
PENYUSUNAN DAN PENGEMBANGAN INSTRUMEN PENELITIAN

Oleh :

Dr.Ir. Pudji Muljono, MSi

**Disampaikan pada Lokakarya Peningkatan Suasana Akademik
Jurusan Ekonomi FIS-UNJ tanggal 5 sampai dengan 9 Agustus 2002**

A. Pendahuluan

Instrumen atau alat pengumpul data adalah alat yang digunakan untuk mengumpulkan data dalam suatu penelitian. Data yang terkumpul dengan menggunakan instrumen tertentu akan dideskripsikan dan dilampirkan atau digunakan untuk menguji hipotesis yang diajukan dalam suatu penelitian.

Instrumen memegang peranan yang sangat penting dalam menentukan mutu suatu penelitian, karena validitas atau kesahihan data yang diperoleh akan sangat ditentukan oleh kualitas instrumen yang digunakan, disamping prosedur pengumpulan data yang ditempuh. Hal ini mudah dipahami karena instrumen berfungsi mengungkapkan fakta menjadi data, sehingga jika instrumen yang digunakan mempunyai kualitas yang memadai dalam arti valid dan reliabel maka data yang diperoleh akan sesuai dengan fakta atau keadaan sesungguhnya di lapangan. Sedang jika kualitas instrumen yang digunakan tidak baik dalam arti mempunyai validitas dan reliabilitas yang rendah, maka data yang diperoleh juga tidak valid atau tidak sesuai dengan fakta di lapangan, sehingga dapat menghasilkan kesimpulan yang keliru.

Untuk mengumpulkan data dalam suatu penelitian, kita dapat menggunakan instrumen yang telah tersedia dan dapat pula menggunakan instrumen yang dibuat sendiri. Instrumen yang telah tersedia pada umumnya adalah instrumen yang sudah dianggap baku untuk mengumpulkan data variabel-variabel tertentu. Dengan demikian, jika instrumen baku telah tersedia untuk mengumpulkan data variabel penelitian maka kita dapat langsung menggunakan instrumen tersebut, dengan catatan bahwa teori yang dijadikan landasan penyusunan instrumen tersebut sesuai dengan teori yang diacu dalam penelitian kita. Selain itu, konstruk variabel yang diukur oleh instrumen tersebut juga sama dengan konstruk variabel yang hendak kita ukur dalam penelitian kita. Akan tetapi, jika instrumen yang baku belum tersedia untuk mengumpulkan data variabel penelitian maka instrumen untuk mengumpulkan data variabel tersebut harus dibuat sendiri oleh peneliti.

Dalam rangka memahami tentang pengembangan instrumen, maka berikut ini akan dibahas mengenai beberapa hal yang terkait dengan itu di antaranya langkah-langkah penyusunan dan pengembangan instrumen, teknik penyusunan dan penilaian butir instrumen, proses validasi konsep melalui panel, dan proses validasi empirik melalui ujicoba.

B. Langkah-langkah Penyusunan dan Pengembangan Instrumen

Untuk memahami konsep penyusunan dan pengembangan instrumen, maka di bawah ini akan disajikan proses atau langkah-langkah yang ditempuh dalam penyusunan instrumen dilengkapi dengan bagan proses penyusunan item-item instrumen suatu penelitian. Secara garis besar langkah-langkah penyusunan dan pengembangan instrumen adalah sebagai berikut :

1. Berdasarkan sintesis dari teori-teori yang dikaji tentang suatu konsep dari variabel yang hendak diukur, kemudian dirumuskan konstruk dari variabel tersebut. Konstruk pada dasarnya adalah bangun pengertian dari suatu konsep yang dirumuskan oleh peneliti.
2. Berdasarkan konstruk tersebut dikembangkan dimensi dan indikator variabel yang sesungguhnya telah tertuang secara eksplisit pada rumusan konstruk variabel pada langkah 1.
3. Membuat kisi-kisi instrumen dalam bentuk tabel spesifikasi yang memuat dimensi, indikator, nomor butir dan jumlah butir untuk setiap dimensi dan indikator.
4. Menetapkan besaran atau parameter yang bergerak dalam suatu rentangan kontinum dari suatu kutub ke kutub lain yang berlawanan, misalnya dari rendah ke tinggi, dari negatif ke positif, dari otoriter ke demokratik, dari dependen ke independen, dan sebagainya.
5. Menulis butir-butir instrumen yang dapat berbentuk pernyataan atau pertanyaan. Biasanya butir instrumen yang dibuat terdiri atas dua kelompok yaitu kelompok butir positif dan kelompok butir negatif. Butir positif adalah pernyataan mengenai ciri atau keadaan, sikap atau persepsi yang positif atau mendekati ke kutub positif, sedang butir negatif adalah pernyataan mengenai ciri atau keadaan, persepsi atau sikap negatif atau mendekati ke kutub negatif.
6. Butir-butir yang telah ditulis merupakan konsep instrumen yang harus melalui proses validasi, baik validasi teoretik maupun validasi empirik.
7. Tahap validasi pertama yang ditempuh adalah validasi teoretik, yaitu melalui pemeriksaan pakar atau melalui panel yang pada dasarnya menelaah seberapa jauh dimensi merupakan jabaran yang tepat dari konstruk, seberapa jauh

indikator merupakan jабaran yang tepat dari dimensi, dan seberapa jauh butir-butir instrumen yang dibuat secara tepat dapat mengukur indikator.

8. Revisi atau perbaikan berdasarkan saran dari pakar atau berdasarkan hasil panel.
9. Setelah konsep instrumen dianggap valid secara teoretik atau secara konseptual, dilakukanlah pengandaan instrumen secara terbatas untuk keperluan ujicoba.
10. Ujicoba instrumen di lapangan merupakan bagian dari proses validasi empirik. Melalui ujicoba tersebut, instrumen diberikan kepada sejumlah responden sebagai sampel uji-coba yang mempunyai karakteristik sama atau ekuivalen dengan karakteristik populasi penelitian. Jawaban atau respon dari sampel uji-coba merupakan data empiris yang akan dianalisis untuk menguji validitas empiris atau validitas kriteria dari instrumen yang dikembangkan.
11. Pengujian validitas dilakukan dengan menggunakan kriteria baik kriteria internal maupun kriteria eksternal. Kriteria internal, adalah instrumen itu sendiri sebagai suatu kesatuan yang dijadikan kriteria sedangkan kriteria eksternal, adalah instrumen atau hasil ukur tertentu di luar instrumen yang dijadikan sebagai kriteria.
12. Berdasarkan kriteria tersebut diperoleh kesimpulan mengenai valid atau tidaknya sebuah butir atau sebuah perangkat instrumen. Jika kita menggunakan kriteria internal, yaitu skor total instrumen sebagai kriteria maka keputusan pengujian adalah mengenai valid atau tidaknya butir instrumen dan proses pengujiannya biasa disebut analisis butir. Dalam kasus lainnya, yakni jika kita menggunakan kriteria eksternal, yaitu instrumen atau ukuran lain di luar instrumen yang dibuat yang dijadikan kriteria maka keputusan pengujiannya adalah mengenai valid atau tidaknya perangkat instrumen sebagai suatu kesatuan.
13. Untuk kriteria internal atau validitas internal, berdasarkan hasil analisis butir maka butir-butir yang tidak valid dikeluarkan atau diperbaiki untuk diujicoba ulang, sedang butir-butir yang valid dirakit kembali menjadi sebuah perangkat instrumen untuk melihat kembali validitas kontennya berdasarkan kisi-kisi. Jika secara konten butir-butir yang valid tersebut dianggap valid atau memenuhi syarat, maka perangkat instrumen yang terakhir ini menjadi instrumen final yang akan digunakan untuk mengukur variabel penelitian kita.

14. Selanjutnya dihitung koefisien reliabilitas. Koefisien reliabilitas dengan rentangan nilai (0-1) adalah besaran yang menunjukkan kualitas atau konsistensi hasil ukur instrumen. Makin tinggi koefisien reliabilitas makin tinggi pula kualitas instrumen tersebut. Mengenai batas nilai koefisien reliabilitas yang dianggap layak tergantung pada presisi yang dikehendaki oleh suatu penelitian. Untuk itu kita dapat merujuk pendapat-pendapat yang sudah ada, karena secara eksak tidak ada tabel atau distribusi statistik mengenai angka reliabilitas yang dapat dijadikan rujukan.
15. Perakitan butir-butir instrumen yang valid untuk dijadikan instrumen final.

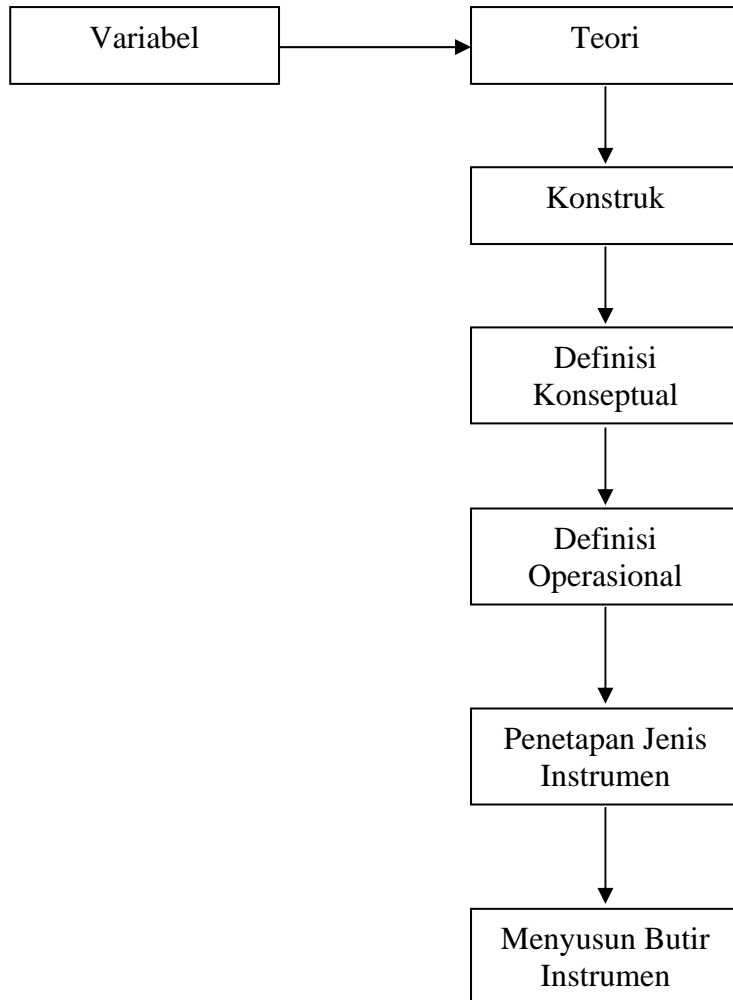
Alur tahapan penyusunan dan pengembangan instrumen dapat dilihat pada Gambar 1 berikut.

Dari bagan tersebut terlihat bahwa untuk keperluan penyusunan dan pengembangan instrumen pertama-tama adalah penetapan konstruk variabel penelitian yang merupakan sintesis dari teori-teori yang telah dibahas dan dianalisis yang penyajiannya diuraikan dalam pengkajian teoretik atau tinjauan pustaka. Konstruk tersebut dijelaskan dalam definisi konseptual variabel, yang di dalamnya tercakup dimensi dan indikator dari variabel yang hendak diukur. Berdasarkan konstruk tersebut ditetapkan indikator-indikator yang akan diukur dari variabel tersebut.

Selanjutnya item-item instrumen dibuat untuk mengukur indikator-indikator yang telah ditetapkan dengan cara seperti telah dikemukakan pada proses penyusunan dan pengembangan instrumen point 4 dan 5. Karena bentuk item-item instrumen yang akan dibuat harus sesuai dengan instrumen yang dipilih, maka sebelum menulis item-item instrumen terlebih dahulu peneliti harus memilih jenis instrumen apa yang sesuai untuk mengukur indikator dari variabel yang akan diteliti.

Ada beberapa jenis instrumen yang biasa digunakan dalam penelitian, antara lain kuesioner, skala (skala sikap atau skala penilaian), tes, dan lain-lain.

Kuesioner adalah alat pengumpul data yang berbentuk pertanyaan yang akan diisi atau dijawab oleh responden. Beberapa alasan digunakannya kuesioner adalah : (1) kuesioner terutama dipakai untuk mengukur variabel yang bersifat faktual, (2) untuk memperoleh informasi yang relevan dengan tujuan penelitian, dan (3) untuk memperoleh informasi dengan validitas dan reliabilitas setinggi mungkin.



Gambar 1. Alur Penyusunan dan Pengembangan Instrumen

Skala adalah alat pengumpul data untuk memperoleh gambaran kuantitatif aspek-aspek tertentu dari suatu barang, atau sifat-sifat seseorang dalam bentuk skala yang sifatnya ordinal, misalnya sangat baik, baik, sedang, tidak baik, dan sangat tidak baik; atau sangat setuju, setuju, netral, tidak setuju, sangat tidak setuju; atau sangat sering, sering, kadang-kadang, jarang, dan tidak pernah. Skala dapat berbentuk skala sikap yang biasanya ditujukan untuk mengukur variabel yang bersifat internal psikologis dan diisi oleh responden yang bersangkutan. Selain itu, skala dapat pula berbentuk skala penilaian yakni apabila skala tersebut ditujukan untuk mengukur variabel yang indikator-indikatornya dapat diamati oleh

orang lain, sehingga skala penilaian bukan diberikan kepada unit analisis penelitian (yang bersangkutan) tetapi diberikan atau diisi oleh orang lain yang mempunyai pengetahuan atau pengalaman yang cukup memadai tentang keadaan subyek yang menjadi unit analisis dalam kaitannya dengan variabel yang akan diukur.

Tes adalah prosedur sistematis yang dibuat dalam bentuk tugas-tugas yang distandardisasikan dan diberikan kepada individu atau kelompok untuk dikerjakan, dijawab, atau direspons, baik dalam bentuk tertulis, lisan maupun perbuatan.

Secara khusus untuk keperluan pengukuran dan penyesuaian dengan jenis instrumen, maka variabel-variabel yang akan diukur atau diteliti dibedakan atas dua kelompok yaitu variabel konseptual dan variabel faktual. Variabel konseptual dapat dibedakan lagi atas dua macam, yaitu variabel yang sifatnya konstruk seperti sikap, motivasi, kreativitas, gaya kepemimpinan, konsep diri, kecemasan, dan lain-lain; serta variabel yang sifatnya konten atau bersifat pengetahuan, yaitu berupa penguasaan responden terhadap seperangkat konten atau pengetahuan yang semestinya dikuasai atau diujikan dalam suatu tes atau ujian.

C. Teknik Penyusunan dan Penilaian Butir Instrumen

Secara umum ada beberapa hal yang perlu diperhatikan dalam menulis butir instrumen, baik instrumen dalam bentuk skala sikap, skala penilaian, maupun tes. Hal-hal yang perlu diperhatikan di antaranya :

1. Butir harus langsung mengukur indikator, yaitu penanda konsep yang berupa sesuatu kenyataan atau fakta (*das solen*) seperti keadaan, perasaan, pikiran, kualitas, kesediaan, dan sebagainya.
2. Jawaban terhadap butir instrumen dapat mengindikasikan ukuran indikator apakah keadaan responden berada atau dekat ke kutub positif atau keadaan responden berada atau dekat ke kutub negatif, misalnya jika berada atau dekat ke kutub positif menandakan sikap positif, menandakan motivasi tinggi, menandakan kepemimpinan yang efektif, menandakan intensitas tinggi, menandakan produktivitas tinggi, menandakan gaya kepemimpinan demokratik, menandakan iklim kerja yang kondusif, dan sebagainya. Sedang jika berada atau dekat ke kutub negatif menandakan sikap negatif, menandakan motivasi rendah, menandakan kepemimpinan yang tidak efektif, menandakan intensitas rendah, menandakan produktivitas rendah, menandakan gaya kepemimpinan otoriter, menandakan iklim kerja yang tidak kondusif, dan sebagainya.

3. Butir dapat berbentuk pertanyaan atau pernyataan dengan menggunakan bahasa yang sederhana, jelas, tidak mengandung tafsiran ganda, singkat, dan komunikatif.
4. Opsi dari setiap pertanyaan atau pernyataan itu harus relevan menjawab pertanyaan atau pernyataan tersebut.
5. Banyaknya opsi menunjukkan panjang skala yang secara konseptual kontinum. Karena distribusi jawaban responden secara teoretik mendekati distribusi normal untuk jumlah populasi cukup besar, maka sebaiknya menggunakan skala ganjil.

Penulisan Butir Tes

Tipe Pilihan Ganda

1. Item hendaklah menanyakan hal yang penting untuk diketahui.
2. Tulislah item yang berisi pernyataan pasti.
3. Utamakan item yang mengandung pernyataan umum yang bertahan lama.
4. Buatlah item yang berisi hanya satu gagasan saja.
5. Buatlah item yang menyatakan inti pertanyaan dengan jelas. Gunakan kalimat sederhana dan tidak berlebih-lebihan.
6. Sebaiknya item tidak didasari oleh pernyataan negatif.
7. Gunakan bahasa yang jelas, kata yang sederhana, dan pernyataan yang langsung.
8. Item harus memberikan alternatif bagi isi pernyataan yang paling penting.
9. Berikan alternatif jawaban yang jelas berbeda.
10. Alternatif yang ditawarkan hendaknya mempunyai struktur dan arti yang sejajar atau dalam satu kategori.

11. Penggunaan alternatif yang semata-mata meniadakan atau bertentangan dengan alternatif yang lain, haruslah dihindari.
12. Bilamana mungkin, susunlah alternatif jawaban dalam urutan besarnya atau urutan logisnya.
13. Penggunaan alternatif "bukan salah-satu di atas" atau "semua yang di atas" hanya baik apabila kebenaran bersifat mutlak dan bukan semata-mata masalah lebih dan kurang baik atau masalah kebenaran relatif.
14. Jangan menjebak siswa dengan menanyakan hal yang tidak ada jawabannya.
15. Hindari penggunaan kata-kata yang dapat dijadikan petunjuk oleh siswa dalam menjawab.

Tipe Benar-Salah

Kaidah atau petunjuk penulisan item tipe benar-salah telah dikemukakan oleh Ebel (1979) sebagaimana berikut ini.

1. Item haruslah mengungkap ide atau gagasan yang penting.
2. Item tipe benar-salah hendaknya menguji pemahaman, jangan hanya mengungkap ingatan mengenai suatu fakta atau hafalan.
3. Kebenaran atau ketidakbenaran suatu item haruslah bersifat mutlak.
4. Item harus menguji pengetahuan yang spesifik dan jawabannya tidak jelas bagi semua orang, kecuali bagi mereka yang menguasai pelajaran.
5. Item harus dinyatakan secara jelas.

Tipe Jawaban Pendek

1. Pernyataan atau pertanyaan item harus ditulis dengan hati-hati sehingga dapat dijawab dengan hanya satu jawaban yang pasti.
2. Sebaiknya rumuskan jawabannya lebih dahulu baru kemudian menulis pertanyaannya.

3. Gunakan pertanyaan langsung, kecuali bilamana model kalimat tak selesai akan memungkinkan jawaban yang lebih jelas.
4. Usahakan agar dalam pertanyaan tidak terdapat petunjuk yang mungkin digunakan oleh subjek dalam menjawab item.
5. Jangan menggunakan kata atau kalimat yang langsung dikutip dari buku.

Tipe Pasangan

1. Premis dan respons hendaknya dibuat dalam jumlah yang tidak sama.
2. Baik premis maupun respons haruslah berisi hal yang homogen, yaitu dari sejenis kategori isi.
3. Usahakan agar premis dan responsnya berisi kalimat-kalimat atau kata yang pendek.
4. Buatlah petunjuk pemasangan yang jelas, sehingga penjawab soal atau pertanyaan mengetahui dasar apakah yang harus digunakan dalam memasangkan premis dan responsnya.
5. Sedapat mungkin susunlah premis dan respons masing-masing secara alfabetik atau menurut besaran kuantitatifnya.

Tipe Karangan (Esai)

1. Berikan pertanyaan atau tugas yang mengarahkan penjawab pertanyaan (siswa) agar dapat menunjukkan penguasaan pengetahuan yang penting.
2. Buatlah pertanyaan yang arah jawabannya jelas, sehingga para ahli dapat setuju bahwa satu jawaban akan lebih baik daripada yang lainnya.
3. Jangan menanyakan sikap atau pendapat.
4. Sebaiknya pertanyaan diawali oleh kata-kata seperti, "Bandingkan ...", "Berikan alasan ...", "Jelaskan mengapa ...", "Beri contoh ...", dan sebagainya.

5. Jangan memberi kesempatan kepada penjawab soal untuk memilih dan menjawab hanya sebagian di antara nomor pertanyaan yang disediakan.
6. Sebaiknya, tulis lebih dahulu satu jawaban ideal yang dikehendaki, baru kemudian menyusun pertanyaannya.

Penulisan Butir untuk Skala Sikap Model Likert

Untuk menulis pernyataan sikap yang bermutu, penyusun skala harus menuruti suatu kaidah atau pedoman penulisan pernyataan agar ciri-ciri pernyataan sikap tidak terlupakan dan agar setiap pernyataan mempunyai kemampuan membedakan antara kelompok responden yang setuju dengan kelompok responden yang tidak setuju terhadap objek sikap.

Beberapa petunjuk untuk menyusun skala Likert di antaranya :

1. Tentukan objek yang dituju, kemudian tetapkan variabel yang akan diukur dengan skala tersebut.
2. Lakukan analisis variabel tersebut menjadi beberapa sub variabel atau dimensi variabel, lalu kembangkan indikator setiap dimensi tersebut.
3. Dari setiap indikator di atas, tentukan ruang lingkup pernyataan sikap yang berkenaan dengan aspek kognisi, afeksi, dan konasi terhadap objek sikap.
4. Susunlah pernyataan untuk masing-masing aspek tersebut dalam dua kategori, yakni pernyataan positif dan pernyataan negatif, secara seimbang banyaknya.

Sementara itu Edwards (1957) telah meramu berbagai saran dan petunjuk dari para ahli menjadi suatu pedoman atau kriteria penulisan pernyataan sikap. Beberapa kriteria yang dimaksud adalah sebagai berikut.

1. Jangan menulis pernyataan yang membicarakan mengenai kejadian yang telah lewat kecuali kalau objek sikapnya berkaitan dengan masa lalu.
2. Jangan menulis pernyataan yang berupa fakta atau dapat ditafsirkan sebagai fakta.
3. Jangan menulis pernyataan yang dapat menimbulkan lebih dari satu penafsiran.

4. Jangan menulis pernyataan yang tidak relevan dengan objek psikologisnya.
5. Jangan menulis pernyataan yang sangat besar kemungkinannya akan disetujui oleh hampir semua orang atau bahkan hampir tak seorang pun yang akan menyetujuinya.
6. Pilihlah pernyataan-pernyataan yang diperkirakan akan mencakup keseluruhan liputan skala afektif yang diinginkan.
7. Usahakan agar setiap pernyataan ditulis dalam bahasa yang sederhana, jelas, dan langsung. Jangan menuliskan pernyataan dengan menggunakan kalimat-kalimat yang rumit.
8. Setiap pernyataan hendaknya ditulis ringkas dengan menghindari kata-kata yang tidak diperlukan dan yang tidak akan memperjelas isi pernyataan.
9. Setiap pernyataan harus berisi hanya satu ide (gagasan) yang lengkap.
10. Pernyataan yang berisi unsur universal seperti "tidak pernah", "semuanya", "selalu", "tak seorang pun", dan semacamnya, seringkali menimbulkan penafsiran yang berbeda-beda dan karenanya sedapat mungkin hendaklah dihindari.
11. Kata-kata seperti "hanya", "sekedar", "semata-mata", dan semacamnya harus digunakan seperlunya untuk menghindari kesalahan penafsiran isi pernyataan.
12. Jangan menggunakan kata atau istilah yang mungkin tidak dapat dimengerti oleh para responden.
13. Hindarilah pernyataan yang berisi kata negatif ganda.

Penulisan Butir untuk Skala Penilaian

Pada prinsipnya, penyusunan butir untuk skala penilaian hampir sama dengan penyusunan butir untuk skala sikap. Perbedaannya terletak pada konteks pernyataan, yaitu untuk skala sikap mengenai keadaan atau perasaan atau penilaian yang bersangkutan terhadap obyek sikap sedang skala penilaian mengenai keadaan, kemampuan, penampilan, atau kinerja orang lain berdasarkan penilaian orang yang

mengisi skala penilaian tersebut. Selanjutnya, seperti halnya juga instrumen yang lain, penyusunan skala penilaian hendaknya memperhatikan hal-hal berikut ini.

1. Tentukan tujuan yang akan dicapai dari skala penilaian tersebut sehingga jelas apa yang seharusnya dinilai.
2. Berdasarkan tujuan tersebut, tentukan aspek atau variabel yang akan diungkap melalui instrumen ini.
3. Tetapkan bentuk rentangan nilai yang akan digunakan, misalnya nilai angka atau kategori.
4. Buatlah item-item pernyataan yang akan dinilai dalam kalimat yang singkat tetapi bermakna secara logis dan sistematis.
5. Ada baiknya menetapkan pedoman mengolah dan menafsirkan hasil yang diperoleh dari penilaian tersebut.

Skala penilaian dalam pelaksanaannya dapat digunakan oleh dua orang penilai atau lebih, dalam menilai subjek yang sama. Maksudnya adalah agar diperoleh hasil penilaian yang objektif mengenai perilaku subjek yang dinilai.

Penulisan Butir untuk Kuesioner

Cara menyusun kuesioner beserta butir-butir yang tercantum di dalamnya haruslah tetap mengacu pada pedoman penyusunan instrumen secara umum, sehingga berlaku pula langkah-langkah sebagaimana telah dijelaskan di bagian terdahulu. Dimulai dengan analisis variabel, pembuatan kisi-kisi, dan kemudian sampai pada penyusunan pertanyaan untuk kuesioner.

Secara lebih teknis, petunjuk untuk membuat kuesioner adalah sebagai berikut.

1. Mulai dengan pengantar yang isinya berupa permohonan mengisi kuesioner sambil menjelaskan maksud dan tujuannya.
2. Jelaskan petunjuk atau cara mengisinya supaya tidak salah. Kalau perlu, berikan contoh pengisiannya.

3. Mulai dengan pertanyaan untuk mengungkapkan identitas responden. Dalam identitas ini sebaiknya tidak diminta mengisi nama. Identitas cukup mengungkapkan jenis kelamin, usia, pendidikan, pekerjaan, pengalaman, dan lain-lain yang ada kaitannya dengan tujuan kuesioner.
4. Isi pertanyaan sebaiknya dibuat beberapa kategori atau bagian sesuai dengan variabel yang diungkapkan, sehingga mudah mengolahnya.
5. Rumusan pertanyaan dibuat singkat, tetapi jelas sehingga tidak membingungkan dan menimbulkan salah penafsiran.
6. Hubungan antara pertanyaan yang satu dengan pertanyaan lainnya harus dijaga sehingga tampak keterkaitan logikanya dalam satu rangkaian yang sistematis. Hindari penggolongan pertanyaan terhadap indikator atau persoalan yang sama.
7. Usahakan agar jawaban, yakni kalimat atau rumusannya tidak lebih panjang daripada pertanyaan.
8. Kuesioner yang terlalu banyak atau terlalu panjang akan melelahkan dan membosankan responden sehingga pengisiannya tidak objektif lagi.
9. Ada baiknya kuesioner diakhiri dengan tanda tangan si pengisi untuk menjamin keabsahan jawabannya.
10. Untuk melihat validitas jawaban kuesioner, ada baiknya kuesioner diberikan kepada beberapa responden secara acak dan dilakukan wawancara dengan pertanyaan yang identik dengan isi kuesioner yang telah diisinya.

D. Proses Validasi Konsep Melalui Panel

1. Memeriksa instrumen mulai dari konstruk sampai penyusunan butir

Dalam kaitan ini, beberapa hal yang perlu diperhatikan antara lain :

- ◆ Apakah dimensi yang dirumuskan sudah merupakan jabaran yang tepat dari konstruk yang telah dirumuskan dan sesuai untuk mengukur konstruk dari variabel yang hendak diukur ?

- ◆ Apakah indikator yang dirumuskan sudah merupakan jabaran yang tepat dari dimensi yang telah dirumuskan dan sesuai untuk mengukur konstruk dari variabel yang hendak diukur ?
- ◆ Apakah butir-butir instrumen yang dibuat telah sesuai untuk mengukur indikator-indikator dari variabel yang hendak diukur ?

2. Menilai butir

Butir yang sudah dibuat diberikan kepada sekelompok panel untuk dinilai dengan tetap mengacu pada tolok ukur di atas. Metode penilaian butir dapat dilakukan dengan beberapa cara, misalnya dengan Metode Thurstone dan Pair Comparison.

Teknik penilaian butir dengan Metode Thurstone

I. Langkah-langkah

1. Membuat sejumlah pernyataan (sekitar 40 -50 butir) yang relevan untuk variabel yang hendak diukur.
2. Membentuk panel yang terdiri dari sejumlah ahli (20-40 orang) untuk menilai relevansi pernyataan (item) yang telah dibuat.

Tentukan skala penilaian (1-11) atau (1-13). Jika dipakai skala penilaian (1-11), maka skala tersebut menunjukkan bahwa :

Skala 1 → untuk pernyataan yang sangat tidak relevan

Skala 11 → untuk pernyataan yang sangat relevan

Misalnya :

Satu pernyataan dinilai oleh 20 orang ahli sebagai panel.

Skala	1	2	3	4	5	6	7	8	9	10	11
<i>f</i> Penilai	0	0	0	0	1	3	7	5	3	1	0

Kemudian diubah menjadi frekuensi kumulatif (*fk*) sebagaimana berikut ini :

Skala	1	2	3	4	5	6	7	8	9	10	11
<i>fk</i>	0	0	0	0	1	4	11	16	19	20	20

Selanjutnya dihitung *Median* (*Md*) dan *Kuartil* (*Q*) dari data tersebut sehingga diperoleh :

$$Q_1 = 6,2 \quad Q_2 = Md = 6,9 \quad Q_3 = 7,8$$

Dari hasil tersebut kemudian dapat diinter-pretasikan bahwa :

- ◆ Semakin tinggi nilai *Md*, berarti pernyataan semakin baik atau semakin relevan dengan konstruk variabel yang hendak diukur.
 - ◆ Semakin kecil nilai $Q - (Q_3 - Q_1)$, berarti semakin kuat persetujuan Panel.
3. Menyusun atau menentukan jenjang item atau pernyataan dengan berdasarkan pada nilai *Md* hasil panel.
 4. Menentukan skor reponden dengan cara melakukan pembobotan atas dasar nilai *Md*.

Misalnya :

Ada 5 item yang dipilih untuk mengukur sikap terhadap konsep A. Jika skala item (1-11) dan hasil serta bobotnya (*Md*) seperti tabel berikut, maka skor responden adalah :

<i>No.</i>	<i>Hasil</i>	<i>Bobot</i>	<i>Skor</i>
1	8	9,0	72
2	5	8,5	42,5
3	9	8	72
4	10	7	70
5	5	6,5	32,5

Skor responden = 289
 Skor netral = 234
 Atau :
 Skor responden = 7,4
 Skor netral = 6,0

II. Contoh

Hasil panel terhadap 10 butir skala adalah sebagai berikut :

<i>No. Butir</i>	<i>Median</i>	<i>Q₃ - Q₁</i>
1	6,9	1,6
2	9,5	1,2
3	5,2	1,4
4	6,5	1,5
5	10,2	1,2
6	7,4	2,4
7	9,2	1,5
8	8,5	1,4
9	5,0	1,5
10	7,8	1,7

Data hasil penelitian :

<i>Skor butir</i>	<i>Bobot</i>	<i>Skor</i>
1	6,9	6,9
4	9,5	38,0
3	5,2	15,6
4	6,5	26,0
5	10,2	51,0
3	7,4	22,2
4	9,2	36,8
5	8,5	42,5
2	5,0	10,0
3	7,8	23,4
34	76,2	272,4

Dikoreksi :
Skor total = 272,4
Rata-rata = 3,575

Tidak dikoreksi :
Skor total = 259,08
Rata-rata = 3,4.

Teknik penilaian butir dengan Pair Comparison

Pair comparison (perbandingan berpasangan) merupakan salah satu teknik penilaian butir yang bertujuan untuk mengukur sikap kelompok terhadap beberapa butir yang kemungkinan menjadi pilihan. Metode *pair comparison* dapat juga digunakan untuk menentukan bobot relevansi berdasarkan pendapat sekelompok orang.

Caranya adalah semua butir dipasangkan dua-dua, kemudian sekelompok orang diminta menentukan manakah butir-butir yang lebih baik.

Jumlah pasangan yang dapat disusun bisa dirumuskan sebagai berikut :

$$\frac{n(n-1)}{2}$$

dimana : n = banyaknya butir (obyek /pilihan).

Langkah-langkah :

1. Membuat pertanyaan sebanyak pasangan butir yang kemungkinan dapat disusun.
Misalkan jika terdapat butir 1, 2, 3 dan 4; maka dapat disusun pertanyaan sebanyak 6 buah, yaitu :
 - ♦ Manakah yang lebih baik antara butir 1 dan butir 2 ?
 - ♦ Manakah yang lebih baik antara butir 1 dan butir 3 ?
 - ♦ Manakah yang lebih baik antara butir 1 dan butir 4 ?
 - ♦ Manakah yang lebih baik antara butir 2 dan butir 3 ?
 - ♦ Manakah yang lebih baik antara butir 2 dan butir 4 ?
 - ♦ Manakah yang lebih baik antara butir 3 dan butir 4 ?

2. Pertanyaan diberikan kepada responden (sekelompok orang) untuk diisi (misalnya 120 orang)
3. Hasilnya dirangkum dalam matrik frekuensi, sebagai berikut :

Butir	1	2	3	4
1	-	72	78	96
2	48	-	54	84
3	42	66	-	66
4	24	36	54	-

4. Matriks frekuensi ditransformasi menjadi matriks proporsi seperti berikut ini

Butir	1	2	3	4
1	0,50	0,60	0,65	0,80
2	0,40	0,50	0,45	0,70
3	0,35	0,55	0,50	0,55
4	0,20	0,30	0,45	0,50

5. Matriks proporsi ditransformasi menjadi matriks Z (lihat Tabel Z)

Butir	1	2	3	4
1	0,000	0,253	0,385	0,842
2	-0,253	0,000	-0,126	0,524
3	-0,385	0,126	0,000	0,126
4	-0,842	-0,524	-0,126	0,000

6. Selanjutnya dari matriks Z dihitung jumlah menurut kolom, rata-rata, dan bobotnya.

	1	2	3	4
Jumlah	-1,480	-0,145	0,133	1,492
Rata-rata	-0,370	-0,036	0,033	0,373
Penyesuaian	0,000	0,334	0,403	0,743
Bobot	0	334	403	743

Jika akan ditentukan dua butir terbaik dari hasil perhitungan tersebut, maka yang terpilih adalah butir 4 dan 3 karena masing-masing memiliki bobot yang tertinggi yaitu 743 dan 403.

E. Proses Validasi Empirik Melalui Ujicoba

Setelah suatu instrumen dianggap valid secara konseptual maka langkah berikutnya adalah instrumen tersebut diujicobakan pada sekelompok responden yang merupakan sampel ujicoba. Dari Jawaban atau respon dari sampel ujicoba tersebut diperoleh data yang akan dianalisis untuk menguji validitas instrumen dengan menggunakan validitas internal.

Analisis data hasil ujicoba atau analisis butir pada pokoknya dimaksudkan untuk menguji validitas butir-butir instrumen atau soal-soal tes secara empiris atau berdasarkan data empiris yang diperoleh dari ujicoba. Dalam pembahasan ini validitas yang akan diuji adalah validitas butir atau validitas soal dengan menggunakan kriteria internal, yaitu skor total tes. Skor total instrumen atau tes dapat ditetapkan sebagai kriteria untuk menentukan validitas butir instrumen atau soal tes karena secara konsep atau konten perangkat instrumen atau tes yang telah dibuat dan diujicobakan sudah dianggap valid.

Pengujian validitas butir instrumen atau soal tes dilakukan dengan menghitung koefisien korelasi antara skor butir instrumen atau soal tes dengan skor total instrumen atau tes. Butir atau soal yang dianggap valid adalah butir instrumen atau soal tes yang skornya mempunyai koefisien korelasi yang signifikan dengan skor total instrumen atau tes.

Jika skor butir instrumen atau soal tes kontinum (misalnya skala sikap atau soal bentuk uraian dengan skor butir 1- 5 atau skor soal 0 - 10) dan diberi simbol

X_i dan skor total instrumen atau tes diberi simbol X_t , maka rumus yang digunakan untuk menghitung koefisien korelasi antara skor butir instrumen atau soal dengan skor total instrumen atau skor total tes adalah :

$$r_{it} = \frac{X_i X_t}{x_i^2 x_t^2}$$

Keterangan :

- r_{it} = koefisien korelasi antara skor butir soal dengan skor total
- x_i = jumlah kuadrat deviasi skor dari X_i
- x_t = jumlah kuadrat deviasi skor dari X_t

Contoh Perhitungan Korelasi Butir untuk Soal Bentuk Uraian dengan Skor Butir Kontinum

Data hasil uji coba adalah sebagai berikut :

Butir

Nomor	1	2	3	4	5	6	7	Jumlah
1	5	4	3	5	3	5	3	28
2	5	4	3	4	3	4	3	26
3	4	4	2	4	3	4	3	24
4	4	3	3	3	4	3	4	24
5	5	5	3	4	5	5	4	31
6	3	3	2	3	2	3	1	17
7	3	3	2	3	2	2	2	17
8	3	2	2	3	2	2	2	16
9	2	2	1	2	1	2	1	11
10	2	1	1	1	1	1	1	8
Jumlah	36	31	22	32	26	31	24	202

Penyelesaian :

$$\sum X_t = 202$$

$$\sum X_t^2 = 4592$$

$$\sum x_i^2 = 511,6$$

$$1. \sum X_1 = 36$$

$$\sum X_1^2 = 142$$

$$\sum x_1^2 = 142 - \frac{36^2}{10} = 12,4$$

$$\sum X_1 X_t = 805$$

$$\sum x_1 x_t = 805 - \frac{202 \times 36}{10} = 77,8$$

$$r_{1t} = \frac{77,8}{\sqrt{511,6 \times 12,4}} = 0,98$$

$$2. \sum X_2 = 31$$

$$\sum X_2^2 = 109$$

$$\sum x_2^2 = 109 - \frac{31^2}{10} = 12,9$$

$$\sum X_2 X_t = 703$$

$$\sum x_2 x_t = 703 - \frac{202 \times 31}{10} = 76,8$$

$$r_{2t} = \frac{76,8}{\sqrt{511,6 \times 12,9}} = 0,95$$

$$3. \sum X_3 = 22$$

$$\sum X_3^2 = 54$$

$$\sum x_3^2 = 54 - \frac{22^2}{10} = 5,6$$

$$\sum X_3 X_t = 494$$

$$\sum x_3 x_t = 494 - \frac{202 \times 22}{10} = 49,6$$

$$r_{3t} = \frac{49,6}{\sqrt{511,6 \times 5,6}} = 0,93$$

$$4. \sum X_4 = 32$$

$$\sum X_4^2 = 114$$

$$\sum x_4^2 = 114 - \frac{32^2}{10} = 11,6$$

$$\sum X_4 X_t = 716$$

$$\sum x_4 x_t = 716 - \frac{202 \times 32}{10} = 69,6$$

$$r_{4t} = \frac{69,6}{\sqrt{511,6 \times 11,6}} = 0,90$$

$$5. \sum X_5 = 26$$

$$\sum X_5^2 = 82$$

$$\sum X_5 X_t = 604$$

$$\sum x_5 x_t = 604 - \frac{202 \times 26}{10} = 78,8$$

$$\sum x_5^2 = 82 - \frac{26^2}{10} = 14,4$$

$$r_{5t} = \frac{78,8}{\sqrt{511,6 \times 14,4}} = 0,92$$

$$6. \sum X_6 = 31$$

$$\sum X_6 X_t = 714$$

$$\sum X_6^2 = 113$$

$$\sum x_6 x_t = 714 - \frac{202 \times 31}{10} = 87,8$$

$$r_{6t} = \frac{87,8}{\sqrt{511,6 \times 16,9}} = 0,94$$

$$7. \sum X_7 = 24$$

$$\sum X_7 X_t = 556$$

$$\sum X_7^2 = 70$$

$$\sum x_7 x_t = 556 - \frac{202 \times 24}{10} = 71,2$$

$$\sum x_7^2 = 70 - \frac{24^2}{10} = 12,4$$

$$r_{7t} = \frac{71,2}{\sqrt{511,6 \times 12,4}} = 0,89$$

Untuk $n = 10$ dengan α sebesar $0,05$ didapat nilai tabel $r = 0,631$. Karena nilai koefisien korelasi antara skor butir dengan skor total untuk semua butir lebih besar dari $0,631$, maka semua butir mempunyai korelasi signifikan dengan skor total tes. Dengan demikian maka semua butir tes dianggap valid atau dapat digunakan untuk mengukur hasil belajar. Selanjutnya akan dihitung koefisien reliabilitas dengan menggunakan rumus koefisien Alpha, yaitu :

$$r_{ii} = \frac{k}{k-1} \left(1 - \frac{\sum s_i^2}{s_t^2} \right)$$

Keterangan :

r_{ii} = koefisien reliabilitas tes

k = cacah butir

s_i^2 = varians skor butir

s_t^2 = varians skor total

Koefisien reliabilitas dari contoh di atas adalah :

Pertama-tama dihitung varians butir sebagai berikut:

Nomor butir	Varians butir
1	1,24
2	1,29
3	0,56
4	1,16
5	1,44
6	1,69
7	1,24
Jumlah	8,62

$$\sum s_i^2 = 8,62$$

$$s_t^2 = 51,16$$

$$r_{ii} = \frac{7}{6} \left(1 - \frac{8,62}{51,16} \right) = 0,97$$

Jadi koefisien reliabilitas tes (dengan 7 butir) pada contoh di atas adalah 0,97.

Jika skor butir soal dis-kontinum (misalnya soal bentuk obyektif dengan skor butir soal 0 atau 1) maka kita menggunakan koefisien korelasi biserial dan rumus yang digunakan untuk menghitung koefisien korelasi biserial antara skor butir soal dengan skor total tes adalah :

$$r_{bis(i)} = \frac{X_i - X_t}{S_t} \sqrt{\frac{p_i}{q_i}}$$

Keterangan :

- $r_{bis(i)}$ = koefisien korelasi biserial antara skor butir soal nomor i dengan skor total
 X_i = rata-rata skor total responden yang menjawab benar butir soal nomor i
 X_t = rata-rata skor total semua responden
 S_t = standar deviasi skor total semua responden
 p_i = proporsi jawaban yang benar untuk butir soal nomor i
 q_i = proporsi jawaban yang salah untuk butir soal nomor i

Contoh Perhitungan Korelasi Butir untuk Soal Bentuk Obyektif

Data hasil uji coba adalah sebagai berikut :

Butir

Nomor	1	2	3	4	5	6	7	Jumlah
1	1	1	1	1	0	0	0	4
2	1	1	0	1	1	1	0	5
3	0	1	1	1	0	0	0	3
4	1	1	0	0	0	0	0	2
5	0	1	0	0	0	0	0	1
6	1	1	1	1	1	1	1	7
7	1	1	1	1	1	1	0	6
8	0	0	0	0	0	0	0	0
9	1	1	0	0	1	0	0	3
10	1	1	1	1	1	0	0	5
Jumlah	7	9	5	6	5	3	1	36

$$X_t = 3,60$$

$$S_t = 2,107$$

$$1. \bar{X}_1 = \frac{32}{7} = 4,57 \rightarrow r_{bis1} = \frac{4,57 - 3,6}{2,107} \sqrt{\frac{0,7}{0,3}} = 0,70$$

$$2. \bar{X}_2 = \frac{36}{9} = 4 \rightarrow r_{bis2} = \frac{4 - 3,6}{2,107} \sqrt{\frac{0,9}{0,1}} = 0,57$$

$$3. \bar{X}_3 = \frac{25}{5} = 5 \rightarrow r_{bis3} = \frac{5 - 3,6}{2,107} \sqrt{\frac{0,5}{0,5}} = 0,66$$

$$4. \bar{X}_4 = \frac{30}{6} = 5 \rightarrow r_{bis4} = \frac{5 - 3,6}{2,107} \sqrt{\frac{0,6}{0,4}} = 0,81$$

$$5. \bar{X}_5 = \frac{26}{5} = 5,2 \rightarrow r_{bis5} = \frac{5,2 - 3,6}{2,107} \sqrt{\frac{0,5}{0,5}} = 0,76$$

$$6. \bar{X}_6 = \frac{18}{3} = 6 \rightarrow r_{bis6} = \frac{6 - 3,6}{2,107} \sqrt{\frac{0,3}{0,7}} = 0,75$$

$$7. \bar{X}_7 = \frac{7}{1} = 7 \rightarrow r_{bis7} = \frac{7 - 3,6}{2,107} \sqrt{\frac{0,1}{0,9}} = 0,54$$

Rekapitulasi koefisien korelasi dan status butir :

Nomor butir	r-butir	r-tabel	Status
1	0,70	0,63	Valid
2	0,57	0,63	Tidak valid
3	0,66	0,63	Valid
4	0,81	0,63	Valid
5	0,76	0,63	Valid
6	0,75	0,63	Valid
7	0,54	0,63	Tidak valid

Ternyata dari tujuh butir soal tes ada 5 butir yang valid dan dua butir tidak valid. Oleh karena itu perlu dilakukan perhitungan tahap kedua untuk menghitung koefisien korelasi antara skor butir dengan skor total baru (5 butir), sebagai berikut.

Data hasil uji coba adalah sebagai berikut :

Butir

Nomor	1	3	4	5	6	Jumlah
1	1	1	1	0	0	3
2	1	0	1	1	1	4
3	0	1	1	0	0	2
4	1	0	0	0	0	1
5	0	0	0	0	0	0
6	1	1	1	1	1	5
7	1	1	1	1	1	5
8	0	0	0	0	0	0
9	1	0	0	1	0	2
10	1	1	1	1	0	4
Jumlah	7	5	6	5	3	26

$$X_t = 2,6$$

$$S_t = 1,8$$

$$1. \bar{X}_1 = \frac{24}{7} = 3,43 \rightarrow r_{bis(1)} = \frac{3,43 - 2,6}{1,8} \sqrt{\frac{0,7}{0,3}} = 0,70$$

$$2. \bar{X}_3 = \frac{19}{5} = 3,8 \rightarrow r_{bis(3)} = \frac{3,8 - 2,6}{1,8} \sqrt{\frac{0,5}{0,5}} = 0,67$$

$$3. \bar{X}_4 = \frac{23}{6} = 3,83 \rightarrow r_{bis(4)} = \frac{3,83 - 2,6}{1,8} \sqrt{\frac{0,6}{0,4}} = 0,84$$

$$4. \bar{X}_5 = \frac{20}{5} = 4 \rightarrow r_{bis(5)} = \frac{4 - 2,6}{1,8} \sqrt{\frac{0,5}{0,5}} = 0,78$$

$$5. \bar{X}_6 = \frac{14}{3} = 4,67 \rightarrow r_{bis(6)} = \frac{4,67 - 2,6}{1,8} \sqrt{\frac{0,3}{0,7}} = 0,75$$

Untuk $n = 10$ dengan alpha sebesar 0,05 didapat nilai tabel $r = 0,631$. Karena nilai koefisien korelasi biserial antara skor butir dengan skor total untuk semua butir lebih besar dari 0,631, maka semua butir mempunyai korelasi biserial yang signifikan dengan skor total tes. Dengan demikian maka semua butir tes (5 butir) dianggap valid atau dapat digunakan untuk mengukur hasil belajar. Selanjutnya akan dihitung koefisien reliabilitas dengan menggunakan rumus KR-20, sebagai berikut :

$$r_{ii} = \frac{k}{k-1} \left(1 - \frac{\sum p_i q_i}{s_t^2} \right)$$

Keterangan :

r_{ii} = koefisien reliabilitas tes

k = cacah butir

$p_i q_i$ = varians skor butir

p_i = proporsi jawaban yang benar untuk butir nomor i

q_i = proporsi jawaban yang salah untuk butir nomor i

s_t^2 = varians skor total

Koefisien reliabilitas dari contoh di atas adalah :

Pertama-tama dihitung varians butir ($p_i q_i$) sebagai berikut:

Nomor butir	p_i	q_i	$p_i q_i$
1	0,7	0,3	0,21
3	0,5	0,5	0,25
4	0,6	0,4	0,24
5	0,5	0,5	0,25
6	0,3	0,7	0,21
Jumlah			1,16

$$\sum p_i q_i = 1,16$$

$$s_t^2 = 3,24$$

$$r_{ii} = \frac{5}{4} \left(1 - \frac{1,16}{3,24} \right) = 0,80$$

Jadi koefisien reliabilitas tes (dengan 5 butir) pada contoh di atas adalah 0,80.