**Pertemuan 4, Online 3**

**VALIDITAS DALAM PENELITIAN EKSPERIMENTAL**

**Sumber:**

**Seniati, L., Yulianto, A., & Setiadi, B.N. (2015). Psikologi Eksperimen. Jakarta: PT. Indeks.**

 Pengertian validitas secara umum menyangkut dua hal, yaitu validitas alat ukur dan validitas penelitian. Validitas alat ukur berkaitan dengan seberapa besar suatu alat ukur mampu mengukur apa yang ingin diukur,berkaitan dengan hubungan sebab akibat yang dihasilkan. Berbeda dengan validitas alat ukur, validitas penelitian berkaitan dengan kontrol terhadap variabel sekunder.

 Ada dua jenis validitas dalam penelitian, yaitu validitas internal dan validitas eksternal. **Validitas internal** berkaitan dengan sejauhmana hubungan sebab-akibat antara VB dan VT yang ditemukan dalam penelitian. Semakin kuat hubungan sebab-akibat antara VB dan VT maka semakin besar validitas internal suatu penelitian. **Validitas eksternal** berkaitan dengan generalisasi hasil penelitian, yaitu sejauh mana hasil penelitian dapat diterapkan pada subjek, situasi, dan waktu di luar situasi penelitian. Pada penelitian eksperimental, validitas yang ingin dicapai adalah validitas internal karena penelitian eksperimental merupakan penelitian yang memberikan VB untuk dilihat pengaruhnya terhadap VT.

**FAKTOR-FAKTOR YANG MEMPENGARUHI VALIDITAS INTERNAL**

 Ada beberapa faktor yang dapat mempengaruhi kekuatan hubungan sebab-akibat antara VB dengan VT, yaitu (Robinson, 1981, Christensen, 2001 ):

***Proactive history***

 Faktor ini merupakan faktor perbedaan individual yang dibawa ke dalam penelitian. Yang termasuk *proactive history* antara lain: usia, jenis kelamin, kepribadian, sikap, inteligensi, dan sebagainya. Faktor-faktor ini dapat mempengaruhi VT, baik secara sendiri maupun berinteraksi dengan VB. Hal ini mungkin saja terjadi pada penelitian eksperimental. Oleh karena itu dapat muncul pertanyaan: “Bila ada perbedaan nilai VT antara KK dengan KE, apakah perbedaan tersebut diakibatkan oleh VB ataukah karena proactive history yang tidak dikontrol?”

 Tidak semua faktor yang termasuk *proactive history* dapat mempengaruhi VT, hanya faktor-faktor tertentu yang relevan saja dengan penelitian. Misalnya, penelitian mengenai pengaruh metode mengingat terhadap ingatan; inteligensi dan usia menjadi faktor yang lebih relevan dibandingkan jenis kelamin atau sikap. Bila para subjek pada KK dan KE memiliki inteligensi atau usia yang jauh berbeda, maka perbedaan VT pada kedua kelompok tersebut mungkin disebabkan oleh inteligensi atau usia dan bukan karena metode mengingat. Karena faktor-faktor ini merupakan faktor di luar VB yang mempengaruhi VT, maka dapat mengurangi kekuatan hubungan sebab akibat antara VB dengan VT. Karena itu *proactive history* merupakan variabel sekunder dan perlu dikontrol. Kontrol terhadap inteligensi dan usia pada contoh di atas dapat dilakukan dengan konstansi, yaitu subjek yang digunakan pada kedua kelompok memiliki tingkat inteligensi dan tingkat usia yang sama.

***Retroactive history***

 Beberapa buku menyebut faktor ini sebagai *history* saja (Martin, 1991; Christensen, 2001). Faktor ini hanya terjadi pada penelitian yang menggunakan *pretest-posttest*, di mana setiap subjek mengalami pengukuran VT sebanyak dua kali, yaitu sebelum dan sesudah dilakukannya eksperimen, dan juga ada jarak waktu di antaraa pengukuran tersebut. Perubahan atau pengaruh yang dialami subjek di antara waktu pemberian *pretest* dengan *posttest* tersebut dapat mempengaruhi perubahan VT. Tidak ada faktor-faktor umum, semuanya tergantung apa yang terjadi di antara *posttest* dengan *pretest*.

 Misalnya, penelitian mengenai pengaruh kampanye HIV/AIDS terhadap perubahan perilaku seks mahasiswa. Setelah diberi *pretest* (berupa skala sikap), ternyata ada seorang teman mereka yang meninggal akibat HIV/AIDS. Kejadian ini dapat membuat mereka mengurangi perilaku seks mereka agar tidak tertular HIV/AIDS. Dua minggu kemudian mereka diberi skala sikap kembali (*posttest*) untuk mengetahui perubahan perilaku mereka setelah diberikan kampanye. Dari hasil analisis statistik ternyata ada perbedaan perilaku pada KB dan KK, sehingga disimpulkan kampanye HIV/AIDS berpengaruh pada perubahan perilaku seks. Padahal kenyataannya perubahan perilaku tersebut bukan disebabkan oleh kampanye HIV/AIDS, melainkan kematian seorang rekan akibat HIV / AIDS. Karena itu retroactive history merupakan variabel sekunder yang perlu dikontrol.

***Maturation***

 Maturation atau kematangan adalah perubahan biologis dan atau perubahan psikologis yang sistematis pada organisme dalam suatu waktu tertentu. Faktor ini lebih mungkin teijadi pada penelitian jangka panjang (longitudinal). Faktor ini juga lebih sering terjadi pada penelitian yang menggunakan anak-anak, karena perubahan biologis dan psikologis mereka terjadi lebih cepat. Perubahan biologis maupun psikologis yang terjadi pada subjek dapat mempengaruhi VT. Karena faktor ini bukan VB maka dapat mempengaruhi validitas internal. *Maturation* dapat diatasi dengan menggunakan kelompok kontrol, yaitu menggunakan kelompok subjek lain yang tidak diberikan perlakuan VB.

 Contohnya, penelitian tentang pengaruh aktivitas bermain di luar ruang terhadap kemampuan motorik anak TK. Di masa awal sekolah, para guru melihat kemampuan motorik sebagian besar siswa kurang baik. Untuk itu, para guru memutuskan untuk lebih banyak melibatkan anak pada aktivitas bermain di luar ruangan dibandingkan di dalam ruangan. Di akhir tahun ajaran, para guru melihat kemampuan motorik para siswa lebih baik dibandingkan kemampuan sebelumnya. Mereka menyimpulkan penyebabnya adalah aktivitas luar ruangan yang dilakukan selama satu tahun. Namun kesimpulan ini bisa saja salah karena mungkin saja tanpa aktivitas luar ruang yang dilakukan di TK pun, kemampuan motorik para siswa lebih baik dari sebelumnya akibat perkembangan biologis yang cepat.

***Testing***

 Dalam melakukan penelitian, peneliti dapat memberikan pretest dan posttest kepada subjek untuk melihat perbedaan sebelum dan sesudah pemberian perlakuan. Seringkali tes yang diberikan pada dua waktu yang berbeda tersebut merupakan tes yang sama. Dengan kondisi ini kemungkinan skor yang diperoleh subjek pada *posttest* akan berbeda, baik lebih tinggi ataupun lebih rendah, dibandingkan skor pada *pretest*. Hal ini terjadi karena subjek berusaha untuk mengingat kembali atau mempelajari hal-hal yang berkaitan dengan soal atau jawaban yang diberikan, sehingga respons yang terukur bukan merupakan pengaruh dari pemberian VB. Dengan kata lain, tanpa diberikan VB pun, skor *posttest* subjek akan berbeda dari skor *pretest*. Dengan demikian testing dapat menurunkan kekuatan hubungan sebab-akibat antara VB dengan VT.

 Contoh gagal/tidak sesuai, pada penelitian pengaruh film kekerasan terhadap perubahan sikap, subjek diminta untuk mengisi kuesioner mengenai sikap terhadap kekerasan sebelum dipertontonkan film. Pemberian kuesioner ini menyebabkan subjek menjadi sensitif dan selektif ketika menonton film tersebut. Akibatnya ketika berespons dalam *posttest*, jawaban yang diberikan tidak sesuai dengan pengaruh film tersebut.

 Untuk mengatasi masalah yang berkaitan dengan testing, ada beberapa cara yang dapat dilakukan, yaitu:

* Bila tidak terlalu dibutuhkan, jangan menggunakan pretest.
* Gunakan *posttest* yang tidak sama tapi setara dengan *pretest*. Setara maksudnya adalah sama dalam hal: apa yang diukur, banyaknya soal, cara penyajian soal, cara berespons yang diminta, batas waktu pengerjaan, tapi dengan bentuk soal yang berbeda. Hal ini agar subjek tidak mengingat kembali jawaban atau berespons dari hasil belajar, diluar hasil perlakuan VB.
* Sisipkan pernyataan atau pertanyaan pengecoh pada *pretest* sehingga tidak mengarahkan subjek untuk menduga permasalahan penelitian atau perlakuan yang diteliti.
* Gunakan desain penelitian yang melibatkan kelompok yang tidak diberi pretest(misalnya *the Solomon fourgroup design*).

***Statistical regression***

 Dalam hukum statistik, pengukuran yang dilakukan secara berulangulang akan menyebabkan nilai ekstrem, yaitu nilai tertinggi dan nilai terendah, cenderung mendekati nilai rata-rata, meskipun tidak diberikan perlakuan apapun. Skor pada kelompok tertinggi cenderung akan menjadi lebih rendah, sedangkan skor pada kelompok terendah cenderung akan menjadi lebih tinggi dibandingkan sebelumnya. Karena itu faktor statistical regression hanya terjadi pada penelitian yang melibatkan *pretest* dan *posttest*.

 *Statistical regression* dapat terjadi apabila alat ukur yang digunakan tidak reliabel, sehingga menyebabkan ketidak konsistenan skor subjek antara *pretest* dan *posttest*. Ketidak konsistenan ini menyebabkan pengukuran yang tidak akurat, sehingga merendahkan validitas internal dari penelitian. Dalam penelitian eksperimental, *statistical regression* ini dapat dihindari apabila subjek yang digunakan hanya berasal dari satu kelompok ekstrim saja, yaitu hanya kelompok subjek dengan skor tinggi atau hanya kelompok subjek dengan skor rendah.

***Experimental mortality***

 Pada penelitian eksperimental yang melibatkan *pretest-posttest* dalam jangka waktu cukup lama ataupun pada penelitian *Within-subject*, seringkali jumlah subjek pada akhir penelitian berkurang dibandingkan dengan ketika awal penelitian. Hal ini mungkin saja disebabkan ada subjek yang meninggal, menderita sakit, mengalami kecelakaan, atau tidak bersedia mengikuti penelitian hingga selesai. Hal ini menyebabkan penurunan jumlah subjek, sehingga berpengaruh pada analisis statistik yang dilakukan. Semakin banyak jumlah subjek, semakin besar kemungkinan Ho ditolak (ada pengaruh VB terhadap VT). Bila ada perbedaan jumlah subjek pada *pretest* dan *posttest*, maka yang digunakan adalah jumlah subjek pada *posttest* karena merekalah yang memiliki data skor yang lengkap, yaitu skor *pretest* dan *postest*, yang akan digunakan dalam perhitungan statistik. Dengan demikian, subjek yang tidak memiliki skor *posttest* tidak diikutsertakan dalam perhitungan statistik. Validitas internal penelitian menjadi rendah seiring semakin berkurangnya subjek yang digunakan, karena memperkecil kemungkinan signifikansi hasil penelitian. Tidak ada cara yang dapat dilakukan untuk mengatasi faktor ini.

**Interaction effect**

 Faktor ini dapat terjadi pada desain penelitian eksperimental withinsubject, di mana setiap subjek mendapat lebih dari satu perlakuan. Pengaruh dari perlakuan yang diterima subjek sebelumnya belum hilang benar sehingga dapat berinteraksi dengan perlakuan selanjutnya dalam mempengaruhi VT. Dengan kondisi ini, hubungan sebab-akibat yang dihasilkan akan lemah, seningga mengurangi validitas internal penelitian. Dengan kata lain, dapat dikatakan bahwa urutan perlakuan tertentu dari variasi VB yang diberikan pada setiap subjek dapat menimbulkan perbedaan hasil pengukuran VT dibandingkan apabila diberikan urutan perlakuan yang berbeda. Karena itu *interaction effect* disebut juga sequencing effect (Christensen, 2001).

 Pada penelitian pengaruh jenis musik terhadap keadaan emosi, misalnya, jenis musik divariasikan menjadi musik klasik, musik dangdut, dan musik rock. Semua jenis musik ini yang akan diperdengarkan pada setiap subjek. Bila musik rock diberikan pertama kali, kemudian musik klasik, kemungkinan saat diperdengarkan musik klasik subjek masih terbawa emosi ketika mendengarkan musik rock.

 Untuk mengatasi *interaction effect*, dilakukan *counterbalancing*, yaitu memberikan urutan variasi VB yang berbeda pada subjek penelitian. Untuk contoh di atas, dilakukan *counterbalancing* dengan cara: beberapa subjek diperdengarkan dengan urutan: musik dangdut, musik rock, dan musik klasik; subjek lainnya diperdengarkan musik rock, musik klasik, lalu musik dangdut; dan beberapa subjek sisanya diperdengarkan musik klasik, musik dangdut, dan musik rock. Perhatikan tabel di bawah ini.

|  |  |  |  |
| --- | --- | --- | --- |
| Subjek | Perlakuan 1 | Perlakuan 2 | Perlakuan 3 |
| A | Musik dangdut | Musik rock | Musik klasik |
| B | Musik dangdut | Musik rock | Musik klasik |
| C | Musik dangdut | Musik rock | Musik klasik |
| D | Musik rock | Musik klasik | Musik dangdut |
| E | Musik rock | Musik klasik | Musik dangdut |
| F | Musik rock | Musik klasik | Musik dangdut |
| G | Musik klasik | Musik dangdut | Musik rock |
| H | Musik klasik | Musik dangdut | Musik rock |
| I | Musik klasik | Musik dangdut | Musik rock |

***lnstrumentation effect***

 Alat ukur yang digunakan dalam penelitian dapat turut mempengaruhi validitas internal penelitian. Alat ukur yang tidak akurat akan menurunkan validitas internal dari penelitian eksperimental yang dilakukan. Dalam penelitian psikologi, alat ukur yang digunakan biasanya berupa tes, kuesioner/skala, ataupun lembar observasi/ wawancara. Agar penelitian memiliki validitas internal yang tinggi, alat ukur yang digunakan harus valid dan reliable.

 Keahlian atau keterampilan dari pengadministrasi tes (*tester*), pengamat ataupun pewawancara juga turut mempengaruhi validitas internal suatu penelitian. Untuk mencegah terjadi faktor instrumentation, maka tester, pengamat, atau pewawancara perlu diberikan pelatihan terlebih dahulu sebelum melakukan pengambilan data.

***Experimenfer effect***

 Dalam suatu penelitian yang melibatkan manusia, interaksi antara eksperimenter dengan subjek penelitian turut mempengaruhi validitas internal penelitian. Baik eksperimenter maupun subjek penelitian, saling memiliki harapan berkaitan dengan perannya, sehingga perilaku ataupun pikiran dari kedua belah pihak tersebut, baik secara sengaja atau tidak sengaja, dapat mempengaruhi keakuratan dari penelitian yang dilakukan. Berkaitan dengan hal ini, maka ada dua sumber yang mempengaruhi validitas internal penelitian, yaitu eksperimenter dan subjek penelitian.

**Bias eksperimenter**

 Bias eksperimenter ini disebut juga sebagai *experimenter effect* (Christensen, 2001). Seringkali perilaku ataupun karakteristik yang ada pada eksperimenter dapat menimbulkan bias, sehingga dapat mempengaruhi hasil penelitian. Ada dua bentuk bias yang dipengaruhi eksperimenter, yaitu **atribut eksperimenter** dan **harapan eksperimenter**.

**Atribut eksperimenter**

 Yang dimaksud dengan atribut eksperimenter adalah karakteristik fisik dan psikologis dari eksperimenter yang mungkin berinteraksi dengan VB dalam mempengaruhi VT. Karakteristik tersebut meliputi atribut biososial (misalnya usia, jenis kelamin, ras, dan agama); atribut psikososial (misalnya kehangatan, kecerdasan, atau agresivitas), dan faktor situasional.

Contohnya penelitian pengaruh kehadiran pengasuh (VB) terhadap rasa aman (VT) di hari-hari pertama sekolah pada anak usia balita. Penelitian dilakukan pada dua sekolah TK, di mana sekolah pertama mengizinkan sedangkan sekolah kedua tidak mengizinkan pengasuh hadir di kelas bersama anak. Namun, ternyata atribut biososial guru TK kedua sekolah berbeda. Sebagian besar guru pada sekolah pertama adalah wanita berusia lebih dari 40 tahun, sedangkan pada sekolah kedua lebih banyak guru wanita dan pria berusia 20-25 tahun. Akibatnya, rasa aman yang diukur tidak hanya dipengaruhi oleh kehadiran pengasuh tetapi juga oleh jenis kelamin dan usia guru yang berpartisipasi dalam penelitian.

**Harapan eksperimenter**

 Bias ini diakibatkan oleh harapan dari eksperimenter terhadap hasil penelitian yang dilakukan. Harapan ini dapat mengarahkan eksperimenter secara tidak sengaja untuk berperilaku tertentu sehingga dapat menyebabkan bias dalam penelitian, karena dapat mempengaruhi baik eksperimenter sendiri maupun subjek.

 Misalnya penelitian tentang pengaruh tayangan agresif terhadap agresivitas anak memiliki hipotesis bahwa tayangan agresif akan menyebabkan perilaku agresif pada anak. Saat melakukan observasi pada anak yang diberi tayangan agresif, eksperimenter lebih sering menganggap sebagian besar perilaku anak sebagai perilaku agresif, sedangkan pada anak yang tidak diberi tayangan agresif, perilaku yang sama tidak dianggap sebagai perilaku agresif. Harapan eksperimenter dapat dikontrol dengan melakukan *doubIe-blind procedure*, yaitu eksperimenter tidak mengetahui kelompok subjek manakah yang akan diberi manipulasi VB. *DoubIe-blind* artinya baik eksperimenter dan juga subjek tidak mengetahui mengenai perlakuan yang akan diberikan dalam penelitian.

**Efek Partisipan**

 Sebagai makhluk yang memiliki kemampuan kognitif yang tinggi, subjek penelitian seringkali berusaha untuk mencari tahu apa yang akan mereka alami, apa yang harus mereka lakukan, serta merencanakan respons yang akan diberikan. Hal ini dapat mereka ketahui dari cara eksperimenter menyambut mereka, dari instruksi yang diberikan, tugas yang harus mereka kerjakan, dari setting penelitian dan peralatan yang digunakan, dan juga informasi yang mereka dengar tentang penelitian tersebut. Hal-hal ini disebut sebagai demand characteristics (Christensen, 2001), yang membuat subjek termotivasi untuk berespons tertentu. Respons subjek ini bukan disebabkan manipulasi VB yang diberikan dan dapat mempengaruhi hasil penelitian, sehingga menurunkan validitas internal dari penelitian tersebut.

 Contohnya pada penelitian mengenai pengaruh jenis musik terhadap agresivitas yang dilakukan pada sebuah ruangan yang dilengkapi cermin satu arah (one-way mirror), beberapa subjek menyadari bahwa ketika mereka sendirian di dalam ruangan tersebut, sebenarnya eksperimenter mengamati dari ruangan lain di balik cermin. Karena itu mereka berusaha untuk berperilaku wajar atau biasa saja ketika di dalam ruangan tersebut, meskipun peneliti memberikan perlakuan berupa jenis musik yang diperkirakan akan meningkatkan agresivitas.

***Participant sophistication***

 Pengetahuan dan familiaritas subjek penelitian terhadap topik penelitian atau metode eksperimental yang dilakukan dapat mempengaruhi hasil penelitian. Karena itu, penelitian psikologi sering dilakukan pada mahasiswa tingkat awal atau orang awam yang pengetahuannya masih terbatas mengenai penelitian eksperimental.

 Contohnya penelitian tentang legitimasi kepatuhan terhadap agresivitas yang dilakukan oleh Milgram. Subjek diminta untuk memberikan sengatan listrik pada seorang murid setiap kali melakukan kesalahan dalam proses belajar. Sengatan listrik yang diberikan semakin meningkat, mulai dari 15 volt hingga 450 volt, setiap kali murid tersebut melakukan kesalahan dalam menjawab pertanyaan. Murid tersebut sebenarnya merupakan bagian dari peneliti (*confederate*), sehingga ketika subjek penelitian memberikan sengatan listrik, murid tersebut tampak berpura-pura mengalami rasa sakit karena memang sesungguhnya tidak ada sengatan listrik. Hasilnya cukup mengejutkan, di mana 65% subjek penelitian terus memberikan sengatan listrik sampai 450 volt. Padahal, dengan sengatan 220 volt seperti di rumah saja seseorang bisa mengalami kematian. Hasil penelitian ini dimungkinkan karena subjek tidak memiliki pengetahuan tentang voltage listrik. jika subjek memiliki pengetahuan, manipulasi tersebut tidak akan memberikan hasil yang diharapkan.

**FAKTOR-FAKTOR YANG MEMPENGARUHI VALIDITAS EKSPERIMENTAL**

 Secara umum ada tiga faktor yang dapat mempengaruhi validitas eksternal, yaitu: validitas populasi, validitas ekologis, dan validitas temporal (Christensen, 2001).

**Validitas populasi**

 Validitas populasi berkaitan dengan kemampuan hasil suatu penelitian untuk digeneralisasikan dari sampel penelitian kepada populasi yang lebih besar. Validitas populasi berkaitan dengan teknik pengambilan sampel, apakah dilakukan secara acak atau tidak. Bila dilakukan secara acak (*random sampling*), maka validitas populasi akan semakin tinggi. Dalam kenyataannya, *random sampling* sulit untuk dilakukan karena tidak diketahui jumlah anggota sebenarnya atau karena anggota populasi sulit dijangkau oleh peneliti.

 Ada dua sumber sampel yang perlu dibedakan, yaitu **populasi target** dan **populasi yang dapat dijangkau peneliti (*experimentally accessible population*)**. Populasi target adalah populasi lebih besar di mana hasil penelitian akan digeneralisasikan, sedangkan populasi yang dapat dijangkau peneliti adalah kelompok populasi yang tersedia untuk peneliti.

 Misalnya penelitian mengenai pengaruh media pengajaran terhadap prestasi mahasiswa. Pada penelitian ini, yang menjadi populasi target adalah semua mahasiswa di lndonesia, sedangkan populasi yang dapat dijangkau adalah mahasiswa Fakultas Psikologi Universitas lndonesia karena peneliti lebih mudah dalam mengambil subjeknya. Berdasarkan dua sumber sampel ini, maka generalisasi hasil penelitian memiliki dua tingkatan. Pertama, generalisasi dari sampel (subjek penelitian) kepada *experimentally accessible population*. Kedua, generalisasi dari *experimentally accessible population* kepada populasi target.

 Karena berkaitan dengan pengambilan sampel, maka validitas populasi dipengaruhi oleh bias seleksi. Bias seleksi merupakan kesalahan dalam mengambil sampel yang tidak sesuai dengan karakteristik dari subjek penelitian. Karena subjek penelitiannya tidak mewakili populasi maka validitas eksternal penelitian menjadi rendah.

 Misalnya, penelitian mengenai pengaruh metode pengajaran terhadap prestasi belajar siswa SMU dilakukan pada suatu SMU swasta di Jakarta, dimana semua siswanya berjenis kelamin perempuan dan berstatus sosial ekonomi tinggi. Bila hasil penelitian ini akan diterapkan pada SMU negeri, maka kemungkinannya kecil untuk diperoleh hasil yang sama karena ada perbedaan antara sampel penelitian yang di gunakan dengan populasi. Hasil penelitian tersebut hanya dapat diterapkan pada SMU yang juga siswanya berjenis kelamin dan berstatus sosial ekonomi tinggi pula.

**Validitas ekologis**

 Bila validitas populasi berkaitan dengan subjek, maka validitas ekologis berkaitan dengan situasi atau kondisi lingkungan. Kemampuan hasil penelitian untuk digeneralisasikan pada situasi atau kondisi lingkungan yang berbeda disebut validitas ekologis. Validitas ekologis suatu penelitian dapat menjadi tinggi apabila pengaruh dari manipulasi VB tidak terkait dengan setting penelitian tersebut. Dengan kata lain, hasil penelitian dapat diterapkan walaupun pada situasi yang berbeda dengan situasi penelitian. Ada empat faktor yang dapat mempengaruhi validitas ekologis, yaitu:

***Multipe-treatment interference***

 Faktor ini berkaitan dengan pengaruh perlakuan yang diberikan sebelumnya terhadap perlakuan lain yang akan diberikan selanjutnya. Selain mempengaruhi validitas internal, urutan perlakuan juga turut mempengaruhi validitas eksternal karena generalisasi hasil penelitian ini terbatas hanya pada situasi di mana urutan perlakuan sama seperti pada kondisi dilakukannya penelitian.

 Misalnya, pada contoh penelitian sebelumnya mengenai pengaruh jenis musik terhadap keadaan emosional, setiap subjek diberi perlakuan dengan diperdengarkan secara berurutan musik klasik, musik dangdut, dan musik rock. Hasil ini sulit untuk diterapkan pada situasi lain, di mana subjek diberikan urutan yang berbeda, misalnya musik dangdut, musik klasik, kemudian musik rock.

 Faktor ini tidak hanya dapat terjadi pada penelitian *within-subject* di mana setiap subjek mendapat lebih dari satu perlakuan, tetapi juga dapat terjadi pada penelitian *between-subject* apabila keikutsertaan subjek dalam penelitian sebelumnya dapat mempengaruhi respons pada penelitian yang sedang diikuti.

 Misalnya pada penelitian mengenai pengaruh musik klasik terhadap ingatan jangka pendek, subjek diperdengarkan musik klasik dan diberikan beberapa kata, setelah itu mereka diminta untuk menuliskan kata-kata yang diberikan sebelumnya. Ternyata sebelum mengikuti penelitian ini, sebagian besar subjek telah mengikuti penelitian mengenai pengaruh metode mengingat terhadap ingatan, sehingga subjek dapat mengingat kata-kata dengan mudah, walaupun tidak diperdengarkan musik klasik.

***Hawthorne effect***

 Faktor ini terjadi ketika subjek menyadari bahwa ia sedang diteliti sehingga ia menampilkan tingkah laku tertentu. Dengan demikian, respons yang diberikan oleh subjek bukanlah disebabkan oleh manipulasi VB. Oleh karena itu, hasilnya hanya berlaku pada penelitian tersebut, sehingga validitas eksternalnya menjadi rendah.

 Nama *Hawthorne* effect ini diperoleh dari penelitian di pembangkit listrik *Hawthorne* milik *Western Electric Company*, di mana dilakukan penelitian pada sekelompok pekerja perempuan untuk melihat pengaruh dari perubahan kondisi kerja (cahaya, jam kerja, suara musik, dan sebagainya) terhadap produktivitas kerja. Apapun perubahan kondisi kerja yang dilakukan, produktivitas kerja subjek meningkat. Akhirnya, peneliti menyimpulkan bahwa peningkatan produktivitas disebabkan berkaitan dengan perlakuan khusus yang diberikan tim peneliti.

 *Hawthome effect* dapat dicegah dengan *singIe-blind procedure*, yaitu membatasi pengetahuan subjek mengenai perlakuan yang diberikan. Dengan kata lain, subjek penelitian tidak mengetahui bahwa ia sedang diteliti.

***Experimenter effect***

 *Experimenter effect* membatasi generalisasi hasil penelitian karena dihasilkan dari interaksi dengan atribut ataupun harapan dari eksperimenter. Dengan kata lain, generalisasi hanya dapat dilakukan pada situasi yang mirip dengan situasi saat penelitian dilakukan.

***Pretesting effect***

 Faktor ini disebabkan oleh pemberian *pretest*. Selain dapat mempengaruhi validitas internal, pemberian *pretest* juga dapat mempengaruhi validitas eksternal karena generalisasi hasil penelitian hanya terbatas pada populasi yang diberikan *pretest* sebelumnya. Faktor ini dicegah dengan menggunakan alat ukur yang tidak membuat subjek menyadari mengenai apa yang diukur atau menduga apa yang sedang diteliti.

**Validitas temporal**

 Validitas eksternal ini berkaitan dengan generalisasi hasil penelitian pada waktu yang berbeda. Seperti telah diketahui, hasil sebuah penelitian eksperimental bergantung dari lamanya rentang waktu antara pemberian VB dengan pengukuran VT. Tidak dipertimbangkannya variabel waktu dapat mengancam validitas eksternal sebuah penelitian. Ada tiga variasi waktu yang mempengaruhi validitas ekternal, yaitu:

**Variasi musiman (*Seasonal variation*)**

 Variasi ini berkaitan dengan kejadian yang secara umum atau biasa terjadi sepanjang waktu dalam populasi. Ada dua bentuk dari variasi musiman, yaitu: *fixed-time variation* dan *variabIe-time variation*.

***Fixed-time variation***

 Varian ini terjadi apabila perubahan terjadi pada waktu tertentu atau waktu yang dapat diramalkan (misalnya, kemacetan di Jakarta terjadi pada hari kerja, kemacetan di daerah Puncak Bogor terjadi pada setiap akhir pekan, atau banjir terjadi di Jakarta pada musim penghujan). Misalnya, penelitian eksperimental mengenai pengaruh jumlah polisi terhadap kemacetan di Jakarta. Bila penelitian dilakukan pada hari libur di Jakarta, maka hasil penelitian ini hanya berlaku pada hari libur-hari libur dan tidak berlaku pada hari kerja.

***Variable-time variation***

 Varian ini terjadi apabila perubahan tidak dapat diramalkan waktu terjadinya namun dapat diketahui kejadiannya (misalnya, waktu seseorang putus cinta tidak dapat diketahui, namun ketika itu terjadi, kita dapat memperkirakan dampak psikologis yang timbul). Misalnya penelitian mengenai pengaruh terapi stres terhadap penurunan stress. Dari hasil penelitian ini diketahui terapi stress berpengaruh terhadap penurunan stress. Namun, hasil ini tidak selalu dapat diterapkan pada individu yang mengalami stress setiap waktu.

**Variasi siklus (*Cyclical variation*)**

 Variasi ini merupakan bentuk dari variasi musiman, namun terjadi di dalam diri manusia dan makhluk lainnya. Siklus pada diri manusia dan makhluk lainnya dapat mengubah pengaruh VB atau berinteraksi dengan VB mempengaruhi VT. Contohnya, penelitian pengaruh musik terhadap konsentrasi dilakukan pada siang hari. Telah diketahui bahwa pada siang hari, kondisi fisik dan kognitif manusia biasanya menurun. Oleh karena itu, hasil penelitian ini hanya bisa digeneralisasikan pada siang hari. Dengan kata lain, validitas eksternal penelitian menjadi tinggi ketika digeneralisasikan pada siang hari.

**Variasi personal (*Personological variation*)**

 Variasi ini merupakan variasi dari karakteristik individu sepanjang waktu. Walaupun secara umum karakteristik seseorang cenderung stabil, namun beberapa karakteristik cenderung berubah pada waktu tertentu. Misalnya, penelitian mengenai pengaruh media kampanye terhadap partisipasi mengikuti pemilu dilakukan menjelang masa pemilu. Padahal kesadaran berpolitik masyarakat cenderung meningkat pada masa menjelang Pemilu, dan cenderung turun beberapa waktu setelah Pemilu. Bila demikian, maka hasil penelitian ini hanya dapat digeneralisasikan pada masa-masa menjelang pemilu.

**HUBUNGAN VALlDlTAS INTERNAL DAN VALIDITAS EKSTERNAL**

 Telah dikemukakan di awal bahwa penelitian eksperimental lebih mementingkan validitas internal dibandingkan validitas eksternal. Hal ini terjadi karena pada penelitian eksperimental, peneliti ingin mengetahui hubungan sebab-akibat antara VB dengan VT sehingga peneliti berusaha untuk mengontrol VS, menggunakan subjek penelitian yang relatif homogen, dan hanya dilakukan pada suatu waktu tertentu. Dengan mengontrol VS maka pengaruh dari variabel-variabel di luar VB terhadap VT dihilangkan. Akibatnya, kondisi penelitian menjadi tidak alamiah. Dengan demikian, validitas internal menjadi tinggi karena situasi penelitiannya terkontrol. Namun di pihak lain, validitas eksternalnya (khususnya validitas ekologis) menjadi rendah karena hasil penelitiannya sulit untuk diterapkan dalam kehidupan sehari-hari.

 Karena penelitian eksperimental lebih mementingkan randomisasi dibandingkan *random* *sampling*, maka subjek penelitian yang di gunakan tidak mewakili populasi. Dengan kata lain, subjek penelitian yang digunakan menjadi homogen. Karena subjek penelitiannya memiliki karakteristik seragam atau homogen, maka pengaruh VB terhadap VT dalam penelitian cenderung tinggi. Namun, hasil penelitian ini tidak dapat digeneralisasikan kepada individu di luar subjek penelitian, walaupun dalam populasi yang sama. Ini berarti, subjek penelitian yang homogen menyebabkan tingginya validitas internal. Namun, validitas eksternalnya menjadi rendah (khususnya validitas populasi) karena sulit diterapkan pada individu lain.

 Pemberian VB kepada subjek penelitian untuk dilihat pengaruhnya terhadap VT dalam penelitian eksperimental lebih sering hanya dilakukan pada waktu tertentu. Karena hanya melibatkan suatu waktu tertentu, maka kekuatan pengaruh VB terhadap VT lebih besar apabila dibandingkan pemberian VB pada beberapa waktu yang berbeda. Oleh karena itu, validitas internalnya menjadi tinggi, sedangkan validitas eksternalnya (khususnya validitas temporal) menjadi rendah karena hanya berlaku pada waktu yang sama seperti ketika penelitian dHakukan.

 Dari pembahasan di atas, dapat disimpulkan bahwa validitas penelitian memiliki hubungan terbalik; penelitian yang memiliki validitas internal rendah, maka validitas eksternalnya tinggi. Begitu sebaliknya, penelitian yang memiliki validitas internal yang tinggi, maka validitas eksternalnya rendah. Penelitian eksperimental bertipe lapangan memiliki validitas ekologis yang lebih tinggi dibandingkan penelitian eksperimental laboratorium. Namun, validitas populasi dan validitas temporal pada kedua tipe penelitian eksperimental tidak terlalu berbeda. Oleh karena itu, dapat dikatakan bahwa penelitian eksperimental lapangan cenderung memiliki validitas eksternal yang lebih tinggi dibandingkan penelitian eksperimental laboratorium.

**RINGKASAN**

* Validitas penelitian meliputi validitas internal dan validitas eksternal. Validitas internal berkaitan dengan sejauhmana hubungan sebab-akibat antara VB dan VT; sedangkan validitas eksternal berkaitan dengan sejauhmana hasil penelitian dapat digeneralisasikan pada subjek, situasi, dan situasi yang berbeda.
* Validitas internal dipengaruhi oleh faktor-faktor: proactive history, retroactive history, maturation, testing, statistical regression, experimental mortality, interaction effect, instrumentation effect, experimenter effect, dan participant sophistication; sedangkan validitas eksternal dipengaruhi oleh: validitas populasi, validitas ekologis, dan validitas temporal.
* Dalam penelitian eksperimental, validitas internal lebih dipentingkan daripada validitas eksternal karena penelitian eksperimental lebih melihat pengaruh VB terhadap VT dengan melakukan kontrol yang ketat terhadap VS sehingga hasil penelitian belum tentu dapat dengan mudah digeneralisasikan.