

BAGIAN I

PENDAHULUAN

Makalah ini disusun berdasar pada cepter 12 dan 13 dari buku *Research Design in Counseling*, karya P. Paul Heppner, Bruce E. Wampold dan Dennis M. Kivlighan, diterbitkan tahun 2007 di Belmont, USA, oleh Thomson Brooks/Cole.

Pada cepter 12 dibahas **Mendesain dan Mengevaluasi Variabel Bebas**, berisikan (1) Mengoperasionalkan variabel bebas, dilakukan sebelum memulai penelitian; (2) Mendeskripsikan metode untuk menguji atau memverifikasi data variabel bebas atau pengujian data, dilakukan selama penelitian berlangsung; (3) Memperkirakan hasil penelitian, apakah data sesuai atau tidak sesuai dengan variabel bebas, dilakukan setelah penelitian dilaksanakan; dan (4) mendiskusikan variabel bebas, sehingga dapat ditentukan kedudukan/posisi variabel tersebut.

Pada Cepter 13 dibahas **Mendesain dan Memilih Variabel Terikat**. variabel terikat (terkadang disebut ukuran terikat) adalah untuk mengukur konstruk yang dihipotesakan sebagai efek/akibat (disebut sebagai konstruk efek). Secara rinci, cepter ini membahas: (1) Mengoperasional variabel terikat, mengupas isu-isu reliabilitas dan validitas; dan (2) Metode pengumpulan data, terdiri atas metode pelaporan diri, penilaian terhadap orang lain dan peristiwa, observasi behavioral, indeks fisiologis, wawancara, teknik projektif, dan penelitian tidak menonjol.

Pada bagian pembahasan cepter ini, dilakukan kajian analisis-komparatif dengan bahasan materi sejenis dari berbagai sumber yang relevan.

BAGIAN II

RANGKUMAN ISI CEPTER

A. Mendesain dan Mengevaluasi Variabel Bebas

Tujuan utama riset adalah memperoleh kejelasan mengenai hubungan kausalitas antara variabel bebas dengan variabel terikat. Oleh karena itu, variabel bebas perlu dirancang dan dievaluasi secara tepat.

Pemilihan, perancangan dan penilaian terhadap variabel bebas sangat penting dalam memahami dan menafsirkan relasi sebab-akibat dalam suatu penelitian. Jika variabel bebas tidak dirancang secara baik, maka penelitian tidak akan menghasilkan sesuatu, tidak akan diperoleh dampak yang diharapkan, kesimpulan bias atau kurang bermakna. Perancangan variabel bebas yang jelek, menyebabkan variabel tersebut tidak kokoh, bias, dan terasa asing.

Beberapa isu tentang pengembangan dan pemilihan variabel bebas, yaitu: (1) mengoperasionalkan variabel bebas, peneliti harus merancang dan merumuskan variabel bebas secara cermat, ini dilakukan sebelum memulai penelitian; (2) mendeskripsikan metode untuk menguji atau memverifikasi data variabel bebas, yang sering dikenal sebagai pengujian data, dilakukan selama penelitian berlangsung; (3) memperkirakan hasil penelitian, apakah data sesuai atau tidak sesuai dengan variabel bebas, dilakukan setelah penelitian dilaksanakan; dan (4) mendiskusikan variabel bebas yang tidak dapat dimanipulasi, sehingga dapat ditentukan kedudukan variabel tersebut.

1. Mengoperasionalkan Variabel Bebas

Sebelum memulai penelitian, peneliti perlu memusatkan perhatian hal-hal berikut untuk mengoperasionalkan variabel bebas: (a) menentukan kondisi atau level variabel bebas; (b) merefleksikan secara tepat konstruk variabel yang dirancang ke dalam pertanyaan penelitian; (c) membatasi keragaman dengan pengkondisian; dan (d) merumuskan pentingnya keragaman dalam kondisi.

a. Menentukan Kondisi

Dalam penelitian konseling, model variabel bebas terdiri atas kondisi yang beragam. Dalam variabel bebas mungkin terdapat dua kondisi, seperti melakukan atau tidak melakukan tindakan. Dalam hal ini, suatu variabel bebas dapat mengandung banyak kondisi.

Pelaksanaan penelitian dapat menguji beberapa tindakan dari keseluruhan kondisi variabel, atau suatu kelompok penelitian dapat dibandingkan dengan suatu kelompok kontrol.

Ada dua acuan dalam mendiskusikan kondisi variabel bebas. Pertama, menggunakan istilah kondisi yang berindikasi terhadap konstitusi variabel bebas. Dalam hal ini, level variabel bebas, pengelompokan (*groups*), kategorisasi (*categories*), dan perlakuan (*treatments*) adalah makna lain yang saling terkait dan digunakan di dalam pembahasan desain penelitian. Dalam konteks ini, tritmen mengacu pada pengkondisian secara umum dan bukan intervensi psikologis. Kedua, variabel bebas dapat dirumuskan sebagai variabel kategori, yakni setiap kategori deskrit (*level, condition, group, atau treatment*) berbeda. Variabel bebas mungkin saja dapat menjadi *categories*.

Dalam desain regresi secara umum termasuk variabel bebas secara kuantitatif, pengkondisian ini menekankan gambaran sesuatu secara beragam (Wampold & Drew, 1990). Untuk mudahnya, dalam suatu penelitian tentang minuman, variabel bebasnya dapat

mengacu pada level 'dosis yang berbeda' (sebagai contoh, tidak minum, 2 cc, 4 cc, dan 6 cc) atau, dalam penelitian tritmen psikoterapi, variabel bebasnya 'penugasan pekerjaan rumah yang banyak'. Variabel bebas dalam penelitian kuantitatif jarang digunakan dalam penelitian konseling.

Dalam desain eksperimental, pengkondisian variabel bebas harus dirumuskan dalam penelitian. Rumusan itu menjadi acuan dalam memanipulasi eksperimen, sebab peneliti sangat perlu memanipulasi variabel bebas untuk menentukan pengaruh atau efek variabel bebas tersebut. Dengan kata lain, variabel bebas akan kelihatan pengaruhnya atau efeknya setelah dikorelasikan dengan variabel terikat.

b. Merefleksikan Konstruk secara Tepat

Variabel bebas dirancang untuk merefleksikan konstruk yang didesain secara kausal dalam pertanyaan penelitian. Variabel bebas harus dirumuskan secara tepat atau operasional. Jika konstruk sebab-akibat dioperasionalkan tidak secara tepat, kemungkinan penafsiran terhadap kesimpulan menjadi bias dan menyesatkan.

Untuk ilustrasi pentingnya merefleksikan secara tepat konstruk hubungan yang dirancang, seperti studi yang dilakukan Malkiewich dan Merluzzi (1980) untuk menguji model kesesuaian klien-tritmen. Hipotesis penelitian menetapkan pemikiran-pemikiran konseptual bahwa akan memperoleh keuntungan tinggi dari konseling secara relatif tidak terstruktur, dan akan memperoleh keuntungan tingkat rendah dari konseling secara relatif terstruktur. Dalam studi ini, struktur konseling adalah salah satu dari variabel bebas, oleh karena itu struktur konseling harus dirumuskan secara operasional termasuk tiga kondisi variabel bebas, yaitu kondisi disensitisasi, restrukturisasi logis, dan kelompok kendali. Kondisi disensitisasi merepresentasikan struktur tinggi, dan restrukturisasi

logis sebagai struktur rendah. Dalam penelitian ini hasilnya hipotesis nol diterima, hal ini mungkin dikarenakan variabel bebas itu tidak menyediakan contoh-contoh, baik dari konseling terstruktur maupun yang tidak terstruktur.

Kedua tritmen konseling itu kurang jelas karena tidak cukup mewakili struktur-struktur berbeda, sebab kedua intervensi agak terstruktur. Hal ini diakibatkan adanya keterpautan tingkat suatu konstruk dengan konstruk lainnya, yang menjadi suatu ancaman terhadap validitas konstruk. Untuk pengujian yang lebih baik atas variabel bebas (struktur konseling) itu, akan berguna bila menyediakan suatu konseling yang secara jelas tidak terstruktur, juga menyediakan suatu cakupan struktur konseling yang lebih luas. Sebagai contoh, konseling *client-centered* (yang sering ditandai sebagai yang tidak terstruktur) mungkin dapat digunakan untuk mewakili truktur yang rendah.

Konstruk ditunjukkan oleh suatu variabel bebas yang operasional dengan variasi contoh-contoh yang terseleksi, atau dengan menggunakan variasi stimulus (suatu saat disebut sampling stimulus). Variabilitas di antara stimuli akan meningkatkan kemampuan generalisasi terhadap kesimpulan. Prinsip ini dapat dipahami dengan menguji hanya satu stimulus atau contoh yang digunakan. Misalnya seorang peneliti berminat di dalam keterbukaan diri (*self-disclosure*) atas proses konseling, setiap konselor hendaknya menggunakan keterbukaan diri yang sama.

c. Perbedaan antar Kondisi Variabel

Dalam seleksi kondisi variabel bebas dapat berbeda hanya sepanjang diinginkan. Jika kondisi memungkinkan untuk berbeda atas dimensi lain, kondisi-kondisi tambahan dapat mengacaukan. Untuk ilustrasi, penelitian sebelumnya memandang bahwa kredibilitas yang dirasakan konselor Eropa-Amerika dan Meksiko-

Amerika sebagai fungsi gaya dan kebudayaan konseling (Ponce & Atkinson, 1989). Dalam hal ini variabel bebas dapat dipertimbangkan dalam suatu rancangan faktorial, yang memfokuskan diri atas keterkaitan antar variabel bebas pada etnisitas konselor. Banyak cara yang dapat ditempuh untuk mengoperasionalkan etnisitas konselor. Dalam penelitian ini etnisitas dapat dimanipulasi dengan menunjukkan partisipasi fotografer terhadap konselor dan dengan menggunakan deskripsi pengantar.

Dalam suatu kondisi, seorang partisipan menunjukkan foto konselor Mexico-Amerika, dalam pengantarnya menggunakan nama panggilan dan tempat kelahiran untuk menggambarkan etnisitas Mexico-Amerika (sebagai contoh, Chavez dan Meksiko). Di dalam kondisi lain, partisipan menunjukkan foto konselor Eropa-Amerika, dan dalam pengantar menggunakan nama panggilan dan tempat kelahiran yang menggambarkan etnisitas Eropa-Amerika (sebagai contoh, Sanders dan Canada). Jelas, dalam pengaturan etnisitas konselor secara operasional; pertanyaan mengacu pada dua kondisi yang berbeda atas banyak kondisi lain. Mengapa Ponce dan Atkinson memilih menggunakan foto-foto, di sana ada kemungkinan konselor Meksiko-Amerika dan Eropa-Amerika dalam foto-foto juga berbeda di dalam daya tarik pribadi, yang akan menyediakan suatu penjelasan alternatif untuk hasil-hasil yang bersinggungan pada variabel bebas ini. Dalam hal ini, peringkat tinggi diberikan kepada konselor Meksiko-Amerika oleh partisipan Meksiko-Amerika, dapat terjadi karena etnisitas konselor atau daya tarik pribadi. Untunglah, Ponce dan Atkinson menyadari potensi ini mengacaukan dan mengontrol, oleh karena itu dengan memastikan bahwa foto konselor dapat dibandingkan dengan menghargai daya tarik pribadi (dan dengan menghargai usia, mungkin mengacaukan yang lain).

Seperti metode-metode sampling, stimulus dideskripsikan sebelum penelitian diselenggarakan, potensi mengacaukan harus dipertimbangkan. Dalam hal ini tidak ada cara yang mungkin dapat

mengeliminasi pengacau-pengacau. Peneliti perlu mengidentifikasi (contoh, dengan memeriksa manipulasi) sebelum memulai penelitian yang memungkinkan peneliti dapat meminimalkan pengacau-pengacau sedini mungkin. Sangat sulit untuk menemukan pengacau-pengacau utama setelah data dikumpulkan.

Ketika peneliti menemukan pengacau itu, perlu dijelaskan secara logika bahwa sesuatu yang mengacaukan kecil kemungkinan terjadi. Daya tarik pribadi tampaknya dapat menjadi variabel penting dalam literatur kredibilitas konselor (lihat Corringan, Dell, Lewis, & Schmidt, 1980). Bagaimanapun, Ponce dan Atkinson tidak mengesampingkan kemungkinan hasil-hasil dari penelitian itu disebabkan oleh pendapat anti-Canadian karena kesukaran yang ada tentang negosiasi suatu perjanjian dagang antara Canada dan United States. Meski penjelasan ini tidak dapat dikesampingkan pada desain eksperimen, tidak ada bukti hubungan-hubungan politis dengan negara asal seorang konselor yang manapun mempengaruhi atau tidak mempengaruhi kredibilitas.

Beberapa pengacau yang menyulitkan tritmen penelitian bersifat unik, satu di antaranya adalah konselor. Mengesampingkan konselor yang mengacaukan dapat tercapai dengan pemilikan tritmen silang konselor secara konstan; dalam hal ini konselor-konselor yang sama akan mengadministrasi semua tritmen. Bagaimana pun, sejumlah konselor dapat menjadi lebih terampil dalam tritmennya dibanding dengan yang lain, atau konselor memiliki suatu kesetiaan dalam suatu tritmen dan tritmen yang lainnya. Karenanya, superioritas dalam suatu tritmen tidak akan terjadi pada tritmen yang sama, tetapi sebagai gantinya dengan keterampilan atau kepatuhan konselor.

Untuk memiliki keahlian dalam pengadministrasian tritmen tertentu, perlu strategi yang dapat menjelaskan bahwa yang mungkin mengacaukan itu berhubungan dengan pengalaman, pelatihan dan sebagainya. Kemungkinan lainnya untuk memilih

konselor secara relatif (sebagai contoh, tingkat pendidikan siswa dalam konseling), secara random dirancang dalam tritmen, dan memberi mereka pelatihan yang sama di dalam masing-masing tritmen. Hal ini merujuk pada validitas eksternal dalam penelitian, sebab suatu simpulan tidak dapat digeneralisasikan oleh konselor yang kurang berpengalaman.

d. Menetapkan Kejelasan Sumber Perbedaan

Keragaman di antara kondisi-kondisi atas kondisi yang diinginkan harus jelas. Sebagai contoh, Ponce dan Atkinson (1989) dapat menggunakan nama panggilan atau tempat kelahiran konselor untuk mengoperasionalkan etnisitas, dapat dieliminasi dengan daya tarik pribadi yang mengacaukan dan menjadikan penelitian lebih sederna; termasuk di dalamnya foto untuk meningkatkan kejelasan dari etnisitas. Meski muncul perbedaan atas dimensi penting variabel bebas ini terhadap validitas penelitian, akan berbahaya ketika perbedaan di dalam terlalu besar. Jika para peserta menyimpulkan hipotesis penelitian dari prosedur penelitian, ada kemungkinan respon akan menjadi bias. Transparansi yang jelas menciptakan suatu situasi di mana peserta itu boleh bereaksi terhadap situasi eksperimental. Inferensi lain terhadap hipotesis sebagai dasar atas pernyataan tujuan penelitian dan berbagai prosedur, seperti halnya kejelasan atas manipulasi eksperimental. Kiranya, para peserta yang menebak hipotesis riset cenderung untuk menanggapi dengan cara-cara yang menyenangkan peneliti, dan hal seperti itu mengkonfirmasi hipotesis riset. Reaksi terhadap situasi eksperimental perlu didiskusikan, karena mungkin dapat menjadi ancaman terhadap penelitian kredibilitas konselor.

Kondisi-kondisi variabel bebas perlu sesuai dengan dimensi yang diharapkan, bukan dengan dimensi-dimensi lain. Dimensi yang diharapkan hendaknya merefleksikan pertanyaan penelitian yang

diinginkan. Keragaman kondisi-kondisi eksperimental atas kondisi yang diharapkan harus jelas, tetapi tidak transparan. Para peserta di dalam suatu kondisi hendaknya menyadari akan komponen yang kritis dari kondisi itu, tetapi tidak perlu menyimpulkan atau menduga hipotesis penelitian. Tentu saja sangat sulit membuat keputusan dengan jelas dan transparan, dan ini keterampilan yang diharapkan diperoleh dari penelitian.

2. Menguji Manipulasi Variabel

Ketika peneliti sudah berusaha sungguh-sungguh untuk merumuskan dan mengoperasionalkan variabel bebas, bukan jaminan bahwa manipulasi eksperimental akan mencapai tujuannya. Ini mungkin terjadi pada peneliti yang salah menilai hal penting dari variabel bebas. Untuk memverifikasi suatu manipulasi variabel perlu dirancang secara tepat. Hal-hal berikut sebaiknya digunakan untuk menguji manipulasi variabel : (1) kondisi-kondisi itu sebagai dimensi yang sangat diharapkan; (2) kondisi-kondisi itu tidak mencerminkan dimensi lain; dan (3) tritmen dapat diimplementasikan di dalam petunjuk yang diharapkan.

Untuk menentukan apakah kondisi-kondisi itu sebagai dimensi yang diharapkan, jadmén atas keterkaitan karakteristik dimensi perlu dikroscek dengan kondisi-kondisi yang berbeda. Penentuan ini dapat dibuat dengan banyak cara, di antaranya pemeriksaan-pemeriksaan dapat dibuat dari peserta-peserta mereka sendiri. Sebagai contoh, Jones & Gelso (1988) dalam suatu studi tentang dampak dari gaya-gaya penafsiran, gaya dimanipulasikan dengan peserta-peserta mampu mendengarkan audiotapes dari suatu sesi konseling. Dalam suatu kondisi, konselor menafsirkan secara tentatif ungkapan-ungkapan dan mengakhiri dengan suatu pertanyaan.

3. Penafsiran Hasil

Tujuan utama mendesain penelitian eksperimental adalah menetapkan hubungan timbal balik antara variabel bebas dengan variabel terikat. Yang sama-sama pentingnya terkait dengan penelitian eksperimen adalah penafsiran hasil eksperimen. Kegiatan ini akan memberikan banyak informasi berdasarkan kesimpulan hasil penelitian yang telah dilakukan.

a. Hasil yang secara Statistik Signifikan

Indikasi statistik signifikan adalah apabila hasil-hasil yang diperoleh mengindikasikan untuk setiap kondisi cukup berbeda dan konsekuensinya hipotesis nol untuk ketidakberbedaan ditolak. Dapat dikatakan bahwa tampak nyata terdapat perbedaan terhadap suatu kondisi. Sebagai contoh, dalam studi perbedaan perlakuan, hasil statistik signifikan menunjukkan bahwa terdapat perlakuan yang lebih efektif dibanding dengan perlakuan yang lain. Hasil ini menggambarkan bahwa kumpulan hipotesis nol dari ketidakberbedaan di antara perlakuan ditolak.

Sekilas tampak bahwa hasil statistik signifikan ini mudah untuk ditafsirkan, padahal sebenarnya di dalamnya banyak ruang yang dapat membuat peneliti menjadi bingung. Sebagaimana telah diskusikan sebelumnya bahwa hasil penelitian mungkin saja akan membaurkan; oleh karena itu akan sangat mungkin dihasilkan penjelasan yang lain di luar dari yang diharapkan. Dalam penelitian tentang perlakuan, eksperimen yang dilakukan oleh seorang terapis dapat menghasilkan data yang membaurkan. Walaupun peneliti sudah dengan sangat hati-hati menjaga kondisi variabel tersebut, memang pada akhirnya tidak ada eksperimen yang sempurna. Oleh karena itu sangat penting memikirkan kemungkinan-kemungkinan terjadinya pembauran. Sekalipun demikian, manipulasi tanda pemeriksaan dapat dilakukan untuk menyingkirkan sisa alternatif,

walaupun pemberian tanda juga dapat menjadikan kondisi semakin membingungkan.

Untuk meminimalkan reaksi, seorang peneliti hendaknya mempertimbangkan pengelolaan pemberian tanda setelah mengukur variabel terikat (sudah pasti, bahwa kemudian pemberian tanda akan mempengaruhi pengukuran variabel terikat).

b. Hasil yang secara Statistik Tidak Signifikan

Dari perspektif pilosofi keilmuan, hasil hipotesis nol atau tidak signifikan sangat banyak memberikan informasi. Meskipun begitu, hasil tidak signifikan dapat juga menjadikan faktor penyebab kurang jelasnya pengaruh variabel, termasuk tidak adekuasinya statistik, kurang jelasnya instrumen, kegagalan asumsi tes statistik, prosedur yang ceroboh dan cenderung bias.

Kegagalan mendeteksi interkasi yang diharapkan dalam memberikan perlakuan terhadap klien menjadi salah satu faktor penyebab tidak adekuatnya disain variabel bebas (Malkiewich dan Merluzzi's, 1980). Manipulasi eksperimental dikatakan sukses apabila dapat membedakan kondisi dari variabel yang sedang diteliti; apabila tidak ditemukan perbedaaan tersebut dapat dikatakan bahwa hipotesis yang dirancang tidak sesuai dengan kondisi kesekarangan. Jones dan Gelso's (1988) melakukan studi penafsiran dalam konseling dengan tidak membuat interaksi yang diharapkan antara tipe klien dengan gaya penafsiran; tanpa memeriksa manipulasi, ini akan mempermudah atribusi hasil nol sehingga penafsiran dapat menjadi beragam.

Pada sisi lain, temuan yang tidak signifikan dapat juga terjadi karena manipulasi yang salah, sebagaimana hasil pemeriksaan menunjukkan kondisi yang ada tidak berbeda dengan kondisi awal dan perbedaan yang diharapkan pada variabel tidak terikat tidak ditemukan.

4. Status Variabel

Status variabel menekankan pada perlunya perhatian dari seorang peneliti tentang variabel bebas. Dengan mendesain variabel bebas pada kondisi tertentu, peneliti berusaha menguji pengaruh terhadap variabel terikat. Penggunaan istilah ‘manipulasi’ untuk menggambarkan karakteristik proses kesengajaan tersebut.

Variabel bebas dapat dimanipulasi dan pengaruhnya terhadap variabel terikatlah yang akan dinilai; apabila semua berjalan dengan baik, maka hubungan timbal baliknya dapat dikatakan mantap. Yang lebih jelas, status variabel tidak dapat dimanipulasi dan uji statistik dipergunakan untuk mendeteksi asosiasi tersebut.

Tidak perlu membuat perbedaan antara variabel bebas dengan status variabel, oleh karena itu salah satunya dapat menjadi yang terbaik sedangkan yang lainnya bersifat inferior. Dalam hal ini, yang terpenting adalah bahwa variabel bebas dimanipulasi, oleh sebab itu kesimpulan sebagai dampak sebab akibat dapat dibuat secara langsung. Bagaimanapun juga kesimpulan dalam suatu kasus akan terbentuk menjadi beragam bentuk dan biasanya akan menjadi semakin berat. Misalnya penelitian tentang merokok dan kesehatan. Perilaku merokok secara etis tidak dapat dimanipulasi, artinya tidak dapat subjek penelitian dipaksa untuk merokok atau sebaliknya yaitu untuk berhenti merokok. Bagi yang tidak merokok, mereka yakin bahwa perilaku tersebut akan mengganggu kesehatan, oleh karena itu mereka tidak mau mencoba. Sebaliknya bagi yang sudah terbiasa merokok, mereka mungkin saja yakin bahwa merokok akan mengganggu kesehatan akan tetapi mereka sulit untuk menghentikan perilaku dalam bentuk kebiasaan merokok tersebut.

Akhirnya, dapat dikatakan bahwa secara logis variabel yang tidak dapat dimanipulasi tentunya tidak akan memberikan pengaruh

(Holland, 1986). Sebagai contoh, karena merupakan sesuatu yang tidak mungkin melibatkan jenis kelamin dalam suatu penelitian, maka tidak mungkin dikatakan bahwa jenis kelamin turut memberikan pengaruh. Selain itu, secara tidak langsung dapat disimpulkan bahwa variabel yang dapat dimanipulasi adalah variabel yang terkait dengan aktivitas-aktivitas biologis dan kultural, sedangkan yang sulit untuk dimanipulasi adalah yang terkait dengan aspek psikis dan jenis kelamin.

B. Mendesain dan Memilih Variabel Terikat

Tujuan dasar dari variabel terikat (terkadang disebut ukuran terikat) adalah untuk mengukur konstruk yang dihipotesiskan sebagai efek/akibat (disebut sebagai konstruk efek). Jadi, pemilihan atau pendesainan variabel-variabel terikat dan metode-metode pengumpulan data berkenaan dengan variabel-variabel terikat itu merupakan aktivitas-aktivitas kritis bagi para peneliti. Perhatian yang ekstrim harus dilakukan dalam proses ini, karena pilihan terhadap variabel-variabel terikat dapat menjadi penting/kritis untuk kebaikan penelitian tersebut. Misalnya, laporan seorang ibu dapat digunakan untuk menilai perilaku anak-anaknya, namun penilaian-penilaiannya terhadap perilaku putranya dapat lebih banyak dipengaruhi oleh psikopatologisnya dibandingkan oleh perilaku aktual anaknya.

1. Mengoperasionalkan Variabel Terikat

Memilih atau mendesain variabel-variabel terikat merupakan operasionalisasi yang memadai dari konstruk-konstruk, yang dituangkan ke dalam pertanyaan penelitian itu merupakan sebuah langkah kritis/penting di dalam penelitian. Variabel-variabel terikat harus dirancang atau dipilih untuk merefleksikan konstruk-konstruk yang dimasukkan ke dalam pertanyaan penelitian.

a. Isu-isu Psikometrik

Satu pertanyaan penting mengenai operasionalisasi dari sebuah konstruk melibatkan properti-properti psikometrik dari variabel terikat. Para peneliti harus mengetahui pada tataran apa variabel-variabel terikat yang dipilih untuk mengoperasionalkan sebuah konstruk yang handal dan valid. Jika taksiran-taksiran terhadap reliabilitas dan validitas itu buruk, maka operasionalisasi konstruk itu kemungkinan tidak memadai.

1) Reliabilitas

Supaya informatif, skor-skor pada variabel terikat harus bervariasi di antara para peserta suatu penelitian. Jika setiap orang memperoleh skor yang sama pada sebuah variabel tidak ada hal yang dapat dipelajari mengenai individu; kendatipun demikian, ketika skor para peserta berbeda, kita mulai mempelajari sesuatu mengenai bagaimana para peserta itu berbeda. Diharapkan bahwa perbedaan-perbedaan antara kedua skor itu disebabkan perbedaan-perbedaan yang nyata pada level karakteristik yang diminati; yakni varian dalam skor-skor harus merefleksikan varian pada para respondennya.

Faktor vital untuk mempertimbangkan varian dalam skor-skor itu berkaitan dengan konstruk sentral yang sedang diukur. Di dalam teori tes, dikatakan bahwa untuk masing-masing individu sebuah skor sejati muncul yang merefleksikan tingkatan aktual dari konstruk yang diperhatikan. Tingkatan dimana diperoleh skor-skor yang merefleksikan skor-skor sejati untuk individu-individu adalah reliabilitas skor. Secara lebih teknis, reliabilitas adalah varian dalam skor-skor yang disebabkan perbedaan-perbedaan sejati di antara individu.

Biasanya, koefisien reliabilitas untuk variabel X digambarkan dengan simbol r_{xx} . Sebuah koefisien r_{xx} yang sama dengan 0,80 mengindikasikan bahwa 80% dari varian di dalam skor-skor tersebut disebabkan perbedaan-perbedaan sejati, dan bahwa 20% disebabkan oleh faktor-faktor lain. Berkenaan dengan reliabilitas, perlu terlebih dahulu mengkaji beberapa sumber *error* 'kekeliruan' dalam pengukuran: *error* respon acak, *error* spesifik, *error* sementara, ketidaksepakatan *interrater*, penyekoran dan perekaman *errors*, dan pembauran (*compounding*).

a) *Error Respon Acak.*

Seringkali terdapat suatu kekeliruan dalam respon yang dibuat peserta. Contoh yang paling jelas dari kekeliruan ini muncul dalam merespon item-item tertulis pada sebuah instrumen *paper-and-pencil test*, namun *error* respon acak muncul dalam pengukuran semua jenis. Seorang peserta mungkin membaca kata "ever" sebagai "never" dan memberi respon dengan itu.

Berkenaan dengan jenis *error* ini ada beberapa poin penting. Pertama, penilaian dari hampir semua karakteristik yang bermakna dari individu dan situasi berisi *error* respon acak. Kedua, instrumen-instrumen biasanya berisi banyak item yang mengukur ciri yang sama, sehingga sebuah respon acak tunggal tidak terlalu mempengaruhi skor totalnya.

b) *Error Spesifik*

Error spesifik adalah kekeliruan yang dihasilkan oleh sesuatu yang unik terhadap instrumen yang berbeda dari apa yang dimaksud peneliti. Misalnya, di dalam sebuah instrumen yang dirancang untuk mengukur depresi, pertanyaan-pertanyaannya disusun dalam suatu cara tertentu sehingga para peserta tahu dengan baik bahwa respon-respon pada pertanyaan-pertanyaan tersebut bervariasi dalam tingkatan hasrat/keinginan sosial; dalam kasus semacam itu,

respon-respon para peserta itu ditentukan pada satu tataran dengan derajat dimana mereka ingin tampil seperti yang diharapkan secara sosial (suatu konstruk yang *legitimate* dengan sendirinya), dan juga pada tingkatan dimana mereka terdepresikan. *Error* spesifik itu campuran karena skor-skor pada instrumen ini mengukur baik itu depresi maupun hasrat sosial.

c) *Error Sementara*

Error sementara muncul ketika seorang peneliti sedang mengukur sebuah sifat yang stabil pada suatu poin tunggal sekaligus atau dalam merespon suatu stimulus tunggal dalam suatu cara dimana kondisi-kondisi tersebut pada saat itu atau dengan stimulus tertentu mempengaruhi pengukuran dari sifatnya. Perhatikan pengukuran depresi, perwujudannya dapat dipengaruhi oleh keadaan suasana hati yang sementara: respon-respon seorang mahasiswa perguruan tinggi yang terdepresi pada sebuah inventori depresi, misalnya, akan dipengaruhi dengan penerimaan sebuah nilai ujian yang gagal sejam sebelum penilaian; respon para peserta lain akan dipengaruhi oleh suasana hati yang diciptakan oleh peristiwa-peristiwa saat ini.

d) *Ketidaksepakatan antar Penilai*

Di dalam penelitian konseling, para penilai seringkali digunakan untuk memperoleh penilaian-penilaian. Perhatikan sebuah penelitian tentang perilaku antisosial anak-anak sekolah yang melibatkan observasi-observasi naturalistik terhadap perilaku anak-anak di dalam latar sosial. Sekalipun para penilai perilaku anak-anak tersebut dilatih untuk mengikuti suatu sistem pengkodean, beberapa varians pada penilaian para pengamat mungkin disebabkan oleh pengamatnya itu sendiri bukannya

perilaku tersebut (misalnya, beberapa penilai itu lebih peka dengan perilaku-perilaku negatif).

e) Penyekoran dan Perekaman Error

Error dalam penilaian dapat diciptakan oleh para peneliti melalui penyekoran dan perekaman *error*. Dalam hal ini, beberapa *error* diciptakan dalam suatu cara tertentu dengan memanipulasi data di dalam proses penyekoran sebuah pedoman pada upaya mempersiapkan data untuk analisis statistik. Kekeliruan-kekeliruan ini berfungsi sebagai kekeliruan respon acak mengaburkan varian skor sejati.

f) Pembauran Errors

Kekeliruan-kekeliruan yang telah disebutkan dapat dibaurkan untuk membentuk sebuah penilaian dengan reliabilitas yang bukan kepalang. Perhatikan skenario kasus terburuk: beberapa pengamat, masing-masing mengobservasi seorang peserta, menilai suatu karakteristik hanya satu waktu dalam merespon sebuah stimulus tunggal yang menggunakan satu item, instrumen tertulis, kemudian merekam respon tersebut, yang akan dimasukkan ke dalam komputer. Hal ini menjadi masalah manakala seorang instruktur praktikum diminta untuk menilai, pada sebuah skala 1 sampai 100, tingkat keterampilan dari seorang instruktur praktikum dengan seorang klien tertentu di dalam suatu sesi tertentu. Operasionalisasi dari keterampilan konseling ini mengangkat banyak sumber kekeliruan. Pertama, terdapat variansi yang tidak dikenal di antara para instruktur praktikum. Kedua, hanya sebuah item tunggal bermakna ganda yang digunakan untuk menilai konstruk tersebut. Ketiga, tingkat keterampilan yang diperlihatkan dalam suatu sesi tunggal tunduk pada kekeliruan sementara yang disebabkan oleh karakteristik dari klien, pada kehadiran faktor-

faktor tertentu yang mempengaruhi kesuksesan dari sesi tersebut, dan pada faktor-faktor lain. Keempat, kesempatan-kesempatan penyekoran dan perekaman kekeliruan itu tidak diminimalisir.

g) Menginterpretasi Taksiran-taksiran Reliabilitas

Penentuan reliabilitas sebuah instrumen penelitian melibatkan banyak pertimbangan. Pertama, setiap koefisien reliabilitas adalah suatu taksiran terhadap reliabilitas sejati, dalam cara yang sama sebuah rata-rata dari sebuah sampel adalah sebuah taksiran terhadap rata-rata populasi tersebut. Kedua, reliabilitas merefleksikan varian yang disebabkan skor-skor sejati, namun itu tidak mengindikasikan skor-skor sejati yang sedang diukur. Ketiga, adalah bahwa reliabilitas itu didasarkan pada skor-skor dan tidak pada instrumen dari mana skor itu diperoleh.

Sebuah instrumen dapat bekerja dengan memadai untuk satu tipe peserta namun tidak untuk tipe lainnya, atau di bawah sekumpulan kondisi namun tidak di bawah yang lainnya. Misalnya, sebuah pengukuran kecemasan yang menghasilkan taksiran-taksiran reliabilitas yang memadai dengan para mahasiswa S1 ketika diberikan dalam sebuah kelas yang mungkin benar-benar tidak bermanfaat untuk mengukur *agrophobics* di dalam sebuah latar laboratorium. Dengan kata lain, instrumen itu mungkin sangat peka untuk perbedaan rentang tengah dalam kecemasan namun tidak peka pada rentang atas. Ini dinamakan sebuah *ceiling effect* 'efek atap'; semua *agrophobic* dapat mencetak skor maksimum atau mendekati. Tentu saja, permasalahan ini juga dapat terwujud pada dasar dari rentang tersebut, yang menciptakan sebuah *floor effect* 'efek lantai'. Reliabilitas juga terikat pada karakteristik-karakteristik para peserta, seperti kemampuan membaca dan usia.

Faktor-faktor lain yang perlu dipertimbangkan, termasuk validitas, waktu yang diperlukan untuk menyelesaikan instrumen tersebut, dan biaya.

h) Menghitung Taksiran-taksiran Reliabilitas

Terdapat banyak cara untuk menaksir reliabilitas skor-skor, masing-masing peka pada satu atau lebih kekeliruan-kekeliruan (*errors*). Jika berbagai item dari sebuah instrumen itu mengukur konstruk yang sama, maka skor-skor pada item-item tersebut akan cenderung untuk bervariasi; yakni seseorang yang memiliki sebuah tingkatan tinggi atas konstruk tersebut (misalnya, gelisah/cemas) akan cenderung menjawab semua item tersebut dalam satu arah, sementara seseorang yang memiliki sebuah tingkatan rendah terhadap sebuah konstruk (misalnya, tidak gelisah) akan cenderung menjawab semua item dalam cara lain. *Konsistensi internal* merujuk pada homogenitas dari item-item tersebut. Ketika skor-skor untuk berbagai item itu tinggi, konsistensi internal pun tinggi.

Indeks-indeks yang memperhitungkan pengukuran-pengukuran yang diambil pada waktu-waktu berbeda atau dibuat untuk merespon stimuli berbeda itu peka dengan efek-efek sementara. Yang paling lazim dari indeks semacam itu adalah korelasi tes-rites. Jika sebuah konstruk diharapkan tetap stabil selama suatu periode waktu, dan jika instrumen itu tidak tunduk pada kekeliruan sementara atau kekeliruan respon acak, maka korelasi-korelasi tes-rites itu haruslah tinggi. Jika konsistensi internal tinggi namun koefisien tes-rites itu relatif rendah dan konstruknya diharapkan stabil selama periode waktu itu, maka skor-skor tersebut merefleksikan efek-efek sementara.

Satu permasalahan dengan koefisien tes-rites adalah bahwa koefisien itu menaksir berlebihan reliabilitas karena koefisien tersebut tidak peka dengan kekeliruan spesifik. Jika sesuatu yang unik diukur dengan sebuah instrumen, maka karakteristik unik ini akan terukur pada pemberian instrumen kedua ini juga. Satu cara untuk menangani permasalahan ini adalah menggunakan bentuk-bentuk paralel pada dua waktu tersebut.

i) Efek Ketidakreliabelan pada Hubungan di antara Variabel

Instrumen-instrumen harus menghasilkan ukuran-ukuran yang reliabel/ajeg agar dapat bermanfaat dalam penelitian konseling. Perhatikan dua konstruk, A dan B, dan dua ukuran dari konstruk tersebut, X dan Y, secara berturut-turut. Jika semua sumber kekeliruan untuk kedua konstruk ini setara sekitar 30%; yakni, $r_{xx} = 0,70$ dan $r_{yy} = 0,70$. Di dalam contoh ini, peneliti mengklaim bahwa dua konstruk hadir, dan bahwa interpretasi-interpretasi dapat dibuat mengenai konstruk-konstruk itu dari variabel-variabel X dan Y. Namun perlu diingat bahwa korelasi 0,70 adalah korelasi dari ukuran X dan Y, bukan korelasi dari konstruk A dan B. Karena kekeliruan tersebut pada masing-masing ukuran tidak dapat berkaitan secara sistematis, korelasi yang diperoleh dari pengukuran itu kurang dari korelasi dari konstruk-konstruk tersebut, dan kita mengatakan bahwa korelasi dari konstruk-konstruk tersebut telah ditipiskan/dikecilkan dengan ketidak-reliabilitas-an dari ukuran-ukuran tersebut.

Teori tes klasik memberikan sebuah rumus untuk mengoreksi pengukuran tersebut:

$$r_{AB} = \frac{r_{xy}}{r_{xx} r_{yy}}$$

Korelasi antara konstruk-konstruk tersebut setara dengan korelasi yang diperoleh antara ukuran-ukuran yang dibagi dengan akar pangkat dua dari produk reliabilitas dari ukuran tersebut. Korelasi antara konstruk-konstruk berdasarkan contoh di atas menjadi:

$$r_{AAB} = 0,70 / \sqrt{(0,70)(0,70)} = 100$$

Korelasi antara konstruk-konstruk tersebut berkorelasi secara sempurna, korelasi yang diperolehnya akan secara dramatis dikecilkan dengan ketidakreliabelan.

2) Validitas

Dari banyak tipe validitas, tipe yang paling penting untuk tujuan penelitian adalah *validitas konstruk* – tingkatan yang skor-skoranya merefleksikan konstruk yang diharapkan bukannya konstruk lain. Jelasnya, skor-skor yang tidak ajeg (*unreliable*) tidak dapat memiliki validitas konstruk karena skor-skor itu kebanyakan disebabkan kekeliruan acak (*random error*). Kendatipun demikian, skor-skor yang ajeg dapat merefleksikan satu atau lebih konstruk yang lain dari pada konstruk yang dispesifikan. Secara khusus, skor-skor dapat cukup *reliable* namun kurang validitas konstruk. Jika korelasi-korelasi yang diharapkan ditemukan, maka akan muncul *validitas konvergen*.

a) *Analisis Faktor dan Penggunaan Skala suatu Instrumen*

Terkadang analisis faktor digunakan untuk mengembangkan skala-skala. Strategi ini melibatkan item-item analisis faktor bukannya variabel. Sejumlah item itu tunduk dengan sebuah analisis faktor, item-item disegregasikan dengan pembebanan pada faktor-faktor, deskriptor yang diberikan pada faktor-faktor, dan skor-skor subskala itu dihitung berdasarkan segregasi. Secara umum, ini bukanlah sebuah prosedur yang menghasilkan hasil-hasil yang memuaskan. Terdapat tiga permasalahan: (1) metode tersebut bersifat teoretis dan dapat mengarah pada faktor-faktor yang memiliki basis psikologis kecil dan disokong dengan data; (2) sekalipun jika analisis faktor menggunakan sebuah metode yang menghasilkan faktor-faktor independen, skor-skor subskala kemungkinan akan berkorelasi tinggi, karena item-item membebani semua faktor pada derajat tertentu; dan (3) reliabilitas dari item-item tunggal itu rendah, sehingga hasil-hasil dari analisis-analisis faktor seringkali tidak stabil, yang di dalamnya validasi lintas kasus itu penting.

Sebuah pengembangan pada strategi analisis faktor eksplorasi adalah mengembangkan item-item yang secara spesifik mengukur faktor-faktor sebuah konstruk. Strategi ini digunakan untuk mengembangkan satu dari instrumen-instrumen konseling yang digunakan secara paling luas, *the Counselor Rating Form* (CRF; Barak & LaCrosse, 1975). CRF adalah sebuah skala 36 item yang didesain untuk mengukur tiga karakteristik konselor yang berkaitan dengan pengaruh sosial yang dimiliki konselor terhadap klien: kepercayaan, kemenarikan, dan keahlian.

Sekalipun analisis-analisis faktor CRF telah memverifikasi eksistensi dari tiga faktor tersebut, korelasi di antara faktor-faktor tersebut itu tinggi (umumnya dalam rentang 0,60 sampai dengan 0,80), yang mengungkapkan bahwa satu faktor umum beroperasi.

Faktor umum ini, diberi nama faktor "orang baik" yang mengungkapkan bahwa respon-respon pada CRF itu pada pokoknya disebabkan suatu opini umum mengenai konselor.

Pembahasan terdahulu terhadap CRF dan subskalanya mengangkat sebuah isu: apakah orang harus menggunakan skor total dari sebuah instrumen atau skor-skor subskalanya. Pilihan itu eksklusif; yakni, orang tidak boleh menggunakan skor total dan satu atau lebih skor-skor subskala sekaligus dalam analisis yang sama, karena keduanya itu secara linear terikat dan menyebabkan solusi-solusi noneksisten atau tidak bermakna dalam analisis-analisis statistik. Keputusan untuk menggunakan skor-skor subskala atau skor-skor total pada pokoknya berkaitan dengan hipotesis-hipotesis dari penelitian tersebut, namun secara parsial berkaitan dengan psikometrik juga.

b) *Ukuran-ukuran Jamak dari sebuah Konstruk untuk*

Meningkatkan Validitas Konstruk

Penggunaan variabel-variabel terikat yang jamak itu seringkali direkomendasikan. Tidak ada variabel yang dapat secara memadai mengoperasionalkan sebuah konstruk, karena beberapa varians dalam variabel ini disebabkan konstruk-konstruk lain (varian spesifik) dan sebagian disebabkan oleh kekeliruan. Dengan menggunakan beberapa variabel itu dapat secara lebih memadai merepresentasikan konstruk tersebut karena satu variabel akan peka dengan aspek dari konstruk yang tidak ada dalam variabel-variabel lain. Ketumpangtindihan dari variabel-variabel ini merefleksikan esensi dari konstruk (Bagan 13.1).

Alasan lain memasukkan ukuran-ukuran jamak adalah pengharapan bahwa konstruk-konstruk berbeda menghasilkan hasil-hasil berbeda.

Ukuran-ukuran jamak dari konstruk-konstruk juga dapat digunakan untuk menghindari penyurutan korelasi antara konstruk-konstruk dan dapat memperhitungkan variansi metode.

c) *Menghitung Korelasi-korelasi Antar Konstruk yang tidak*

Dikecilkan oleh Ketidakreliabelan

Ketidakreliabelan mengecilkan ukuran-ukuran asosiasi, seperti korelasi. Ukuran-ukuran jamak dari sebuah konstruk dapat digunakan untuk mendeteksi hubungan-hubungan antara konstruk-konstruk yang murni dengan ketidakreliabelan/ketidakejegan. Sebagai contoh, modeling persamaan struktural dapat digunakan untuk mendeteksi hubungan-hubungan yang sejati di antara konstruk-konstruk. Pemodelan persamaan struktural adalah sebuah metode statistik yang menguji hubungan antara konstruk-konstruk dengan menggunakan beberapa ukuran yang teramati untuk mengoperasionalkan konstruk tersebut (Cole, 1986).

Pada contoh tersebut, dua konstruk penting – depresi dan kecemasan – oleh Tanaka-Matsumi dan Kameoka (1986). Tanaka-Matsumi dan Kameoka memberikan tiga ukuran depresi yang lazim digunakan dan enam ukuran kecemasan yang lazim dipakai; korelasi-korelasi antar ukuran-ukuran ini dihadirkan pada Tabel 13.1. Beberapa observasi dapat dibuat dari tabel ini. Tampak bahwa ukuran-ukuran dari konstruk yang sama tersebut secara moderat tinggi, menunjukkan suatu validitas konvergen. Konstruk kecemasan dan depresi tampak berkaitan karena korelasi-korelasi yang diperoleh di antara ukuran-ukuran depresi dan kecemasan terentang

dari 0,33 sampai 0,74. Perlu diingat bahwa semua korelasi dalam tabel tersebut diperkecil oleh ketidakreliabilan. Pemodelan persamaan struktural memberikan sebuah alat penaksiran korelasi dari konstruk-konstruk depresi dan kegelisahan, yang mempertimbangkan ketidakreliabelan ini.

Hasil-hasil dari pemodelan persamaan struktural tersebut dihadirkan pada Bagan 13.2. Terlebih dahulu catat anak panah – anak panah dari elips "Depresi" pada Zung D, BDI, dan DACL (variabel-variabel yang teramati dalam persegi panjang), yang mengindikasikan bahwa konstruk (atau variabel yang *latent*) dari depresi membebani ketiga instrumen ini. Variabel tersembunyi "Depresi" adalah suatu entitas statistik yang merepresentasikan konstruk yang dioperasionakan dengan tiga ukuran depresi. Demikian pula dengan konstruk kecemasan adalah pengukuran yang dikembangkan secara statistik. Konstruk kecemasan dari enam ukuran yang teramati, dengan bobot faktor terentang dari 0,55 sampai 0,90.

Korelasi dari konstruk-konstruk depresi dan kecemasan kemudian ditaksirkan dari variabel-variabel tersembunyi dari kecemasan dan depresi. Anak panah ganda yang berupa kurva/lengkungan antara "Depresi" dan "kecemasan" menggambarkan korelasi ini, yang terhitung 0,98. Ini mengatakan bahwa taksiran korelasi konstruk depresi dan kecemasan, seperti yang diukur oleh tiga ukuran depresi dan enam ukuran kecemasan adalah 0,98. Korelasi ini tidak dikecilkan oleh ketidakreliabelan. Kesimpulannya adalah bahwa konstruk-konstruk depresi tidak nyata.

d) *Menghilangkan Variansi Metode*

Pada contoh-contoh sebelumnya, semua ukuran kecemasan dan depresi adalah pengukur bentuk tulisan (*pencil-and-paper*). Validitas konstruk itu terikat pada penilaian-penilaian yang menggunakan metode berbeda. Kemungkinan terjadi bahwa sesuatu dalam instrumen-instrumen ini mempengaruhi respon-respon para peserta, namun tidak berkaitan baik dengan depresi maupun dengan kecemasan. Satu kemungkinan adalah negativitas sifat/ciri, suatu kecenderungan umum untuk mengevaluasi diri secara negatif pada semua dimensi; para responden ini akan tampak lebih terdepresi dan lebih cemas daripada kasus nyatanya.

Kemungkinan lainnya adalah sebuah keadaan suasana hati (*mood*) yang sementara yang dapat mempengaruhi respon-respon pada instrumen tersebut. Para mahasiswa yang menghadiri sesi pengetesan beberapa saat setelah menerima nilai ujian tengah semester dapat mengalami perasaan-perasaan sementara yang dimunculkan oleh hasil-hasil ujian tersebut. Hanya karena satu metode yang digunakan, kemungkinan-kemungkinan ini mempengaruhi respon-respon semua instrumen secara sama, meningkatkan korelasi-korelasi di antara respon tersebut.

Variansi yang lazim pada semua ukuran yang menggunakan metode sama itu disebut variansi metode. Variansi metode menggembungkan hubungan-hubungan di antara variabel; yakni, hubungan antara dua ukuran yang disebabkan tidak hanya oleh sebuah hubungan konseptual dalam konstruk-konstruk yang dimaksud, melainkan juga pada suatu hubungan berkenaan dengan bagaimana konstruk-konstruk itu diukur. Sementara ketidakreliabelan mengecilkan korelasi, variansi metode menggembungkan korelasi.

Variansi metode seringkali muncul dalam penelitian konseling ketika berbagai aspek konseling dinilai dari perspektif yang sama.

Misalnya, jika para supervisor menilai baik kompetensi kultural dari konselor maupun kemajuan terapeutik klien, maka korelasi antara kompetensi kultural dan hasil dipengaruhi sebagian oleh perspektif penilaian. Jika supervisor memiliki sikap yang secara umum mendukung pada konselor, maka supervisor itu akan cenderung menilai semua aspek dari konselor dan klien tersebut sebagai hal yang positif.

e) *Ukuran jamak – Pertimbangan Akhir*

Bagian-bagian sebelumnya dapat dirangkum dengan enam poin berikut:

- (1) Suatu operasi tunggal (yakni, sebuah skala atau instrumen tunggal) hampir akan selalu merepresentasikan sebuah konstruk secara buruk.
- (2) Korelasi antara dua konstruk dikecilkan dengan ketidak-reliabilitias-an.
- (3) Ketidak-reliabilitas-an selalu membuatnya lebih sulit untuk mendeteksi efek-efek yang nyata karena kekuatan statistik tereduksi.
- (4) Korelasi antara dua ukuran yang menggunakan metode yang sama digembungkan dengan variansi metode.
- (5) Jika dimungkinkan, ukuran-ukuran beragam yang menggunakan metode jamak hendaknya digunakan untuk mengoperasionalkan sebuah konstruk.
- (6) Biasanya, interpretasi-interpretasi harus dibuat pada tingkatan konstruk, karena jarang kita tertarik dalam ukuran-ukuran itu semata. Pengetahuan tentang efek-efek ketidak-reliabilitas-an dan variansi metode itu penting untuk menarik kesimpulan-kesimpulan yang tepat.

f) *Generalisabilitas*

Dalam pembahasan validitas konstruk, seperti halnya reliabilitas, validitas merupakan suatu properti skor-skor dan bukan properti instrumen. Selain itu, tingkatan dimana variabel-variabel memproduksi skor-skor yang secara memadai merefleksikan sebuah konstruk itu tergantung pada tipe peserta yang digunakan dalam penelitian tersebut. Hubungan-hubungan yang diperlihatkan oleh pemodelan persamaan struktural itu hanya dapat digeneralisir pada orang-orang yang mirip dengan mereka yang digunakan untuk mengumpulkan data. Ponterotto dan Casas (1991) menganalisis penelitian multikultural dalam jurnal-jurnal konseling dan menemukan bahwa hanya 25% dari instrumen-instrumen tersebut yang digunakan dalam penelitian itu dikembangkan menggunakan populasi minoritas rasial dan etnis. Validitas konstruk dari ke-75%-nya dikembangkan pada kelompok-kelompok lain (pada pokoknya orang Amerika Eropa kelas menengah) itu dipertanyakan. Ponterotto dan Casas (1991) berkesimpulan bahwa kurangnya instrumentasi berbasis minoritas itu merupakan satu dari keterbatasan terbesar penelitian multikultural yang mereka telaah. Sekalipun saat ini lebih banyak instrumen yang berbasis minoritas, asumsi generalisabilitas antara kelompok-kelompok terus menjadi sebuah isu utama dalam bidang ini.

b. Reaktivitas

Variabel terikat harus peka dengan karakteristik peserta, namun proses penilaian itu sendiri tidak boleh mempengaruhi karakteristiknya secara langsung; yakni, ukuran terikat harus mengindikasikan bagaimana peserta itu berfungsi secara normal. Terkadang, sesuatu mengenai pemerolehan skor-skor yang kurang terikat mengubah situasinya sehingga bacaan-bacaan 'yang keliru' diperoleh. Variabel-variabel yang mempengaruhi karakteristik-

karakteristik dari para peserta yang berupaya diukur itu dikatakan *reaktif*. Misalnya, seorang peserta tes yang cemas dapat melaporkan peningkatan kecemasan pada sebuah instrumen pelaporan diri karena penyelesaian instrumen itu seperti pengambilan sebuah tes; seorang anak yang agresif dapat memperlihatkan penurunan perilaku agresif ketika diobservasi/diamati oleh orang dewasa dibandingkan pada waktu lain. Jelasnya, hakikat reaktif dari variabel terikat harus dipertimbangkan dalam mendesain penelitian; lagi-lagi pengetahuan tentang bidang substantif itu vital.

c. Pertimbangan Prosedural

Sejumlah isu prosedural harus dipertimbangkan ketika memilih atau mendesain variabel terikat. Seringkali, waktu yang dipakai pada penilaian itu penting bagi kesuksesan penelitian. Para peserta akan enggan untuk ikut penelitian yang menuntut waktu yang panjang dalam menyelesaikan isian dan instrumen, atau jika mereka ikut pun mereka mungkin memberi respon secara ceroboh pada item-itemnya (meningkatkan variansi *error*), khususnya pada akhir dari sebuah periode penilaian yang panjang. Seperti yang disebutkan sebelumnya, kemampuan baca dari instrumen itu penting untuk performansi psikometriknnya. Setiap instrumen yang diberikan harus diperiksa untuk memastikan bahwa para peserta dapat membaca secara memadai bahan-bahannya.

Susunan dari pemberian instrumen dapat memiliki efek pada respon; satu instrumen dapat membuat peka atau bahkan mempengaruhi respon-respon pada instrumen lain. Sebuah instrumen yang menarik perhatian patologi peserta sendiri (misalnya, *the Minnesota Multiphasic Personality Inventory*) dapat mempengaruhi bagaimana peserta memberi nilai pengukur-pengukur lain (misalnya, konselor pada CRF). Susunan itu juga penting ketika

instrumen yang sama itu diberikan secara berulang. Performansi pada suatu saat dapat disebabkan oleh respon sebelumnya dan bukan pada jumlah karakteristiknya.

Ketika ukuran-ukuran berulang digunakan di dalam sebuah penelitian, penggunaan bentuk-bentuk alternatif itu diharapkan jika efek-efek pengujian diantisipasi. Bentuk-bentuk alternatif memungkinkan peneliti memberikan sebuah *pretest* dan *posttest* tanpa harus menggunakan instrumen yang identik.

Para mahasiswa seringkali mengajukan pertanyaan-pertanyaan tentang apakah yang harus dimasukkan ketika menggambarkan properti-properti psikometrik sebuah inventori untuk sesuatu seperti bagian metode dari tesis, disertasi, atau artikel jurnal mereka. Sekalipun semua isu-isu konseptual itu merupakan pertimbangan-pertimbangan esensial dalam desain penelitian, ketika menggambarkan properti-properti psikometrik sebuah instrumen yang akan digunakan, para penulis biasanya memberikan sebuah deskripsi tentang inventori, langsung dalam susunan berikut ini:

- 1) Deskripsi pengukur itu sendiri.
 - a) nama instrumen
 - b) akronim
 - c) penulis
 - d) rujukan kunci
 - e) deskripsi singkat dari konstruk yang diukur instrumen
 - f) laporan diri, observasi behavioral, wawancara, atau internet
 - g) jumlah item dan contoh item
 - h) tipe item (misalnya, item Likert)
 - i) faktor-faktor atau subskala, dan definisinya
 - j) indikasi dari arahan penyekoran, dan apa yang dimaksud dengan skor tinggi

- 2) Taksiran validitas
 - a) validitas konvergen dan diskriminan
 - b) sampel-sampel dimana pengukuran divalidasikan
- 3) Taksiran Reliabilitas
 - a) koefisien alfa Cronbach
 - b) Test- retest (jika dapat diterapkan)
 - c) Reliabilitas adalah sebuah properti skor-skor yang dihasilkan dari suatu pemberian sebuah tes, karena itu para peneliti harus melaporkan taksiran-taksiran reliabilitas untuk kumpulan data saat itu.

2. Metode Pengumpulan Data

a. Pelaporan Diri

Pada pengukur pelaporan diri (*self-report*), peserta menilai tingkatan dimana suatu karakteristik itu hadir atau dimana suatu perilaku itu muncul. Pelaporan diri dapat dicapai dengan memberi respon pada item-item dalam sebuah inventori, mengisi sebuah log (catatan harian), atau menulis sebuah jurnal. Peserta itu sendiri dalam hal ini membuat observasi atau laporan. Umumnya, asumsi dibuat sehingga laporan tersebut secara akurat merefleksikan keadaan yang sebenarnya dari persoalan – bahwa para peserta memberi respon secara jujur dan akurat.

1) Keuntungan pelaporan diri

Sekalipun pelaporan diri dapat berupa banyak bentuk, metode ini memiliki beberapa keuntungan umum yang membuatnya menjadi alat penilaian yang paling populer dalam penelitian konseling. Pelaporan diri itu relatif mudah diberikan. Kemudian pelaporan diri juga dapat digunakan untuk mengakses fenomena yang sangat sulit atau tidak mungkin diukur. Pelaporan diri juga

cocok dengan pandangan fenomenologis dari konseling dan psikoterapi.

2) Kerugian pelaporan diri

Kerugian yang paling nampak dan mengganggu dari pelaporan diri adalah: (1) metode ini rapuh dengan distorsi (yang disengaja ataupun tidak) oleh peserta; (2) peserta juga mungkin tidak tahu karakteristik dari yang diukur; dan (3) berkaitan dengan kongruensi antara sebuah perspektif fenomenologis dan pelaporan diri, bahwa pelaporan diri itu kurang dihargai oleh beberapa perspektif teoretis lain.

3) Kasus Khusus: Pengumpulan data internet

Peningkatan akses internet bagi banyak individu telah mengarahkan internet menjadi sebuah alat penelitian yang semakin layak. Penggunaan skala-skala pelaporan diri yang diadaptasikan untuk pengumpulan data *online* dapat mempermudah penyelidikan pertanyaan-pertanyaan penelitian yang mungkin sebelumnya secara logistik tidak layak atau tidak praktis (misalnya, pemerolehan data pelaporan diri dari sebuah sampel para konselor dalam pendidikan yang beragam secara geografis) dan memiliki sejumlah keuntungan pragmatik lain (misalnya, reduksi data yang hilang, reduksi kekeliruan; lihat Stanton, 1988). Penggunaan teknologi internet memberikan sebuah jalan yang mudah bagi para responden untuk mengisi survey, yang tentunya meningkatkan penilaian sebuah respon. Keuntungan praktis dari pengumpulan data internet adalah bahwa data dapat secara langsung dimasukkan ke dalam sebuah paket *software* statistik.

Kerugian dari pengukuran pelaporan diri diperburuk dalam konteks pengumpulan data internet, yaitu: (1) penelitian yang

menggunakan pengumpulan data *online* terbatas pada para peserta yang memiliki akses pada komputer, (2) dalam hal validitas dan reliabilitas, apakah internet itu merupakan sarana yang tepat untuk mengumpulkan data penelitian.

b. Penilaian (*Ratings*) terhadap Orang Lain dan Peristiwa

Penelitian konseling seringkali mengandalkan pada penilaian (*ratings*) yang dibuat terhadap orang lain atau peristiwa. Prosedur-prosedur tersebut mirip dengan prosedur untuk pelaporan diri, kecuali para responden memberi nilai karakteristik dari peserta atau peristiwa tersebut. Seringkali para responden adalah para ahli, dan penilaian mereka diasumsikan merefleksikan secara akurat karakteristik dari orang atau peristiwa. Misalnya, dalam studi-studi *treatment*, terapis atau orang yang signifikan dapat menilai tingkatan disfungsi atau peningkatan dari seorang klien. Sebuah telaahan terhadap literatur dalam konseling memperlihatkan bahwa penilaian langsung dari para peserta itu jarang digunakan. Kendatipun demikian, banyak penelitian memperoleh variabel-variabel dari penilaian peristiwa, khususnya sesi-sesi konseling. Misalnya, *the Sessiona Evaluation Questionnaire* (SEQ; Stiles, 1980) dirancang untuk mengukur kedalaman dan kehalusan dari sesi konseling.

Penilaian terhadap orang lain atau peristiwa membagi banyak keuntungan dengan pelaporan diri, khususnya kemudahan pemberiannya dan fleksibilitas. Ketika *raters* itu para ahli, penilaian mereka pada pokoknya berharga karena penilaian tersebut dibuat dengan suatu latar belakang dan pemahaman yang mendalam.

Permasalahan pokok dengan *rating* terhadap orang lain dan peristiwa adalah bahwa *rating* dapat secara sistematis bias. Ini pada dasarnya merupakan suatu masalah ketika *raters* mengetahui hipotesis dan kondisi-kondisi kognitif yang dimiliki para peserta.

Ketika *raters* digunakan untuk membuat penilaian mengenai peristiwa, penilaian-penilaian itu dapat merefleksikan karakteristik dari penilai dan karakteristik peristiwa. Ketika para peserta (konselor dan klien) menilai kedalaman dan kehalusan dari wawancara tentang SEQ, mereka sebenarnya melaporkan persepsi-persepsi mereka tentang wawancara, dan berkenaan dengan itu penilaian mereka merupakan pelaporan diri. Jadi, ketika menginterpretasikan *ratings* dari peristiwa (atau orang lain) para peneliti harus hati-hati memisahkan varians karena perbedaan dalam *event* dari varian yang disebabkan oleh *raters* itu sendiri.

Satu strategi untuk menguji *varians* karena *raters* adalah menggunakan pengamat yang netral atau banyak, dan kemudian menguji perbedaan di antara para *raters*.

Permasalahan lain dengan *ratings* adalah karena *ratings* seringkali relatif umum, dan hal itu tidak mungkin menentukan apa yang mengarahkan *rating* tersebut.

c. Observasi Behavioral

Ukuran-ukuran behavioral diperoleh dari observasi-observasi terhadap perilaku yang jelas, biasanya kebanyakan oleh seorang observer terlatih. Psikologi behavioral telah menekankan pentingnya perilaku yang nampak dan tidak memberikan tekanan pada fenomena intrapsikis. Karena itu, mengamati dan merekam perilaku adalah komponen kunci dari analisis perilaku yang diterapkan. Pada pokoknya, observasi behavioral itu sama dengan *ratings* terhadap orang atau peristiwa lain, kecuali berkenaan dengan ukuran-ukuran behavioral yang memfokuskan pada perilaku yang tampak, dapat diamati dan tidak menggantungkan diri pada inferensi para penilai.

Keuntungan umum dari observasi behavioral adalah bahwa observasi itu merupakan ukuran-ukuran langsung dan objektif. Sekalipun ada bias sistematis dalam observasi dan perekaman perilaku yang tampak, ukuran-ukuran behavioral biasanya tidak tunduk pada bias-bias personal yang menjadi sifat pelaporan diri. Keuntungan lain dari ukuran behavioral adalah bahwa para peserta dapat dinilai dalam berbagai lingkungan.

Diantara kerugian observasi behavioral adalah fakta bahwa permasalahan dan kepedulian para klien seringkali tidak terpusat pada perilaku yang dapat diamati. Kepuasan pernikahan adalah sebuah konstruk yang sulit dioperasionalkan secara behavioral. Kerugian lain dari observasi behavioral adalah berkenaan dengan kerepresentatifan.

Isu-isu yang berkaitan dengan reliabilitas itu bersifat problematis bagi penilaian behavioral. Keputusan seorang observer bahwa suatu perilaku tertentu muncul mungkin bersifat idiosinkratis bagi observer itu. Dalam konteks penilaian behavioral, isu reliabilitas dinilai dengan menghitung indeks-indeks kesepakatan; yakni, seberapa baik para observer menyepakati mengenai kemunculan perilaku yang ditargetkan? Seperti halnya penilaian tradisional, kesepakatan inter-observer itu adalah sebuah topik yang kompleks.

d. Indeks Fisiologis

Respon-respon biologis dari para peserta seringkali digunakan untuk menyimpulkan keadaan-keadaan psikologis. Banyak fenomena psikologis memiliki korelasi-korelasi fisiologis yang dapat digunakan sebagai variabel-variabel terikat. Pada kenyataannya, respon-respon fisiologis seringkali dapat dianggap sebagai ukuran-

ukuran langsung dari sebuah konstruk. Misalnya, sementara pelaporan diri dari kecemasan dapat dibiaskan dengan sejumlah faktor, ukuran-ukuran perkembangan fisiologis dapat dibuat secara langsung dan dapat dianggap bebas dari bias. Kendatipun demikian, sekalipun pengembangan fisiologis itu merupakan sebuah fokus dalam konseptualisasi teoretis dari kecemasan, hubungan antara keadaan-keadaan fisiologis dan fenomena psikologis itu tidak selangsung seperti yang diantisipasi dalam tahun-tahun awal dari penelitian ini. Lebih jauh lagi, pengukur-pengukur fisiologis itu mahal, memerlukan keahlian, mungkin reaktif, dan tunduk pada kekeliruan yang disebabkan sejumlah faktor mekanis dan elektronis. Akibatnya, pengukur-pengukur fisiologis jarang digunakan dalam penelitian konseling. Kendatipun demikian, pengembangan teknik pengukuran fisiologis yang semakin canggih sangat menjanjikan bagi psikologi konseling dan mungkin menjadi sebuah bidang yang semakin memfokus untuk generasi peneliti yang akan datang.

e. Wawancara

Wawancara itu merupakan sebuah alat langsung untuk memperoleh informasi dari para peserta. Pada dasarnya, proses penggunaan wawancara untuk memperoleh data berkenaan dengan sebuah variabel terikat itu sama, kecuali bahwa tujuannya adalah untuk menghitung suatu konstruk. Dalam kehidupan sehari-hari wawancara merupakan suatu aktivitas yang pervasif; kita hanya bertanya pada orang-orang untuk memberikan informasi. Wawancara biasanya melibatkan sebuah interaksi interpersonal antara pewawancara dan yang diwawancarai atau peserta. Kerlinger (1986) mendukung penggunaan wawancara personal karena kontrol yang lebih besar dan kedalaman dari informasi dapat diperoleh.

Kedalaman informasi seringkali merupakan hasil dari perencanaan dan pengembangan jadwal wawancara yang cermat.

Wawancara personal memungkinkan fleksibilitas dalam desain kuesioner; pewawancara dapat memberikan penjelasan, membuat keputusan selama wawancara mengenai ketepatan dari suatu respon tertentu, dan mengevaluasi motivasi peserta.

Kendatipun demikian, wawancara itu memerlukan biaya dalam bentuk uang dan waktu. Jika topik-topiknya sensitif, maka para peserta mungkin ragu untuk memberikan informasi dibandingkan kalau mereka dibiarkan memberi respon secara anonim pada sebuah kuesioner.

f. Teknik Proyektif

Rasionalisasi dibalik teknik-teknik proyektif adalah bahwa respon-respon peserta pada stimuli ambigu akan memperlihatkan segi kepribadian mereka. *The Thematic Apperception Test* dan *the Rorschach* itu mungkin dua tes proyektif yang paling terkenal. Kendatipun demikian, beragam kemungkinan muncul, termasuk gambar-gambar, menulis esay, melengkapi kalimat, memainkan mainan, mengasosiasikan kata-kata, dan sebagainya. Asumsinya adalah bahwa karena metode itu tidak langsung, para peserta tidak akan menyensor sendiri. Pada gilirannya, respon-respon peserta itu merupakan ukuran-ukuran tidak langsung dan perlu ditafsirkan dengan cara tertentu. Penyeoran tes-tes proyektif itu biasanya subyektif, sekalipun ada beberapa sistem yang sangat objektif untuk menyeor seperti sistem *Exner* untuk penyeoran respon *Rorschach*.

Secara historis, teknik-teknik proyektif diasosiasikan dengan pendekatan-pendekatan psikodinamika untuk memahami perilaku manusia. Kendatipun demikian, seiring dengan popularitas

pendekatan psikodinamika yang menurun, demikian juga dengan penggunaan teknik proyektif.

g. Penelitian Tidak Menonjol

Untuk menghilangkan reaktivitas, seringkali dimungkinkan mengumpulkan data para peserta tanpa kepekaan terhadap proses ini. Pengukuran-pengukuran digunakan dalam suatu cara sehingga para peserta tidak menyadari prosedur penilaian, hal ini dikenal sebagai pengukur yang tidak menonjol. Dimungkinkan untuk mengamati peserta (dalam naturalistik) tanpa mereka tahu, untuk mengamati peserta dalam situasi direncanakan, mengumpulkan data dari arsip atau sumber lain, atau meneliti jalur-jalur fisik. Kebanyakan psikolog sangat tertarik pada sumber-sumber data yang tidak menonjol.

Keuntungan pengukuran yang tidak menonjol/mencolok mata adalah bahwa pengukur-pengukur tersebut non-reaktif. Karena para peserta itu tidak tahu bahwa data sedang dikumpulkan, mereka tidak mengubah responnya. Pengukuran-pengukuran yang tidak menonjol itu seringkali sangat akurat. Kendatipun demikian, terdapat sejumlah keterbatasan pengukuran-pengukuran yang tidak menonjol. Tipe-tipe tertentu dari pengukuran-pengukuran yang tidak menonjol itu tidak etis, sulit dan/atau mahal untuk diperoleh. Selain itu, jika data sekali diperoleh, interpretasi atau klasifikasi itu seringkali diperlukan.

BAB III

PEMBAHASAN DAN IMPLIKASI

A. Pembahasan

Istilah variabel atau peubah merupakan salah satu konsep kunci dalam penelitian kuantitatif. Setiap variabel yang dikaji harus diidentifikasi dan didefinisikan secara gamblang sampai ke tingkat yang operasional, sehingga dapat diukur (*measurable*) (Furqon, 1997: 10). Pemahaman dan pemaknaan atas variabel bebas yang diteliti sangat mendasari teknik pengukurannya, yang pada akhirnya akan mempengaruhi ketepatan informasi atau data yang diperoleh.

Makalah ini mengkaji variabel bebas dan variabel terikat dalam penelitian. Kedua variabel ini merupakan indikasi dari pendekatan kuantitatif, terutama penelitian eksperimental (*experimental research*). Penelitian eksperimental merupakan penelitian yang benar-benar mengkaji hubungan sebab-akibat atau kausalitas antara dua atau lebih variabel. Variabel bebas (*independent variable*) yakni variabel yang mempengaruhi variabel lain. Variabel bebas ini dikenal dengan berbagai sebutan, seperti variabel pendahulu, variabel masukan (input), variabel prediktor, dan *treatment* dalam penelitian eksperimental (Furqon, 1997: 11). Perlakuan (*treatment*) atau manipulasi terhadap variabel bebas dapat dilihat hasil atau pengaruhnya terhadap variabel terikat. Manipulasi variabel bebas yang disebut perlakuan atau *treatment* inilah yang membedakan penelitian eksperimental dari penelitian lainnya (Gay dalam Furqon, 1997: 10).

Perlakuan peneliti terhadap variabel bebas dalam suatu penelitian eksperimental, bergantung pada bagaimana variabel bebas itu didesain atau dikonstruksikan secara operasional. Oleh karena itu, variabel bebas perlu dipilih, didesain dan dievaluasi, sehingga

dapat digunakan untuk memahami dan menafsirkan relasi sebab-akibat dalam suatu penelitian eksperimental.

1. Mendesain dan Mengevaluasi variabel Bebas

Pada cepter 12 dikemukakan isu-isu tentang pengembangan dan pemilihan variabel bebas atau bagaimana variabel bebas itu didesain dan dievaluasi secara tepat, sehingga menjadi variabel yang fisibel untuk diteltiti. Ketiia mendesain variabel bebas, paling tidak ada empat isu yang harus menjadi perhatian penelitian, yaitu: (1) mengoperasionalkan variabel bebas, peneliti harus merancang dan merumuskan variabel bebas secara cermat, ini dilakukan sebelum memulai penelitian; (2) mendeskripsikan metode untuk menguji atau memverifikasi data variabel bebas, yang sering dikenal sebagai pengujian data, dilakukan selama penelitian berlangsung; (3) memperkirakan hasil penelitian, apakah data sesuai atau tidak sesuai dengan variabel bebas, dilakukan setelah penelitian dilaksanakan; dan (4) mendiskusikan variabel bebas yang tidak dapat dimanipulasi, sehingga dapat ditentukan kedudukan variabel tersebut.

Dalam mengoperasionalkan variabel bebas, peneliti harus mendesain dan merumuskan variabel bebas secara cermat sebelum memulai penelitian. Untuk mengoperasionalkan variabel bebas harus menentukan kondisi atau level variabel bebas, merefleksikan konstruk secara tepat, mengidentifikasi perbedaan antar kondisi variabel, dan menetapkan kejelasan sumber perbedaan di antara variabel bebas yang akan dipilih dalam penelitian.

Kondisi atau level variabel bebas dalam penelitian bimbingan dan konseling perlu diidentifikasi dan dirumuskan secara cermat. Model variabel bebas dalam penelitian konseling terdiri atas kondisi yang beragam, suatu variabel bebas dapat mengandung banyak kondisi. Pelaksanaan penelitian dapat menguji beberapa perlakuan

dari keseluruhan kondisi variabel, atau suatu kelompok penelitian dapat dibandingkan dengan suatu kelompok kontrol.

Dalam desain eksperimental, pengkondisian variabel bebas harus dirumuskan secara cermat, karena menjadi acuan dalam memanipulasi eksperimen. Peneliti sangat perlu memanipulasi variabel bebas untuk menentukan pengaruh atau efek variabel bebas tersebut, yang akan kelihatan setelah dikorelasikan dengan variabel terikat.

Ada dua acuan menentukan kondisi variabel bebas. Pertama, menggunakan istilah kondisi yang berindikasi terhadap konstitusi variabel bebas, yaitu level variabel, pengelompokan (*groups*), kategorisasi (*categories*), dan perlakuan (*treatments*) adalah makna lain yang saling terkait dan digunakan di dalam pembahasan desain penelitian. Kedua, variabel bebas dapat dirumuskan sebagai variabel kategori, yakni setiap kategori deskript (*level, condition, group, atau treatment*) berbeda. Variabel bebas mungkin saja dapat menjadi *categories*.

Variabel bebas dirancang untuk merefleksikan konstruk yang didesain secara kausal ke dalam pertanyaan penelitian. Variabel bebas harus dirumuskan secara tepat atau operasional. Jika konstruk sebab-akibat dioperasionalkan tidak secara tepat, kemungkinan penafsiran terhadap kesimpulan menjadi bias dan menyesatkan.

Pada cepter 12 ini diilustrasikan dalam menguji model kesesuaian klien-tritmen. Hipotesis penelitian menetapkan pemikiran-pemikiran konseptual bahwa akan memperoleh keuntungan tinggi dari konseling secara relatif tidak terstruktur, dan akan memperoleh keuntungan tingkat rendah dari konseling secara relatif terstruktur. Struktur konseling adalah salah satu dari variabel bebas, oleh karena itu struktur konseling harus dirumuskan secara operasional termasuk tiga kondisi variabel bebas, yaitu kondisi disensitisasi, restrukturisasi logis, dan kelompok kendali. Kondisi

disensitisasi merepresentasikan struktur tinggi, dan restrukturisasi logis sebagai struktur rendah. Dalam penelitian ini hasilnya hipotesis nol diterima, hal ini mungkin dikarenakan variabel bebas itu tidak menyediakan contoh-contoh, baik dari konseling terstruktur maupun yang tidak terstruktur.

Kedua tritmen konseling itu kurang jelas karena tidak cukup mewakili struktur-struktur berbeda, sebab kedua intervensi agak terstruktur. Hal ini diakibatkan adanya keterpautan tingkat suatu konstruk dengan konstruk lainnya, yang menjadi suatu ancaman terhadap validitas konstruk. Oleh karena itu, konstruk variabel harus secara operasional dirumuskan, sehingga tampak jelas bedanya dari konstruk variabel bebas lainnya.

Untuk pengujian yang lebih baik atas variabel bebas (struktur konseling) itu, akan berguna bila menyediakan suatu konseling yang secara jelas tidak terstruktur, juga menyediakan suatu cakupan struktur konseling yang lebih luas.

Perbedaan antar kondisi variabel perlu dirumuskan secara jelas. Kondisi variabel bebas dapat berbeda hanya sepanjang diinginkan. Bila perbedaan antar kondisi variabel tidak dirumuskan secara jelas, maka dapat mengacaukan hasil penelitian. Ketika peneliti menemukan pengacau dari perbedaan kondisi variabel, perlu dijelaskan secara logika bahwa sesuatu yang mengacaukan kecil kemungkinan terjadi. Pada ilustrasi (Heppner, 2007: 301-302) penelitian Ponce dan Atkinson (1989) tentang pengaruh hubungan politis negara asal konselor (Meksiko-Amerika dan Eropa-Amerika) terhadap kredibilitas konselor, daya tarik pribadi tampaknya dapat menjadi variabel penting dalam literatur kredibilitas konselor (Corrigan, Dell, Lewis, & Schmidt, 1980). Pada desain eksperimen tersebut, tidak ada bukti hubungan-hubungan politis dengan negara asal seorang konselor yang manapun mempengaruhi atau tidak mempengaruhi kredibilitas.

Beberapa pengacau yang menyulitkan tritmen penelitian bersifat unik, satu di antaranya adalah konselor. Mengesampingkan hal yang mengacaukan konselor dapat tercapai dengan pemilikan tritmen silang konselor secara konstan; dalam hal ini konselor-konselor yang sama akan mengadministrasi semua tritmen. Bagaimana pun, sejumlah konselor dapat menjadi lebih terampil dalam tritmennya dibanding dengan yang lain, atau konselor memiliki suatu kesetiaan dalam suatu tritmen dan tritmen yang lainnya. Karenanya, superioritas dalam suatu tritmen tidak akan terjadi pada tritmen yang sama, tetapi sebagai gantinya dengan keterampilan atau kepatuhan konselor.

Untuk memiliki keahlian dalam pengadministrasian tritmen tertentu, perlu strategi yang dapat menjelaskan bahwa yang mungkin mengacaukan itu berhubungan dengan pengalaman, pelatihan dan sebagainya. Kemungkinan lainnya untuk memilih konselor secara relatif (sebagai contoh, tingkat pendidikan siswa dalam konseling), secara random dirancang dalam tritmen, dan memberi mereka pelatihan yang sama di dalam masing-masing tritmen. Hal ini merujuk pada validitas eksternal dalam penelitian, sebab suatu simpulan tidak dapat digeneralisasikan oleh konselor yang kurang berpengalaman.

Kejelasan sumber perbedaan perlu diidentifikasi dan ditetapkan pada saat mendesain variabel bebas. Jadi, keragaman di antara kondisi-kondisi atas kondisi yang diinginkan harus jelas. Sebagai contoh, Ponce dan Atkinson (1989) dapat menggunakan nama panggilan atau tempat kelahiran konselor untuk mengoperasionalkan etnisitas, dapat dieliminasi dengan daya tarik pribadi yang mengacaukan dan menjadikan penelitian lebih sederna; termasuk di dalamnya foto untuk meningkatkan kejelasan dari etnisitas. Meski muncul perbedaan atas dimensi penting variabel bebas ini terhadap validitas penelitian, akan berbahaya ketika perbedaan di dalam terlalu besar. Jika para peserta menyimpulkan

hipotesis penelitian dari prosedur penelitian, ada kemungkinan respon akan menjadi bias. Transparansi yang jelas menciptakan suatu situasi di mana peserta itu boleh bereaksi terhadap situasi eksperimental. Inferensi lain terhadap hipotesis sebagai dasar atas pernyataan tujuan penelitian dan berbagai prosedur, seperti halnya kejelasan atas manipulasi eksperimental. Kiranya, para peserta yang menebak hipotesis riset cenderung untuk menanggapi dengan cara-cara yang menyenangkan peneliti, dan hal seperti itu mengkonfirmasi hipotesis riset. Reaksi terhadap situasi eksperimental perlu didiskusikan, karena mungkin dapat menjadi ancaman terhadap penelitian kredibilitas konselor.

Kondisi-kondisi variabel bebas perlu sesuai dengan dimensi yang diharapkan, bukan dengan dimensi-dimensi lain. Dimensi yang diharapkan hendaknya merefleksikan pertanyaan penelitian yang diinginkan. Keragaman kondisi-kondisi eksperimental atas kondisi yang diharapkan harus jelas, tetapi tidak transparan.

Mendeskripsikan metode untuk menguji atau memverifikasi data variabel bebas, sering dikenal sebagai pengujian manipulasi variabel atau pengujian data, yang dilakukan selama penelitian berlangsung. Ketika peneliti sudah berusaha sungguh-sungguh untuk merumuskan dan mengoperasionalkan variabel bebas, bukan jaminan bahwa manipulasi eksperimental akan mencapai tujuannya. Ini mungkin terjadi pada peneliti yang salah menilai hal penting dari variabel bebas. Untuk memverifikasi suatu manipulasi variabel perlu dirancang secara tepat. Hal-hal berikut sebaiknya digunakan untuk menguji manipulasi variabel : (1) kondisi-kondisi itu sebagai dimensi yang sangat diharapkan; (2) kondisi-kondisi itu tidak mencerminkan dimensi lain; dan (3) tritmen dapat diimplementasikan di dalam petunjuk yang diharapkan.

Untuk menentukan apakah kondisi-kondisi itu sebagai dimensi yang diharapkan, jadmén atas keterkaitan karakteristik dimensi perlu dikroscek dengan kondisi-kondisi yang berbeda. Penentuan ini dapat dibuat dengan banyak cara, di antaranya pemeriksaan-pemeriksaan dapat dibuat dari peserta-peserta mereka sendiri. Sebagai contoh, Jones & Gelso (1988) dalam suatu studi tentang dampak dari gaya-gaya penafsiran, gaya dimanipulasikan dengan peserta-peserta mampu mendengarkan audiotapes dari suatu sesi konseling. Dalam suatu kondisi, konselor menafsirkan secara tentatif ungkapan-ungkapan dan mengakhiri dengan suatu pertanyaan.

Tujuan utama mendesain penelitian eksperimental adalah menetapkan hubungan timbal balik antara variabel bebas dengan variabel terikat. Yang sama-sama pentingnya terkait dengan penelitian eksperimen adalah penafsiran hasil eksperimen. Kegiatan ini akan memberikan banyak informasi berdasarkan kesimpulan hasil penelitian yang telah dilakukan. Hasil-hasil penelitian yang diperoleh secara statistik dapat signifikan atau tidak signifikan.

Indikasi statistik signifikan adalah apabila hasil-hasil yang diperoleh mengindikasikan untuk setiap kondisi cukup berbeda dan konsekuensinya hipotesis nol untuk ketidakberbedaan ditolak. Dapat dikatakan bahwa tampak nyata terdapat perbedaan terhadap suatu kondisi. Sebagai contoh, dalam studi perbedaan perlakuan, hasil statistik signifikan menunjukkan bahwa terdapat perlakuan yang lebih efektif dibanding dengan perlakuan yang lain. Hasil ini menggambarkan bahwa kumpulan hipotesis nol dari ketidakberbedaan di antara perlakuan ditolak.

Dari perspektif pilosofi keilmuan, hasil nol atau tidak signifikan sangat banyak memberikan informasi. Meskipun begitu, hasil tidak signifikan dapat juga menjadikan faktor penyebab kurang jelasnya pengaruh variabel, termasuk tidak adekuasi

statistik, kurang jelasnya instrumen, kegagalan asumsi tes statistik, prosedur yang ceroboh dan cenderung bias.

Kegagalan mendeteksi interkasi yang diharapkan dalam memberikan perlakuan terhadap klien menjadi salah satu faktor penyebab tidak adekuatnya disain variabel bebas (Malkiewich dan Merluzzi's, 1980). Manipulasi eksperimental dikatakan sukses apabila dapat membedakan kondisi dari variabel yang sedang diteliti; apabila tidak ditemukan perbedaaan tersebut dapat dikatakan bahwa hipotesis yang dirancang tidak sesuai dengan kondisi kesekarang.

Status variabel menekankan pada perlunya perhatian dari seorang peneliti tentang variabel bebas. Dengan mendesain variabel bebas pada kondisi tertentu, peneliti berusaha menguji pengaruh terhadap variabel terikat. Penggunaan istilah 'manipulasi' untuk menggambarkan karakteristik proses kesengajaan tersebut.

Variabel bebas dapat dimanipulasi dan pengaruhnya terhadap variabel terikatlah yang akan dinilai; apabila semua berjalan dengan baik, maka hubungan timbal baliknya dapat dikatakan mantap. Yang lebih jelas, status variabel tidak dapat dimanipulasi dan tes statistik dipergunakan untuk mendeteksi asosiasi tersebut.

2. Mendesain dan Memilih Variabel Terikat

Pendesainan dan pemilihan variabel-variabel terikat dan metode-metode pengumpulan data berhadapan merupakan aktivitas-aktivitas kritis bagi para peneliti. Perhatian yang ekstra hati-hati harus dilakukan dalam proses ini, karena pilihan terhadap variabel-variabel terikat dapat menjadi penting untuk kebaikan penelitian tersebut. Tujuan dasar dari variabel terikat (ukuran terikat) adalah untuk mengukur konstruk yang dihipotesakan sebagai efek/akibat (konstruk efek).

Memilih atau mendesain variabel-variabel terikat merupakan operasionalisasi yang memadai dari konstruk-konstruk. Variabel-variabel terikat harus dirancang atau dipilih untuk merefleksikan konstruk-konstruk yang dirumuskan dalam pertanyaan penelitian. Operasionalisasi dari sebuah konstruk melibatkan properti-properti psikometrik dari variabel terikat. Para peneliti harus mengetahui pada tataran apa variabel-variabel terikat yang dipilih untuk mengoperasionalkan sebuah konstruk yang handal dan valid. Jika taksiran-taksiran terhadap reliabilitas dan validitas itu buruk, maka operasionalisasi konstruk itu kemungkinan tidak memadai.

Biasanya, koefisien reliabilitas untuk variabel X digambarkan dengan simbol r_{xx} . Sebuah koefisien r_{xx} yang sama dengan 0,80 mengindikasikan bahwa 80% dari varian di dalam skor-skor tersebut disebabkan perbedaan-perbedaan sejati, dan bahwa 20% disebabkan oleh faktor-faktor lain. Dalam penelitian dapat terjadi kekeliruan (*error*) reliabilitas, oleh karena itu perlu mengkaji beberapa sumber *error* 'kekeliruan' dalam pengukuran: (1) *Error respon acak*, yaitu kekeliruan dalam merespon item-item tertulis dalam semua jenis pengukuran, dapat diatasi dengan menilai semua karakteristik yang bermakna dari individu/situasi dan instrumen penelitian biasanya berisi banyak item yang mengukur ciri yang sama; (2) *Error spesifik*, yaitu kekeliruan yang dihasilkan oleh sesuatu yang unik bagi instrumen yang berbeda dari apa yang dimaksud peneliti, respon-respon para peserta itu ditentukan pada satu tataran dengan derajat dimana mereka ingin tampil seperti yang diharapkan secara sosial atau hasrat sosial; (3) *Error sementara*, yaitu kekeliruan karena kondisi-kondisi atau stimulus tertentu mempengaruhi pengukuran; (4) *Ketidaksepakatan antar penilai*, yaitu timbulnya varians pada penilaian para pengamat yang disebabkan oleh pengamatnya; (5) *Error* dalam penilaian yang diciptakan oleh peneliti melalui penyekoran dan perekaman data; dan (5) *Pembauran errors*, yaitu

pembauran/pencampuran kekeliruan-kekeliruan di atas, sehingga menimbulkan penilaian dengan reliabilitas yang luar biasa. Misalnya, ketika seorang instruktur praktikum konseling diminta untuk menilai praktikum keterampilan konseling dengan seorang klien dalam suatu sesi tertentu, dengan menggunakan skala 1 sampai 100. Operasionalisasi dari keterampilan konseling ini mengangkat banyak sumber kekeliruan. Pertama, terdapat variansi yang tidak dikenal di antara para instruktur praktikum. Kedua, hanya sebuah item tunggal bermakna ganda yang digunakan untuk menilai konstruk tersebut. Ketiga, tingkat keterampilan yang diperlihatkan dalam sesi tunggal dipengaruhi kekeliruan sementara yang mempengaruhi kesuksesan dari sesi tersebut. Keempat, kesempatan-kesempatan kekeliruan penyekoran dan perekaman tidak diminimalisir.

Dalam menginterpretasi taksiran-taksiran reliabilitas sebuah instrumen penelitian melibatkan banyak pertimbangan. Pertama, setiap koefisien reliabilitas merupakan taksiran terhadap reliabilitas sejati, yakni rata-rata dari sebuah sampel adalah taksiran terhadap rata-rata populasi. Kedua, reliabilitas merefleksikan varian yang disebabkan skor-skor sejati, namun itu tidak mengindikasikan skor-skor sejati yang sedang diukur. Ketiga, reliabilitas itu didasarkan pada skor-skor dan tidak pada instrumen dari mana skor itu diperoleh. Faktor-faktor lain yang perlu dipertimbangkan, termasuk validitas, waktu yang diperlukan untuk menyelesaikan instrumen tersebut, dan biaya.

Terdapat banyak cara untuk menaksir reliabilitas skor-skor, masing-masing peka pada satu atau lebih kekeliruan-kekeliruan (*errors*) (Heppner dkk., 2007: 319). Jika berbagai item dari sebuah instrumen itu mengukur konstruk yang sama, maka skor-skor pada item-item tersebut akan cenderung untuk bervariasi; yakni seseorang yang memiliki sebuah tingkatan tinggi atas konstruk

tersebut (misalnya, gelisah/cemas) akan cenderung menjawab semua item tersebut dalam satu arah, sementara seseorang yang memiliki sebuah tingkatan rendah terhadap sebuah konstruk (misalnya, tidak gelisah) akan cenderung menjawab semua item dalam cara lain. *Konsistensi internal* merujuk pada homogenitas dari item-item tersebut. Ketika skor-skor untuk berbagai item itu tinggi, konsistensi internal pun tinggi.

Indeks-indeks yang memperhitungkan pengukuran yang paling lazim digunakan adalah korelasi tes-rites. Jika sebuah konstruk diharapkan tetap stabil selama suatu periode waktu, dan jika instrumen itu tidak tunduk pada kekeliruan sementara atau kekeliruan respon acak, maka korelasi-korelasi tes-rites itu haruslah tinggi. Jika konsistensi internal tinggi namun koefisien tes-rites itu relatif rendah dan konstruknya diharapkan stabil selama periode waktu itu, maka skor-skor tersebut merefleksikan efek-efek sementara (Heppner dkk., 2007: 319).

Satu permasalahan dengan koefisien tes-rites adalah bahwa koefisien itu menaksir berlebihan reliabilitas karena koefisien tersebut tidak peka dengan kekeliruan spesifik (Heppner dkk., 2007: 320). Jika sesuatu yang unik diukur dengan sebuah instrumen, maka karakteristik unik ini akan terukur pada pemberian instrumen ini untuk kedua kalinya. Satu cara untuk menangani permasalahan ini adalah menggunakan bentuk-bentuk paralel pada dua waktu tersebut.

Dari banyak tipe validitas, tipe yang paling penting untuk tujuan penelitian adalah *validitas konstruk* – tingkatan yang skor-skornya merefleksikan konstruk yang diharapkan bukannya konstruk lain. Jelasnya, skor-skor yang tidak ajeg (*unreliable*) tidak dapat memiliki validitas konstruk karena skor-skor itu kebanyakan disebabkan kekeliruan acak (*random error*). Kendatipun demikian,

skor-skor yang ajeg dapat merefleksikan satu atau lebih konstruk yang lain dari pada konstruk yang dispesifikan. Secara khusus, skor-skor dapat cukup *reliable* namun kurang validitas konstruk. Jika korelasi-korelasi yang diharapkan ditemukan, maka akan muncul *validitas konvergen* (Heppner, 2007: 322).

Mutu suatu penelitian terutama dinilai dari validitas hasil yang diperoleh (Furqon, 1997: 12). Validitas penelitian terdiri atas validitas internal dan validitas eksternal. Validitas internal berkaitan dengan keyakinan peneliti tentang ketepatan atau kesahihan hasil penelitian. Dengan kata lain validitas internal adalah validitas yang berkenaan dengan keabsahan atau validitas hasil suatu percobaan (Ruseffendi, 2005: 56). Selanjutnya ditegaskan bahwa validitas internal dapat dirusak karena: adanya peristiwa, pengetesan, materi tes, perlakuan yang berbaur, keterlibatan peneliti/petugas, regresi statistik, kekeliruan statistik, pemiliha subjek, subjek hilang, dan subjek mendewasa.

Validitas eksternal berkaitan dengan tingkat generalisasi hasil penelitian yang diperoleh (Furqon, 1997: 12). Validitas eksternal adalah validitas yang berkenaan dengan dapat tidaknya hasil penelitian diperluas penerapannya untuk subjek dan lingkungan lain (Ruseffendi, 2005: 61). Dengan demikian kuat lemahnya validitas eksternal suatu eksperimen bergantung pada seberapa jauh generalisasi dapat dibuat dari hasil eksperimen tersebut.

Analisis faktor dapat digunakan untuk mengembangkan skala-skala. Strategi ini melibatkan item-item analisis faktor bukannya variabel. Sejumlah item itu tunduk dengan sebuah analisis faktor, item-item disegregasikan dengan pembebanan pada faktor-faktor, deskriptor yang diberikan pada faktor-faktor, dan skor-skor subskala itu dihitung berdasarkan segregasi.

Sebuah pengembangan strategi analisis faktor adalah mengembangkan item-item yang secara spesifik mengukur faktor-

faktor sebuah konstruk. Strategi ini digunakan untuk mengembangkan satu dari instrumen-instrumen konseling yang digunakan secara paling luas, *the Counselor Rating Form* (CRF; Barak & LaCrosse, 1975). CRF adalah sebuah skala yang didesain untuk mengukur tiga karakteristik konselor yang berkaitan dengan pengaruh sosial yang dimiliki konselor terhadap klien: kepercayaan, kemenarikan, dan keahlian.

Penggunaan variabel-variabel terikat yang jamak seringkali direkomendasikan untuk meningkatkan validitas konstruk. Tidak ada variabel yang dapat secara memadai mengoperasionalkan sebuah konstruk, karena beberapa varians dalam variabel ini disebabkan konstruk-konstruk lain (varian spesifik) dan sebagian disebabkan oleh kekeliruan. Dengan menggunakan beberapa variabel itu dapat secara lebih memadai merepresentasikan konstruk tersebut karena satu variabel akan peka dengan aspek dari konstruk yang tidak ada dalam variabel-variabel lain. Memasukkan ukuran-ukuran jamak diharapkan agar konstruk-konstruk berbeda dapat menghasilkan hasil-hasil berbeda. Ukuran-ukuran jamak dari konstruk-konstruk juga dapat digunakan untuk menghindari penyurutan korelasi antara konstruk-konstruk dan dapat memperhitungkan variansi metode.

Ukuran-ukuran jamak dari sebuah konstruk dapat digunakan untuk mendeteksi hubungan-hubungan antara konstruk-konstruk yang murni dengan ketidakreliabelan/ketidakajegan. Sebagai contoh, modeling persamaan struktural dapat digunakan untuk mendeteksi hubungan-hubungan yang sejati di antara konstruk-konstruk. Pemodelan persamaan struktural adalah sebuah metode statistik yang menguji hubungan antara konstruk-konstruk dengan menggunakan beberapa ukuran yang teramati untuk mengoperasionalkan konstruk tersebut (Cole, 1986).

Pada contoh tersebut, dua konstruk penting – depresi dan kecemasan – oleh Tanaka-Matsumi dan Kameoka (1986). Tanaka-Matsumi dan Kameoka memberikan tiga ukuran depresi yang lazim digunakan dan enam ukuran kecemasan yang lazim dipakai; korelasi-korelasi antar ukuran-ukuran ini dihadirkan pada Tabel 13.1. Beberapa observasi dapat dibuat dari tabel ini. Tampak bahwa ukuran-ukuran dari konstruk yang sama tersebut secara moderat tinggi, menunjukkan suatu validitas konvergen. Konstruk kecemasan dan depresi tampak berkaitan karena korelasi-korelasi yang diperoleh di antara ukuran-ukuran depresi dan kecemasan terentang dari 0,33 sampai 0,74. Perlu diingat bahwa semua korelasi dalam tabel tersebut diperkecil oleh ketidakreliabilisan. Pemodelan persamaan struktural memberikan sebuah alat penaksiran korelasi dari konstruk-konstruk depresi dan kegelisahan, yang mempertimbangkan ketidakreliabelan ini.

Korelasi dari konstruk-konstruk depresi dan kecemasan kemudian ditaksirkan dari variabel-variabel tersembunyi dari kecemasan dan depresi. Anak panah ganda yang berupa kurva/lengkungan antara "Depresi" dan "kecemasan" menggambarkan korelasi ini, yang terhitung 0,98. Ini mengatakan bahwa taksiran korelasi konstruk depresi dan kecemasan, seperti yang diukur oleh tiga ukuran depresi dan enam ukuran kecemasan adalah 0,98. Korelasi ini tidak dikecilkan oleh ketidakreliabelan. Kesimpulannya adalah bahwa konstruk-konstruk depresi tidak nyata.

Variansi yang lazim pada semua ukuran yang menggunakan metode sama itu disebut variansi metode. Variansi metode menggabungkan hubungan-hubungan di antara variabel; yakni, hubungan antara dua ukuran yang disebabkan tidak hanya oleh sebuah hubungan konseptual dalam konstruk-konstruk yang dimaksud, melainkan juga pada suatu hubungan berkenaan dengan

bagaimana konstruk-konstruk itu diukur. Sementara ketidak-reliabelan mengecilkan korelasi, variansi metode menggembungkan korelasi. Variansi metode seringkali muncul dalam penelitian konseling ketika berbagai aspek konseling dinilai dari perspektif yang sama. Misalnya, jika para supervisor menilai baik kompetensi kultural dari konselor maupun kemajuan terapeutik klien, maka korelasi antara kompetensi kultural dan hasil dipengaruhi sebagian oleh perspektif penilaian. Jika supervisor memiliki sikap yang secara umum mendukung pada konselor, maka supervisor itu akan cenderung menilai semua aspek dari konselor dan klien tersebut sebagai hal yang positif.

Berdasarkan paparan di atas, ada enam isu sentral pengukuran dalam penelitian, yaitu:

- 1) Suatu operasi tunggal (yakni, sebuah skala atau instrumen tunggal) hampir akan selalu merepresentasikan sebuah konstruk secara buruk.
- 2) Korelasi antara dua konstruk dikecilkan dengan ketidak-reliabilitiasan.
- 3) Ketidak-reliabilitas-an selalu membuatnya lebih sulit untuk mendeteksi efek-efek yang nyata karena kekuatan statistik tereduksi.
- 4) Korelasi antara dua ukuran yang menggunakan metode yang sama digembungkan dengan variansi metode.
- 5) Jika dimungkinkan, ukuran-ukuran beragam yang menggunakan metode jamak hendaknya digunakan untuk mengoperasionalkan sebuah konstruk.
- 6) Biasanya, interpretasi-interpretasi harus dibuat pada tingkatan konstruk, karena jarang kita tertarik dalam ukuran-ukuran itu semata. Pengetahuan tentang efek-efek ketidak-reliabilitas-an dan variansi metode itu penting untuk menarik kesimpulan-kesimpulan yang tepat.

Validitas merupakan suatu properti skor-skor dan bukan properti instrumen. Tingkatan dimana variabel-variabel memproduksi skor-skor yang secara memadai merefleksikan sebuah konstruk itu tergantung pada tipe peserta dalam penelitian tersebut. Hubungan-hubungan yang diperlihatkan oleh pemodelan persamaan struktural itu hanya dapat digeneralisir pada orang-orang yang mirip dengan mereka yang digunakan untuk mengumpulkan data. Ponterotto dan Casas (1991) menganalisis penelitian multikultural dalam jurnal-jurnal konseling dan menemukan bahwa hanya 25% dari instrumen-instrumen tersebut yang digunakan dalam penelitian itu dikembangkan menggunakan populasi minoritas rasial dan etnis. Validitas konstruk dari ke-75%-nya dikembangkan pada kelompok-kelompok lain. Ponterotto dan Casas (1991) berkesimpulan bahwa kurangnya instrumentasi berbasis minoritas itu merupakan satu dari keterbatasan terbesar penelitian multikultural yang mereka telaah.

Variabel terikat harus peka dengan karakteristik peserta, namun proses penilaian itu sendiri tidak boleh mempengaruhi karakteristiknya secara langsung; yakni, ukuran terikat harus mengindikasikan bagaimana peserta itu berfungsi secara normal. Variabel-variabel yang mempengaruhi karakteristik-karakteristik dari para peserta yang berupaya diukur itu dikatakan *reaktif*. Hakikat reaktif dari variabel terikat harus dipertimbangkan dalam mendesain penelitian.

Sejumlah isu prosedural harus dipertimbangkan ketika memilih atau mendesain variabel terikat. Seringkali, waktu yang dipakai pada penilaian itu penting bagi kesuksesan penelitian. Para peserta akan enggan untuk ikut penelitian yang menuntut waktu yang panjang dalam menyelesaikan isian dan instrumen, atau jika mereka ikut pun mereka mungkin memberi respon secara ceroboh pada item-itemnya (meningkatkan variansi *error*), khususnya pada

akhir dari sebuah periode penilaian yang panjang. Seperti yang disebutkan sebelumnya, kemampuan baca dari instrumen itu penting untuk performansi psikometriknya. Setiap instrumen yang diberikan harus diperiksa untuk memastikan bahwa para peserta dapat membaca secara memadai bahan-bahannya.

Metode pengumpulan data dapat mempengaruhi reliabilitas dan validitas hasil penelitian. Beberapa metode yang dapat digunakan dalam penelitian, yaitu pelaporan diri, penilaian (*rating*) terhadap orang lain dan peristiwa, observasi behavioral, indeks fisiologis, wawancara, dan teknik projektif.

Pada pengukur pelaporan diri (*self-report*), peserta menilai tingkatan dimana suatu karakteristik itu hadir atau dimana suatu perilaku itu muncul. Pelaporan diri dapat dicapai dengan memberi respon pada item-item dalam sebuah inventori, mengisi sebuah log (catatan harian), atau menulis sebuah jurnal. Peserta itu sendiri ini membuat observasi atau laporan. Diasumsi bahwa laporan tersebut secara akurat merefleksikan keadaan yang sebenarnya dari persoalan atau peserta memberi respon secara jujur dan akurat. Metode ini memiliki beberapa keuntungan: (1) yang membuatnya menjadi alat penilaian; (2) relatif mudah diberikan; (3) dapat digunakan untuk mengakses fenomena yang sangat sulit atau tidak mungkin diukur; dan (4) sesuai dengan pandangan fenomenologis dari konseling dan psikoterapi. Adapun kerugian pelaporan diri adalah: (1) metode ini rapuh dengan distorsi (yang disengaja ataupun tidak) oleh peserta; (2) peserta juga mungkin tidak tahu karakteristik dari yang diukur; dan (3) berkaitan dengan kongruensi antara sebuah perspektif fenomenologis dan pelaporan diri, bahwa pelaporan diri itu kurang dihargai oleh beberapa perspektif teoretis lain. Pengumpulan data pelaporan diri melalui internet (*online*) dapat mengatasi hambatan logistik, karena data dapat secara langsung dimasukkan ke dalam sebuah paket *software* statistik. Namun

kerugiannya yaitu: (1) terbatas pada para peserta yang memiliki akses pada komputer, (2) dalam hal validitas dan reliabilitas, apakah internet itu merupakan sarana yang tepat untuk mengumpulkan data penelitian.

Penelitian konseling seringkali mengandalkan pada penilaian (*ratings*) yang dibuat terhadap orang lain atau peristiwa. Prosedur-prosedur tersebut mirip dengan prosedur untuk pelaporan diri, kecuali para responden memberi nilai karakteristik dari peserta atau peristiwa tersebut. Seringkali para responden adalah para ahli, dan penilaian mereka diasumsikan merefleksikan secara akurat karakteristik dari orang atau peristiwa. Penilaian terhadap orang lain atau peristiwa mudah memberikannya dan fleksibilitas. Ketika *raters* itu para ahli, penilaian mereka dipandang berharga karena dibuat dengan suatu latar belakang dan pemahaman yang mendalam. Namun, *rating* seringkali bersifat umum dan dapat secara sistematis bias, terutama ketika *raters* mengetahui hipotesis dan kondisi-kondisi kognitif yang dimiliki para peserta. Strategi untuk menguji *varians* karena *raters* adalah menggunakan penilai yang netral atau banyak, dan kemudian menguji perbedaan di antara para *raters*.

Ukuran-ukuran behavioral dapat diperoleh dari observasi-observasi terhadap perilaku yang jelas, biasanya dilakukan oleh seorang observer terlatih. Psikologi behavioral telah menekankan pentingnya perilaku yang nampak dan tidak memberikan tekanan pada fenomena intrapsikis. Karena itu, mengamati dan merekam perilaku adalah komponen kunci dari analisis perilaku yang diterapkan. Observasi behavioral itu sama dengan *ratings* terhadap orang atau peristiwa lain, kecuali berkenaan dengan ukuran-ukuran behavioral yang memfokuskan pada perilaku yang tampak, dapat diamati dan tidak menggantungkan diri pada inferensi para penilai. Keuntungan observasi behavioral adalah bahwa observasi itu merupakan ukuran-ukuran langsung dan objektif, serta para peserta

dapat dinilai dalam berbagai lingkungan. Sedangkan kerugiannya adalah kurang representatif, dan fakta bahwa permasalahan dan kepedulian para klien seringkali tidak terpusat pada perilaku yang dapat diamati.

Respon-respon biologis dari para peserta seringkali digunakan untuk menyimpulkan keadaan-keadaan psikologis. Banyak fenomena psikologis memiliki korelasi-korelasi fisiologis yang dapat digunakan sebagai variabel-variabel terikat. Pada kenyataannya, respon-respon fisiologis seringkali dapat dianggap sebagai ukuran-ukuran langsung dari sebuah konstruk. Pengukur-pengukur fisiologis itu mahal, memerlukan keahlian, mungkin reaktif, dan tunduk pada kekeliruan yang disebabkan sejumlah faktor mekanis dan elektronis. Akibatnya, pengukur-pengukur fisiologis jarang digunakan dalam penelitian konseling. Kendatipun demikian, pengembangan teknik pengukuran fisiologis yang semakin canggih sangat menjanjikan bagi psikologi konseling dan mungkin menjadi sebuah bidang yang semakin memfokus untuk generasi peneliti yang akan datang.

Wawancara itu merupakan sebuah alat langsung untuk memperoleh informasi dari para peserta. Pada dasarnya, proses penggunaan wawancara untuk memperoleh data berkenaan dengan sebuah variabel terikat itu sama, kecuali bahwa tujuannya adalah untuk menghitung suatu konstruk. Kerlinger (1986) mendukung penggunaan wawancara personal karena kontrol yang lebih besar dan kedalaman dari informasi dapat diperoleh. Kedalaman informasi seringkali merupakan hasil dari perencanaan dan pengembangan jadwal wawancara yang cermat.

Rasionalisasi dibalik teknik-teknik proyektif adalah bahwa respon-respon peserta pada stimuli akan memperlihatkan segi kepribadian mereka. Secara historis, teknik-teknik proyektif

diasosiasikan dengan pendekatan-pendekatan psikodinamika untuk memahami perilaku manusia. Kendatipun demikian, seiring dengan popularitas pendekatan psikodinamika yang menurun, demikian juga dengan penggunaan teknik proyektif.

Untuk menghilangkan reaktivitas, seringkali dimungkinkan mengumpulkan data para peserta tanpa kepekaan terhadap proses ini. Pengukuran-pengukuran digunakan dalam suatu cara sehingga para peserta tidak menyadari prosedur penilaian, hal ini dikenal sebagai pengukur yang tidak menonjol. Dimungkinkan untuk mengamati peserta (dalam naturalistik) tanpa mereka tahu, untuk mengamati peserta dalam situasi direncanakan, mengumpulkan data dari arsip atau sumber lain, atau meneliti jalur-jalur fisik. Kebanyakan psikolog sangat tertarik pada sumber-sumber data yang tidak menonjol.

B. Implikasi

Mengacu pada rangkuman isi cepter dan pembahasan di atas, dapat dirumuskan beberapa implikasi untuk meningkatkan kualitas proses dan hasil penelitian di bidang bimbingan dan konseling.

1. Mendesain dan mengevaluasi variabel bebas, merupakan langkah vital dalam proses penelitian kuantitatif terutama penelitian eksperimental. Apalagi penelitian dalam bidang bimbingan dan konseling, yang senantiasa berkaitan dengan keperilakuan konselor dan konseli. Oleh karena itu, ketika mendesain dan mengevaluasi variabel bebas akan menyangkut aspek-aspek perilaku individu yang unik dan substansi yang khas dari bimbingan dan konseling.

Dalam mengidentifikasi, mendesain dan mengevaluasi variabel bebas perlu didukung oleh penguasaan konseptual/teoretis

yang memadai berkenaan dengan aspek-aspek perilaku individu (konselor dan konseli) dan substansi bimbingan dan konseling yang dijadikan variabel bebas. Hal ini penting, karena variabel bebas yang dipilih harus didesain konstruknya secara operasional. Konstruk variabel bebas suatu penelitian dipandang operasional apabila merujuk pada minimal dua landasan konseptual yang kokoh, sehingga indikator-indikatornya dapat diamati (*observable*) dan dapat diukur (*measurable*).

Mengoperasionalkan variabel bebas harus dilakukan secara cermat sebelum memulai penelitian. Agar rumusan atau konstruk variabel bebas benar-benar operasional, maka peneliti perlu: (1) menentukan kondisi atau level variabel bebas secara jelas; (2) merefleksikan atau menjabarkan secara tepat konstruk variabel yang dirancang ke dalam pertanyaan penelitian; (c) membatasi keragaman dengan pengkondisian, sehingga tidak memunculkan kondisi-kondisi lain yang dapat mengacaukan penelitian; dan (d) menetapkan kejelasan sumber-sumber perbedaan dengan merumuskan keragaman dalam kondisi-kondisi variabel, merancang dan menjelaskan secara logis bahwa kondisi-kondisi lain tidak akan membiaskan penelitian.

Bila peneliti sudah berusaha secara sungguh-sungguh dalam mengidentifikasi, merumuskan dan mengoperasionalkan variabel bebas, maka peneliti perlu memverifikasi manipulasi variabel eksperimental agar dapat mencapai tujuannya. Untuk menguji manipulasi variabel, perlu jelas betul apakah kondisi-kondisi itu sebagai dimensi yang sangat diharapkan dan tidak mencerminkan dimensi lain; serta apakah tritmen dapat diimplementasikan sesuai petunjuk yang diharapkan. Jika manipulasi variabel bebas telah benar-benar teruji, maka diharapkan diperoleh penafsiran secara tepat atas hasil-hasil eksperimen, sehingga kesimpulan pun dapat dirumuskan secara akurat. Pada akhirnya, diharapkan diperoleh

hasil-hasil penelitian yang secara statistik signifikan, yang memiliki validitas internal dan validitas eksternal secara memadai.

2. Mengidentifikasi, mendesain dan memilih variabel terikat harus dilakukan oleh peneliti secara cermat, karena akan menjadi dasar pengukuran konstruk yang dihipotesiskan sebagai akibat dari manipulasi variabel bebas. Oleh karena itu, pendesainan variabel-variabel terikat dan ketepatan metode-metode pengumpulan datanya merupakan aktivitas-aktivitas kritis bagi para peneliti sebelum melakukan penelitian.

3. Untuk meningkatkan reliabilitas dan validitas hasil-hasil penelitian, terutama dalam penelitian eksperimental, perlu diidentifikasi sejak dini beberapa sumber kekeliruan (*errors*) dalam pengukuran, sehingga tidak menurunkan kualitas temuan-temuan penelitian.

4. Metode pengumpulan data dapat mempengaruhi reliabilitas dan validitas hasil penelitian. Setiap metode dalam pengumpulan data penelitian memiliki kekurangan dan kelebihan, oleh karena itu peneliti perlu menetapkan metode penelitian yang sesuai dengan desain penelitian. Dengan demikian, instrumen penelitian harus sesuai dengan konstruk variabel atau definisi operasional variabel penelitian yang dikembangkan.

DAFTAR PUSTAKA

- Borg, W.R. & Gall, M.D. (1989). ***Educational Research: An Introduction***. London: Longman, Inc.
- Bogdan, R. C. & Biklen, S. K. (1982). ***Qualitative Research for Education***. Boston: Allyn and Bacon, Inc.
- Ellis, A.K. & Fouts, J.T. (1993). ***Research on Educational Innovations***. Prenceton Junction, New Jersey: Eye on Education.
- Fraenkel, Jack R & Wallen, Normal E. (1993). ***How to Design and Evaluate Research in Education***. New York: McGraw-Hill, Inc.
- Furqon. (2002). ***Statistika Terapan untuk Penelitian***. Bandung: Alfabeta.
- Furqon. (2006). Penilaian Hasil Belajar untuk Meningkatkan Mutu Pendidikan. ***Makalah***. Disampaikan pada Pengukuhan Jabatan Guru Besar Tetap dalam Bidang Ilmu Penelitian dan Evaluasi pada Fakultas Ilmu Pendidikan Universitas Pendidikan Indonesia.
- Gall, M.D., Gall, J.P. & Borg, W.R. (2003). ***Educational Research: An Introduction***. Boston: Pearson Education, Inc.
- Heppner, P. P., Wampold, B.E. & Kivlighan, D.M. (2008). ***Research Design in Counseling***. Belmont, California: Thomson Brooks/Cole.
- Janda, Louis H. (1999). ***Career Tests***. Massachusetts: Adams Media Corporation.
- Kerlinger, F. N. (1976). ***Fundamental of Behavior Research***. New York: Holt, Rinehart and Winston.
- Mouly, G. J. (1978). ***Educational Research***. Boston: Allyn and Bacon, Inc.
- Prince, Jeffrey P. & Heiser, Lisa J. (2000). ***Essentials of Career Interest Assessment***. New York: John Wiley & Sons, Inc.

- Ruseffendi. (2005). ***Dasar-dasar Penelitian pendidikan dan Bidang Non-Eksakta Lainnya***. Bandung: Tarsito.
- Seligman, L. (1994). ***Development Career Counseling and Assessment***. London: Sage Publications, Inc.
- Sudjana. (1982). ***Metoda Statistika***. Bandung: Tarsito.
- Tuckman, B. W. (1978). ***Constructing Educational Research***. San Diego: HBJ Publishers.
- Winarno Surakhmad. (1982). ***Pengantar Penelitian Ilmiah: Dasar, Metode dan Teknik***. Bandung: Tarsito.