

Perbandingan Inferensi Kausal Versi Donald Campbell dengan Donald Rubin

T. Dicky Hastjarjo¹

Fakultas Psikologi Universitas Gadjah Mada

Abstract

In a psychological experiment the manipulation of independent variable is deliberately introduced to observe its effects on the dependent variable. The cause-effect relationship then is inferred. This article explained briefly two models of causal inferences. One model developed by Donald Campbell are more focused on design elements in ruling out threats to validity. On the other hand, Rubin's model of causal inference emphasized in mathematical precision of the causal effects. The comparison of the two different but complementary models involves both randomized experiment and observational study.

Keywords: causality; causal model; causal inference; experimental validity; threats to validity

Pengantar

Penelitian psikologi dengan metode kuantitatif, baik noneksperimen maupun eksperimen, pada umumnya menguji hubungan antara dua variabel atau lebih. Hubungan antar variabel dinyatakan dengan pernyataan jika X maka Y. Hubungan antara variabel X dengan Y dalam penelitian eksperimental dapat mencerminkan hubungan sebab-akibat atau kausal (*cause-effect relationship*). Sebuah penelitian eksperimental mampu menyimpulkan hubungan kausal oleh karena variabel penyebab dengan sengaja dimunculkan oleh peneliti atau bahasa metodologisnya terdapat manipulasi variabel independen. Kausalitas merupakan salah satu kajian utama dalam psikologi eksperimen.

Donald Campbell (1957) dan bersama kolega (1966, 1979, 2002) telah merintis serta menyempurnakan penjelasan mengenai

kausalitas selama lebih dari setengah abad. Namun demikian kausalitas juga diterangkan dari perspektif lain secara intensif oleh Donald Rubin (1974, 1986, 2004, 2005). Hastjarjo (2011) telah menulis panjang lebar prinsip kausalitas menurut tradisi yang dikembangkan oleh Donald Campbell, kali ini akan dibandingkan dengan pandangan Donald Rubin mengenai kausalitas. Terdapat persamaan antara keduanya namun juga ada perbedaan tajam antara pendekatan Campbell dengan Rubin (Shadish, 2010). Uraian mengenai latar belakang kedua tokoh ini akan membantu memberi gambaran perbedaan perspektif mereka tentang kausalitas.

Donald Campbell berlatar belakang pendidikan psikologi mulai dari sarjana sampai doktor dan berkarir di dunia pendidikan tinggi sampai meninggalnya di bulan Mei 1996. Hasil pemikirannya sangat dikenal di kalangan psikologi dan pendidikan dengan konsep seperti eksperimen-kuasi (*quasi-experiment*), validitas internal

¹ Korespondensi mengenai artikel ini dapat dilakukan melalui: dickyh@ugm.ac.id

(*internal validity*), validitas eksternal (*external validity*) dan ancaman terhadap validitas (*threats to validity*). Sementara itu, Donald Rubin adalah pemegang gelar sarjana psikologi namun meneruskan kuliah di program statistika sehingga mendapatkan gelar doktor statistika. Sesudah bertugas sebentar di lembaga ternama *Educational Testing Service (ETS)* Rubin kembali ke dunia perguruan tinggi sebagai pakar statistika. Rubin terkenal dengan pemikirannya mengenai topik eksperimen nonacak (*nonrandomized experiment*) serta metode praktis untuk mengatasi problem eksperimen lapangan nonacak. Rubin mempopulerkan metode seperti analisis skor propensitas (*propensity score analysis*), dan model kausal Rubin. Hasil pemikiran Rubin kurang dikenal di kalangan psikologi namun lebih dikenal dalam bidang statistika, kedokteran dan kesehatan masyarakat. Ironisnya, publikasi pertama Rubin mengenai kausalitas diterbitkan ditahun 1974 di jurnal psikologi, yaitu *Journal of Educational Psychology* (Rubin, 1974; Shadish, 2010). Kedua tradisi yang menjelaskan kausalitas ini berkembang terpisah dan kedua pakar jarang mengutip pemikiran satu sama lain. Artikel ini akan mengupas teori tentang kausalitas dari dua perspektif yang berbeda yaitu Donald Campbell dengan Donald Rubin. Perbandingan inferensi kausal yang dikembangkan Campbell dengan yang dikembangkan Rubin dalam artikel ini akan berdasarkan sistematika tulisan West dan Thoemmes (2010).

Pembahasan

Model Kausal Campbell (MKC)

Model Kausal Campbell (selanjutnya disingkat MKC) mengembangkan sebuah pendekatan praktis mengenai inferensi kausal dalam rancangan pra-eksperimen,

eksperimen-kuasi serta eksperimen acak yang digunakan di berbagai ranah ilmu pengetahuan berperilaku. Selama hampir lima puluh tahun Campbell dan kolega (Campbell & Stanley, 1966; Cook & Campbell, 1979; Shadish, Cook & Campbell, 2002) mengumpulkan serta menyempurnakan daftar ancaman terhadap validitas eksperimen (Shadish, 2010; West & Thoemmes, 2010). Konsep kuncinya adalah gagasan mengenai ancaman yang masuk akal (*plausible*) terhadap validitas, yakni faktor-faktor yang mungkin dapat berpotensi melemahkan sejumlah aspek proses inferensi kausal dalam setting penelitian tertentu. Tugas peneliti adalah melakukan identifikasi mengenai ancaman yang masuk akal terhadap validitas dan fitur-fitur yang dapat berpotensi meniadakan ancaman serta memasukkan fitur tersebut kedalam unsur rancangan eksperimen. Campbell (1957, 1966) merumuskan dikotomi penting sewaktu peneliti membuat inferensi kausal dari eksperimen, yaitu validitas internal dan validitas eksternal. Validitas internal berkaitan dengan pertanyaan "Apakah benar perlakuan eksperimental membuat perbedaan pada variabel eksperimen tertentu" (1957, h. 297; 1966, h. 5). Validitas internal adalah dasar minimal yang harus dipenuhi sebab jika tidak dipenuhi maka sebuah eksperimen menjadi tidak bisa ditafsirkan.

Tradisi MKC mutakhir (Cook & Campbell, 1979; Shadish, Cook & Campbell, 2002) membedakan empat tipe validitas eksperimen: validitas konklusi statistik, validitas internal, validitas konstruk serta validitas eksternal (lihat juga Hastjarjo, 2014). Kajian West dan Thoemmes (2010) menitikberatkan permasalahan validitas internal sebagai pertimbangan utama mengenai inferensi kausal. Mereka memberi contoh dengan menggambarkan X sebagai indikator perlakuan (misal, 1 = perlakuan/

treatment [T]; 0 = Kontrol/*Control* (C) serta Y sebagai variabel dampak (*outcome variable*) (variabel dependen). Validitas internal mempermasalahkan apakah hubungan antara perlakuan dengan dampak adalah bersifat kausal pada populasi penelitian. Apakah manipulasi variabel independen X menghasilkan perubahan pada Y? Memang sejak awal Campbell (1975) lebih menekankan validitas internal dengan menulis “validitas internal adalah pertimbangan lebih dulu dan sangat diperlukan” dan “dasar terendah agar setiap eksperimen dapat diinterpretasikan sehingga diberi sebutan “*the sine qua non*” (Campbell & Stanley, 1966). West dan Thoemmes (2010) perlu memberi catatan bahwa validitas internal tidak mempermasalahkan aspek khusus/tertentu dari perlakuan yang menghasilkan perubahan maupun aspek khusus/tertentu dari dampak yang mengalami perubahan tersebut.

Validitas Internal

MKC menitikberatkan pada konsep ancaman terhadap validitas internal yang membuat identifikasi mengenai alasan tertentu mengapa peneliti bisa salah dalam mengambil inferensi kausal. Campbell (1957) serta Campbell & Stanley (1966) pada awalnya menyajikan delapan ancaman terhadap validitas internal, lalu berkembang di buku kedua (Cook & Campbell, 1979) menjadi 13 ancaman serta disempurnakan lagi di buku ketiga (Shadish *et al.*, 2002) menjadi 9 ancaman terhadap validitas internal. West dan Thoemmes (2010) memfokuskan pada rancangan eksperimen yang membandingkan kelompok T dan C yang mengandung baik pengukuran *baseline* (pra perlakuan) serta dampak (pasca perlakuan). Rancangan inilah yang mereka gunakan untuk membandingkan perspektif

MKC (Model Kausal Campbell) dengan MKR (Model Kausal Rubin).

Dengan berasumsi bahwa (a) semua partisipan mengalami prosedur eksperimental yang sama kecuali kondisi perlakuan serta (b) semua partisipan diukur baik pada waktu *baseline* maupun dampak, maka menurut MKC ancaman utama terhadap validitas internal seperti maturasi, sejarah serta seleksi dapat diabaikan. Sebuah ancaman akan menimbulkan persoalan jika ancaman tersebut bekerja secara berbeda pada kelompok T dan C. Ancaman tersebut terutama adalah ancaman yang merupakan interaksi antara sebuah ancaman lain dengan ancaman karena seleksi yang disebabkan kemungkinan partisipan pada kelompok T dan C sudah berbeda pada saat awal eksperimen. West dan Thoemmes (2010) memberi contoh penelitian mengenai efek dua metode mengajar pada peningkatan prestasi matematika. Dua metode mengajar tadi adalah metode mengajar dengan bantuan komputer dibandingkan dengan metode pengajaran baku di kelas. Kelompok perlakuan T mendapatkan metode pengajaran dengan bantuan komputer sedangkan kelompok kontrol C mendapatkan metode pengajaran baku. Semua siswa mendapatkan pengukuran *baseline* di awal masa sekolah dan pengukuran dampak perlakuan di masa akhir sekolah serta mereka semua mengalami prosedur eksperimental yang identik terkecuali perlakuan. Ancaman terhadap validitas internal akan merupakan interaksi antara (1) *seleksi* dengan *maturasi*. Siswa di kelompok T dan C mungkin secara alamiah bertumbuh-kembang dengan laju kecepatan yang berbeda bahkan dalam ketiadaan perlakuan. Misalnya, siswa yang berbakat (*gifted*) akan lebih banyak memperoleh peningkatan prestasi dibanding siswa berkecerdasan rata-rata bahkan seandainya mereka mendapatkan metode mengajar

yang sama, (2) *seleksi dengan sejarah*. Partisipan di kelompok T dan C mungkin mengalami kejadian historis yang berbeda. Misalnya, di satu kelompok diubah menjadi kelas kecil oleh aturan kepala sekolah, sementara di kelompok lain tetap menjadi kelas besar. Kelompok kelas kecil mungkin memperoleh peningkatan prestasi matematika lebih besar dibanding kelompok kelas besar, (3) *seleksi dengan instrumentasi*. Partisipan kelompok T dan C mungkin mendapatkan pengukuran dampak yang berbeda atau alat ukur dampak mungkin mempunyai properti pengukuran berbeda pada kedua kelompok (Misalnya, terjadi efek lantai/*floor effect* atau efek atap/*ceiling effect*, atau struktur faktor yang berbeda). Satu kelompok mungkin berubah karena diukur dengan pengukuran pascaperlakuan yang lebih mudah yang menyebabkan prestasi matematika menjadi lebih tinggi walaupun bukan karena efek kurikulum baru tersebut, (d) *seleksi dengan regresi statistik*. Satu kelompok mungkin diseleksi menurut skor *baseline* yang ekstrim sedangkan di kelompok lain diseleksi menurut nilai rata-rata. Dalam kasus ini maka besarnya regresi ke arah rerata pada dua kelompok tersebut akan berbeda serta menghasilkan artefak yang seakan-akan efek perlakuan.

Cara Mengatasi Ancaman terhadap Validitas Internal: Unsur Rancangan.

Sesudah ancaman terhadap validitas internal teridentifikasi maka MKC akan mengidentifikasi prosedur yang memperkecil kemungkinan terjadinya ancaman terhadap validitas internal serta mengidentifikasi unsur khusus yang dapat ditambahkan pada rancangan sehingga ancaman tersebut hilang. Misalnya, penempatan secara acak partisipan ke dalam kondisi perlakuan merupakan elemen umum rancangan. Jika maturasi merupakan ancaman

maka pengukuran pra perlakuan atau pretes tambahan dapat digunakan untuk mengestimasi kecenderungan maturasional di setiap kelompok sebelum perlakuan diberikan. Prioritas MKC adalah lebih pada penyempurnaan rancangan dibanding penyesuaian analisis statistika.

Shadish (2010) menyatakan bahwa tradisi MKC yang memprioritaskan validitas internal tentu lebih berfokus pada cara merancang eksperimen yang mampu meningkatkan validitas internal daripada tipe validitas lain. Cara pertama membangun inferensi kausal yang kuat adalah dengan merancang eksperimen yang mampu mengurangi "jumlah hipotesis tandingan yang masuk akal untuk menjelaskan data". Cara kedua adalah melakukan asesmen ancaman terhadap validitas eksperimen sesudah penelitian selesai dilakukan. Cara kedua ini sebenarnya kurang meyakinkan namun terkadang merupakan satu-satunya pilihan yang ada ketika rancangan yang lebih kuat tidak dapat dipakai atau untuk keperluan mengkritisi penelitian yang telah selesai. Inilah sebabnya MKC selalu berusaha mencari rancangan baru yang mampu meningkatkan inferensi kausal. MKC lebih menekankan pentingnya rancangan daripada analisis statistika. Lebih baik mempunyai sebuah rancangan eksperimen yang meniadakan ancaman terhadap validitas sebelum penelitian dilakukan ketimbang mendasarkan diri pada penilaian manusia yang bisa salah serta analisis statistika untuk menentukan apakah ancaman terhadap validitas telah terjadi sesudah penelitian selesai. Mengutip Light, Singer & Willet yang menulis "anda tak bisa memperbaiki dengan analisis apa yang telah anda kerjakan secara ceroboh dengan rancangan", Shadish dan Cook (1999) menegaskan "Jika tiba pada inferensi kausal dari eksperimenkuasi, rancangan berkuasa, bukan

statistika". Rancangan diibaratkan sebagai seorang tuan sedang analisis statistika adalah abdinya tulis mereka berdua (lihat juga Hastjarjo, 2010).

Model Kausal Rubin (MKR)

Model kausal Rubin (selanjutnya disingkat MKR) atau sering disebut model dampak potensial menjelaskan inferensi kausal dari perspektif matematika/statistika formal (West & Thoemmes, 2010). MKR akan memulai penjelasan dengan merumuskan secara jelas pengertian efek kausal, memeringi asumsi yang diperlukan untuk membuat inferensi kausal untuk setiap rancangan, serta memberikan notasi matematika yang jelas juga. Selain presisi matematis, kegunaan perspektif MKR pada bidang terapan terletak pada nilai heuristik dari konsep dampak potensial (*potential outcomes*). Konsep dampak potensial ini telah terbukti memberikan cara berpikir yang generatif mengenai bagaimana memperoleh estimasi yang tepat mengenai besaran efek kausal yang diinginkan bagi permasalahan penelitian yang sulit.

Model Dasar

Uraian mengenai MKR diawali dengan konsep *the causal estimand*, yakni besaran perbedaan dalam Y yang disebabkan hanya oleh karena perlakuan. West dan Thoemmes (2010) merumuskan *the causal estimand* dengan memulai dari sebuah unit yang diukur tanpa kesalahan (*error*). Sebuah unit khas dalam psikologi adalah seseorang partisipan manusia meskipun sebuah unit bisa saja seekor hewan, kelompok, komunitas dan sebagainya. Perlakuan T diberikan kepada partisipan dan variabel dampak Y diamati. Idealnya, perlakuan perbandingan/kontrol C diberikan kepada partisipan yang sama pada saat yang sama dan dalam konteks yang sama, serta kemudian respon diamati. Efek kausal setiap partisipan atau

efek perlakuan individual dirumuskan sebagai $Y_T(u) - Y_C(u)$, dimana $Y_T(u)$ menunjukkan respon Y dari unit u terhadap perlakuan T dan $Y_C(u)$ menunjukkan respon unit u terhadap perlakuan C. Perbandingan dua dampak tadi memberikan rancangan ideal bagi efek kausal.

Rancangan ideal di atas memang memberi rumusan teoretis yang bermanfaat mengenai efek kausal yang bersifat heuristik. Namun dalam dunia nyata rumusan teoretis tersebut adalah sesuatu yang mustahil dicapai. Misalnya, kalau kembali pada contoh penelitian mengenai efek pengajaran dengan bantuan komputer (T) dibandingkan pengajaran baku di kelas (C) untuk meningkatkan prestasi matematika. Peneliti tidak mungkin memberikan T dan mengukur dampaknya, menempatkan lagi partisipan pada waktu dan tempat yang sama/identik kemudian memberikan C dan mengukur dampaknya (lihat juga Hastjarjo, 2011, hal 4-5 mengenai kemustahilan mengamati secara langsung prinsip kontrafaktual dalam menentukan efek/akibat). Akibatnya, peneliti harus memilih rancangan yang mendekati rancangan ideal tersebut. Rancangan yang pada umumnya digunakan adalah rancangan eksperimen acak (dengan asumsi tambahan) yang akan memberikan estimasi yang tidak bias mengenai efek kausal rata-rata dalam populasi, $U_T - U_C$.

Terdapat tiga implikasi penting dari rumusan efek kausal tadi. *Pertama*, Model tersebut membuat jelas bahwa peneliti membandingkan dua perlakuan. Penelitian psikologi biasanya membedakan antara kelompok eksperimen dengan kelompok kontrol, yang kemungkinan dua-duanya efektif memengaruhi variabel dampak. Bahkan dalam kasus kelompok plasebo pada penelitian farmasi mungkin saja menghasilkan efek farmakologis atau kognitif yang memengaruhi dampak pada

variabel dependen. *Kedua*, model tersebut membutuhkan sebuah pernyataan yang tepat mengenai dua perlakuan yang dibandingkan. Misalnya, pernyataan tersebut digambarkan dalam rumusan operasional mengenai kedua perlakuan itu oleh peneliti. Contoh lain adalah kedua perlakuan dapat dibandingkan hanya bagi subset orang yang berpotensi mendapatkan salah satu perlakuan. Contoh kasus lain lagi adalah bahwa peneliti harus secara seksama mengembangkan konseptualisasi mengenai perlakuan alternatif yang berpotensi diberikan kepada seorang individu (misal dibandingkan dengan perlakuan apa?). *Ketiga*, model tersebut membuat jelas bahwa peneliti mampu mengamati dua kumpulan partisipan, yaitu (a) kelompok A yang mendapat T dan (b) kelompok B yang mendapat C. Kelompok A dan B mungkin kelompok yang sudah terbentuk (misal dua komunitas) atau mereka mungkin kumpulan partisipan yang terseleksi atau telah ditugaskan untuk menerima kondisi T atau C. Sebuah konsep yang paling penting lagi adalah konsep dampak potensial (*potential outcomes*) dari dua kelompok hipotetis, yakni (c) kelompok A yang mendapat C dan (d) kelompok B yang mendapat T. Bayangkan bahwa peneliti akan harus membandingkan rerata dampak dari kedua perlakuan tadi. Secara statistika, menurut rancangan ideal maka apa yang peneliti idealnya ingin miliki adalah estimasi baik dari:

$$\mu\tau(A) - \mu\epsilon(A) \quad \text{atau} \quad (1A)$$

$$\mu\tau(B) - \mu\epsilon(B) \quad (1B)$$

A dan B merupakan kelompok yang mendapatkan perlakuan. Baik persamaan 1A dan 1B menggambarkan efek kausal rata-rata. Penting untuk dicatat bahwa persamaan 1A dan 1B mungkin bukan menunjukkan efek kausal rata-rata yang sama sebab kelompok A dan kelompok B

mungkin berasal dari populasi berbeda. Pada kenyataannya peneliti mempunyai estimasi tentang

$$\mu\tau(A) - \mu\epsilon(B) \quad (2)$$

Analisis konseptual Rubin membutuhkan asumsi tambahan jika perbandingan antara dua kelompok yang ditunjukkan di persamaan 2 merupakan perbandingan ideal seperti yang ditunjukkan oleh persamaan 1A atau 1B.

Eksperimen Acak

Para peneliti menilai rancangan eksperimen acak adalah rancangan "baku emas (*golden standard*)". Proses penempatan secara acak yang tepat terhadap partisipan ke dalam kondisi perlakuan menghasilkan harapan bahwa kondisi perlakuan akan secara statistik independen dari *baseline* kovariat dalam populasi. Independensi statistik menunjukkan bahwa distribusi kovariat yang terukur maupun tak terukur, secara merata, akan seimbang (misal, mempunyai distribusi sama pada *baseline* dibawah kondisi T dan C). Dua implikasi dari independensi statistik adalah (a) rerata kondisi T dan C akan sama pada semua kemungkinan kovariat yang terukur maupun tak terukur pada saat *baseline*, (b) kovarians penempatan pada perlakuan dan kovariat yang terukur atau tidak terukur pada *baseline* akan bernilai sama dengan 0. Harapan ini berimplikasi bahwa tidak terdapat perbedaan pada kovariat potensial antara kelompok T dan C pada saat awal eksperimen dalam populasi. Harapan ini merupakan landasan kuat bagi inferensi kausal, meskipun dalam sebuah eksperimen tidak ada jaminan bahwa rerata level pengukuran pra-perlakuan dari dua kondisi perlakuan akan tidak berbeda pada kovariat tertentu.

Pandangan Model Kausal Campbell (MKC) terhadap Eksperimen Acak

Dengan berasumsi bahwa semua prosedur eksperimen adalah sama kecuali hanya berbeda perlakuan serta tidak ada atrisi, maka terdapat empat ancaman terhadap validitas internal yang harus ditiadakan pada rancangan yang membandingkan kelompok T dengan C yang melibatkan pengukuran *baseline* dan dampak pasca perlakuan kepada semua partisipan. Ke empat ancaman terhadap validitas internal tersebut telah diuraikan di atas (lihat uraian *interaksi* antara faktor seleksi dengan ancaman lain). Penempatan secara acak membuat faktor seleksi tidak masuk akal pada semua variabel yang mungkin terjadi, baik yang terukur atau yang tidak diukur, sehingga sekaligus meniadakan semua interaksi dengan seleksi.

Perspektif MKC mutakhir (Shadish *et al.*, 2002) menambahkan dua ancaman terhadap validitas internal yang tidak mampu dihilangkan oleh penempatan secara acak, yakni (1) Fitur yang tidak konsisten dari eksperimen. Fitur protokol penelitian dalam situasi eksperimen -- misal *setting*, instruksi, jarak waktu antara pengukuran *baseline* dengan dampak, sifat pemberian tes -- tidak dibuat konstan. Sebagai contoh, partisipan di kelompok kontrol diberi alat pengukuran dampak yang bentuknya lebih pendek, atau tidak diberikan pengukuran *baseline* sebab mereka diprediksi tidak bersifat kooperatif dibanding kelompok eksperimen yang mendapatkan perlakuan menyenangkan, (2) Atrisi terkait dengan eksperimen. Atrisi menunjukkan berkurangnya partisipan saat pengukuran pasca perlakuan baik karena alasan menolak mengisi kuesioner, tidak hadir saat sesi pengukuran, pindah alamat dan sebagainya. Perhatian khusus pada validitas internal akan timbul jika atrisi berkaitan dengan penempatan dalam

perlakuan. Misalnya, dengan contoh mengenai prestasi matematika, siswa dengan kemampuan matematika rendah mungkin akan mengundurkan diri dari program inovatif baru sedangkan siswa berkemampuan matematika tinggi akan mengundurkan diri dari program baku.

MKC memfokuskan pada usaha pencegahan ancaman terhadap validitas internal daripada melakukan koreksi pada ancaman tersebut. Shadish *et al.* (2002) misalnya akan meminimalkan fitur yang tidak diinginkan dari kondisi perlakuan dan proses pengukuran, memonitor atrisi selama eksperimen, serta mengembangkan prosedur untuk mempertahankan dan melacak partisipan.

Pandangan Model Kausal Rubin (MKR) terhadap Eksperimen Acak

Model Rubin membutuhkan asumsi tambahan untuk mendapatkan estimasi yang tidak bias dalam eksperimen acak terutama di ranah terapan. Ada empat asumsi tambahan. Pertama, ketaatan terhadap perlakuan keseluruhan. Semua partisipan dalam rancangan eksperimen acak harus menyelesaikan keseluruhan protokol perlakuan, tidak boleh ada partisipan yang mengundurkan diri dari perlakuan atau pindah ke perlakuan lain. Misalnya, dalam eksperimen acak oleh Baker seperti dikutip West dan Thoemmes (2010) mengenai pengaruh skrining mamografi untuk mendeteksi kanker payudara pada wanita berusia di atas 50 tahun, dilaporkan sepertiga dari wanita yang diberi perlakuan mamografi tidak memperoleh skrining, dan banyak wanita dalam kelompok kontrol tanpa skrining mamografi akan memperoleh skrining di luar eksperimen.

Kedua, tidak terjadi atrisi dalam pengukuran pasca perlakuan. Semua partisipan eksperimen acak menyelesaikan pengukuran dampak pasca perlakuan.

Kegagalan memenuhi asumsi ini yang umum terjadi adalah partisipan menolak menjawab pertanyaan, mengundurkan diri sebelum pengukuran pasca perlakuan, atau pindah alamat dalam penelitian longitudinal. Dengan adanya ketaatan penuh terhadap perlakuan, maka hanya jika atrisi berbeda antara kondisi T dengan C yang akan secara teoretis menimbulkan problem.

Ketiga, *Stable Unit Treatment Value Assumption* yang disingkat SUTVA. Model Kausal Rubin bertumpu pada asumsi utama yaitu SUTVA yang menyatakan bahwa representasi dampak potensial serta efek mencerminkan semua nilai yang dapat diobservasi dalam sebuah eksperimen (Shadish, 2010). SUTVA adalah asumsi apriori bahwa nilai Y akan sama untuk unit u jika diberikan perlakuan t tidak peduli mekanisme apa yang digunakan untuk menempatkan perlakuan t kepada unit u serta tidak peduli perlakuan apa yang diberikan kepada unit lain serta hal ini berlaku bagi semua $u = 1, \dots, N$ dan semua $t = 1, \dots, T$ (Rubin, 1986, h. 961). Tujuan asumsi SUTVA adalah menjamin bahwa sebuah nilai respon akan diamati untuk seseorang partisipan di bawah T dan sebuah nilai respon (diharapkan berbeda) teramati di bawah C. Misalnya, partisipan dalam eksperimen prevensi HIV mungkin memiliki penurunan risiko ketularan HIV jika partner seksualnya juga seseorang yang dikenakan kondisi perlakuan efektif daripada seorang yang dikenakan perlakuan tidak efektif. Contoh lain misalnya mengenai eksperimen tentang efek cara pengajaran baru yang pernah diuraikan di atas, siswa yang dikenakan instruksi pengajaran baku mungkin akan mendapatkan bahan metode pengajaran dengan bantuan komputer dari siswa temannya yang dikenakan pengajaran berbantuan komputer. Rubin (2005) menjelaskan asumsi

SUTVA mempunyai dua subasumsi, yakni (a) tidak ada interferensi antara unit-unit, yakni dampak yang diobservasi pada sebuah unit tidak dipengaruhi oleh perlakuan yang didapatkan oleh unit lain, yakni bahwa baik $Y_i(1)$ atau $Y_i(0)$ tidak dipengaruhi oleh perlakuan apapun yang diterima oleh unit lain, serta (b) tidak ada versi tersembunyi dari perlakuan; tidak peduli bagaimana unit i menerima perlakuan 1, dampak yang akan diobservasi akan menjadi $Y_i(1)$ serta secara sama bagi unit yang mendapat perlakuan 0. Keempat, *Efek perlakuan yang konstan* (opsional). Dengan mempertimbangkan bahwa penempatan acak dan tiga asumsi di atas telah terpenuhi, maka MKR memberikan estimasi tidak bias tentang efek kausal rata-rata, yakni $\mu_T - \mu_C$ sebuah efek yang mencirikan populasi partisipan yang dapat dirandomisasi. MKR menitikberatkan pada efek kausal rata-rata ini. Akan tetapi efek kausal rata-rata ini tidak dapat dikhususkan pada seseorang partisipan, $\mu_T(u) - \mu_C(u)$ efek kausal individual jika eksperimen ideal dapat dilakukan pada partisipan yang sama pada saat yang sama pada konteks sama. Estimasi tidak bias terhadap efek kausal individual dalam konteks rancangan eksperimen antar-subjek yang ditempatkan secara acak akan dapat tercapai dengan menambahkan asumsi bahwa efek perlakuan akan konstan sebuah nilai tunggal akan mencirikan semua partisipan.

Bagaimana jika asumsi-asumsi di atas tidak dipenuhi atau dilanggar? MKR menyodorkan konsep *LATE (Local Average Treatment Effect, Efek perlakuan rata-rata lokal)*. Sebelum membahas *LATE*, perlu memahami dulu kategorisasi partisipan menjadi empat (4) kelompok: *Adheres, Always takers, Never takers, dan Defiers*. Partisipan taat (*Adheres*) adalah partisipan yang patuh pada kondisi perlakuan yang dikenakan kepada mereka. *Always takers*

adalah partisipan yang tidak patuh pada kondisi perlakuan yang didapat sebab meski berada di kelompok kontrol mereka akan mencari perlakuan. *Never takers* adalah partisipan yang selalu menolak perlakuan. *Defiers* adalah partisipan yang mencari perlakuan saat di kelompok kontrol dan sebaliknya mencari kondisi kontrol saat di kelompok perlakuan. Efek perlakuan rata-rata lokal (*LATE*) adalah perbandingan antara partisipan taat yang dikenai kondisi perlakuan dengan partisipan taat yang dikenai kondisi kontrol. Efek perlakuan rata-rata lokal dapat dihitung dengan rumus:

$$LATE = \frac{\bar{Y}_T - \bar{Y}_C}{\hat{\Pi}_T}$$

Dimana $\hat{\Pi}_T$ merupakan proporsi partisipan patuh dalam kondisi perlakuan, \bar{Y}_T adalah rerata kelompok T, \bar{Y}_C adalah rerata kelompok C dari analisis *ITT (Intent to Treat)*.

Membandingkan Dua Perspektif MKC dengan MKR mengenai Eksperimen Acak

West dan Thoemmes (2010) menyatakan bahwa kedua perspektif secara umum banyak memiliki kesepakatan dalam mengkaji eksperimen acak. Kedua perspektif berpendapat bahwa penempatan secara acak akan menghasilkan estimasi yang tidak bias tentang besaran serta arah efek kausal. Model MKC lebih menitik-beratkan metode praktis untuk mencegah ancaman terhadap validitas internal sedangkan MKR memfokuskan pada pengembangan asumsi tambahan yang mampu menghasilkan penyesuaian analitis yang tepat bagi permasalahan sehingga besaran yang tepat mengenai efek kausal dapat diestimasi. MKC masih mau memakai prosedur penyesuaian statistik Rubin seperti dalam tulisan Shadish *et al.* (2002, bab 10) dilain pihak MKR juga tidak

memberi peringatan terhadap penggunaan cara praktis oleh Campbell untuk melakukan prevensi ancaman terhadap validitas internal. Kedua perspektif bersifat komplementer.

West dan Thoemmes (2010) menguraikan dua perbedaan antara perspektif MKC dengan MKR, yani MKR lebih menekankan faktor kepatuhan penuh terhadap perlakuan serta asumsi SUTVA dibandingkan MKC. Perbedaan tradisi Rubin dengan Campbell berkaitan dengan kepatuhan terhadap perlakuan serta asumsi SUTVA tampaknya berkaitan dengan perbedaan dalam penekanan di teori mereka. Pertama, Rubin bertumpu pada pengerjaan yang tepat untuk merumuskan perlakuan, tapi bagi Campbell rumusan operasional hanya merupakan contoh dari konstruk yang berkaitan dengan perlakuan yang digunakan dalam sebuah eksperimen tertentu. Kedua, Rubin mengandalkan pada presisi kerangka dampak potensial, sedangkan tradisi Campbell lebih menggunakan prinsip kontrafaktual secara umum. Perbedaan ini tampaknya mencerminkan penekanan dalam tradisi Rubin merumuskan kondisi yang cukup untuk mendapatkan estimasi tepat besaran efek kausal sedangkan tradisi Campbell tampak lebih menekankan konklusi yang jelas mengenai arah efek kausal.

Studi Observasional

Studi observasional adalah penelitian yang bertujuan menentukan inferensi kausal, namun perlakuan tidak bisa dikenakan secara random oleh peneliti. Khususnya, sebuah rancangan dengan kelompok T dan kelompok C dengan pengukuran *baseline* kovariat, seringkali terdapat pengukuran pra perlakuan mengenai dampak. Rancangan ini oleh Campbell dan Stanley (1966) disebut *nonequivalent control group design*. Studi observasional mewarisi semua

isu yang diidentifikasi oleh baik Campbell maupun Rubin untuk rancangan eksperimen. Permasalahan utama tambahan bagi studi observasional adalah bahwa aturan subjek ditempatkan kedalam kelompok eksperimen dan kontrol tidak diketahui dan harus diasumsikan tidak acak. Akibatnya, partisipan di kelompok T dan C mungkin berbeda pada *baseline* dalam hal kovariat yang terukur maupun tak terukur. Kovariat ini menjadi penjelasan alternatif yang potensial (pencemar yang terkait seleksi) untuk setiap efek perlakuan yang dapat diamati.

Perspektif MKC

Perspektif Campbell menekankan empat ancaman terhadap validitas internal seperti yang telah diuraikan di atas berkaitan dengan eksperimen acak, yakni ---seleksi x maturasi, seleksi x sejarah, seleksi - instrumentasi serta seleksi x regresi statistik. Tugas peneliti adalah menggunakan pengetahuan ilmiah saat ini untuk membuat identifikasi ancaman yang paling penting terhadap validitas internal serta meminimalkannya kedalam unsur rancangan untuk meniadakan ancaman tersebut.

Shadish *et al.* (2002) menyajikan daftar ekstensif dari unsur rancangan sasaran yang dapat berpotensi digunakan untuk meniadakan keempat ancaman terhadap validitas internal tersebut. Tiga elemen rancangan akan diuraikan serta diikuti oleh penggambaran penggunaan untuk memperkuat inferensi kausal. *Pertama, Matching.* Jika partisipan kelompok T dan C dapat disesuaikan dengan sukses dalam semua kovariat penting yang berhubungan dengan baik perlakuan maupun dampak, maka bias seleksi dapat dieliminasi. Strategi ini sering akan, meski tidak selalu, mengurangi ancaman interaksi dengan seleksi. *Kedua, Nonequivalents dependent variables.* Variabel dependen yang tidak setara (*Nonequivalents*

dependent variables) dirumuskan sebagai sebuah variabel yang diharapkan dapat dipengaruhi oleh ancaman yang sama dengan yang memengaruhi variabel dampak namun tidak terpengaruh oleh perlakuan. Jika variabel dampak menunjukkan dipengaruhi oleh perlakuan sedangkan variabel dependen tak setara menunjukkan tidak dipengaruhi perlakuan, maka keyakinan mengenai inferensi kausal akan meningkat. Dengan contoh terdahulu, pengajaran berbasis bantuan komputer ternyata meningkatkan prestasi matematika, dan tidak meningkatkan mata pelajaran lainnya, dibandingkan pengajaran baku, maka pola hasil eksperimen ini akan membantu meniadakan ancaman interaksi seleksi x sejarah. *Ketiga, Penggunaan pengukuran pra perlakuan berulang-ulang.* Jika pengukuran pra perlakuan dapat dilakukan beberapa kali, pola maturasi dapat diestimasi pada kelompok T dan C sebelum perlakuan serta kemungkinan dilakukan ekstrapolasi pada periode pasca perlakuan, memecahkan ancaman interaksi seleksi dengan maturasi.

Ringkasan

Perspektif MKC akan menambahkan unsur rancangan ke dalam rancangan dasar untuk meniadakan ancaman terhadap validitas dalam sebuah eksperimen tertentu sehingga membuat sulit untuk mengidentifikasi ancaman terhadap validitas (faktor pencemar) yang berpotensi menjelaskan pola hasil yang teramati. Tujuan utama adalah meniadakan ancaman validitas internal dengan menambahkan unsur rancangan. Namun demikian, setiap unsur rancangan masih dapat diberikan kritikan. Misalnya, variabel dependen tak setara (*Nonequivalents dependent variables*) mungkin berpotensi kurang reliabel.

Perspektif MKR

Perspektif Rubin menekankan pencocokan (*matching*) sebagai strategi paling disukai dalam studi observasional. Secara konseptual, sejauh mana partisipan kelompok kontrol dapat diidentifikasi identik dengan kelompok perlakuan pada saat *baseline*, sebuah estimasi yang tidak bias mengenai efek kausal individual dapat dicapai. Namun dalam praktek ketika seseorang mengestimasi efek kausal rata-rata perbandingan satu kelompok T dengan sekelompok C, seseorang itu tidak yakin bahwa kedua kelompok tidak berbeda dalam pengukuran variabel kunci saat *baseline* pada populasi. Dari persamaan kedua di atas $Y_T(A) - Y_C(A)$ adalah estimasi naif/*prima facie* dari efek kausal. Dari perspektif dampak potensial, apa yang diharapkan diestimasi itu adalah sebuah kombinasi terbobot dari persamaan IA dan IB, $\pi[\mu_T(A) - \mu_C(A)] + \pi[\mu_T(B) - \mu_C(B)]$ di mana A menunjukkan kelompok yang mendapat perlakuan, B menunjukkan kelompok kontrol serta π adalah proporsi populasi yang ada di kelompok perlakuan. Dengan mempertimbangkan bahwa estimasi dampak potensial tidak dapat diobservasi dan bahwa efek kausal sejati dari perlakuan relatif terhadap kontrol mungkin berbeda jika diterapkan pada kelompok A dan B, maka efek kausal naif/*prima facie* akab bersifat bias. Dengan kata lain, sejauh bahwa sejumlah variabel *baseline* juga berkaitan dengan dampak, maka hasilnya akan bersifat bias juga. Semua bagian dari “efek perlakuan” yang teramati mungkin bukan disebabkan perlakuan melainkan oleh karena perbedaan kelompok dalam hal variabel latar belakang.

MKR menggunakan satu kovariat kunci sebagai faktor penting dalam proses penempatan kedalam kelompok eksperimen. Jika dapat dilakukan pencocokan yang

tepat mengenai kovariat *baseline* setiap partisipan di kelompok T dengan partisipan di kelompok C, maka $Y_{Ti} | COV_i - Y_{Ci} | COV_i$ menunjukkan estimasi tidak bias dari efek kausal individual. Secara umum jika COV mempunyai distribusi identik pada kelompok T dan C maka kovariat itu tidak dapat menjadi penyebab efek perlakuan teramati. Salah satu hambatan utama agar tercapai pencocokan yang berhasil kebutuhan akan kovariat yang jamak. Bahkan dengan sampel yang banyak cukup sulit untuk mengidentifikasi kecocokan yang tepat jika banyak variabel dilibatkan. MKR mengembangkan metode skor propensitas (*propensity scores*) untuk memberi solusi bagi proses pencocokan (*matching*) yang melibatkan banyak kovariat.

Sebuah skor propensitas menunjukkan probabilitas seorang subjek akan ditempatkan ke dalam kelompok T adalah $0 < P(T) < 1$. Skor propensitas diestimasi berdasar kovariat yang diukur pada saat *baseline*, misalnya dengan cara regresi logistik. Jika pencocokan kelompok T dan C mencapai keseimbangan dalam hal skor propensitas, maka kelompok T dan C diharapkan akan seimbang pada kovariat apapun yang diukur darimana skor propensitas dikonstruksi. Dengan kata lain, jika semua kovariat yang berkaitan dengan baik penempatan perlakuan serta dampak telah diukur dan semua skor propensitas ada diantara angka $0 < P(T) > 1$, maka mungkin mencapai estimasi yang tidak bias mengenai efek kausal T. Perbedaan rerata kelompok T dan C yang cocok pada skor propensitas akan menunjukkan efek kausal rerata, $E[(Y_T - Y_C) | P(T)]$. Akan tetapi dengan sampel kecil sampai sedang akan ada beberapa persoalan untuk mencapai hasil ideal yang diturunkan dari teori statistika.

Pembandingan Perspektif MKC dengan MKR mengenai Studi Observasional

Perspektif Campbell menekankan identifikasi ancaman terhadap validitas dan memakai fitur rancangan tertentu untuk meniadakan ancaman terhadap validitas yang mengancam rancangan tertentu. Strategi ini mendasarkan diri pada pengetahuan ilmiah mutakhir. Strategi ini mengembangkan sekumpulan hipotesis yang kompleks mengenai pola hasil yang akan seharusnya terlihat di berbagai rangkaian banyak elemen rancangan dan membandingkan hasil amatan dengan hasil yang seharusnya. Konteks penelitian tertentu akan menentukan ancaman terhadap validitas tertentu serta elemen rancangan yang akan dijadikan sasaran. Tidak ada rancangan manapun yang menjadi lebih diutamakan terkecuali rancangan acak. Sejauh hasil cocok dengan pola yang dihipotesiskan maka elemen rancangan akan meniadakan variabel pencemar yang berpotensi mampu memberi penjelasan lain terhadap hasil eksperimen. Kegagalan untuk mencocokkan dengan pola hipotesis akan menciptakan ketidakpastian mengenai apakah ada efek kausal yang sebenarnya.

Ada tiga alasan utama mengapa perspektif Campbell akan dapat mengalami permasalahan (West & Thoemess, 2010). Pertama, tidak semua upaya ilmiah bahkan pengetahuan ilmiah mutakhir bebas dari kesalahan sehingga mungkin upaya ilmiah tersebut tidak akurat. Masih belum ada kepastian bahwa semua ancaman terhadap validitas telah berhasil diidentifikasi. Kedua, sebagian pola hasil yang dihipotesiskan menggambarkan prediksi "tak ada perbedaan", misalnya dua kelompok kontrol yang berbeda diharapkan tidak berbeda hasilnya satu sama lain; sebuah variabel dependen tidak setara diharapkan tidak dipengaruhi oleh perlakuan. Namun perlu diperhatikan untuk mendapatkan

besar sampel yang memadai dan penggunaan pengukuran yang reliabel agar mampu meminimalkan kemungkinan tidak terdeteksinya efek perlakuan yang sesungguhnya. Ketiga, tradisi MKC meniadakan ancaman terhadap validitas eksperimen secara dikotomis: ya atau tidak samasekali (*all or none*), yakni ancaman dihilangkan atau tidak dapat menghilangkan ancaman samasekali. Reichardt sebagaimana dikutip West & Thoemess (2010) mengembangkan prosedur memerinci besaran porsi ancaman yang akan dihilangkan, meski prosedur ini masih dalam pengembangan lebih lanjut.

Tradisi Rubin menekankan perkembangan teori matematis yang elegan yang mencoba menghasilkan pencocokan berbagai pengukuran kovariat antara kelompok perlakuan dan kontrol. Pencocokan skor propensitas dapat menghasilkan estimasi yang tidak bias dari efek kausal. Kunci utama keberhasilan model skor propensitas terletak pada kualitas input ilmiah yang tidak salah terkait pemilihan kovariat pada *baseline*. Tradisi Campbell skeptis terhadap pencocokan dan tidak memperlakukan pencocokan lebih utama dibandingkan elemen rancangan. Ada empat keberatan utama yang patut diajukan pada tradisi Rubin menurut West & Thoemess (2010) mengutip Campbell & Erlebacher. Pertama, kemungkinan kovariat penting malah tidak disertakan. Kedua, Pengukuran kovariat yang penting mungkin reliabilitasnya rendah sehingga penyesuaian yang dilakukan tidak sempurna. Ketiga, dalam beberapa situasi penelitian, partisipan di kelompok T dan C kemungkinan berkembang dengan laju yang berbeda, sebuah pencemar yang mungkin tidak terdeteksi oleh penggunaan kovariat pada *baseline*. Keempat, kemungkinan tidak terdapat tumpang tindih yang cukup memadai dalam hal variabel kovariat antara kelompok T dengan C.

Penutup

Banyak model menjelaskan kausalitas atau hubungan sebab-akibat sebagai salah satu kajian psikologi. Tulisan ini membandingkan model kausal Campbell (MKC) dengan model yang dikembangkan oleh Rubin (MKR). Model Campbell lebih menekankan pada rancangan beserta unsur rancangan yang dapat mencegah ancaman terhadap validitas eksperimen sedangkan model Rubin lebih menekankan pada metode membuat estimasi efek kausal secara tepat dengan prinsip matematika dan statistika. Kedua model kausal tersebut menurut West & Thoemmes (2010) bersifat saling melengkapi meski terdapat perbedaan. Sebenarnya masih ada model kausalitas lain yakni menurut Pearl (2000, 2009) dan Steyer (2017). Steyer berkesempatan memberikan kuliah model kausal di Fakultas Psikologi UGM akhir tahun 2017. Model Steyer lebih maju dari model kausal Rubin serta menyediakan perangkat lunak untuk menghitung estimasi dampak potensial serta efek kausal. Dua model kausal terakhir akan diuraikan pada kesempatan yang akan datang.

Daftar Pustaka

- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 34(4), 297-312.
- Campbell, D. T. & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Rand McNally & Co: Chicago.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin Co: Boston.
- Hastjarjo, T. D. (2010). *Eksperimen-kuasi dan generalisasi inferensi kausal*. Prosiding Konferensi Nasional mengenai Psikologi Eksperimen, Yogyakarta 27 Januari 2010.
- Hastjarjo, T. D. (2011). Kausalitas menurut tradisi Donald Campbell. *Buletin Psikologi*, 19(1), 1-15. doi: 10.22146/bpsi.11542
- Hastjarjo, T. D. (2014). Rancangan eksperimen acak. *Buletin Psikologi*, 22(2), 73-86. doi: 10.22146/bpsi.11455
- Pearl, Y. (2000). *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.
- Pearl, Y. (2009). *Causality: Models, reasoning, and inference (2nd Ed)*. New York: Cambridge University Press.
- Rubin, B. D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 56(5), 688-701.
- Rubin, B. D. (1986). Which ifs have causal answers. *Journal of the American Statistical Association*, 81(396), 961-962.
- Rubin, B. D. (2004). Teaching statistical inference for causal effects in experiments and observational studies. *Journal of Educational and Behavioral Statistics*, 29(3), 343-367. doi: 10.3102/10769986029003343
- Rubin, B. D. (2005). Causal inference using potential outcomes: Design, modelling, decisions. *Journal of the American Statistical Association*, 100(469), 322-334. doi: 10.1198/016214504000001880
- Shadish, W. R. (2010). Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods*, 15(1), 3-17. doi: 10.1037/a0015916.
- Shadish, W. R., & Cook, T. D. (1999). Comment---design rules: More steps toward a complete theory of quasi-

- experimentation. *Statistical Science*, 14, 294-300.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Co: Boston.
- Steyer, R. (2017). *Theory of causal effects*. Manuscript in preparation. FSU Jena, Jena
- West, S. G., & Thoemmes, F. (2010). Campbell's and Rubin's perspectives on causal inferences. *Psychological Methods*, 15(1), 18-37. doi: 10.1037/a0015917.