

Systematic Literature Review:

Validitas Internal dan Eksternal dalam Desain Ekperimen

Siska Merrydian¹, Indah Mustika Rini², Wardani Rahayu³, Riyadi⁴

^{1,2,3,4}Pascasarjana Universitas Negeri Jakarta

siska.merrydian@mhs.unj.ac.id, indah.mustika.rini@mhs.unj.ac.id,

wardani.rahayu@unj.ac.id, riyadi@unj.ac.id

ABSTRACT

In experimental research must have external and internal validity in order to control secondary variables. External validity relates to the generalization of research results, while internal validity relates to the extent of the causal relationship between the independent variable and the dependent variable. Therefore, it is important to examine internal and external validity in experimental design. The method used in this study is a systematic literature review. The database used for literature search utilizes two digital libraries, namely Google Scholar and Scopus. The literature search will take place in May 2024. In order for the selected literature to be relevant, the search uses the keyword "internal external validity". The analysis technique used is a narrative method by grouping the data being explored. This research is expected to provide information related to internal and external validity in experimental design.

Keywords: *Internal Validity; External Validity; Experiment Design*

ABSTRAK

Dalam penelitian eksperimen harus mempunyai validitas eksternal dan internal agar dapat mengontrol variabel sekunder. Validitas eksternal berkaitan dengan generalisasi hasil penelitian, sedangkan validitas internal berkaitan dengan sejauh mana hubungan sebab akibat antara variabel independen dengan variabel dependen. Oleh karena itu, menjadi penting untuk mengkaji validitas internal dan eksternal dalam desain eksperimen. Metode yang digunakan dalam penelitian ini adalah tinjauan pustaka secara sistematis. Basis data yang digunakan untuk pencarian literatur memanfaatkan dua perpustakaan digital yaitu *Google Scholar* dan *Scopus*. Penelusuran literatur dilakukan pada bulan Mei 2024. Agar literatur yang dipilih relevan, pencarian menggunakan kata kunci "*internal external validity*". Teknik analisis yang digunakan adalah metode naratif dengan mengelompokkan data yang digali. Penelitian ini diharapkan dapat memberikan informasi terkait validitas internal dan eksternal dalam desain eksperimen.

Kata kunci: **Validitas Internal; Validitas Eksternal; Desain Eksperimen**

PENDAHULUAN

Dalam beberapa dekade terakhir di berbagai bidang, seperti ketenagakerjaan, pendidikan, pembangunan, dan kesehatan, penggunaan metode eksperimen dan kuasi eksperimen telah meluas. Penekanannya pada eksperimental dan kuasi eksperimental didorong untuk menghasilkan hasil yang valid secara internal. Pada saat yang sama, skala global eksperimen menunjukkan hal yang kurang ditekankan namun menjadi perhatian utama yaitu validitas eksternal. Dalam mengevaluasi validitas eksternal serangkaian eksperimen, seseorang mengajukan pertanyaan, "pada populasi apa, pengaturan, dan variabel, bisakah efek ini digeneralisasikan?" (Campbell, 1957). Dengan kata lain, validitas eksternal dapat berupa diukur dalam hal kesalahan dalam prediksi untuk populasi baru di luar yang tercakup dalam bukti basis. Dengan satu atau segelintir penelitian dalam rentang terbatas konteks, validitas eksternal sebagian besar merupakan masalah teoritis spekulasi. Namun dengan jumlah valid internal yang banyak studi di berbagai konteks, masuk akal untuk berharap bahwa peneliti mengumpulkan pengetahuan yang dapat digeneralisasikan, artinya, bukan sekedar belajar tentang waktu dan tempat tertentu di mana penelitian dijalankan tetapi tentang apa yang akan terjadi jika intervensi serupa dilaksanakan di lain waktu atau tempat (Dehejia *et al.*, 2021).

Para ilmuwan telah mengembangkan dan menerapkan sejumlah metode statistik untuk membuat kesimpulan kausal yang valid, dikenal sebagai revolusi kredibilitas. Tren ini telah berfokus pada validitas internal, dimana peneliti berusaha untuk memperkirakan dampak sebab akibat secara tidak memihak, tanpa membuat asumsi yang kuat. Salah satu dari perdebatan metodologis yang paling penting dan sudah berlangsung lama adalah tentang validitas eksternal, bagaimana ilmuwan dapat menggeneralisasi temuan kausal di luar studi tertentu (Egami & Hartman, 2023).

Validitas mengacu pada perkiraan kebenaran atau kegunaan suatu kesimpulan (Trochim & Donnelly, 2006). Validitas bukanlah sebuah properti dari sebuah teori atau desain melainkan sebuah kesimpulan dari suatu penelitian. Artinya, kesimpulan mungkin valid pada beberapa penelitian, namun tidak pada penelitian lain, meskipun penelitian tersebut mempunyai teori yang persis sama atau desain (Shadish *et al.*, 2002).

Validitas eksternal menunjukkan sejauh mana kesimpulan diambil dari sampel penelitian tertentu berlaku untuk populasi yang lebih luas atau populasi sasaran lainnya (Findley *et al.*, 2015). Terminologi validitas eksternal yang terus berkembang telah menciptakan kebingungan yang signifikan di masa lalu 60 tahun. Sangat dipengaruhi oleh Brunswik (1947) pada desain representatif, Campbell (1957) menciptakan perbedaan antara validitas eksternal dan internal.

Validitas eksternal menguji apakah temuan penelitian dapat digeneralisasikan ke konteks lain. Studi dilakukan terhadap sampel, dan jika pengambilan sampel dilakukan secara acak, sampel mewakili populasi, dan sebagainya hasil suatu penelitian dapat digeneralisasikan secara valid kepada populasi dari mana sampel itu diambil. Tetapi hasilnya mungkin tidak dapat

digeneralisasikan ke populasi lain. Validitas internal menguji apakah cara dalam tempat penelitian dirancang, dilaksanakan, dan dianalisis memungkinkan jawaban yang dapat dipercaya terhadap pertanyaan penelitian. Validitas internal menguji sejauh mana kesalahan sistematis (bias) ada. Kesalahan sistematis seperti itu dapat timbul melalui bias seleksi, bias kinerja, bias deteksi, dan bias gesekan. Jika validitas internal dikompromikan, kadang-kadang dapat diperbaiki, misalnya dengan rencana analisis yang dimodifikasi (Andrade, 2018).

Meskipun konsep validitas eksternal tersebar luas dibahas dalam ilmu-ilmu sosial, hanya ada sedikit secara eksplisit menggabungkan validitas eksternal ke dalam desain atau analisis. Hanya 11% dari semua studi eksperimental dan 13% dari semua observasional studi kausal yang diterbitkan dalam *American Political Science Review* tahun 2015 hingga 2019 berisi analisis formal validitas eksternal dalam teks utama, dan tidak ada yang membahas kondisi di mana generalisasi dapat dipercaya. Kurangnya pendekatan empiris untuk validitas eksternal tetap ada, mungkin karena studi ilmu sosial memiliki beragam tujuan dan kekhawatiran seputar eksternal validitasnya, namun sebagian besar metodologi yang ada memilikinya berfokus terutama pada subkelompok ancaman itu secara statistik lebih mudah diatur. Dalam banyak aplikasi, kekhawatiran penting tentang validitas eksternal tidak menerima evaluasi empiris (Egami & Hartman, 2023).

Sangat penting dalam studi kasus yang bersifat menunjukkan hubungan dan penyebab dan efek antar elemen. Validitas internallah yang meyakinkan peneliti agar hasilnya dapat diterima berdasarkan desain penelitian. Namun, itu relevansinya tidak signifikan untuk penelitian deskriptif dan eksploratif. Di dalam validitas internal, isu kuncinya adalah memastikan peneliti dapat mengkonstruksi argumen kausal yang masuk akal dan cukup kuat untuk mendukung penelitian hasil (Quintao *et al.*, 2020).

Ancaman terhadap validitas internal dan eksternal selalu ada menimbulkan masalah bagi para peneliti, terutama setelah Donald Campbell mengakui dan mendefinisikan istilah-istilah ini lebih dari satu setengah abad yang lalu (Campbell *et al.*, 1963; Campbell & Fiske, 1959). Saat merancang eksperimen dan menafsirkan temuan, ketegangan antara tujuan eksperimen dan validitas menjadi jelas. Eksperimen memberikan hasil yang paling langsung untuk menentukan efek sebab akibat dan menguji teori karena memaksimalkan kontrol dan validitas internal dengan menyederhanakan, mengisolasi, dan membuat hal yang paling penurut fenomena kompleks (Manzi, 2012; Pearl & Mackenzie, 2018), namun konsesi ini dibuat dengan mengorbankan mengurangi validitas eksternal atau generalisasi dari temuan.

Pada prinsipnya, eksperimen yang ideal adalah eksperimen dengan nilai tinggi validitas internal dan eksternal. Namun dalam praktiknya, hal ini terjadi seringkali tidak dapat dicapai. Kapan tujuannya adalah untuk menguji teori dan menyimpulkan hipotesis, validitas internal lebih penting daripada validitas eksternal karena dapat diasumsikan bahwa hipotesisnya adalah diharapkan untuk diuji dalam kondisi yang sangat spesifik dan serba bisa dalam kondisi yang setara

(Guala & Mittone, 2005). Misalnya untuk menguji teori dan hipotesis bagaimana konteks mempengaruhi representasi nilai subjektif dan pilihan (Gluth *et al.*, 2020; Khaw *et al.*, 2017), seseorang harus mampu memanipulasi variabel tersebut dengan tepat sebagai konteks dan nilai subjektif sambil memastikan faktor asing lain tidak membiaskan kesimpulan sebab akibat. Demikian pula, teori fungsi *cingulate anterior* membuat prediksi yang harus diuji dengan paradigma eksperimental yang memanipulasi dan mengesampingkan variabel yang relevan (Fromer *et al.*, 2021; Shenhav *et al.*, 2020; Vassena *et al.*, 2020). Dalam hal ini, pengendalian eksperimen itu penting.

Menjadi kebaruan dalam telaah ini adalah melengkapi artikel terdahulu terkait validitas internal dan eksternal dalam desain eksperimen.

METODE PENELITIAN

Artikel yang digunakan dalam tinjauan pustaka ini adalah artikel yang diperoleh dengan menggunakan *Google Scholar* dan *Scopus*. Pencarian literatur penelitian yang relevan dengan topik penelitian ini dilakukan dengan kata kunci "*internal external validity*". Penelusuran literatur dilakukan pada bulan Mei 2024. Metode naratif bertujuan untuk mendeskripsikan validitas internal eksternal dalam desain eksperimen.

HASIL DAN PEMBAHASAN

Artikel Lin *et al.*, (2021) menggali kompleksitas penelitian psikologi, menyoroti perlunya pendekatan yang lebih kohesif dan dapat ditiru dalam membangun teori yang melingkupi pembahasan terkait bagaimana perkembangan bidang psikologi selama bertahun-tahun dalam hal perkembangan teori dan temuan empiris, kritik utama terhadap praktik penelitian psikologi saat ini, dan cara mengatasinya serta cara peneliti mengatasi ketegangan antara tujuan eksperimen dan kekhawatiran validitas untuk meningkatkan kekuatan teori psikologi.

Bidang psikologi telah mengalami evolusi signifikan dalam perkembangan teori dan temuan empiris selama bertahun-tahun. Berikut adalah beberapa poin penting yang menyoroti evolusi ini (Lin *et al.*, 2021):

- a. Landasan Sejarah: psikologi berakar pada filsafat dan fisiologi, dengan teori awal berfokus pada introspeksi dan strukturalisme. Hal ini berkembang menjadi fungsionalisme, yang menekankan fungsi adaptif perilaku.
- b. Behaviorisme: awal abad ke-20 menyaksikan kebangkitan behaviorisme, yang berfokus pada perilaku yang dapat diamati dan menolak studi tentang proses mental. Pendekatan ini mendominasi psikologi selama beberapa dekade, mengarah pada fokus pada penelitian eksperimental dan teori pengkondisian.
- c. Revolusi Kognitif: pada tahun 1950-an dan 1960-an, revolusi kognitif menantang behaviorisme dengan menekankan studi tentang proses mental seperti memori, persepsi, dan pemecahan masalah. Pergeseran ini menyebabkan perkembangan psikologi kognitif sebagai subbidang utama.

- d. Perspektif Biologis dan Evolusi: dengan kemajuan dalam ilmu saraf dan genetika, psikologi semakin banyak memasukkan perspektif biologis dan evolusi. Hal ini menghasilkan pemahaman yang lebih baik tentang bagaimana struktur dan fungsi otak mempengaruhi perilaku dan kognisi.
- e. Pengaruh Sosial dan Budaya: psikologi juga telah berkembang dengan mempertimbangkan faktor sosial dan budaya dalam membentuk perilaku. Ini mencakup penelitian tentang topik-topik seperti pengaruh sosial, perbedaan budaya, dan dampak norma-norma masyarakat terhadap perilaku individu.
- f. Pendekatan Interdisipliner: dalam beberapa tahun terakhir, ada trend yang berkembang menuju penelitian interdisipliner di bidang psikologi. Hal ini mencakup kolaborasi dengan bidang-bidang seperti ilmu saraf, ekonomi, dan ilmu komputer untuk mendapatkan wawasan dan pendekatan baru dalam mempelajari perilaku.
- g. Penelitian Besar: ada pergeseran ke arah penelitian besar-besaran yang ambisius yang bertujuan untuk mengatasi masalah-masalah sosial yang kompleks seperti perubahan iklim, misinformasi, dan moralitas. Pendekatan ini melibatkan triangulasi teoretis dan integrasi data lintas disiplin ilmu untuk meningkatkan relevansi dan dampak penelitian psikologi.
- h. Tantangan dan Peluang: meskipun psikologi telah mencapai kemajuan signifikan dalam pengembangan teori dan temuan empiris, terdapat tantangan seperti masalah replikasi dan teori yang terfragmentasi. Dengan mengenali tantangan-tantangan ini dan mengadopsi pendekatan seperti triangulasi, peneliti dapat meningkatkan kekuatan dan penerapan teori-teori psikologi.

Singkatnya, bidang psikologi telah berkembang dari akar filosofis awalnya menjadi ilmu multidisiplin yang menggabungkan beragam perspektif dan metodologi. Dengan merangkul kolaborasi interdisipliner dan mengatasi masalah validitas, psikologi terus memajukan pemahaman kita tentang perilaku dan kognisi manusia. Praktik penelitian psikologi saat ini menghadapi beberapa kritik, yang menyoroti perlunya perbaikan dalam metodologi, reproduktifitas, dan integrasi teoretis. Berikut adalah beberapa kritik utama dan solusi potensial (Lin *et al.*, 2021):

- a. Masalah Replikasi: salah satu kritik utama adalah kurangnya replikasi dari banyak penelitian psikologi, sehingga menimbulkan kekhawatiran tentang keandalan temuan penelitian. Untuk mengatasi hal ini, para peneliti dapat memprioritaskan melakukan studi replikasi dan melakukan pra-registrasi protokol penelitian untuk meningkatkan transparansi dan reproduktifitas.
- b. Teori Terfragmentasi: bidang psikologi telah dikritik karena memiliki teori-teori yang terfragmentasi yang mungkin tidak selalu selaras atau terintegrasi dengan baik satu sama lain. Untuk mengatasi hal ini, peneliti dapat melakukan triangulasi teori, yang melibatkan menjembatani berbagai teori dan sumber data untuk menciptakan kerangka teori yang lebih kohesif dan komprehensif.
- c. Masalah Validitas: ada ketegangan antara tujuan eksperimen dalam psikologi, seperti pengujian teori dan eksplorasi fenomena, yang dapat menimbulkan masalah validitas. Para peneliti dapat mengatasi permasalahan ini dengan

mengenali masalah validitas internal bersama dan menggunakan metode triangulasi untuk memastikan bahwa eksperimen dan teori didasarkan pada fenomena dunia nyata.

- d. Keterbatasan Metodologis: beberapa kritikus berpendapat bahwa penelitian psikologis mungkin dibatasi oleh pendekatan metodologis yang sempit. Peneliti dapat mengatasi keterbatasan ini dengan menerapkan pendekatan multimetode yang menggabungkan teknik penelitian yang berbeda, seperti eksperimen perilaku, neurofisiologi, dan studi lapangan, untuk memberikan pemahaman yang lebih komprehensif tentang fenomena psikologis.
- e. Putusan Interdisipliner: psikologi terkadang dapat beroperasi secara terpisah, terputus dari disiplin ilmu lain yang dapat menawarkan wawasan berharga. Dengan mendorong kolaborasi interdisipliner dan terlibat dalam bidang-bidang seperti ilmu saraf, ekonomi, dan sosiologi, para peneliti dapat memperkaya penelitian dan mengatasi masalah-masalah sosial yang kompleks dengan lebih efektif.
- f. Bias Publikasi: ada kekhawatiran mengenai bias publikasi dalam psikologi, yaitu studi dengan hasil positif lebih besar kemungkinannya untuk dipublikasikan dibandingkan studi dengan hasil nihil atau negatif. Untuk melawan bias ini, para peneliti dapat mendukung inisiatif praktik sains terbuka, seperti menerbitkan penelitian yang telah didaftarkan sebelumnya, berbagi data secara terbuka, dan mendorong transparansi dalam melaporkan temuan penelitian.

Dengan mengatasi kritik ini melalui peningkatan praktik penelitian, transparansi, dan kolaborasi, bidang psikologi dapat meningkatkan kualitas dan dampak penelitiannya, sehingga menghasilkan teori dan temuan yang lebih kuat yang memberikan kontribusi yang berarti terhadap pemahaman tentang perilaku dan kognisi manusia. Peneliti dapat mengatasi ketegangan antara tujuan eksperimen dan kekhawatiran validitas untuk meningkatkan kekuatan teori psikologi dengan mengadopsi strategi berikut (Lin *et al.*, 2021):

1. Kejelasan Sasaran: peneliti harus mendefinisikan dengan jelas tujuan eksperimen, apakah itu pengujian teori, eksplorasi fenomena, atau pengembangan teori. Memahami tujuan spesifik penelitian membantu dalam memprioritaskan validitas internal atau eksternal berdasarkan tujuan penelitian.
2. Evaluasi Validitas: peneliti harus mengevaluasi validitas internal dan eksternal eksperimen sehubungan dengan tujuan. Untuk pengujian teori, validitas internal sangat penting, sedangkan untuk generalisasi dan penerapan di dunia nyata, validitas eksternal menjadi lebih penting. Dengan menyelaraskan kekhawatiran validitas dengan tujuan eksperimen, peneliti dapat memastikan keandalan dan relevansi temuan.
3. Triangulasi: penggunaan metode triangulasi, seperti triangulasi data lintas waktu, budaya, peneliti, dan laboratorium, dapat membantu meningkatkan kekuatan temuan penelitian. Dengan mengintegrasikan berbagai sumber data

dan perspektif, peneliti dapat memperkuat validitas dan reliabilitas kesimpulannya.

4. **Triangulasi Teoritis:** peneliti dapat melakukan triangulasi teoretis dengan mengintegrasikan ide-ide dari berbagai disiplin ilmu dan kerangka teoretis. Pendekatan ini membantu menghubungkan eksperimen dan teori dengan fenomena dunia nyata, mendorong pemahaman proses psikologis yang lebih komprehensif dan terintegrasi.
5. **Pendekatan Multimetode:** memanfaatkan pendekatan multimetode, seperti menggabungkan eksperimen perilaku, neurofisiologi, dan studi lapangan, dapat memberikan pandangan yang lebih holistik tentang fenomena psikologis. Dengan menggunakan metode yang beragam, peneliti dapat mengurangi keterbatasan pendekatan tunggal dan meningkatkan validitas dan generalisasi temuan.
6. **Kolaborasi Interdisipliner:** berkolaborasi dengan peneliti dari berbagai disiplin ilmu, seperti ilmu saraf, ekonomi, dan sosiologi, dapat memperkaya penelitian psikologi dengan menggabungkan perspektif dan metodologi yang berbeda. Kolaborasi interdisipliner mendorong inovasi, integrasi teoritis, dan pemahaman yang lebih komprehensif tentang fenomena psikologis yang kompleks.
7. **Praktik Sains Terbuka:** menerapkan praktik sains terbuka, seperti prapendaftaran, berbagi data, dan pelaporan transparan, dapat meningkatkan transparansi dan reproduktifitas penelitian psikologis. Dengan mengedepankan keterbukaan dan akuntabilitas, peneliti dapat memperkuat validitas dan kredibilitas penelitian.

Dengan mengintegrasikan strategi-strategi ini ke dalam praktik penelitian, para psikolog dapat secara efektif mengatasi ketegangan antara tujuan eksperimen dan masalah validitas, sehingga mengarah pada teori-teori psikologis yang lebih kuat dan berdampak yang memajukan pemahaman tentang perilaku dan kognisi manusia.

Validitas Internal

Validitas internal menunjukkan sejauh mana suatu eksperimen dapat menunjukkan hubungan sebab akibat antara variabel independen dan variabel dependen dengan mengurangi sumber kesalahan atau bias dan dengan mengendalikan variabel perancu. Validitas internal merujuk pada validitas inferensi mengenai apakah kovariansi yang teramati antara A dan B mencerminkan sebuah hubungan kausal dari A ke B sebagaimana variabel tersebut dimanipulasi atau diukur. Validitas internal akan didukung jika peneliti dapat menunjukkan bahwa A mendahului B, bahwa A berkorelasi dengan B yang ditentukan oleh validitas konklusi statistik, serta tidak ada kemungkinan penjelasan lain yang dapat menerangkan hubungan antara A dan B.

Berikut adalah daftar ancaman terhadap validitas internal, yaitu alasan-alasan mengapa inferensi bahwa ada hubungan kausal antara dua variabel mungkin tidak benar (Hastjarjo, 2011).

- a. Presedensi temporal yang kabur: kurangnya kejelasan variabel mana yang terjadi lebih dahulu mungkin akan menghasilkan kebingungan mengenai variabel mana yang menjadi penyebab dan mana yang menjadi efek/akibat.
- b. Seleksi: perbedaan sistematis ciri responden diantara kelompok eksperimen dan kontrol yang dapat juga menghasilkan efek yang teramati.
- c. Sejarah: kejadian yang berlangsung pada saat bersamaan dengan perlakuan dapat menghasilkan efek yang teramati.
- d. Maturasi: perubahan yang terjadi secara alamiah sepanjang waktu dapat keliru dikira sebagai efek perlakuan.
- e. Regresi: jika unit diseleksi berdasar skor ekstrim, seringkali akan mempunyai skor yang kurang ekstrim pada variabel lain (termasuk skor pada pemberian tes ulang), sebuah kejadian yang dapat keliru dikira sebagai efek perlakuan.
- f. Atrisi atau mortalitas: mundurnya/hilangnya responden saat perlakuan atau saat pengukuran dapat menghasilkan efek artifaktual jika peristiwa kehilangan tersebut secara sistematis berkorelasi dengan kondisi perlakuan.
- g. Pengujian: paparan sebuah tes dapat mengubah skor pada paparan tes selanjutnya, sebuah peristiwa yang dapat keliru dikira sebagai efek perlakuan.
- h. Instrumentasi: sifat dasar sebuah instrumen pengukuran mungkin berubah dalam cara tertentu sesuai dengan perubahan waktu atau kondisi sehingga perubahan tersebut dapat keliru dikira sebagai efek perlakuan.
- i. Efek aditif dan interaktif ancaman terhadap validitas internal: pengaruh sebuah ancaman dapat ditambahkan dengan pengaruh ancaman lain.

Validitas Eksternal

Validitas eksternal mengacu pada keterwakilan dan relevansi sampel, latar belakang, dan prosedur eksperimen. Ini menentukan seberapa umum atau relevan hasil eksperimen untuk konteks, populasi, atau situasi lain. Sebagai contoh ingin menguji bagaimana model pembelajaran mempengaruhi hasil belajar siswa, harus dipertimbangkan apakah siswa yang terlibat dalam eksperimen adalah siswa yang diinginkan, apakah lingkungan kelasnya sebanding dengan lingkungan kelas lainnya, dan apakah model pembelajaran tersebut layak dan sesuai untuk guru dan mata pelajaran lain.

Validitas eksperimen diukur berdasarkan seberapa dapat dipercaya dan bermakna hasilnya. Eksperimen tidak dapat memberikan informasi atau bukti yang dapat diandalkan tentang pertanyaan atau masalah penelitian jika tidak valid. Desain eksperimen, pemilihan peserta, metode pengumpulan dan analisis data, dan interpretasi hasilnya adalah beberapa faktor yang dapat memengaruhi validitas eksperimen.

Terdapat empat komponen validitas eksternal yang dikenal sebagai validitas X-, T-, Y-, dan C. Validitas X mengacu pada sejauh mana temuan sebab akibat dapat

digeneralisasikan ke populasi lain berdasarkan kovariat sebelum perlakuan. Validitas T berfokus pada generalisasi temuan kausal terhadap pengobatan atau intervensi lain. Validitas Y memeriksa kemampuan generalisasi temuan kausal terhadap hasil lain atau variabel terikat yang menjadi perhatian. Terakhir validitas C mempertimbangkan generalisasi temuan kausal pada konteks atau setting lain di mana penelitian dapat diterapkan. Keempat dimensi ini memberikan kerangka komprehensif untuk menilai validitas eksternal dari temuan kausal dalam penelitian ilmu sosial (Egami & Hartman, 2023).

Generalisasi efek (*effect-generalization*) dan generalisasi tanda (*sign-generalization*) berbeda dalam hal tujuan terkait konteks analisis validitas eksternal. Generalisasi efek berfokus pada generalisasi besaran efek sebab akibat pada populasi atau konteks yang lebih luas. Generalisasi efek sangat penting dalam studi yang mempunyai implikasi kebijakan dimana pemahaman mengenai besaran efek sebenarnya sangatlah penting. Para peneliti bertujuan untuk menilai bagaimana dampak kausal yang diamati dalam penelitian tertentu dapat diterapkan pada populasi, perlakuan, hasil, atau konteks yang berbeda. Sedangkan Generalisasi tanda bertujuan untuk mengevaluasi apakah arah akibat sebab akibat dapat digeneralisasikan pada berbagai latar atau populasi yang berbeda. Tujuan ini relevan ketika peneliti menguji mekanisme teoritis dan teori substantif yang mempunyai implikasi yang dapat diamati hanya pada arah efek pengobatan, bukan pada besaran pastinya. Generalisasi tanda dapat menjadi kompromi praktis ketika generalisasi efek tidak dapat dilakukan karena diperlukan asumsi yang lebih kuat. Dengan membedakan kedua tujuan tersebut, peneliti dapat menyesuaikan analisis validitas eksternal untuk fokus pada generalisasi besaran dampak atau menilai arah dampak berdasarkan tujuan penelitian spesifik dan data yang tersedia (Egami & Hartman, 2023).

Pada generalisasi efek terdapat tiga kelas penduga untuk memperkirakan Efek Perlakuan Rata-Rata Populasi Target (T-PATE). Penduga ini digunakan untuk menggeneralisasi besarnya dampak sebab akibat pada populasi atau konteks yang lebih luas. Pertama, estimator berbasis pembobotan. Estimator ini melibatkan pemberian bobot pada observasi dalam sampel agar lebih mewakili populasi target. Penduga berbasis pembobotan menyesuaikan perbedaan antara sampel dan populasi sasaran untuk memperkirakan dampak pengobatan rata-rata. Kedua, penduga berbasis hasil. Penduga berbasis hasil menggunakan informasi tentang hasil yang diinginkan untuk menyesuaikan potensi bias dalam memperkirakan dampak pengobatan. Dengan berfokus pada hasil, para penduga ini bertujuan untuk memberikan perkiraan yang lebih akurat mengenai dampak sebab akibat pada populasi sasaran. Terakhir, penduga yang kuat ganda. Penduga ini menggabungkan unsur-unsur pendekatan berbasis bobot dan berbasis hasil untuk meningkatkan ketahanan estimasi. Penduga ini memberikan perkiraan yang lebih andal dengan menggabungkan informasi dari penugasan pengobatan dan hasilnya. Dengan memanfaatkan ketiga kelas estimator ini, peneliti dapat meningkatkan kemampuan generalisasi efek kausal dari penelitian spesifik ke populasi atau konteks yang lebih

luas, sehingga meningkatkan validitas eksternal dari temuan (Egami & Hartman, 2023).

Dimensi konseptual inti pada validitas eksternal dibingkai dalam kerangka M – STOUT yang mencakup sebagai berikut: *Units* (unit), merujuk pada entitas atau individu yang menjadi tempat generalisasi temuan penelitian; *Treatments* (perlakuan), melibatkan intervensi atau kondisi yang diterapkan dalam penelitian dan kemampuan generalisasinya pada konteks lain; *Outcomes* (Hasil), berfokus pada hasil atau dampak yang diamati dalam penelitian dan penerapannya pada situasi yang berbeda; *Settings* (Latar), berkaitan dengan faktor lingkungan atau kontekstual di mana penelitian dilakukan dan relevansinya dengan latar lain; *Mechanisms* (mekanisme), mempertimbangkan proses atau mekanisme mendasar yang melaluinya pengobatan dapat memberikan hasil; dan *Time* (waktu), mengatasi dimensi temporal, termasuk bagaimana efek dapat berubah seiring waktu dan relevansi waktu dalam membuat kesimpulan (Findley *et al.*, 2015).

Para peneliti yang berorientasi metodologis telah memberikan kontribusi yang signifikan terhadap pemahaman validitas eksternal dalam beberapa tahun terakhir dengan cara menekankan pentingnya validitas eksternal sebagai mitra setara dengan validitas internal/inferensi kausal dalam kemajuan ilmu pengetahuan, mengembangkan kerangka kerja seperti M-STOUT untuk memperjelas dimensi validitas eksternal dan membedakan antara kemampuan generalisasi dan kemampuan pengangkutan, memberikan panduan bagaimana mengatasi kompleksitas dimensi validitas eksternal, termasuk unit, perlakuan, hasil, setting, mekanisme, dan waktu, mengadvokasi dimasukkannya diskusi khusus mengenai validitas eksternal dalam studi penelitian untuk memastikan pelaporan dan evaluasi yang akurat, mengusulkan kriteria evaluatif seperti kegunaan model, masuk akal cakupan, dan kredibilitas spesifikasi untuk menilai kredibilitas kesimpulan terkait kemampuan generalisasi dan pengangkutan, mendorong peneliti untuk melaporkan secara transparan seluruh dimensi M-STOUT dan mengkarakterisasi populasi teoritis, populasi yang dapat diakses, dan sampel untuk meningkatkan kredibilitas kesimpulan validitas eksternal (Findley *et al.*, 2015).

Pada artikel Ramspek *et al.*, (2021) diberikan gambaran umum tentang pentingnya validasi eksternal dalam menentukan reproduktifitas dan generalisasi model prediksi. Cakupan pembahasan dalam artikel ini meliputi pertimbangan metodologis utama ketika menilai atau merancang studi validasi eksternal, pentingnya validasi eksternal untuk implementasi klinis model prediksi yang akurat, dan cara melakukan analisis statistik dan interpretasi hasil validasi eksternal yang dapat membantu meningkatkan pelayanan pasien.

Saat menilai atau merancang studi validasi eksternal untuk model prognostic perlu mempertimbangkan pemilihan kelompok validasi, ukuran sampel, inklusi predictor, reproduksibilitas vs. transportabilitas, heterogenitas dalam efek predictor, validasi eksternal geografis, dan pembaruan validasi reguler. Pertimbangan ini penting untuk memastikan keandalan dan penerapan model prognostik dalam praktik klinis melalui studi validasi eksternal. Validasi eksternal sangat penting

untuk implementasi klinis model prediksi yang akurat karena alasan reproduksibilitas, kemampuan generalisasi, jaminan kualitas, utilitas klinis, mencegah pemborosan penelitian, dan keselamatan pasien. Dengan kata lain, validasi eksternal memainkan peran penting dalam membangun keandalan, generalisasi, dan kegunaan klinis model prediksi, yang pada akhirnya mendukung penerapan model prediksi yang aman dan efektif dalam praktik klinis (Ramspek *et al.*, 2021).

Analisis statistik dan interpretasi hasil validasi eksternal sangat penting untuk menerjemahkan temuan penelitian menjadi wawasan yang dapat ditindaklanjuti dan dapat berdampak langsung pada perawatan pasien. Dengan memanfaatkan model prediksi yang tervalidasi secara efektif, penyedia layanan kesehatan dapat meningkatkan pengambilan keputusan, mengoptimalkan alokasi sumber daya, dan pada akhirnya meningkatkan hasil pasien (Ramspek *et al.*, 2021).

Artikel Polit & Beck (2010) membahas mitos dan strategi seputar topik generalisasi dalam penelitian. Pentingnya mengatasi generalisasi dalam penelitian kuantitatif dan kualitatif untuk memastikan bahwa wawasan diakui sebagai bukti berharga untuk praktik. Makalah ini membahas mengenai tiga model generalisasi, mitos tentang kepatuhan terhadap model-model yang menghalangi peluang untuk melakukan generalisasi dalam penelitian, serta penerapan strategi oleh peneliti perawat untuk meningkatkan kesiapan melakukan generalisasi dalam penelitian mereka.

Terdapat tiga model generalisasi yaitu Model Generalisasi Statistik Klasik, Generalisasi Analitik, dan Model Transfer Kasus-ke-Kasus (Transferabilitas). Model Generalisasi Statistik Klasik melibatkan ekstrapolasi dari sampel ke populasi, yang merupakan model tradisional yang mendasari sebagian besar penelitian kuantitatif. Selanjutnya Model Generalisasi Analitik relevan baik dalam penelitian kualitatif maupun kuantitatif dan berfokus pada penarikan kesimpulan yang lebih luas di luar kasus spesifik yang diteliti. Terakhir Model Transfer Kasus-ke-Kasus (Transferabilitas) dikenal sebagai transferabilitas atau penerjemahan kasus-ke-kasus, melibatkan penggunaan temuan dari satu konteks dan menerapkannya pada kelompok orang atau lingkungan berbeda (Polit & Beck, 2010).

Mitos tentang kepatuhan terhadap model generalisasi dapat menghalangi peluang untuk melakukan generalisasi dalam penelitian dengan menciptakan kesalahpahaman dan hambatan dalam menerapkan temuan penelitian secara efektif. Mitos-mitos ini mungkin membuat peneliti mengabaikan nuansa dan kompleksitas yang ada dalam menggeneralisasi hasil penelitian dalam konteks yang berbeda. Misalnya, kepatuhan yang ketat terhadap model generalisasi statistik klasik dapat membatasi kemampuan peneliti untuk mengeksplorasi implikasi yang lebih luas dari temuan di luar populasi sampel. Demikian pula, kesalahpahaman tentang penerapan penelitian kualitatif dapat menghalangi peneliti untuk mengenali potensi transferabilitas wawasan kualitatif ke berbagai latar belakang. Dengan menghilangkan prasangka mitos-mitos ini dan mendorong pemahaman yang lebih mendalam mengenai model generalisasi, para peneliti dapat meningkatkan kualitas

dan relevansi temuan penelitian untuk penerapan yang lebih luas dalam praktik dan kebijakan (Polit & Beck, 2010).

Peneliti keperawatan dapat mengadopsi strategi untuk meningkatkan kesiapan generalisasi dalam penelitian. Pertama peneliti dapat memanfaatkan tiga model generalisasi. Kedua, menantang dan menghilangkan prasangka mitos seputar generalisasi dalam penelitian kuantitatif dan kualitatif untuk mendorong pemahaman yang lebih beragam tentang bagaimana temuan penelitian dapat diterapkan dalam konteks yang berbeda. Selanjutnya berusaha keras untuk mengembangkan bukti penelitian dengan validitas dan integritas tinggi, karena ini merupakan prasyarat penting untuk mencapai tujuan generalisasi. Berikutnya identifikasi potensi penyimpangan dalam generalisasi dan upayakan untuk memberikan kompensasi guna meningkatkan kualitas keseluruhan dan penerapan bukti penelitian, dan terakhir pastikan bahwa temuan penelitian memiliki implikasi praktis terhadap praktik layanan kesehatan dengan mengeksplorasi bagaimana temuan tersebut dapat diterjemahkan ke dalam rekomendasi yang dapat ditindaklanjuti bagi para praktisi (Polit & Beck, 2010).

Estimasi sejauhmana sebuah hubungan kausal berlaku sepanjang variasi orang, seting, perlakuan dan dampak perlakuan secara konseptual mirip dengan pengujian interaksi statistik. Misalnya jika ada interaksi antara perlakuan pendidikan dengan kelas sosial anak, maka peneliti tidak dapat mengambil kesimpulan bahwa hasil yang sama akan berlaku pada semua kelas sosial anak. Oleh sebab itu ancaman terhadap validitas eksternal dikemas dalam interaksi hubungan kausal dengan unit, perlakuan, dampak perlakuan dan seting. Konsep yang melatarbelakangi interaksi itu lebih penting: mencari cara mengapa sebuah hubungan kausal dapat atau tidak dapat berubah sepanjang variasi orang, perlakuan, dampak perlakuan dan seting (Hastjarjo, 2011).

Ancaman terhadap validitas eksternal adalah alasan-alasan mengapa inferensi mengenai hasil eksperimen dapat berlaku sepanjang variasi orang, seting, perlakuan dan dampak perlakuan mungkin salah. Daftar ancaman terhadap validitas eksternal sebuah eksperimen (Hastjarjo, 2011):

- a. Interaksi antara hubungan kausal dengan unit: sebuah efek yang ditemukan dengan jenis unit tertentu mungkin tidak berlaku seandainya unit lain yang telah diteliti.
- b. Interaksi antara hubungan kausal dengan variasi perlakuan: sebuah efek yang ditemukan dengan sebuah variasi perlakuan tertentu mungkin tidak berlaku dengan variasi lain perlakuan tersebut, atau jika perlakuan itu dikombinasikan dengan perlakuan lain, atau jika hanya sebagian dari perlakuan itu digunakan.
- c. Interaksi antara hubungan kausal dengan dampak perlakuan: sebuah efek yang ditemukan dengan satu jenis observasi dampak perlakuan mungkin tidak berlaku jika jenis lain observasi dampak perlakuan digunakan.
- d. Interaksi antara hubungan kausal dengan seting: sebuah efek yang ditemukan dengan satu jenis seting mungkin tidak berlaku jika jenis seting lain digunakan.

- e. Mediasi tergantung konteks: sebuah mediator yang menjelaskan hubungan kausal dalam satu konteks mungkin tidak akan memediasi dalam konteks lain.

KESIMPULAN DAN SARAN

Validitas internal dan eksternal adalah dua konsep kunci dalam penelitian eksperimental. Hal ini mengacu sejauh mana hubungan sebab akibat antara variabel independen dan variabel dependen; dan sejauh mana hasil suatu eksperimen dapat digeneralisasikan ke lingkungan dan situasi lain. Menyiapkan eksperimen agar memiliki validitas internal dan eksternal yang baik memerlukan perhatian sejak awal terhadap faktor-faktor yang dapat mempengaruhi setiap aspek dalam penelitian.

DAFTAR PUSTAKA

- Andrade, C. (2018). Learning Curve Internal , External , and Ecological Validity in. *Indian J Psychol Med*, 40(4), 498–499. <https://doi.org/10.4103/IJPSYM.IJPSYM>
- Brunswick. (1947). *Systematic and Representative Designs of Psychological Experiments. In Proceedings of The Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: Univ. Calif. Press.
- Campbell. (1957). Factors Relevant to The Validity of Experiments in social Settings. *Psychol. Bull.*
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105. <https://doi.org/10.1037/h0046016>
- Campbell, T. D., & Stanley, J. C. (1963). *Experimental and Quasi Experimental Designs for Research*. Chicago: IL:Rand McNally.
- Dehejia, R., Pop-Eleches, C., & Samii, C. (2021). From Local to Global: External Validity in a Fertility Natural Experiment. *Journal of Business and Economic Statistics*, 39(1), 217–243. <https://doi.org/10.1080/07350015.2019.1639407>
- Egami, N., & Hartman, E. (2023). Elements of External Validity: Framework, Design, and Analysis. *American Political Science Review*, 117(3), 1070–1088. <https://doi.org/10.1017/S0003055422000880>
- Findley, M., Kikuta, K., & Denly, M. (2015). External Validity. *Dictionary of Statistics & Methodology*. <https://doi.org/10.4135/9781412983907.n709>
- Fromer, R., Lin, H., Dean Wolf, C. K., Inzlicht, M., & Shenhav, A. (2021). Expectations of reward and efficacy guide cognitive control allocation. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-21315-z>
- Gluth, Kern, Kortmann, & Vitali. (2020). Value based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*.
- Guala, F., & Mittone, L. (2005). Experiments in economics: External validity and the robustness of phenomena. *Journal of Economic Methodology*, 12(4), 495–515. <https://doi.org/10.1080/13501780500342906>
- Hastjarjo, D. T. (2011). Validitas Eksperimen. *Buletin Psikologi*, 19(2), 70–80.
- Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences of the United States of America*, 114(48), 12696–12701. <https://doi.org/10.1073/pnas.1715293114>
- Lin, H., Werner, K. M., & Inzlicht, M. (2021). Promises and Perils of Experimentation:

EduInovasi: Journal of Basic Educational Studies

Vol 4 No 3 (2024) 987 - 1001 P-ISSN 2774-5058 E-ISSN 2775-7269

DOI: 47467/eduinovasi.v4i3.2734

The Mutual-Internal-Validity Problem. *Perspectives on Psychological Science*, 16(4), 854–863. <https://doi.org/10.1177/1745691620974773>

Manzi. (2012). *Uncontrolled: The surprising payoff of trial and error for business, politics, and society*. Basic Books.

Pearl, & Mackenzie. (2018). *The book of why*. Penguin.

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies*, 47(11), 1451–1458. <https://doi.org/10.1016/j.ijnurstu.2010.06.004>

Quintao, C., Andrade, P., & Almeida, F. (2020). How to Improve the Validity and Reliability of a Case Study Approach. *Journal of Interdisciplinary Studies in Education*, 9(2), 264–275. <https://doi.org/10.32674/jise.v9i2.2026>

Ramspek, C. L., Jager, K. J., Dekker, F. W., Zoccali, C., & Van Diepen, M. (2021). External validation of prognostic models: What, why, how, when and where? *Clinical Kidney Journal*, 14(1), 49–58. <https://doi.org/10.1093/ckj/sfaa188>

Shadish, Cook, & Campbell. (2002). *Experimental and Quasi Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Shenhav, A., Musslick, S., Botvinick, M. M., & Cohen, J. D. (2020). Misdirected vigor: Differentiating the control of value from the value of control. *Angewandte Chemie International Edition*, 6(11), 951–952., (401), 5–24. Diambil dari [http://repo.iain-tulungagung.ac.id/5510/5/BAB 2.pdf](http://repo.iain-tulungagung.ac.id/5510/5/BAB%202.pdf)

Trochim, & Donnelly. (2006). *The Research Methods Knowledge Base*. Cincinnati, OH: Atomic Dog. 3rd ed.

Vassena, E., Deraeve, J., & Alexander, W. H. (2020). Surprise, value and control in anterior cingulate cortex during speeded decision-making. *Nature Human Behaviour*, 4(4), 412–422. <https://doi.org/10.1038/s41562-019-0801-5>