vedgår Carr at det fins enkelte tekniske spørsmål innan pedagogikk (Carr, 2001, s. 470). Han vedgår også at det kan finnast visse sanne metodiske generaliseringar for undervisning, men han hevdar at desse er trivielle og ikkje treng vitskapeleg støtte (“snakk til elevane, ikkje til tavla”). Men dersom Carr sin påstand om at utdanning er eit normativt strukturert felt skal nyttast til å kritisere ei teknisk forståing av utdanning, så burde ein anta at alle generaliseringar var like problematiske. Dette viser at Carr når alt kjem til alt ikkje finn noko prinsipielt galt med slike generaliseringar på utdanningsfeltet, sjølv om dette feltet er normativt strukturert. Det reelle poenget hans ser snarare ut til å vere at det ikkje er så *mange* slike generaliseringar. Men då er spørsmålet korleis me veit det. Dersom det fins nokre slike sanne generaliseringar, kan ein ikkje fastslå *a priori* at det ikkje fins fleire.

Carr har rett både i at det fins overordna spørsmål som ein ikkje kan avgjere vitskapeleg, fordi dei er normative, og i at det fins underordna spørsmål som ein ikkje treng avgjere vitskapeleg, fordi dei er trivielle. Men han hopper bukk over det store området mellom desse to kategoriane, der ein finn spørsmål som er faktuelle, men langt frå banale. Det er her me treng forskning, og det er her evidensforsvararen held seg. Her finn me spørsmål om forskjellen mellom “phonics” og “whole language” i leseopplæring og mellom store og små klassestørrelsar, for å nemne to debattar som det har vore mykje forskning på, også i form av RCTs (Mosteller, 1997; Torgerson, 2006). For dei fleste av oss er ikkje dette moralske spørsmål, men spørsmål om “what works”. Ein skal ikkje overdrive og seie at alle pedagogiske spørsmål er av denne typen, men ein skal heller ikkje underdrive og seie at ingen er det.¹¹

¹¹ Å hevde, slik Biesta gjer, at evidenstillhengjarar *berre* er opptatt av “what works” og at dei reduserer *alle* pedagogiske spørsmål til effektivitetsspørsmål, utan omsyn til om midla er moralsk forsvarlege eller ikkje, er nettopp den typen karikering av motstandarane ein må unngå (2007, s. 6).

Skjønnsargumentet

Utdanningsforbundet er naturleg nok mest opptatt av kva EBU har å seie for lærarane. Dei fryktar at “evidensbasert praksis [kan] true det profesjonelle skjønnen” (Utdanningsforbundet, 2008, s. 7). Utvikling av klare reglar for metodeval i undervisning vil leie til manualisering som undergrev lærarens autonomi. Den same frykta artikulere Kjetil Steinsholt: Evidensbasering vil føre til at læraren si rolle blir passivt og mekanisk å implementere det forskinga har fastsett som effektive undervisningsmetodar (Steinsholt, 2009).

Strengt tatt er dette argumentet dødfødt. Ein forsvarar av EBU treng ikkje nekte for tydinga av skjønn, endå mindre ville gjere skjønnen overflødig. Ein av dei fremste EBU-forkjemparane, David Hargreaves, legg til dømes vekt på at evidens, altså resultat frå systematiske kunnskapsoversikter med vekt på RCTs, ikkje står i motsetning til skjønn, men inngår som ein ekstra komponent i grunnlaget for utøvinga av skjønn (1999, s. 246; 1997, s. 409). Dette er sentralt også i hyppig siterte definisjonar av EBP generelt (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996, s. 71). Å klage evidensbevegelsen for å ville manualisere undervisning og gjere lærarar til maskiner er derfor ein type karikaturbasert kritikk som ein bør unngå. Som Grimen poengterer:

Ingen har, så langt jeg har sett, argumentert for at målet med evidensbasering er å eliminere skjønn i profesjonsutøvelsen. Evidenstenkningen er ett av flere forslag til hvordan profesjonelt skjønn kan gjøres mer tilregnelig og ansvarlig. Det er snakk om å komplettere eller informere skjønn. (Grimen, 2009, s. 211)

Men ønskjer ikkje evidenstillhengjaren i det minste å *reducere* bruken av skjønn? Det kan her vere nyttig å dra eit analytisk skilje mellom to ulike roller som omgrepet skjønn spelar. For det første som noko me gjer bruk av når det ikkje føreligg reglar, noko som me nyttar så å seie i mangel av noko betre. Av denne typen skjønn er det ein undertype som evidenstillhengjaren ønskjer å redusere bruken av, nemleg fordommar (Kvernbekk, 2001). For det andre som noko me gjer bruk av når me skal anvende reglar. Ein har reglar, men må anvende dei med kløkt og forstand i den situasjonen ein står i. Poenget er at sjølv om ein skulle redusere behovet for den første typen skjønn, så er den andre typen uunngåeleg, noko som kan motivere ei oppfatning om at alle praktiske reglar er tommelfingerreglar. Men at skjønn er uunngåeleg bør ikkje få oss til å slutte at me ikkje treng reglar i det heile. Tvert imot, skjønn i denne forstand føreset eksistensen av reglar.

Dette minner oss om at skjønn framleis vil vere sentralt, også etter at gode effektstudiar har sagt sitt, noko som EBU-tilhengjaren heller ikkje treng nekte for. Det er dessutan liten grunn til å tru at ein med denne typen forskning vil kome fram til unntakslause reglar (det gjer ein heller ikkje i medisinen). Reglane, som det slett ikkje er sikkert at ein vil finne mange av, vil vere stokastiske, ikkje deterministiske. Då vil det vere unntak frå regelen, og det vil vere opp til læraren å vurdere om den aktuelle situasjonen er eit av unntaka. Men det gjer ikkje reglar verdilause. Legeyrket krev også skjønn, men ikkje mange vil avvise randomiserte forsøk med medisinar av den grunn.

Carr (2001) nyttar også skjønnsargumentet i sin evidenskritikk. Han skil mellom to typar aktivitetar. På den ein sida har ein aktivitetar basert på tekniske ferdigheiter, som til dømes det å reparere ein bil. Her har ein bestemte prosedyrar som ein må følgje, og som ein kan trene seg opp til å følgje på gode eller mindre gode måtar. På den andre sida har ein

grunnleggande trekk ved det menneskelege tilværet som ikkje let seg kodifisere på same måte. Kjærleik og venskap er to eksempel. Sjølv om ein kan vere ein meir eller mindre god ven, og sjølv om ein kan prøve å bli ein betre ven, så er ikkje det å vere ein ven basert på tekniske ferdigheiter. Ein lærer seg ikkje å bli ein god ven gjennom å øve på å følgje reglar, heller ikkje dersom desse venskapsreglane skulle vere godt belagt gjennom psykologisk forskning. Ingen ropar etter å gjere venskap til ein evidensbasert praksis. Påstanden til Carr er at det å vere ein lærar er meir som den andre typen aktivitet enn den første. Kjærleik, venskap og undervisning er, seier han, pre-teoretiske og pre-tekniske aspekt ved menneskelivet, og det er dette evidensbevegelsen overser.

Dessverre er ikkje dette argumentet så overtydande som det høyrer ut, sjølv om det naturlegvis er forskjellar (og likskapar) mellom bilreparasjon og undervisning. Grunnlaget for Carrs distinksjon ligg nettopp i omgrepet skjønn. Forskjellen går mellom aktivitetar der det å følgje reglar er tilstrekkeleg og aktivitetar der det ikkje er det. Bilmekanikk er av den første typen, medan venskap og lærarverksemd er av den andre, seier han. Ein kan følgje visse reglar for korleis å vere ein god ven eller lærar, utan dermed automatisk å bli ein god ven eller lærar. Reglane må appliserast på ein bestemt måte, altså med skjønn. Problemet med dette skiljet, og måten Carr gjer det på, er at dei fleste tekniske ferdigheiter, med eit unntak frå dei aller mest mekaniske og rutineprega, som jamvel maskiner kan klare, også krev ein viss type skjønn. Å sitje med manualen vil aldri vere tilstrekkeleg for ein lærar (eller lege), men det er heller ikkje tilstrekkeleg for bilmekanikarar eller andre handtverkarar. Forskjellen mellom bilmekanikk og venskap (og undervisning) let seg ikkje fanga inn av forskjellen mellom regelbasert og skjønnsbasert praksis, og dermed fell grunnlaget for den *prinsipielle* avvisinga av relevansen av reglar for undervisning bort.

Det er ikkje umulig at ein kan utnytte skjønnsargumentet på ein måte som representerer ein radikal kritikk av EBU/RCT. Faren er at den kritikken blir så radikal at den blir urimeleg. Dersom ein til dømes tek den ekstreme posisjonen som Ekeland (2009) stundom nærmar seg, nemleg at kunnskap på gruppenivå er *verdilaus* i møtet med enkelttilfellet, så vil ikkje RCT vere til mykje hjelp. Men det vil heller ikkje andre vitsskapelege metodar. Som Kant argumenterte for i *Kritikk av Dømmekraften*, vitsskap føreset at verda viser fram eit minimum av system (1995). Kvar elev er unik, seier ein av og til, og det er både riktig og viktig. Statistiske mønster er nettopp berre mønster, og det fins alltid unntak, som oftast mange. Men samtidig er det lite klokt å sjå vekk frå det ein veit om gjennomsnittseleven eller den typiske eleven i den eller den kategorien. Sjølv om me er unike, er me ikkje så spesielle som me likar å tru. Statistikk har ein tendens til å innhente oss i det lange løp.

Kontekstualitetsargumentet

Eit av dei vanlegaste argumenta mot EBU/RCT er at det undervurderer kontekstualiteten av undervisningsmetodar og -program (Olson, 2004; Phillips, 2009; Smeyers, 2009). Kva dette betyr, er dessverre ikkje alltid klart. Eit utkast til presisering finn me i Utdanningsforbundet sitt temanotat:

[Evidenstillhengerne] har et manglende blikk for betydningen av kontekst. RCT undervurderer at vitenskapelig kunnskap også er påvirket av for eksempel sosiokulturelle faktorer. Politiske og administrative tradisjoner, normer og kulturer, samt lokale forhold utgjør alle rammer som virker inn på om gitte innsatser eller tiltak virker. Spørsmålet som må stilles, er derfor under hvilke vilkår er det noe virker. (2008, s. 16)

Med utgangspunkt i dette kan me forstå kontekstualitet på følgjande måte. La M vere ein gitt undervisningsmetode eller -program. Kontekstualitet vil då seie at den M som verkar på objekt O1 ikkje treng gjere det på O2, at den M som verkar på stad S1, ikkje treng gjere det på S2, og at den M som verkar på tidspunkt T1, ikkje treng gjere det på T2. O kan her stå for enten individ eller grupper (elevar eller klassar), T kan stå for datoar eller periodar, og S kan stå for bydel, land, område, kultur, verdsdel osv.

La oss vidare, for argumentet si skuld, føresetje at undervisning verkeleg er kontekstuel i denne tydinga. Problemet med dette som kritisk argument er at det rammar for vidt. Medisin er også eit kontekstuel fenomen: Det som verkar på eit individ eller ei gruppe, treng ikkje verke på eit anna individ eller ei anna gruppe (Slavin, 2004). Så dersom kontekstualitet i seg sjølv er eit hinder for EBU/RCT, tek ein med seg store deler medisinsk forskning i same slengen. Det er likevel rimeleg å anta at medisin er *mindre* kontekstuel enn pedagogikk, kanskje fordi det er færre fysiologiske og biologiske forskjellar mellom folk på ulike stader til ulike tider enn psykologiske og kulturelle forskjellar.

Det fins altså grader av kontekstualitet, som strekk seg mellom to like urimelege ytterpunkt. Det eine ytterpunktet er at ein kan formulere universelle lover for alt som skjer i undervisning. For kvart læringsmål ville det då vere ei oppskrift, eit metodesett som garanterte måloppnåing, uavhengig av tid, stad og individ, i det minste gitt visse føresetnader om elevane sine evner. Dette er utopisk, men det treng ikkje ein EBU-forsvarar nekte for. At ein ikkje kan få reglar for alt ugyldiggjer ikkje ønsket om fleire reglar, og dessutan vil kvar edrueleg tilhengjar av EBU innsjå at eventuelle reglar vil vere probabilistiske, og kanskje også kontekstuel avgrensa.

Det andre ytterpunktet er at val av undervisningsopplegg avheng *fullstendig* av den partikulære konteksten og at denne alltid er heilt unik. Eit slikt standpunkt vil rett nok skape problem for EBU/RCT, men som har me sett tidlegare er det også få som vil akseptere det. Det ville nemleg bety at metodar og strategiar ikkje hadde nokon overføringsverdi i det heile, noko som ville gjort ikkje berre EBU/RCT umulig, men også skjønne og praktisk kunnskap. I ei slik verd ville kvar situasjon vore ny, og alt vore overlate til rå intuisjon. Men dersom læring berre var eit spørsmål om flaks eller magi, er det også tvilsamt om omgrepet “lærer” ville gi meining.

Sanninga, både for medisin og pedagogikk, ligg ein stad mellom desse to ytterpunkta. Men kvar? Dette spørsmålet viser oss bresten i kontekstualitetsargumentet: Korleis veit me om noko er kontekstuel bestemt eller ikkje? Korleis veit me i kva grad eit bestemt felt er kontekstuel? Påstandar om kontekstualitet må sjølv grunngjevast, og “det ser slik ut” er ikkje grunn god nok. For å finne ut i kva grad “what works” er kontekstuel, må ein forske på “what works”. Kontrollerte eksperiment, og replisering av slike, er nettopp ein måte å finne ut om ting som verkar ein stad ikkje verkar ein annan stad. Å avvise EBU/RCT på basis av kontekstualitet er derfor å setje vogna framfor hesten.

Det same gjeld det relasjonelle perspektivet. Her møter ein gjerne det synet at det er relasjonen mellom lærar og elev som er den viktigaste faktoren i læring: To lærarar som gjer om lag det same, kan likevel oppnå heilt ulike resultat. Dette poenget blir av og til nytta som argument mot EBU: “Det kan være liten forskjell i dokumenterte virkninger mellom ulike pedagogiske metoder, men stor forskjell mellom lærere” (Grimen, 2009, s. 206). Svaret på dette argumentet er ikkje å nekte for at det stemmer. Svaret er å spørje korleis ein veit at det stemmer. Systematiske eksperiment er ein måte å finne ut betydningen av slike “practitioner

effects”. Snarare enn å vere eit argument *mot* EBU/RCT, er dette derfor eit argument *for* slik forskning.

Ein nær slektning av kontekstualitetsargumentet er kompleksitetsargumentet. Dette seier at det pedagogiske feltet, med sitt mangfald av elevar og lærarar, mål og metodar, kortsiktige og langsiktige effektar, er altfor *komplekst* til at ein kan kome fram til kunnskap om kva som verkar og ikkje verkar (Lareau, 2009). Det er rett og slett for mange variablar å holde styr på. Dette kan vere både eit prinsipielt og eit praktisk argument. Det er det første dersom ein seier at kompleksiteten gjer det *umulig* å kome fram til kunnskap om kausalitet i undervisning. Det er det andre dersom ein seier at kompleksiteten gjer det *vanskeleg* å kome fram til slik kunnskap. Sidan denne artikkelen avgrensar seg til filosofiske argument, er det berre det første eg skal ta opp her. Eg kjem likevel tilbake til det andre i konklusjonen.

Problemet med dette argumentet er det same som med kontekstualitetsargumentet. Korleis *veit* me at undervisning er så komplekst eller komplisert? Ein kan ikkje vite kor mange av dei variablane som inngår i eit klasserom som er kausalt effektive i høve til læring *før* ein har gjennomført den nødvendige forskinga. Forsking på kva som verkar er slik prinsipielt primær i forhold til spørsmålet om kor komplekse undervisningssituasjonar er. For å setje det på spissen: Det kan tenkast at læring i klasserommet er styrt av nokre få og enkle reglar, og det kan tenkast at læring i klasserommet er så komplekst at ingen reglar kan formulerast, ikkje ein gong probabilistiske og kontekstuelte avgrensa reglar, men om det faktisk er slik kan ein ikkje tenkje seg fram til. Sannsynlegvis veit me nok til å fastslå at sanninga ligg ein stad mellom desse to ytterpunkta, men akkurat kvar sanninga ligg treng me effektforsking for å fastslå.

Eg vil også legge til at det slett ikkje er opplagt at pedagogikken sitt objekt er så mykje meir komplekst og komplisert enn andre vitenskapar sitt. Davies hevdar til dømes at det å framheve at utdanningsfeltet er så mykje meir komplekst og kontekstuellet enn helsefeltet er å undervurdere kompleksiteten og kontekstualiteten på helsefeltet, og at utdanningsforskning derfor ikkje er i noko særstilling i høve til for eksempel medisinen (1999, s. 112).¹² Dessutan vil ein evidenstilhengjar typisk sjå kompleksitet som ei utfordring, ein knute som det er forskinga si oppgåve å løyse opp, snarare enn som eit uoverstigeleg hinder som stoppar evidensprosjektet frå å kome i gang.

Identifikasjonsargumentet

For å gjennomføre ein RCT er det to føresetnader som må vere oppfylt. For det første må intervensjonen (middelet) og resultatet (målet) operasjonaliserast på ein slik måte at det er rimeleg klart om (og i kva grad) intervensjonen er utført og resultatet oppnådd. For det andre må intervensjon og resultat vere identifiserbare transkontekstuellet, dvs. ein må kunne identifisere *same* intervensjon og resultat på tvers av kontekstar (individ, klassar, skular osv.).

Det eg her skal kalle “identifikasjonsargumentet”, problematiserer desse føresetnadene (Erickson & Gutierrez, 2002; Hammersley, 1997; Olson, 2004). Dette er ein type kontekstualitetsargument, og kjem i ein sterk og ein svak versjon. Kjernen i argumentet er at sjølve meininga til termane som peikar ut intervensjon og resultat er kontekstuellet determinerte (sterk versjon) eller sensitive (svak versjon). Medan kontekstualitetsargumentet i forrige seksjon gjekk ut på at *effekten* av intervensjonen er kontekstuellet avhengig, går identifikasjonsargumentet ut på at *identiteten* til både intervensjon og resultat er kontekstuellet

¹² I sitt berømte, men kontroversielle foredrag om behovet for å gjere undervisning til ein forskingsbasert profesjon reknar Hargreaves kompleksitet som eit argument *for* EBU (2007).

avhengig. Påstanden er altså at det er eit internt eller konstitutivt forhold mellom kontekst og intervensjon/resultat, og at ein dermed ikkje kan identifisere ein viss intervensjon eller resultat som *den same* på tvers av kontekstar.

Ved nærare ettertanke vil nok dei fleste vedgå at den sterke versjonen er for sterk. Då seier ein nemleg at den konstitutive relasjonen er så intim at det er prinsipielt umulig å identifisere den same intervensjonen eller resultatet på tvers av kontekster. Og det er ein meir radikal påstand enn dei fleste vil vere villige til å svelgje. Det vil nemleg bety at ein ikkje kan snakke om metodar som verkar meir eller mindre godt i det heile. Av det følger det også at ein heller ikkje kan hevde at kva som verkar er kontekstuel bestemt, for ein kunne då ikkje ha identifisert dette “kva” som verkar på ulike måtar i ulike kontekstar. Det sterke identifikasjonsargumentet undergrev slik det kausale kontekstualitetsargumentet. Og dersom ein ikkje kan identifisere intervensjon (undervisningsmetode) og resultat (oppnåing av læringsmål) på tvers av kontekstar, er det også tvilsamt om ein i det heile tatt kan snakke om metodar og læring, sidan kva ein lærar gjer og ein elev lærer i ein kontekst (klasse, skule etc.) aldri kan vere “det same” som det ein lærar gjer eller ein elev lærer i ein annan kontekst (klasse, skule etc.).

Ein svakare versjon av argumentet nøyer seg med å seie at det er *vanskeleg* å identifisere intervensjon og resultat på tvers av kontekstar. Problemet med denne versjonen er at den ikkje representerer nokon prinsipiell kritikk av EBU/RCT. Ein evidenstillhengar kan utan vidare gå med på dette, og seie at mykje det same gjeld mange diagnostiske termar i medisin og psykologi. Korleis ein skal definere og operasjonalisere termar er eit praktisk problem ein møter *i* forskning, ikkje eit prinsipielt problem *med* forskning av ein viss type, slik Hammersley presenterer det som. Sjølv om ein ikkje umiddelbart kjem på ein god måte å operasjonalisere

til dømes “kritisk tenking” på, betyr ikkje det at det er umulig. Det er ein del av forskingsprosessen å finne det ut. Det er ikkje sikkert det let seg gjere på ein god måte, men det er ikkje noko ein kan konstatere *a priori*.

Dessutan: Sjølv om ei viss overgripande evne eller dygd, som det er ønskeleg å kultivere gjennom utdanning, ikkje let seg fullgodt og ukontroversielt operasjonalisere, så betyr ikkje det at me ikkje kan bli einige om operasjonelle definisjonar av evner og dygder som inngår som konstituentar i den overgripande evna. På same måte som konflikt om korleis ein skal definere “sunn” ikkje hindrar at ein kan ha rimeleg gode definisjonar av ei rekke sjukdommar, hindrar ikkje konflikt om korleis ein skal definere, identifisere og måle “kritisk tenking” at me kan vere einige om at det å gjenkjenne feilslutningar som *petitio principii* inngår som ein del av det.

Cartwright og Kvernbekk

Det kan verke underleg å konkludere med at det ikkje fins gyldige filosofiske argument mot EBP når ein av verdas fremste vitenskapsfilosofar, Nancy Cartwright, dei siste åra har engasjert seg i ein tilsynelatande kritikk av EBP (2007, 2009) Cartwright trekk fram at det også er ulemper ved RCT's, først og fremst at metoden involverer ein “trade-off” mellom intern og ekstern validitet. I praksis er det dessutan vanskeleg å få til perfekte RCTs (nokon droppar til dømes ut). Følgjeleg hevdar ho at det ikkje fins nokon universelt beste metode: Kva metode ein bør bruke, heng saman med kva ein vil vite, kva ein veit, og kva kontekst ein er i. Vidare legg ho vekt på at det er vanskeleg å slutte frå kausalitet i testpopulasjon til målpopulasjon, og at slike slutningar krev skjønn og aldri kan bli ein reint mekanisk affære. Vegen frå solid evidens til klok implementering er difor lang og kronglet.

Alt dette er riktig og viktig. Men det er også noko som ein evidensforkjempar kan vere heilt einig i. Cartwright nyanserer bildet og viser at ting er meir kompliserte enn dei mest idealiserte framstillingane av EBP/RCT tyder på. Slik sett leverer ho ein kritikk av det me kan kalle ekstrem-EBP, men ikkje av den meir edruelege varianten, som er både den mest vanlege og den eg har konsentrert meg om. Dette viser seg også i den retninga som Cartwright si tenking om EBP utvikla seg i. Medan ho tidlegare klaga evidensbevegelsen for å mangle ein adekvat teori om evidens (Cartwright et al., 2007), har ho no forsøkt å utarbeide ein slik teori (Cartwright & Stegenga, 2008). Då blir det tydeleg at ho strengt tatt ikkje er kritisk til EBP/RCT, men at ho vil utfylle og utdjupe vår forståing av dei omgrepa om evidens og kausalitet som det inneber. Her aksepterer ho til og med RCT som ein “gold standard” (Cartwright & Stegenga, 2008, s. 30). Ho legg riktig nok til at vegen frå RCT til policy er lang, men det er ikkje noko nytt.

Ein annan type vitkapsfilosofisk drøfting av evidensbasering står Tone Kvernbekk for (2011). Ho meiner at debatten om EBU/RCT i større grad må ta til seg generelle vitkapsfilosofiske analysar av evidensomgrepet (ho nemner ikkje Cartwright, som nettopp prøver å bygge bru mellom vitkapsfilosofien og evidensdebatten). Eg er einig i at dette kunne vore interessant, men eg reknar det ikkje som nokon vesentleg mangel ved debatten. Grunnen til det, er at ordet “evidens” kan stå for to heilt ulike relasjonar, der berre den eine fangar inn det særeigne ved evidensbasering. Det kan for det første referere til relasjonen mellom data (empiri) og hypotese (teori) i eit eller fleire forsøk. Dette er ein relasjon innad i vitkapsfilosofien, der målet er kunnskap eller sanning. Det kan for det andre referere til relasjonen mellom resultatet av eit eller fleire forsøk (hypotesen, probabilistisk kvalifisert) og praksisen. Dette er ein relasjon mellom vitkapsfilosofien og ein viss sosial praksis, der målet er vellykka

handling. Disse to relasjonane er svært ulike, og reiser ulike problem, sjølv om visse logiske aspekt gjeld dei begge, til dømes underdeterminering.

Medan det stort sett er den første relasjonen som blir tematisert i standard vitenskapsfilosofi (Cartwright er eit unntak), ligg det særeigne ved EBU i synet på den andre relasjonen, meir spesifikt i kva ein reknar som evidens eller belegg for praksis. Det spesielle og kontroversielle med EBU ligg ikkje i at ein opererer med eit spesielt evidensomgrep. “Evidens” blir der nytta i ein relativt ordinær betydning: Å ha evidens for bruk av ein viss undervisningsmetode er, grovt sett, å ha gode grunnar for å bruke han. Det spesielle og kontroversielle med EBU ligg i *kva ein reknar som gode grunnar*, altså kva ein reknar som evidens for ein bestemt praksis. Det er her det såkalla evidenshierarkiet kjem inn, med meta-analysar av RCT’s på toppen. Så når evidensdebatten i liten grad har tatt til seg generelle vitenskapsfilosofiske diskusjonar om evidens, er det etter mitt syn forståeleg. Det utelukkar ikkje at slik vitenskapsfilosofi kan belyse EBU, men då må ein bygge ei bru mellom desse to områda, slik Cartwright nettopp har gjort (Cartwright & Stegenga, 2008).

Konklusjon

I denne artikkelen har eg vurdert nokre filosofiske argument mot EBU/RCT. Det kan naturlegvis finnast andre og betre argument enn dei eg har diskutert, men min tentative konklusjon er at EBU ikkje er spesielt problematisk filosofisk sett. Men eit filosofisk forsvar for EBU impliserer ikkje optimisme med omsyn til sjansen for å faktisk kome fram til solid kunnskap om kva som verkar. Det kan for det første vere store praktiske hindringar i vegen for å få til ein fullgod RCT. Og sjølv om ein skulle få det til, kan resultatet ha avgrensa ekstern validitet. Her slår eit *praktisk* kompleksitets- og kontekstualitetsargument inn.

Kompleksitet og kontekstualitet umuliggjer ikkje EBU/RCT, men gjer det langt *vanskelegare*

å få brukbare resultat. Ein kan bruke store summar og mange ressursar på eit omfattande forsøk, men sjølv om ein skulle oppnå solide resultat, kan ein måtte repetere forsøket i andre byar, landsdelar, land og kulturar for å sjå om effektane er dei same der, noko dei slett ikkje treng vere (Cartwright, 2009, s. 5-7). Og ikkje nok med det, etter ti år kan ein måtte repetere forsøket på ny. Situasjonen har endra seg, og det som verka før, treng ikkje verke lengre, kva Berliner kallar “decade by findings interactions”:

Social and educational research, as good as it may be at the time it is done, sometimes shows these decade by findings interactions. Solid scientific findings in one decade end up of little use in another decade because of changes in the social environment that invalidate the research or render it irrelevant. (Berliner, 2002, s. 20)

Dei beste argumenta mot EBU er med andre ord praktiske, ikkje filosofiske. Det er rett og slett svært dyrt og tidkrevjande å bygge opp denne typen kunnskap. Eg er derfor langt frå like optimistisk som Slavin (2002), som såg for seg at utdanningsvitskapen i det 21. hundreåret skulle kunne gjere dei same framstega som medisinen gjorde i det 20. hundreåret. Kanskje pedagogikk verkeleg er, som Kant sa og Berliner gjenga, “the hardest science of all” (Berliner, 2002; Kant, 2003, §12). På den andre sida er det vanskeleg å vite det før ein har verkeleg prøvd.

References

- Austin, J. L. (1962). *Sense and Sensibilia* (2nd ed.). Oxford: Oxford University Press.
- Berliner, D. C. (2002). Comment: Educational Research: The Hardest Science of All. *Educational Researcher*, 31(8), 18-20. doi:10.3102/0013189X031008018
- Bhrolcháin, M. N. (2001). "Divorce Effects" and Causality in the Social Sciences. *European Sociological Review*, 17(1), 33. doi:10.1093/esr/17.1.33
- Biesta, G. (2007). Why "What Works" Won't Work: Evidence-Based Practice And The Democratic Deficit In Educational Research. *Educational Theory*, 57(1), 1-22. doi:10.1111/j.1741-5446.2006.00241.x
- Campbell, D. T. (1987). Evolutionary Epistemology. I G. Radnitzky & W. W. Bartley (Eds.), *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge* (s. 47-89). La Salle [Ill.]: Open Court Publishing Company.
- Carr, D. (2001). Educational Philosophy, Theory and Research: A Psychiatric Autobiography. *Journal of Philosophy of Education*, 35(3), 461-476. doi:10.1111/1467-9752.00238
- Cartwright, N. (2007). Are RCTs the Gold Standard? *BioSocieties*, 2(01), 11-20. doi:10.1017/S1745855207005029
- Cartwright, N. (2009). Evidence-Based Policy: What's to be done about Relevance? *Philosophical Studies*. doi:10.1007/s11098-008-9311-4
- Cartwright, N., & Stegenga, J. (2008). *A Theory of Evidence for Evidence-Based Policy*. London School of Economics and Political Science, Contingency and Dissent in Science Project.
- Cartwright, N., Goldfinch, A., & Howick, J. (2007). *Evidence-Based Policy: Where is our Theory of Evidence?* London School of Economics and Political Science, Contingency and Dissent in Science Project.

- Cook, T. D., & Payne, M. R. (2002). Objecting to the Objections to Using Random Assignment in Educational Research. I F. Mosteller & R. F. Boruch (Eds.), *Evidence Matters: Randomized Trials in Education Research* (s. 150-178). Washington, D.C.: Brookings Institution Press.
- Daly, J. (2005). *Evidence-Based Medicine and the Search for a Science of Clinical Care*. Berkeley New York: University of California Press ; Milbank Memorial Fund.
- Davies, P. (1999). What Is Evidence-Based Education? *British Journal of Educational Studies*, 47(2), 108-121. doi:10.1111/1467-8527.00106
- Ekeland, T.-J. (2007). Evidens og nypositivisme. *Forskningsspolitikk*, 30(2), 10.
- Ekeland, T.-J. (2009). Hva er evidensen for evidensbasert praksis? I H. Grimen & L. I. Terum (Eds.), *Evidensbasert profesjonsutøvelse* (s. 145-168). Oslo: Abstrakt.
- Elliott, J. (2001). Making Evidence-Based Practice Educational. *British Educational Research Journal*, 27(5), 555-574.
- Erickson, F., & Gutierrez, K. (2002). Comment: Culture, Rigor, and Science in Educational Research. *Educational Researcher*, 31(8), 21-24. doi:10.3102/0013189X031008021
- Friedman, M. (1999). *Reconsidering Logical Positivism*. New York: Cambridge University Press.
- Grimen, H. (2009). Debatten om evidensbasering - noen utfordringer. I Grimen, Harald & L. I. Terum, (Eds.), *Evidensbasert profesjonsutøvelse* (s. 191-222). Oslo: Abstrakt.
- Grossman, J., & Leach, J. (2008). The Rhetoric of Research Methodology. *Social Epistemology*, 22, 325-331. doi:10.1080/02691720802567365
- Hammersley, M. (1997). Educational Research and Teaching: A Response to David Hargreaves' TTA Lecture. *British Educational Research Journal*, 23(2), 141-161. doi:10.1080/0141192970230203

- Hargreaves, D. (1997). In Defence of Research for Evidence-Based Teaching: A Rejoinder to Martyn Hammersley. *British Educational Research Journal*, 23(4), 405-419.
doi:10.1080/0141192970230402
- Hargreaves, D. (1999). Revitalising Educational Research: Lessons from the Past and Proposals for the Future. *Cambridge Journal of Education*, 29(2), 239-249.
doi:10.1080/0305764990290207
- Hargreaves, D. (2007). Teaching as a Research-Based Profession: Possibilities and Prospects. I M. Hammersley (Ed.), *Educational research and evidence-based practice* (s. 3-17). Los Angeles: SAGE.
- Hattie, J. (2009). *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*. London: Routledge.
- Humphreys, P. (1986). Causation in the Social Sciences: an Overview. *Synthese*, 68(1), 1-12.
- Kant, I. (1995). *Kritikk av dømmekraften (i utvalg)*. (E. Hammer, Trans.). Oslo: Pax.
- Kant, I. (2003). *On education*. Mineola N.Y.: Dover Publications.
- Kunnskapsdepartementet. (2008). *St. prp. nr. 1 (2008-2009)*.
- Kvernbekk, T. (2001). Erfaring, praksis og teori. I T. Kvernbekk (Ed.), *Pedagogikk og lærerprofesjonalitet* (s. 146-163). Oslo: Gyldendal Akademisk.
- Kvernbekk, T. (2011). The Concept of Evidence in Evidence-Based Practice. *Educational Theory*, 61(5), 515-532. doi:10.1111/j.1741-5446.2011.00418.x
- Lareau, A. (2009). Narrow Questions, Narrow Answers: The Limited Value of Randomized Control Trials for Education Research. I P. B. Walters, A. Lareau, & S. H. Ranis (Eds.), *Education Research on Trial: Policy Reform and the Call for Scientific Rigor* (s. 145-162). New York: Routledge.

- Lipworth, W., Carter, S. M., & Kerridge, I. (2008). The “EBM Movement”: Where Did it Come From, Where is it Going, and Why Does it Matter? *Social Epistemology*, 22(4), 425-431. doi:10.1080/02691720802559446
- Mayer, R. (2001). Resisting the Assault on Science: The Case for Evidence-Based Reasoning in Educational Research. *Educational Researcher*, 30(7), 29-30. doi:10.3102/0013189X030007029
- Mosteller, F. (1997). The Tennessee Study of Class Size in the Early School Grades. *Bulletin of the American Academy of Arts and Sciences*, 50(7), 14-25. doi:10.2307/3824562
- Norsk Psykologforening. (2007). *Prinsipperklæring (1) om evidensbasert psykologisk praksis (2)*. Retrieved from <http://www.psykologforeningen.no/Fag-og-profesjon/Forfagutoeverer/Fag/Evidensbasert-praksis/Prinsipperklaering-1-om-evidensbasert-psykologisk-praksis-2>
- Oakley, A. (2001). Making Evidence-Based Practice Educational: A Rejoinder to John Elliott. *British Educational Research Journal*, 27(5), 575-576. doi:10.1080/01411920120095744
- OECD. (2003). *New Challenges for Educational Research* (s. 1-144).
- Olson, D. R. (2004). The Triumph of Hope over Experience in the Search for “What Works”: A Response to Slavin. *Educational Researcher*, 33(1), 24-26. doi:10.3102/0013189X033001024
- Peters, R. S. (1967). What is an Educational Process? In R. S. Peters (Ed.), *The Concept of Education* (s. 1-23). London: Routledge & K. Paul.
- Phillips, D. C. (2009). A Quixotic quest? Philosophical Issues in Assessing the Quality of Education Research. I P. B. Walters, A. Lareau, & S. H. Ranis (Eds.), *Education Research on Trial: Policy Reform and the Call for Scientific Rigor*. New York: Routledge.

- Russo, F. (2008). *Causality and Causal Modelling in the Social Sciences*. Berlin: Springer.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence Based Medicine: What it is and What it isn't. *BMJ*, *312*, 71-72. doi:10.1136/bmj.312.7023.71
- Skjervheim, H. (1992). Eit grunnproblem i pedagogisk filosofi. In E. L. Dale (Ed.), *Pedagogisk filosofi* (s. s. 65-78). Oslo: Ad Notam Gyldendal.
- Skjervheim, H. (1996). Det Instrumentalistiske Mistaket. *Deltakar og tilskodar og andre essays* (s. 241-250). Oslo: Aschehoug.
- Slavin, R. . (2002). Evidence-Based Education Policies: Transforming Educational Practice and Research. *Educational Researcher*, *31*(7), 15. doi:10.3102/0013189X031007015
- Slavin, R. E. (2004). Education Research Can and Must Address "What Works" Questions. *Educational Researcher*, *33*(1), 27-28. doi:10.3102/0013189X033001027
- Smeyers, P. (2009). On the Epistemological Basis of Large-Scale Population Studies and Their Educational Use. I D. Bridges, P. Smeyers, & R. Smith (Eds.), *Evidence-Based Education Policy: What Evidence? What Basis? Whose Policy?* Oxford: Wiley-Blackwell.
- Steinsholt, K. (2009). Evidensbaserte standarder eller profesjonalitet. *Bedre Skole*, (1), 54-62.
- Torgerson, C. (2006). *A Systematic Review of the Research Literature on the Use of Phonics in the Teaching of Reading and Spelling*. Nottingham: DfES Publications.
- Utdanningsforbundet. (2008). *Evidens og evidensdebattens betydning for utdanningssystemet*. Oslo.
- Walters, P. B., Lareau, A., & Ranis, S. H. (Eds.). (2009). *Education Research on Trial: Policy Reform and the Call for Scientific Rigor*. New York: Routledge.