

Forståelser av kausalbegrepet i samfunnsvitenskapene

COMPREHENSIONS OF CAUSALITY IN THE SOCIAL SCIENCES

PER ARNE TUFTE

per-arne.tufte@hioa.no

Forfatteropplysninger:

TUFTE, PER ARNE. Dr.polit. i sosiologi, Universitetet i Oslo 2005. Ansatt ved Senter for profesjonsstudier, Høgskolen i Oslo og Akershus siden 2008, som postdoktor fra 2011.

#Postadresse til utsendelse av ferdig hefte:

Postboks 80

1804 Spydeberg#

Å påvise årsakssammenhenger er et sentralt siktemål for samfunnsvitenskapene. Kausalanalyser danner grunnlaget for å kunne forklare og forstå sosiale fenomener bedre og handle på mer hensiktsmessige måter, for eksempel gjennom å utforme bedre sosialpolitiske og helsepolitiske tiltak (Russo, Wunsch & Mouchart 2011).

Ikke alle er av den oppfatning at kausalforklaringer har en sentral plass i samfunnsvitenskapene: «Samfunnsforskningen består ikke i å stille opp hypoteser om kausalsammenhenger, den er en beskrivende vitenskap, dvs. den opererer ikke med kausalforklarende, men med *klassifiserende* begreper» (Østerberg 2000 [1962]:26). Dette er imidlertid en kunstig motsetning fordi samfunnsforskningen er både beskrivende/klassifiserende og forklarende. Gode forklaringer forutsetter gode beskrivelser og klassifiseringer. Og gode beskrivelser/klassifiseringer vil ofte forutsette gode forklaringsmodeller.

Selv om kausalbegrepet intuitivt gir mening, er det vitenskapsteoretisk sett komplekst. Dette essayet drøfter en del sentrale vitenskapsteoretiske tilnærminger til kausalbegrepet og deres relevans for samfunnsvitenskapene. Spørsmålet som stilles er om det spiller noen rolle for

samfunnsforskningen hvordan vi oppfatter kausalbegrepet. Artikkelen drøfter i noen grad ontologiske spørsmål, men er primært fokusert på epistemologiske og metodologiske sider ved kausalitetsbegrepet.

Debatt om kausalbegrepet

Kausalitetsbegrepet vekker gjerne en forestilling om kraft, innflytelse eller påvirkning fra et fenomen, kalt årsaken, til et annet, kalt virkning. Det klassiske eksemplet er biljardballen som støter borti en annen og får sistnevnte til å rulle. Ifølge Lakoff og Johnson (1999) bygger vår forståelse av kausalbegrepet på en metafor om å bruke kroppslig kraft til å endre ting fysisk i omgivelsene («Causation is Forced Movement»). I samfunnsvitenskapene opererer en også med forestillinger om «sosiale krefter», det vil si påvirkning mellom mennesker eller grupper av mennesker, jf. Webers definisjon av makt som et individs evne til å realisere sin vilje mot andres vilje. En rekke av samfunnsvitenskapenes begreper – som vilje, handling, interaksjon, sanksjoner, sosiale roller, konformitet og sosiale klasser – innebærer ulike former for krefter eller påvirkning.

I vitenskapsfilosofien råder det stor uenighet om hvordan kausalitetsbegrepet skal defineres. Skal det for eksempel forstås som observerbar regularitet mellom fenomener, lov eller lovmessighet, kontrafaktisk avhengighet mellom fenomener, intervensjon, mekanismer, overføring av energi mellom fenomener, eller på andre måter? Beebe, Hitchcock og Menzies (2009) peker på flere årsaker til uenigheten: For det første er temaet *komplekst*. Ingen teorier dekker alle aspekter av kausalitet, og samtlige teorier rammes av diverse moteksempler. Det er kanskje ikke mulig å komme fram til et entydig svar på hva kausalitet er, noe som taler for nødvendigheten av en eklektisisme og kausal pluralisme (Weber 2008). For det andre tillegger filosofer kausalbegrepet *ulik metafysisk status*. For enkelte er kausalitet et helt grunnleggende trekk ved virkeligheten, for andre må kausalitet forstås i lys av andre og mer grunnleggende egenskaper, og for atter andre er kausalitet kun et fortolkningsskjema som projiseres over virkeligheten. For det tredje er den filosofiske debatten om kausalitet i langt større grad enn andre filosofiske debatter *koplet til den vitenskapelige utviklingen*. For det fjerde benyttes begrepet på ulike og ikke nødvendigvis kompatible måter *innenfor ulike fagområder*. Kausalitet kan bety ulike ting i fysikken og sosiologien, for ikke å nevne jussen hvor årsaksspørsmålet også koples til spørsmål om skyld og ansvar.

Kausal realisme og konstruktivisme

Ontologiske spørsmål om kausalitet er spørsmål rundt den eventuelle eksistensen av kausalitet som fenomen. Jeg begrenser meg her til debatten om kausal realisme eller konstruktivisme som vedrører spørsmålet om kausalitet eksisterer uavhengig av vår bevissthet, eller er knyttet til vår

bevissthet og oppfattelse av verden rundt. Disse posisjonene kan betegnes som henholdsvis kausal realisme og kausal konstruktivisme (Weber 2008).

En radikal variant av kausal konstruktivisme avviser at kausalitet eksisterer som sådan i naturen eller det sosiale liv. Kausalitet er en kategori som vi forstår virkeligheten gjennom og kan følgelig bli oppfattet ulikt i ulike samfunn og på ulike tidspunkter (Kern 2004). Evolusjonspsykologene hevder i denne forbindelse at evnen til å se mønstre og trekke slutninger om kausalitet har vært nyttig for overlevelse og reproduksjon av menneskearten og at vi derfor har raffinert denne evnen i betydelig grad (Beck & Forstmeier 2005).

Det er vanskelig å tenke seg en seriøs vitenskapsteoretisk posisjon som avviser at sosiale fenomener fører til eller påvirker andre sosiale fenomener (jf. Mellor 1999), og det er også vanskelig å påvise en slik posisjon i samfunnsvitenskapene. Vi observerer og registrerer jo at handlinger og politiske tiltak har konsekvenser. Skepsisen til årsaksbegrepet har antakelig dreid seg om noe annet. Den kan ha vært rettet mot å betrakte kausalitet i termer av *determinisme* og *universalisme*, men dette rammer i liten grad kausalbegrepet siden det ikke forutsetter noen av disse. Andre avviser at årsaksforklaringer har en sentral plass i samfunnsvitenskapene fordi man fort overser meningsdimensjonens og fortolkningens sentrale rolle i det sosiale liv (Taylor 1971). Argumentet kan imidlertid snus: dersom man ensidig fokuserer på fortolkninger og meningssammenhenger, risikerer man å overse sentrale sosiale årsakssammenhenger. Det er heller ikke uten videre en motsetning mellom årsakssammenhenger og meningssammenhenger i analyser av det sosiale liv: «Even if one restricts social science to the giving of interpretations, causal considerations enter into the specification of the web of meanings of social practices and institutions» (Martin 1993). En tredje innvending kan være at årsaksforklaringer gir grunnlag for uønsket manipulasjon og kontroll over mennesker (jf. Mjøset 1991). Bevisstgjøring om sosiale årsakssammenhenger kan imidlertid like gjerne bidra til å øke enkeltmenneskers handlingsmuligheter og frihet.

De fleste filosofiske teorier om kausalitet innebærer en form for kausal realisme i den forstand at kausale sammenhenger mellom fenomener antas å være uavhengige av vår bevissthet om forholdet. Hva forskere eller andre måtte mene om relasjonen mellom klassebakgrunn og utdanningsvalg, spiller ingen rolle for spørsmålet om klassebakgrunn faktisk påvirker utdanningsvalg eller ikke.

En radikal versjon av kausal realisme innebærer at kausalitet har en selvstendig eksistens i naturen og i det sosiale liv. Fenomener kan ha kausale krefter og muligheter (capacities) som får en annen hendelse til å inntreffe. Innenfor kritisk realisme antas for eksempel hendelser og fenomener å være frambrakt av underliggende generative mekanismer (Bhaskar 2008). Kausale krefter betraktes som en distinkt og grunnleggende ontologisk kategori (Mumford 2009). Når en for eksempel kan observere sosiale ulikheter med hensyn til helse, skyldes det generative kausale egenskaper knyttet til sosial posisjon i form av for eksempel yrkesbelastninger og livsstil (trening, alkoholbruk etc.) som produserer helseforskjeller.

En annen innvending er at kausalitet forstått som krefter og påvirkning er et *metafysisk* begrep som ikke kan studeres empirisk. Ifølge en mer reduksjonistisk oppfatning er kausalitet sammensatt av og må reduseres til mer grunnleggende begreper som ikke er kausale. Mot denne innvendingen kan det hevdes at en fort havner i en uendelig regress på jakten etter stadig mer grunnleggende elementer. Dessuten kan det stilles spørsmål om ikke alternative definisjoner også bygger begreper som impliserer kausalitet, slik som for eksempel persepsjon, nødvendighet og handling (Carroll 2009:285).

Hvorvidt en radikal kausal realisme er meningsfull i samfunnsvitenskapene er et åpent spørsmål. Forestillingen om iboende kausale kapasiteter og krefter er mystisk, men samtidig kan menneskets evne til å handle og derigjennom påvirke verden være et eksempel på dette.

De fleste kausalitetsteorier kan kategoriseres som kausal realisme i en mer moderat form som ikke er urimelig i samfunnsvitenskapelig sammenheng. At samfunnsmessige fenomener ofte kan sies å være sosialt konstruert, er ikke ensbetydende med at den sosiale virkelighet ikke er en realitet som påvirker våre liv (Hacking 1999). Sosiale forestillinger og normer får oss til å handle på typiske måter. Det er rimelig å bygge på en moderat realisme hvor man legger til grunn at årsakssammenhenger er reelle, men at vår kunnskap om dem er ufullstendig og knyttet til modeller av virkeligheten. Epistemologisk representerer de fortolkninger fordi vi ikke har direkte tilgang til den sosiale virkeligheten. Sosiale normer og forventninger, sosiale klasser kan ikke observeres direkte. Samtidig er det ikke grunn til å betvile deres eksistens og at de har innflytelse på våre liv.

Tilnærminger til kausalitetsbegrepet

Fire sentrale tilnærminger til kausalitetsbegrepet er regularitetstilnærmingen (herunder lovtilnærmingen), den kontrafaktiske tilnærmingen, manipulasjonstilnærmingen og mekanismetilnærmingen (jf. Brady 2008). Regularitetstilnærmingen, som har hatt stor innflytelse på kvantitativ samfunnsforskning, definerer kausalitet som konstante, observerbare sammenhenger mellom fenomener. Med Psillos (2009) kan en si at de konkurrerende tilnærmingene derimot definerer kausalitet som noe mer enn bare regularitet:

regularitet + X,

det vil si ved å legge noe til regularitetsdefinisjonen som skal gi den et sterkere fundament. Spørsmålet er om dette lille ekstra kan formuleres uten å trekke inn metafysiske begreper som kraft eller påvirkning.

Kausalitet som regularitet

Å betrakte kausalitet som regularitet går tilbake til Humes definisjon: «We may define a CAUSE to be an object precedent and contiguous to another, and where all objects resembling the former are placed in like relations of precedency and contiguity to those objects that resemble the latter» (Hume 2012 [1777]:35). Kausalitet defineres her kun ut fra ikke-kausale fakta fordi Hume betraktet makt, kraft eller nødvendig forbindelser som metafysiske begreper. Vi kan sette opp som kriterier at C er årsak til E hvis og bare hvis:

- i. C berører E i tid og rom (lokalitet),
- ii. E etterfølger C i tid (asymmetri),
- iii. alle hendelser av type C alltid etterfølges av hendelser av type E (konstant konjunksjon).

Prinsippet om *lokalitet* innebærer en avvisning av årsaksforklaringer hvor det er geografisk eller tidsmessig avstand mellom årsak og virkning uten at det foregår noen prosesser imellom. I samfunnsvitenskapene er det imidlertid ikke urimelig å tenke seg store tidsforsinkelser mellom årsak og virkning. Det virker for eksempel rimelig å hevde at internaliserte verdier i oppveksten kan «lagres» mentalt og påvirke seinere valg i livet. På samme måte innebærer begrepet om det sosiomaterielle at det foregår en lagring av kulturelle verdier i bøker, bygninger etc. og at dette igjen kan påvirke handlinger på senere tidspunkt.

Prinsippet om *asymmetri* er mer rimelig i samfunnsvitenskapene. Handling og samhandling er i større eller mindre grad styrt av ønsker, oppfatninger, normer, følelser som alle foreligger i forkant av handlingen.

Hume retter imidlertid skytset mot det siste kriteriet. Argumentet er at vi ikke empirisk kan avgjøre at det foreligger *konstant konjunksjon*. Det kan alltid oppstå nye situasjoner hvor hendelser av type C ikke etterfølges av hendelser av type E, selv om det kanskje er en nedfelt psykologisk tilbøyelighet hos oss mennesker til å se bort fra denne muligheten.

Konstant konjunksjon observeres sjelden i samfunnsvitenskapene i og med at perfekte korrelasjoner så å si aldri forekommer: Regularitetsdefinisjonen kan imidlertid formuleres probabilistisk: C er kausalt relevant for E hvis og bare hvis $P(E | C) \neq P(E)$, det vil si at den betingede sannsynligheten for at hendelser av type E inntreffer, gitt C, er ulik den ubetingede sannsynligheten for E. Dermed kan en i samfunnsvitenskapene for eksempel påvise forhold som påvirker *handlingstilbøyelighet*, eller *risikofaktorer* som gir økt risiko for helseproblemer.

Innvendinger mot regularitetsdefinisjonen er at spuriøse og tilfeldige sammenhenger fort kan bli definert som kausale. Dette er imidlertid trivielle innvendinger som kan imøtegås med krav om robust regularitet, det vil si regularitet som holder når vi har kontrollert for alle potensielt konfunderende årsaksfaktorer.

Ifølge denne tilnærmingen manifesterer sosiale årsakssammenhenger seg i form av relativt stabile sammenhenger mellom sosiale fenomener. Et eksempel kan være at helseforhold varierer systematisk

mellom grupper med ulik sosioøkonomisk bakgrunn. Nærmere analyser kan avdekke mer detaljerte sammenhenger som antyder mer fingraderte årsakssammenhenger, for eksempel utdanningsvalg som er korrelert med yrkesvalg og karriere og i neste omgang korrelert med yrkesrelatert risikofaktorer.

Forskere som støtter seg på denne tilnærmingen vil ofte argumentere for at alle analyser i utgangspunktet er deskriptive og kun påviser korrelasjoner. Enhver tale om kausalitet er å regne som en fortolkning av de påviste sammenhengene. Et felt som er preget av dette er studier av sosial mobilitet. Som Hjellbrekke og Korsnes (Hjellbrekke & Korsnes 2012) påpeker, vet man etter hvert en del om hvordan den sosiale mobiliteten utspiller seg, men man vet mindre om årsaksprosessene bak.

Et vesentligere argument mot regularitetsdefinisjonen er den lar det være igjen noe som trenger forklaring, nemlig eksistensen av regulariteter som sådan. Det kan være ønskelig med et sterkere fundament som kunne gjøre definisjonen mer robust, fortrinnsvis uten å trekke inn metafysiske størrelser som kausale krefter eller liknende. Jeg ser på noen tilnærminger nedenfor.

Kausalitet som lovmessighet

En måte å befestе kausalitetsbegrepet på er å stille krav om at årsaksforklaringer skal vise til et sett med initialbetingelser samt et utsagn om at hendelsen alltid forekommer (en årsakslov) dersom betingelsene er til stede:

Alltid når hendelsene C_1, C_2, \dots, C_n inntreffer, følger hendelse av type E.

Dette kalles også deduktiv-nomologisk modell (D-N) fordi eksplanandum er en logisk konsekvens av eksplanans (deduksjon) og at eksplanans inneholder minst en lov som utgjør et grunnleggende premiss for utledningen av eksplanandum (nomologisk). Modellen innebærer symmetri mellom forklaring og prediksjon: å forklare et fenomen innebærer å vise at det var å forvente som en logisk konsekvens av initialbetingelser og lover.

Vi kan skille mellom tykke og tynne definisjoner av sosiale lover. En *tynn definisjon* innebærer rett og slett at lover dreier seg om generelle regulariteter i naturen eller samfunnet. Lover er «svarte bokser», de viser til regelmessigheter, men sier ingenting om prosessene som fører til regelmessighetene, og da er vi tilbake til regularitetsforståelsen av kausalitet. En *tykk definisjon* legger derimot mer i lovbegrepet. En lov innebærer at effekten inntreffer som en følge av eller med *nødvendighet* av årsaken og da har vi klart beveget oss et skritt videre enn regularitetsdefinisjonen.

I samfunnsvitenskapene har forklaringer ut fra årsakslover vært kritisert fordi det er tvilsomt om man kan etablere universelle lovmessigheter mellom sosiale fenomener og fordi forklaring ut fra «tynne» kausallover (som regulære sammenhenger) kan gi nokså intetsigende forklaringer (Elster 2007). Hempel (1962) beskriver riktignok to probabilistiske varianter av forklaringsmodellen: deduktiv-statistisk (D-S) forklaring som forklarer generelle, regulære sammenhenger ut fra statistiske

lover, og induktiv-statistisk (I-S) forklaring som forklarer enkelthendelser eller -fakta ut fra statistiske lover. Eksempel på slike statistiske lover kunne være:

Når inntekten øker, synker i de fleste tilfeller *andelen* av inntekten som brukes på mat (Engels lov).

Når prisen på et gode øker, synker med stor sannsynlighet etterspørselen (med unntak for såkalte Veblen-goder).

Begge typer probabilistiske forklaringer forutsetter imidlertid relativt høye sannsynligheter for at lovene gjelder. Avhengig av hva som regnes som høy sannsynlighet, kan dette kravet redusere de statistiske lovenes relevans for samfunnsvitenskapene. Spørsmålet er om vi behøver å forutsette eksistensen av (universelle eller statistiske) lovmessigheter for å kausalforklare sosiale fenomener. Selv om enkelte forfekter at det eksisterer sosiale lover (f.eks. Kincaid 2004), er nok den dominerende oppfatningen at det har liten verdi å operere med lover innenfor samfunnsvitenskapene (f.eks. Elster 2007; Roberts 2004).

Kausalitet som kontrafaktisitet

I *An Enquiry Concerning Human Understanding* gir Hume en alternativ definisjon av kausalbegrepet som innebærer at:

C er årsak til E hvis og bare hvis E ikke ville ha inntruffet dersom C ikke hadde inntruffet.

C er følgelig nødvendig (og tilstrekkelig) for at E skal inntreffe, gitt at det ikke er andre betingelser som er tilstrekkelige for å frambringe E. Den probabilistiske varianten innebærer at C er en kausal faktor som fremmer E når $P(E | C) > P(E)$ og hemmer E når $P(E | C) < P(E)$ (Little 1991). Den kontrafaktiske definisjonen av kausalitet gir kausalbegrepet et sterkere fundament enn regularitetsdefinisjonen, angivelig uten å ty til metafysiske elementer eller sirkulære resonnementer. Filosofisk har denne posisjonen vært forsvart av David Lewis (1973, 2000) som tar utgangspunkt i begrepet om mulige verdener (possible worlds):

We may say that one world is closer to actuality than another if the first resembles our actual world more than the second does, taking account of all the respects of similarity and difference and balancing them off one against another (Lewis 1973:559).

Sannhetskriteriet for den kontrafaktiske definisjon av kausalitet er at utsagnet «hvis C var tilfelle, ville også E være tilfelle» er sant i den faktiske verdenen hvis og bare hvis (i) det ikke finnes noen

verdener hvor C (i så fall er utsagnet innholdsløst), eller (ii) noen mulige C-verdener hvor E også inntreffer er nærmere (det vil si mer lik) den faktiske verden enn noen mulige C-verdener hvor E ikke inntreffer (Lewis 1973:560).

Lewis benytter et begrep om *kausal avhengighet* mellom hendelser: hvorvidt E eller ikke-E finner sted er avhengig av om C eller ikke-C finner sted. Lewis begrunner den temporære asymmetrien mellom årsak og virkning ut fra at hendelser typisk sett har færre determinanter (årsaker til eller spor av hendelsen) i forkant enn i etterkant. Det er altså en asymmetri i kraft av over-determinasjon som innebærer at det er et mindre dramatisk inngrep å viske ut en årsak framfor å viske ut sporene av den.

Den kontrafaktiske definisjonen er potensielt beheftet med problemer. Det kan blant annet være vanskelig å fange opp situasjoner med *over-determinasjon* (flere årsaker som hver for seg er tilstrekkelig til å frambringe effekten). Et helseproblem kan skyldes flere sosiale risikofaktorer (yrkesrisiko, livsstil etc.) som alene er tilstrekkelige for å fremme den sviktende helsen. Selv i det kontrafaktiske tilfellet hvor en av disse ikke er til stede, kan helseproblemer fortsatt oppstå som følge av de resterende risikofaktorene. En annen type situasjon som kan by på problemer er tilfeller hvor årsaker mister sin kausale virkning som følge av at andre årsaker inntreffer først (*preemption*). Hvis en pasient lider av en dødelig sykdom og får en behandling som kurerer sykdommen, men samtidig tar livet av pasienten, fungerer ikke den kontrafaktiske logikken, for konklusjonen er at medisinen ikke tok livet av pasienten (vedkommende dør jo også i den nærmeste mulige verden uten medisin).

Et vesentlig bidrag er Holland-Rubin-modellen (Rubin 1974; Holland 1986) som springer ut av den eksperimentelle tradisjonen og har en kontrafaktisk definisjon av kausalitet som utgangspunkt. I denne modellen er en kausaleffekt definert som forskjellen mellom verdien på en utfallsvariabel, Y, avhengig av om en enhet har vært utsatt for en hendelse eller en intervensjon (Y_t) med verdien på utfallsvariabelen for den samme enheten hvis den ikke hadde vært utsatt for hendelsen eller intervensjonen (Y_c):

$$E = Y_t - Y_c$$

Hvis E er forskjellig fra 0, er det en nettoeffekt av intervensjonen som kan tolkes som en kausaleffekt i tråd med den kontrafaktiske definisjonen av kausalitet. Det er imidlertid to problemer her. Sammenlikningen av de to situasjonene gir ikke i seg selv grunnlag for å avgjøre årsaksretningen. Et større problem er at vi ikke kan observere det samme tilfellet i eksperimentsituasjonen (t) og kontrollsituasjonen (c) på en og samme tid. Den nærmeste mulige verden kan ikke observeres.

Her er det to løsninger. Den første er å *observere det samme tilfellet før intervensjonen og etter intervensjonen*. Pretesten (Y_{1c}^*) representerer kontrollsituasjonen for tilfellet i, mens posttesten (Y_{1t}) er eksperimentsituasjonen. Dette forutsetter *temporær stabilitet*, at Y_{1c}^* gir samme verdi som om vi hadde målt Y i den kontrafaktiske situasjonen (Y_{1c}). Sosiale fenomener har imidlertid en tendens til å være i endring, så det er ikke alltid en rimelig forutsetning. Videre forutsettes det at pretestmålingen er

transient, det vil si at pretesten ikke påvirker resultatet i posttestmålingen. Det er for eksempel ikke sjelden at individer lærer av å være gjennom en pretest og endrer sin atferd slik at vi ikke vet om posttesten skyldes denne læringen eller intervensjonen.

Den andre løsningen er å *sammenlikne to tilfeller*, et kontrolltilfelle som ikke eksponeres for påvirkning og et eksperimenttilfelle som eksponeres for påvirkning. Nettoeffekten er differansen mellom de to tilfellene:

$$E = Y_{1t} - Y_{2c}$$

Dette forutsetter *homogenitet mellom tilfellene* (unit homogeneity), det vil si at de er identiske med hensyn til alle relevante forhold. Det innebærer at begge tilfellene ville hatt samme utfall (Y_c) om de var i kontrollsituasjonen og om de var i eksperimentsituasjonen (Y_t). I så fall er det *uavhengighet* mellom utfall og hvilket tilfelle som fordeles til eksperiment- og kontrollgruppe.

Uavhengighetsforutsetningen er streng, og erstattes ofte med en svakere forutsetning om *betinget uavhengighet* (conditional independence). Hvis man vet at en faktor Z fører til brudd på uavhengighetsforutsetningen, kan man betinge på denne variabelen ved å sammenlikne enheter som har samme verdi på variabelen eller kontrollere for risikoen gjennom randomisering.

Ytterligere en forutsetning som også er viktig for undersøkelsens evne til å avdekke et riktig uttrykk for en eventuell kausalsammenheng, er at effekten for et tilfelle *ikke påvirkes av om andre tilfeller i undersøkelsen blir påvirket eller ikke* (noninterference of units). Forutsetningen er brutt når det er interaksjon mellom kontrollgruppe og eksperimentgruppe, for eksempel dersom kontrollgruppen forsøker å kompensere for at de ikke får behandling (kompensatorisk rivalisering), eller at de blir demoraliserte av samme grunn. Dette kalles gjerne interaksjonseffekt eller smitteeffekt i eksperimentlitteraturen.

Metodologisk har kanskje særlig Pearl (2000) og Morgan og Winship (2007) aktualisert den kontrafaktiske posisjonen. Metodelitteraturen inneholder etter hvert en betydelig meny av metoder for å påvise kausalsammenhenger og hvor det tydelig ligger et kontrafaktisk ideal til grunn: instrumentvariabler (IV), paneldata, fixed effects (FE), difference-in-differences (DID), regression discontinuity design (RD) er blant de statistiske verktøyene som har blitt utviklet med sikte på å påvise kausale sammenhenger. Noen av disse behandles i egne artikler i dette nummeret.

Den kontrafaktiske tilnærmingen innebærer en mer offensiv instilling til kausalitet: «Like most people who work with data for a living, we believe that correlation can sometimes provide pretty good evidence of a causal relation, even when the variable of interest has not been manipulated by a researcher or experimenter» (Angrist & Pischke 2009:113).

Empiriske kontrafaktiske analyser vil sammenlikne en faktisk verden med mulige verdener som ligger så nær opp til den første som mulig. Kittelsen et al. (Kittelsen, Magnussen & Anthun 2007) evaluerte for eksempel sykehusreformen av 2002 (hvor staten tok over ansvaret for

spesialisthelsetjenesten fra fylkene) ved å sammenlikne effektivitetsmål tre år før reformen (1999–2001) og tre år etter reformen (2002–2004). Finland, Sverige og Danmark (som ikke hadde en tilsvarende helseform) fungerte som kontrollgruppe. I tillegg kontrollerte man så langt som mulig for forhold som kunne gjøre pretestperioden ulik posttestperioden. Ved hjelp av paneldataanalyser fant man en effektivitetsøkning av reformen på 3–4 prosentpoeng.

En annen måte å anlegge et kontrafaktisk utgangspunkt på er å sammenlikne observerte data med simulerte data som simulerte under kontrafaktiske forutsetninger. Bygren (2013) undersøker årsaker til kjønns- og innvandrers segregering på arbeidsmarkedet ved å sammenlikne faktisk beholdning og flyt av arbeidstakere mellom arbeidsplasser med kontrafaktiske simulerte data hvor flyten av arbeidstakere er tilfeldig for å fjerne bias i segresjonsmålene som skyldes små arbeidsplasser eller at minoritetsandelen er liten («tilfeldig segregering»). Analysene viser at det er betydelig element av tilfeldig segregering, men også systematisk segregering som synes å skyldes homososial bias (at man velger ansatte som likner dem som allerede er i bedriften). Mikroanalytiske, agentbaserte eller andre former for simuleringer gir muligheter for å generere kontrafaktiske data som kan settes opp mot reelle data eller gjennomføre kontrafaktiske tankeeksperimenter (Axelrod 1997).

Kausalitet som manipulasjon

En tilnærming som er beslektet med den kontrafaktiske er å definere kausalitet med utgangspunkt i intervensjon eller manipulasjon hvor det kun gir mening å snakke om at X påvirker Y i de tilfeller hvor en kan forandre Y gjennom å forandre X: «... a necessary and sufficient condition for X to cause Y or to figure in a causal explanation of Y is that the value of Y would change under some intervention on X in some background circumstances ...» (Woodward 2003:15).

En retning innenfor denne tilnærmingen er *handlingsteorier* (agency theories) som knytter årsaksbegrepet til muligheten for menneskelig manipulasjon av årsaken for å frambringe effekten (Gasking 1955, von Wright 1971, Menzies & Price 1993). En innvending er at handlingsteorier er for snevre i og med at menneskelig intervensjon gjøres til utgangspunktet for kausalitet. Dette virker urimelig siden en kan snakke om kausalitet også i sammenhenger hvor ikke det foregår slik intervensjon. En rekke årsaksrelasjoner i naturvitenskapene involverer ikke menneskelig inngripen. I samfunnsvitenskapene er imidlertid antakelsen mer rimelig siden studieobjektet nettopp er menneskelig handling og samhandling.

En annen retning er *intervensjonsteorier* som søker å unngå den direkte koblingen mellom kausalitet og menneskelig handling ved at intervensjoner ikke nødvendigvis er knyttet til menneskelig handling og at en kan resonnerer ut fra hva som ville skjedd dersom manipulering var mulig eller hadde funnet sted: «Instead all that is required is that if human beings were to exist and to carry out the requisite manipulation of A (e.g. the continental plates), B (whether or not an earthquake occurs) would change» (Woodward 2013).

Det kan innvendes at tilnærmingen er *sirkulær* fordi det er vanskelig å definere manipulerbarhet uten å trekke inn kausalitet. Hvis manipulasjonen fører til C som så fører til E, er det ikke da rimelig å betrakte selve manipulasjonen som årsaken til E? Von Wright (1971) forsøker å omgå dette problemet ved å betrakte sammenhengen mellom en handling og dens resultat som iboende og logisk, men ikke kausal. Resultatet av handlingen er en essensiell del av handlingen, den er ikke forårsaket av handlingen.

Selv om den intervensjonistiske tilnærmingen er aktuell for samfunnsvitenskapene, er det en rekke egenskaper som ikke er manipulerbare, for eksempel kjønn og alder. Å snakke om kjønn som årsaksfaktor er vanskelig å forene med kausalitet som manipulasjon, spesielt den handlingsteoretiske tilnærmingen. Samtidig vil det være merkelig å ekskludere kjønn, alder og foreldres sosiale status som årsaksvariabler bare fordi de ikke kan manipuleres. Ottar Hellevik (2002) skriver om utskiftning eller substitusjon som tankemodell i stedet for manipulasjon. Dersom vi «erstatte» en mann med en kvinne, leder det til andre sosiale prosesser som fører til andre resultater for den avhengige variabelen. Det er mulig at man ut fra dette kan argumentere for at Helleviks argument er i tråd med den intervensjonistiske tilnærmingen til kausalitetsbegrepet.

En annen strategi er å omformulere problemstillingen slik at forskeren manipulerer andre størrelser som kan si noe om effekten av ikke-manipulerbare størrelser. I studier av diskriminering i markeder kan man manipulere beslutningstakernes informasjon om søkerens kjønn og alder for å avdekke disse forholdenes betydning på beslutningene. Dette eksemplet illustrerer jo også hvordan eksperimenter kan anvendes til å undersøke «possible worlds».¹

Fordelen med manipulasjonstilnærmingen er at den løser problemet med årsaksretning. Hvis man ved å manipulere en hendelse får en annen hendelse til å inntreffe, må den førstnevnte være årsak til den sistnevnte (gitt det kontrafaktiske forholdet at sistnevnte ikke hadde inntruffet dersom ikke førstnevnte hadde inntruffet). Selv om eksperimentet gjerne knyttes til den kontrafaktiske tilnærmingen til kausalitet, er det mer naturlig å trekke det inn i forbindelse med manipulasjonstilnærmingen siden forskerens manipulasjon av eksperimentvariabelen (treatment) er det som definerer et eksperiment. «Gullstandarden» er randomiserte, kontrollerte studier (RCT, randomized controlled trial), men dette idealet er ikke alltid lett å tilfredsstillere i samfunnsvitenskapene. Utfordringen er i første omgang å få designet et noenlunde realistisk eksperiment. Mange økonomiske eksperimenter er laboratorieeksperimenter, men er disse også gyldige utenfor laboratoriene? Flere har for eksempel stilt spørsmålet om ikke atferdsøkonomenes påvisning av altruisme og solidaritet i diktator-, ultimatum- og tillitsspill er et artefakt frambrakt av laboratorieeksperimentsituasjonen framfor å være et reelt trekk ved individer (se Levitt & Dubner 2009). Rene felteksperimenter har uansett vært nyttige for å undersøke effekter av bistandspolitikk og tiltak for fattigdomsbekjemping (Banerjee & Duflo 2011; Karlan & Appel 2011).

Kausalitet som prosesser og mekanismer

Mekanismetilnærmingen har vunnet terreng innenfor samfunnsvitenskapene de senere årene, men innen filosofien har mekanismebegrepet en lengre historie. En vesentlig bakgrunn er *kausale prosessteorier* hvor kausale sammenhenger forstås ut fra prosesser og interaksjon mellom prosesser. En kausal prosess er en prosess som er i stand til å endre et kjennetegn eller en egenskap ved en annen prosess (Salmon 1994): oppvekstvilkår (første prosess) former et barns holdninger og verdier (andre prosess) som igjen preger dets livsførsel (tredje prosess) som til slutt har betydning for dets helse seinere i livet (fjerde prosess). En alternativ forståelse er at kausale interaksjoner innebærer utveksling av ressurser eller energier mellom prosesser (Dowe 2009). Samhandling forstått som sosiale bytteforhold kan være et eksempel på dette.

Når det gjelder mekanismetilnærmingen, skiller Glennan (2002) mellom prosessteorier som betrakter årsaksmekanismer som en kjede av kontinuerlige prosesser og systemtilnærmingen som betrakter årsaksmekanismer som systemer av interagerende deler. Som eksempel på *prosesstilnærmingen* definerer Elster (2007) mekanismer som hyppig forekommende og lett gjenkjennelige årsaksmønstre som trigges under generelt ukjente betingelser eller med ubestemte konsekvenser. Selvoppfyllende profetier kan være et konkret eksempel på en kausal årsakskjede hvor et i utgangspunktet usant rykte får handlingskonsekvenser, derigjennom styrkes og til slutt blir sant i sine konsekvenser. Som eksempel på *systemtilnærmingen* definerer Hedström (2005) mekanismer som en konstellasjon av individer og aktiviteter koplet slik at de regelmessig frambringer en bestemt type utfall. Betydningen av sosiale nettverk for spredning av ideer og sosiale omveltninger kan tjene som eksempel på denne tilnærmingen. Et kjennetegn ved alle mekanismeteori er at idealet er å spesifisere og teste teorier som beskriver årsaksprosessen mellom explanans og eksplanandum så detaljert som mulig. I tillegg kjennetegnes tilnærmingen av indeterminisme, det vil si at hvilke sosiale mekanismer som trigges og utfallet av mekanismene sjelden kan predikeres. Dette idealet er forenelig med de andre tilnærmingene som er diskutert i dette essayet.

Et vesentlig spørsmål er om mekanisめforklaringer forutsetter metodologisk individualisme eller om det også gir mening å snakke om makromekanismer. Flere argumenterer for at hendelser må forklares som resultat av endringer på makronivå (jf. Bunge 2004; Tilly 2004), altså i sosiale strukturer, institusjonelle ordninger, kulturelle oppfatninger. Hoover (2001) forsvaret denne posisjonen innenfor makroøkonomi ved å betrakte makrofenomener som konfigurasjoner (supervenience) av underliggende mikrofenomener, men med emergente egenskaper som ikke kan reduseres til mikronivået. At en aksjespekulant får panikk er et mikrofenomen, men når denne panikken sprer om seg får vi et makrofenomen – børskrakk – med økonomiske egenskaper som ikke kan reduseres til den enkeltes panikk. Hvorvidt det finnes dem som insisterer på at forklaringer utelukkende må søkes på makronivået er et åpent spørsmål, men det er antakelig tvilsomt.

Elster (2007) argumenterer derimot for at mekanisめforklaringer alltid må være handlings- og samhandlingsorienterte og bygge på plausible psykologiske prinsipper. Colemans (1990) «badekar-

modell» for samfunnsvitenskapelige forklaringer innebærer at sammenhenger mellom sosiale fenomener på makronivået bør forklares gjennom tre ledd. Først bør vi søke årsaksmekanismer som angir hvordan individers oppfatninger, ønsker og handlingsmuligheter påvirkes av den opprinnelige makrotilstanden. Deretter bør man påvise plausible mekanismer som viser hvordan individers oppfatninger, motiver og handlingsmuligheter fører til bestemte handlinger og samhandling. Til slutt må påvise aggregeringsmekanismer som viser hvordan individenes handlinger og samhandling leder til endringer på makronivå.

Konklusjon

Jeg har drøftet noen sentrale forståelser av kausalbegrepet og deres relevans for samfunnsfagene. Enkelte har forsøkt å rydde opp i denne debatten. Hall (2004) hevder at det er to grunnleggende ulike forståelser av kausalbegrepet: kausalitet som kontrafaktisk avhengighet (at virkningen ikke inntreffer dersom ikke årsaken er til stede) og kausalitet som produksjon eller generativ kraft. Ifølge Gerring (2005) kan debatten kokes ned til en minimumsdefinisjon som alle burde kunne enes om, nemlig at en årsak er en hendelse eller betingelse som øker (forstått som produksjon eller generering) sannsynligheten for at et eller annet utfall inntreffer. Det er foreløpig ikke noe som tyder på at slike framstøt har dempet den vitenskapsteoretiske debatten.

Ut fra diskusjonen konkluderer jeg med at kausalitetsbegrepet har en sentral plass i samfunnsforskningen uten at dette behøver å fortrenge meningssammenhengenes og fortolkningenes sentrale plass. Videre innebærer bruken av kausalitetsbegrepet i samfunnsforskningen i hvert fall en mild variant av kausal realisme. Jeg lar imidlertid spørsmålet stå åpent hvorvidt det å definere kausalitet som krefter og kapasiteter bidrar med noe mer enn de tilnærmingene som er diskutert i dette essayet.

Regularitetsdefinisjonen har vært dominerende innenfor samfunnsvitenskapene, og den vesentlige innsikten fra denne tradisjonen er en inngrodd skepsis til selve kausalbegrepet. Til syvende og sist opererer vi aldri med annet enn korrelasjoner. Lovtilnærmingen synes derimot så langt å ha lite å tilby samfunnsforskningen selv om man opererer med statistiske lover.

Samtidig er det etter min mening mye å vinne på å befeste kausalbegrepet ytterligere gjennom en kontrafaktisk eller intervensjonistisk tilnærming. Orienteringen mot kontrafaktisitet har resultert i et bredt og spennende spekter av kausale forskningsdesign og estimeringsteknikker. Siden handlingsbegrepet og intervensjonsbegrepet er sentralt i samfunnsvitenskapene, er manipulasjonsorienteringen spesielt meningsfull. Tilnærmingen er også den eneste som relativt uproblematisk håndterer spørsmålet om årsaksretning. Sist, men ikke minst bør samfunnsforskere arbeide mer i retning av mekanismeforklaringer, i og med at de andre tilnærmingene er mindre orientert mot selve årsaksprosessene som fører fra årsaksfaktor til konsekvens.

Referanser

- Abbott, A. (1998). The Causal Devolution. *Sociological Methods & Research*, 27(2), 148–181.
- Angrist, J. D. & Pischke, J.-S. (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton & Oxford: Princeton University Press.
- Axelrod, R. (1997). Advancing the Art of Simulation in the Social Sciences: Obtaining, Analyzing, and Sharing Results of Computer Models. *Complexity*, 3(2), 16–22.
- Banerjee, A. & Duflo, E. (2011). *Poor Economics*. New York: Public Affairs.
- Beck, J. & Forstmeier, W. (2005). Superstition and Belief as Inevitable By-Products of an Adaptive Learning Strategy. *Human Nature*, 18(1), 35–46.
- Beebe, H., Hitchcock, C. & Menzies, P. (2009). *The Oxford Handbook of Causation*. Oxford: Oxford University Press.
- Bhaskar, R. (2008). *A Realist Theory of Science*. 2. utg. New York: Routledge.
- Brady, H. E. (2008). Causation and Explanation in Social Science. I J. M. Box-Steffensmeier, H. E. Brady & D. Collier (red.), *The Oxford Handbook of Political Methodology* (s. 217–270). Oxford: Oxford University Press.
- Bunge, M. (2004). How Does It Work? The Search for Explanatory Mechanisms. *Philosophy of the Social Sciences*, 34(2), 182–210.
- Bygren, M. (2013). Unpacking the Causes of Segregation Across Workplaces. *Acta Sociologica*, 56(1), 3–19.
- Carroll, J. W. (2009). Anti-Reductionism. I H. Beebe, C. Hitchcock & P. Menzies (red.), *The Oxford Handbook of Causation*. Oxford: Oxford University Press.
- Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge, Mass.: The Belknap Press of Harvard University Press.
- Dowe, P. (2009). Causal Process Theories. I H. Beebe, C. Hitchcock & P. Menzies (red.), *The Oxford Handbook of Causation*. Oxford: Oxford University Press.
- Elster, J. (2007). *Explaining Social Behavior. More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Gasking, D. (1955). Causation and Recipes. *Mind*, 64(256), 479–487.
- Gerring, J. (2005). Causation: A Unified Framework for the Social Sciences. *Journal of Theoretical Politics*, 17(2), 163–198.
- Glennan, S. (2002). Rethinking Mechanistic Explanation. *Philosophy of Science*, 69(3), 342–353.
- Hacking, I. (1999). *The Social Construction of What?*. Cambridge, Mass.: Harvard University Press.
- Hall, N. (2004). Two Concepts of Causation. I J. Collins, N. Hall & L. A. Paul (red.), *Causation and Counterfactuals* (s. 225–276). Cambridge, Mass.: MIT Press.
- Hedström, P. (2005). *Dissecting the Social. On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.

- Hellevik, O. (2002). *Forskningsmetode i sosiologi og statsvitenskap*. Oslo: Universitetsforlaget.
- Hempel, C. G. (1962). Explanation in Science and History. I R. G. Colodny (red.), *Frontiers of Science and Philosophy* (s. 9-19). Pittsburgh: University of Pittsburgh Press.
- Hjellbrekke, J. & Korsnes, O. (2012). *Sosial mobilitet*. Oslo: Samlaget.
- Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945–960.
- Hoover, K. D. (2001). *Causality in Macroeconomics*. Cambridge: Cambridge University Press.
- Hume, D. (2011 [1739]). *A Treatise of Human Nature*. Seattle: Amazon Digital Services.
- Hume, D. (2012 [1777]). *An Enquiry Concerning Human Understanding*. Seattle: Amazon Digital Services.
- Karlan, D. & Appel, J. (2011). *More than Good Intentions. How a New Economics Is Helping to Solve Global Poverty*. New York: Dutton Adult.
- Kern, S. (2004). *A Cultural History of Causality. Science, Murder Novels, and Systems of Thought*. Princeton: Princeton University Press.
- Kincaid, H. (2004). There are Laws in the Social Sciences. I C. Hitchcock (red.), *Contemporary Debates in Philosophy of Science* (s. 168–185). Malden, Mass.: Blackwell Publishing.
- Kittelsen, S. A. C., Magnussen, J. & Anthun, K. S. (2007). *Sykehusproduktiviteten etter statlig overtakelse: En nordisk komparativ analyse*. Oslo: Universitetet i Oslo.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.
- Levitt, S. D. & Dubner, S. J. (2009). *Superfreakonomics. Global Cooling, Patriotic Prostitutes and Why Suicide Bombers Should Buy Life Insurance*. London: Allen Lane.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70(17), 556–567.
- Lewis, D. (2000). Causation as Influence. *Journal of Philosophy*, 97(4), 182–197.
- Little, D. (1991). *Varieties of Social Explanation*. Boulder, Colorado: Westview Press.
- Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, 12, 245–265.
- Martin, M. (1993). Geertz and the Interpretative Approach in Anthropology. *Synthese*, 97, 269–286.
- Mellor, D. H. (1999). *The Facts of Causation*. New York: Routledge.
- Menzies, P. & Price, H. (1993). Causation as a Secondary Quality. *British Journal for the Philosophy of Science*, 44(2), 187–203.
- Mjøset, L. (1991). *Kontroverser i norsk sosiologi*. Oslo: Universitetsforlaget.
- Morgan, S. L. & Winship, C. (2007). *Counterfactuals and Causal Inference. Methods and Principles for Social Research*. Cambridge: Cambridge University Press.
- Mumford, S. (2009). Causal Powers and Capacities. I H. Beebe, C. Hitchcock & P. Menzies (red.), *The Oxford Handbook of Causation* (s. 265-278). Oxford: Oxford University Press.
- Pearl, J. (2000). *Causality*. Cambridge: Cambridge University Press.

- Psillos, S. (2009). Regularity Theories. I H. Beebe, C. Hitchcock & P. Menzies (red.), *The Oxford Handbook of Causation* (s. 131-157). Oxford: Oxford University Press.
- Roberts, J. T. (2004). There are no Laws of the Social Sciences. I C. Hitchcock (red.), *Contemporary Debates in Philosophy of Science* (s. 151–167). Malden, Mass.: Blackwell Publishing.
- Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5), 688–701.
- Russo, F., Wunsch, G. & Mouchart, M. (2011). Inferring Causality through Counterfactuals in Observational Studies – Some Epistemological Issues. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 111(1), 43–64.
- Salmon, W. C. (1994). Causality without Counterfactuals. *Philosophy of Science*, 61(2), 297–312.
doi:10.2307/188214
- Taylor, C. (1971). Interpretation and the Sciences of Man. *Review of Metaphysics*, 25(1), 3–51.
- Tilly, C. (2004). Social Boundary Mechanisms. *Philosophy of the Social Sciences*, 34(2), 211–236.
- Wright, G. H. von (1971). *Explanation and Understanding*. Ithaca: Cornell University Press.
- Weber, E. (2008). The Debate Between Causal Realism and Causal Constructivism: Metaphilosophical Reflections. *Philosophica (gent)*, 81, 59–71.
- Woodward, J. (2003). *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Woodward, J. (2013). Causation and Manipulability. I E. N. Zalta (red.), *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition). Hentet fra
<http://plato.stanford.edu/archives/sum2013/entries/causation-mani/>
- Østerberg, D. (2000 [1962]). Den sosiale realitet. Magisteravhandling fra Institutt for sosiologi, nyutgave ved *Sosiologisk årbok*. Oslo: Universitetet i Oslo.

Noter

¹ Eksemplet bygger sterkt på tilbakemelding fra en anonym fagfelle.