

B A B VIII

TEKNIK PENGUKURAN

Pada penelitian sosial ataupun pendidikan, proses pengukuran adalah rangkaian dari empat aktivitas, yakni 1) menentukan dimensi konsep penelitian. Konsep dan variabel penelitian sosial maupun pendidikan sering kali memiliki lebih dari satu dimensi. Semakin lengkap dimensi suatu variabel, yang dapat diukur, semakin baik ukuran yang dihasilkan. Kita ambil konsep hasil belajar. Dalam konsep taksonomi Bloom hasil belajar dinyatakan sebagai aspek aspek yang berhubungan dengan ranah kognitif, avaktif dan psikomotorik. Artinya ukuran variabel hasil belajar dapat dikatakan lengkap apabila ketiga ranah tersebut tercakup dalam instrumen pengukuran hasil belajar. 2) Rumusan ukuran untuk masing masing dimensi. Ukuran ini biasanya berbentuk pertanyaan pertanyaan yang relevan dengan dimensi tadi. 3) tentukan tingkat ukuran yang akan digunakan. Dalam penelitian sosial dikenal empat tingkat ukuran , yaitu: nominal, ordinal, interval dan rasio. 4) Tentukan tingkat validitas dan reliabilitas dari alat ukur .Pengujian validitas dan reliabilitas perlu dilakukan bila yang dipakai adalah instrumen penelitian yang baru atau instrumen penelitian buatan sendiri.

1. Tingkat Tingkat Pengukuran

Pengukuran dimaksudkan sebagai penunjukan angka angka (gradasi) pada suatu variabel menurut aturan tertentu (standar) Aturan pertama yang perlu diketahui seorang peneliti agar dqapat mengukur atau memberikan angka atau nilai dengan tepat untuk suatu konsep yang diamatinya adalah mengenai tingkat pengukuran atau skala pengukuran. Tingkat pengukuran yang luas digunakan dalam penelitian sosial adalah yang dikembangkan oleh SS Stevans yang membagi tingkat ukuran kedalam empat katagori yakni : nominal , ordinal, interval dan rasio.

a. Skala Nominal. .

Skala nominal adalah tingkat pengukuran yang paling rendah. Pada ukuran ini tidak ada asumsi tentang jarak maupun urutan antara katagori dalam pengukurannya. Dasar penggolongan hanyalah katagori yang terjadi karena adanya persamaan dan ketidak samaan pada obyek ukur. Penggolongan tersebut tidak tumpang tindih (mutually exclusive) dan tuntas (exhaustive). Angka yang ditunjuk untuk suatu katagori tertentu hanyalah sekedar label atau kode atau ciri, untuk membedakannya dengan obyek lainnya. Pada tingkat skala nominal; ini, peneliti dapat mengelompokkan ke dalam dua katagori atau lebih, menurut variabel tertentu. Misalnya jenis kelamin, kelompok agama, status pekerjaan, asal sekolah dan variabel lainnya yang sejenis. Pada pengukuran yang menggunakan skala nominal bilangan atau angka yang tergradasi pada suatu obyek hanya berfungsi sebagai lambang untuk membedakannya dengan obyek lainnya. Misalnya nomer rumah, nomer telepon, Nomer punggung pemain sepak bola dan sebagainya.

b. Skala Ordinal.

Skala ordinal adalah skala yang memungkinkan peneliti mengurutkan responden dari tingkat (posisi) yang rendah ke tingkat yang paling tinggi menurut aturan tertentu yang berlaku terhadap variabel yang di ukur.

Skala ordinal banyak digunakan pada penelitian sosial terutama untuk mengukur kepentingan sikap atau persepsi. Melalui pengukuran ini peneliti dapat membagi responden kedalam urutan atau ranking atas dasar sikapnya pada suatu obyek atau tindakan tertentu. Dalam skala ordinal angka atau bilangan memiliki dua makna yakni : 1). sebagai lambang untuk membedakan suatu obyek dengan obyek lain dan 2). mengisaratkan peringkat atau kedudukan individu didalam kelompok nya. Contohnya tingkatan kepangkatan pada militer, tingkatan kelas disekolah, urutan kejuaraan pada suatu turnamen olahraga dan sebagainya.

c. Skala interval.

Seperti halnya skala ordinal, skala interval dapat membedakan, dapat menurutkan individu, berdasarkan atribut tertentu. Selain itu skala interval juga memberikan informasi tentang interval antara satu individu dengan individu lainnya. Interval (jarak yang sama) pada skala ini dipandang sebagai mewakili kelompok-kelompok interval tertentu dengan jarak yang sama antara satu kelompok dengan kelompok lainnya. Skala dan indeks sikap biasanya menghasilkan ukuran yang berskala interval. Karena itu skala ini merupakan salah satu ukuran yang paling sering digunakan dalam penelitian sosial. Dalam skala pengukuran interval ini bilangan atau angka memiliki empat fungsi yaitu: 1) sebagai lambang untuk membedakan individu, 2) sebagai peringkat atau urutan individu didalam kelompoknya, 3) menunjukkan interval kedudukan individu dimana tiap interval memiliki jarak yang sama atau tetap (constant interval size) dan 4) titik nol bukan dianggap sebagai nilai mutlak (standar atau acuan titik awal pengukuran) .

d. Skala Rasio.

Skala ini menggunakan suatu interval yang jaraknya tidak dinyatakan sebagai perbedaan nilai antar responden, tetapi jarak antara responden yang diukur dengan titik nol mutlak, sehingga perbandingan rasio dapat dilakukan. Dalam skala ini bilangan atau angka yang tergradasi pada individu memiliki empat makna yaitu: 1) sebagai lambang untuk membedakan individu, 2) sebagai peringkat atau urutan individu didalam kelompoknya, 3) menunjukkan interval kedudukan individu dimana tiap interval memiliki jarak yang sama atau tetap (constant interval size) dan 4) titik nol sebagai nilai mutlak (standar atau acuan titik awal pengukuran) .

B. Skala Pengukuran.

Skala adalah seperangkat nilai angka atau bilangan yang ditetapkan (digradasikan) pada suatu obyek, subyek atau tingkah laku dengan tujuan mengukur karakteristiknya (sifatnya). Skala dapat digunakan untuk mengukur karakteristik nilai-nilai dan perilaku individu. Muller menyatakan bahwa ada empat jenis skala

yang biasa digunakan untuk mengukur sikap sosial: yaitu Skala Likert, Thrustone, Guttman dan diferensial sumantik.

1. Skala Likert.

Skala ini digunakan apabila peneliti menginginkan data tentang pendapat responden mengenai masalah yang diteliti. Bentuk skala ini dapat dilakukan untuk penilaian kuantitatif terhadap keseluruhan /setiap responden . Cara ini dengan menetapkan bobot jawaban terhadap tiap tiap item/sub item yang ditetapkan sebagai indikator penelitian, pertanyaannya berbentuk positif atau negatif.

2. Skala Thrustone.

Skala ini dibagi menjadi tiga bagian yaitu 1) metode perbandingan pasangan 2), metode interval pemunculan sama dan 3) metode interval berurutan (atau ukuran dikhotomi).

Suatu skala bertujuan untuk mengurutkan responden berdasarkan kriteria tertentu. Skala yang disusun menurut metode Thrustone disusun sedemikian rupa sehingga interval antara urutan dalam skala mendekati interval yang sama besarnya . Olehkarena itu skala seperti ini sering disebut equal apparing interval atau equal interval scale (Skala interval sama). Dengan demikian ukuran yang dihasilkan skala ini hampir mendekati ukuran interval sehingga dapat menggunakan analisis statistik untuk pengolahan datanya.

Penafsiran skor pada skala Thrustone sama seperti membaca skor pada skala Bogardus, responden yang mempunyai skor lebih tinggi pada skala sikap tertentu berarti lebih mempunyai sikap positif terhadap suatu aksi atau kejadian.

3. Skala Guttman.

Metode Guttman didasarkan pada kenyataan bahwa relevansi tiap tiap indikator terhadap variabel adalah berbeda. Satu indikator mungkin lebih dapat mengukur variabel tersebut dengan lebih tepat. Yang dipertahankan oleh skala Guttman adalah ketunggalan dimensi (unidimensionality). Artinya, skala sebaiknya

hanya mengukur satu dimensi saja dari variabel yang memiliki beberapa dimensi. Prinsip lain yang terdapat pada skala Guttman adalah seperti yang terdapat pada Skala Thurstone. Pernyataan pernyataan dibuat lebih berbobot dan tiap pernyataan bobotnya berbeda. Apa bila seorang responden menyetujui pernyataan yang lebih berat bobotnya maka dia diharapkan akan menyetujui pernyataan pernyataan yang bobotnya lebih rendah.

Untuk menilai ketunggalan dimensi suatu skala diadakan analisis skalogram untuk mendapatkan koefisien reproduibilitas (coefficient of reproducibility), Kr dan koefisien skalabilitas (coefficient of scalability) Ks.

4. Skala Diferensial Semantik.

Skala ini berusaha mengukur arti obyek atau konsep bagi seorang responden. Responden diminta untuk menilai suatu obyek atau konsep pada skala yang mempunyai ajektif yang bertentangan. Skala bipolar ini mengandung unsur evaluasi dalam rentang ekstrem kiri sampai ekstrem kanan misalnya (bagus, buruk; berguna, tidak berguna;) unsur potensi (besar, kecil; kuat, lemah; berat, ringan) dan unsur aktivitas (aktif, pasif; cepat, lambat; panas, dingin). Menurut Osgood ketiga unsur ini dapat mengukur tiga dimensi sikap, antara lain : 1) evaluasi responden tentang obyek atau konsep yang sedang diukur, 2) persepsi responden tentang potensi obyek atau konsep tersebut dan 3) persepsi responden tentang aktivitas objek.

C. Validitas Alat Ukur.

Alat ukur yang berfungsi dengan baik akan mampu mengukur dengan tepat dan mengena pada gejala gejala sosial tertentu, kualitas alat seperti ini dinyatakan sebagai valid. Menurut Kerlinger (1997: 30) membagi validitas alat ukur penelitian dalam tiga jenis yaitu : validitas isi, validitas kriteria dan validitas konstruk.

1. Validitas isi.

Validitas isi mempersoalkan apakah isi dari alat ukur (bahan, topik, substansi) cukup representatif atau cukup merupakan suatu sampling. Validitas isi dipandu

oleh pertanyaan pertanyaan : apakah isi atau substansi alat ukur yang digunakan merupakan representatif dari isi suatu variabel secara universal dan mengandung sifat sifat yang semestinya diukur untuk data penelitian. Validitas isi secara mendasar adalah merupakan suatu pendapat, baik pendapat sendiri ataupun pendapat pakar . tiap tiap item atau soal dalam ujian perlu dipelajari secara seksama dan kemudian dipertimbangkan tentang sifat representatif isi yang akan diuji. Menurut Singarimbun (1995: 128) bahwa validitas isi suatu alat pengukur ditentukan oleh sejauh mana isi alat ukur tersebut mewakili semua aspek yang dianggap sebagai aspek kerangka konsep suatu variabel.

2. **Validitas pragmatik atau validitas kriteria.**

Validitas ini dilihat dengan cara membandingkan kriteria suatu alat ukur dengan suatu kriteria validitas alat ukur yang telah diketahui atau yang dipercaya dapat digunakan untuk mengukur suatu atribut tertentu. Jika skor atau skala yang diukur dibandingkan dengan satu atau lebih kriteria tertentu atau variabel yang dianggap dapat mengukur hal yang seharusnya diukur, maka yang dikerjakan adalah menetapkan validitas kriteria dari alat ukur tersebut. Validitas prediktif termasuk dalam validitas yang berhubungan dengan kriteria .

3. **Validitas Konstruk.**

Menurut Nasir (1999: 177) konstruk adalah suatu abstrak dan generalisasi khusus dan merupakan suatu konsep yang diciptakan khusus untuk kebutuhan ilmiah dan mempunyai pengertian terbatas. Sedangkan Singarimbun (1995: 125), menyatakan konstruk adalah kerangka dari suatu konsep. Misalkan seorang peneliti ingin mengukur konsep “religiusitas” . Pertama tama yang harus dilakukan oleh peneliti adalah mencari apa saja yang merupakan kerangka dari konsep tersebut. Dengan diketahuinya kerangka dari konsep tersebut seorang peneliti akan dapat menyusun tolok ukur operasional konsep tersebut. Untuk mencari kerangka konsep tersebut dapat ditempuh beberapa cara.

Cara yang lazim digunakan dalam dunia penelitian sosial adalah tiga cara yang akan dibahas sebagai berikut:

- a. Mencari definisi definisi konsep yang dikemukakan para ahli yang tertulis didalam literatur. Definisi tentang suatu konsep biasanya berisikan kerangka konsep tersebut. Terkadang para ahli tidak hanya memberikan definisi , tetapi juga telah memberikan kerangka konsep secara jelas . Kalau sekiranya sudah ada definisi yang jelas dan cukup operasional untuk dijadikan dasar penyusunan alat pengukur variabel, maka definisi tersebut sudah dapat langsung dipakai untuk menyusun pertanyaan dalam kusioner. Tetapi bila definisi yang dikemukakan belum operasional, definisi tersebut harus dijabarkan lebih lanjut agar lebih operasional dan dapat dijadikan dasar penyusunan pertanyaan dalam kuisisioner (angket).
- b. Apabila dalam literatur tidak dapat diperoleh definisi konsep yang ingin diukur, peneliti harus mendefinisikan sendiri konsep tersebut. Untuk membantu penyusunan definisi dan mewujudkan definisi tersebut kedalam bentuk yang operasional, peneliti disarankan untuk mendiskusikan konsep tersebut dengan pakar yang kompeten di bidang konsep yang akan diukur. Kemudian pendapat para pakar dan pendapat peneliti, dicari kesamaannya dan diambil konsep yang dianggap lebih logis dan rasional. Selanjutnya dapat disusun kerangka konsep yang dapat diwujudkan berupa pertanyaan yang akan dimuat di dalam alat ukur misalnya angket.
- c. Menanyakan definisi konsep yang akan diukur kepada calon responden, atau orang-orang yang memiliki karakteristik seperti responden. Misalnya peneliti ingin mengukur konsep “religiusitas”. Dalam mendefinisikan konsep ini peneliti dapat langsung menanyakan kepada beberapa calon responden, kemudian disusun kerangka suatu konsep. Pendekatan ketiga ini akhir-akhir ini banyak dipakai para peneliti sosial. Pendekatan ini dianggap baik karena kerangka suatu konsep dikembangkan berdasarkan pendapat calon responden. Cara yang demikian ini akan dapat menghindarkan dari bias yang sering terjadi bila

definisi operasional suatu konsep dikembangkan dari konsep para ahli dari negara barat yang latar belakang budayanya berbeda.

D. Reliabilitas.

Reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat ukur dapat dipercaya atau dapat diandalkan (Singarimbun 1995 : 140). Jika alat tersebut digunakan dua kali untuk mengukur gejala yang sama dan hasil pengukurannya relatif sama (konsisten) maka alat ukur tersebut reliabel. Dengan kata lain reliabilitas menunjukkan kekonsistenan suatu alat ukur di dalam mengukur gejala yang sama. Ada beberapa teknik yang dapat digunakan untuk menghitung indeks reliabilitas yakni metode /teknik paralel, uji ulang , teknik belah dua, dan Cronbach.

a. Metode Paralel.

Teknik ini disebut juga equivalent form atau alternative form. Pada teknik ini perhitungan reliabilitas dilakukan dengan membuat dua jenis alat ukur yang mengukur aspek yang sama. Kedua alat pengukur tersebut diberikan kepada responden yang sama kemudian dicari validitasnya untuk masing masing jenis (Singarimbun, 1995, 145; Kartono 1990, 127). Untuk menghitung reliabilitas perlu mengkorelasikan skor total dari kedua jenis alat ukur tersebut.

b. Metode Uji Ulang.

Menilai reliabilitas dengan teknik ini adalah dengan cara mengulang penelitian dengan menggunakan alat ukur yang sama dengan menggunakan responden yang sama dan dalam kurun waktu yang tidak lama pula. Hasil penelitian pertama kemudian dibandingkan dengan hasil penelitian kedua , dan kesetabilan dari jawaban dianalisis (Nasir 1999: 168). Hal yang sama juga diungkapkan oleh Kartono (1990: 126), bahwa pada teknik uji ulang ini orang menggunakan alat pengukuran yang sama , diberikan kepada sejumlah subjek yang sama pada saat yang berbeda , namun kondisi pengukuran dapat dianggap relatif sama.

c. Metoda Belah Dua (Split half).

Teknik ini hampir sama dengan teknik paralel hanya saja disini observasi dibagi menjadi dua bagian. Teknik ini berkeinginan untuk menguji reliabilitas pertanyaan pertanyaan yang berbentuk skala, yang mempunyai hubungan satu sama lain. Penilaian reliabilitas ini ditujukan untuk mengukur internal konsistensi pertanyaan atau pernyataan (Nasir 199: 171). Menurut Kartono (1990: 128) ada dua macam prosedur dalam teknik belah dua: 1.) mengumpulkan item item yang bernomer ganjil menjadi satu kelompok sendiri , dan yang bernomer genap dijadikan satu kelompok lain. Teknik ini disebut teknik belah dua ganjil-genap. Kemudian dilakukan perhitungan perhitungan yang menghasilkan koefisien reliabilitas instrumen. dengan membelah item item secara random menjadi dua kelompok yang sama banyaknya. Kemudian menghitung koefisien reliabilitasnya. Teknik ini disebut teknik belah dua random.

d. Metode Cronbach.

Cronbach (1952) menyarankan suatu koefisien reliabilitas yang disebut koefisien alpha menurut Croanbach pada hakekadnya merupakan rata rata dari semua koefisien korelasi belah dua yang mungkin dibuat dari satu alat ukur. Oleh karena itu pedoman yang dikemukakan oleh Balian (1988) tersebut di atas juga dapat dijadikan pedoman untuk koefisien alpha. Program komputer SPSS/PC+ dan CSS dapat melakukan perhitungan koefisien alpha dan koefisien belah dua dengan mudah dan cepat.

E. Analisis Butir Item.

Reliabilitas alat ukur dapat ditingkatkan dengan melakukan analisis butir. Analisis butir item (item analisis) dapat dilakukan dengan dua cara, yaitu dengan :

- a. Mencari korelasi antar nilai setiap butir dengan nilai total yang dikurangi nilai butir yang bersangkutan. Apabila butir konsisten dalam mengukur gejala yang diukur, maka butir tersebut akan mempunyai korelasi yang positif dan tinggi. Apabila koefisien korelasinya rendah, berarti butir tersebut perlu diperbaiki atau

dibuang. Dengan memperbaiki atau membuang item tersebut maka reliabilitas alat ukur akan meningkat.

- b. Discriminating power (daya pembeda) . Dengan cara ini instrumen yang telah diisi disusun menurut besarnya nilai total. Kemudian diambil 25 % instrumen yang mempunyai nilai tertinggi sebagai kelompok atas dan 25% yang mempunyai nilai terendah sebagai kelompok bawah. Untuk setiap butir dicari mean dan simpangan baku untuk kemudian diuji apakah terdapat beda yang signifikan antar kedua kelompok. Jika terdapat perbedaan , berarti butir yang bersangkutan baik, sedagkan jika tidak ada perbedaan , berarti butir tersebut tidak baik, karena tidak dapat membedakan kelompok atas dengan kelompok bawah. Item yang seperti ini perlu direvisi atau dibuang agar tingkat reliabilitas alat ukur meningkat.