BAB 5
PERANCANGAN KUESTIONER

5.1 PENDAHULUAN

Banyak penelitian sistem informasi yang menggunakan teknik survei untuk mengumpulkan datanya. Untuk pengumpulan data lewat survei, kuestioner (daftar pertanyaan) harus dipersiapkan terlebih dahulu. Oleh karena itu, bab ini khusus membahas bagaimana merancang suatu kuestioner untuk survei.

Membuat kuestioner bukan hal yang mudah. Jika belum ada kuestioner sebelumnya, maka peneliti tidak mempunyai pilihan lain kecuali harus membuat kuestioner sendiri. Bab ini akan memberi contoh bagaimana Moore dan Benbasat (1991) dan juga Davis (1989) membangun kuestioner mereka. Jika kuestioner sudah ada dan sudah digunakan di penelitian-penelitian lain, peneliti tidak perlu membangun kuestionernya sendiri. Peneliti dapat menggunakan kuestioner yang sudah teruji dari penelitian-penelitian sebelumnya. Bab ini juga akan memberikan contoh bagaimana menggunakan kuestioner dari penelitian-penelitian sebelumnya.

Item-item kuestioner dibangun untuk membentuk suatu konstruk. Oleh sebab itu, sebelum bab ini membahas bagaimana merancang suatu kuestioner, bab ini akan membahas terlebih dahulu mengenai konstruk dan alat ukurnya, yaitu skala.

5.2. KONSTRUK

Variabel laten atau **konstruk** (***construct***) adalah variabel yang masih belum dapat diukur.

Variabel dapat berupa elemen-elemen data yang sudah dapat diukur besarannya, tetapi juga kadang-kadang masih berupa variabel yang belum dapat diukur secara langsung. Variabel yang masih belum dapat diukur secara langsung ini disebut dengan **variabel laten** atau disebut juga dengan istilah **konstruk** (***construct***). Kuestioner diperlukan untuk digunakan membentuk konstruk-konstruk. Kuestioner biasanya berisi dengan banyak item yang dijadikan sebagai pertanyaan-pertanyaan kuestioner. Untuk dapat diukur, maka nantinya item-item pertanyaan tersebut akan diukur dengan skala pengukuran (lihat 5.4). Kumpulan dari beberapa item pertanyaan membentuk suatu konstruk.

Contoh:

Penelitian Chau dan Hu (2002) menggunakan konstruk-konstruk sebagai berikut ini.

1. Konstruk kegunaan persepsian (*perceived usefulness*) diadaptasi dari penelitian Davis (1989).
2. Konstruk kemudahan penggunaan persepsian (*perceived ease of use*) diadaptasi dari penelitian Davis (1989).
3. Konstruk kompatibilitas (*compability*) diadaptasi dari penelitian Taylor dan Todd (1995).
4. Konstruk pengaruh kolega (*peer influence*) diadaptasi dari penelitian Taylor dan Todd (1995).
5. Konstruk kontrol teknologi persepsian (*perceived technology control*) diadaptasi dari penelitian Taylor dan Todd (1995).
6. Konstruk sikap (*attitude*) diadaptasi dari penelitian Taylor dan Todd (1995).
7. Konstruk niat perilaku (*behavioral intention*) diadaptasi dari penelitian Davis (1998), dan Taylor dan Todd (1995).

5.3. ELEMEN-ELEMEN KUESTIONER

Jika masih berupa suatu konstruk (*construct*) yang belum dapat diukur secara langsung, maka konstruk ini harus diidefinisikan secara naratif terlebih dahulu. Definisi naratif dari konstruk (*construct*) disebut dengan konsep (*concept*). Untuk dapat didefinisikan secara operational, maka konstruk (*construct*) ini harus diuraikan sedemikian rupa menjadi dimensi-dimensi dan elemen-elemen data yang dapat diukur.

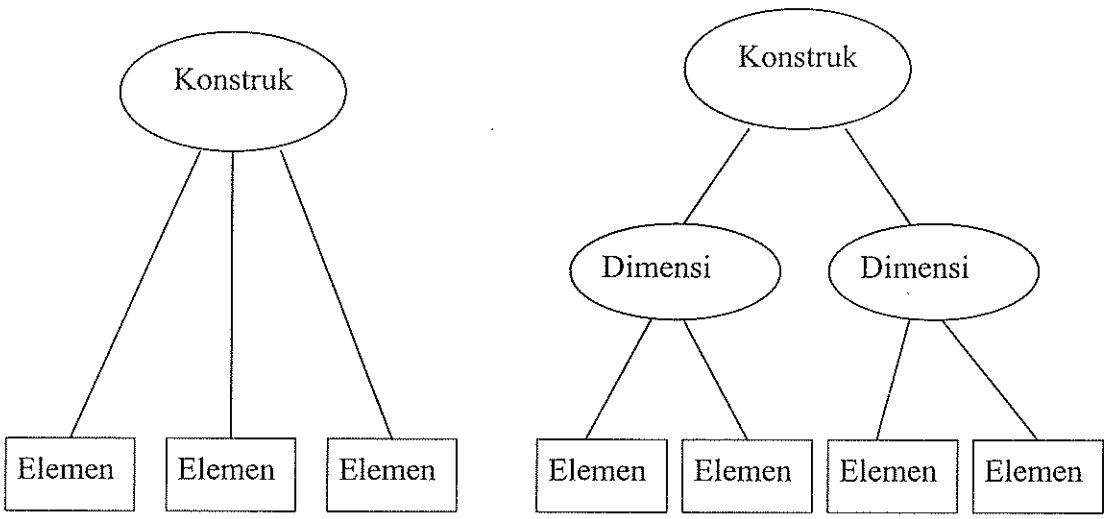
Pengukuran (*measurement*) adalah pemberian nilai properti dari suatu obyek.

Properti (*property*) adalah karakteristik dari obyek.

Pengukuran (*measurement*) adalah pemberian nilai properti dari suatu obyek. Dari definisi ini terlihat bahwa yang diukur adalah properti dari suatu obyek. **Obyek (*object*)** merupakan suatu entitas yang akan diteliti. Obyek (*object*) dapat berupa perusahaan, manusia, karyawan dan lainnya. **Properti (*property*)** adalah karakteristik dari obyek. Properti dapat berupa properti fisik, properti psikologi dan properti sosial. Properti fisik misalnya jika obyeknya adalah manusia, maka properti fisiknya adalah tinggi badan, warna rambut, umur dan lainnya. Jika obyeknya adalah perusahaan, maka properti fisiknya adalah ukuran perusahaan, lokasinya, dan lainnya. Properti psikologis misalnya adalah sikap manusia, kepintaran, motivasi dan lainnya. Properti sosial misalnya adalah status sosial, persepsi masyarakat dan lainnya.

Pengukuran properti psikologi dan sosial lebih sulit diukur karena tidak mudah diobservasi disebabkan properti tersebut masih bersifat konstruk (*construct*) dan abstrak. Misalnya adalah pengukuran konstruk belajar (*learning*). Konstruk ini tidak mudah diobservasi karena masih bersifat abstrak. Teknik yang dapat digunakan untuk mengukurnya adalah memecah konstruk ini kedalam beberapa karakteristik perilaku yang dapat diobservasi. Beberapa konstruk dapat langsung dipecah

dan ditemukan elemen-elemen perilaku yang dapat diukurnya, tetapi banyak konstruk yang tidak dapat langsung ditemukan elemen-elemen perilakunya, tetapi lewat beberapa dimensi terlebih dahulu sebagai berikut ini.



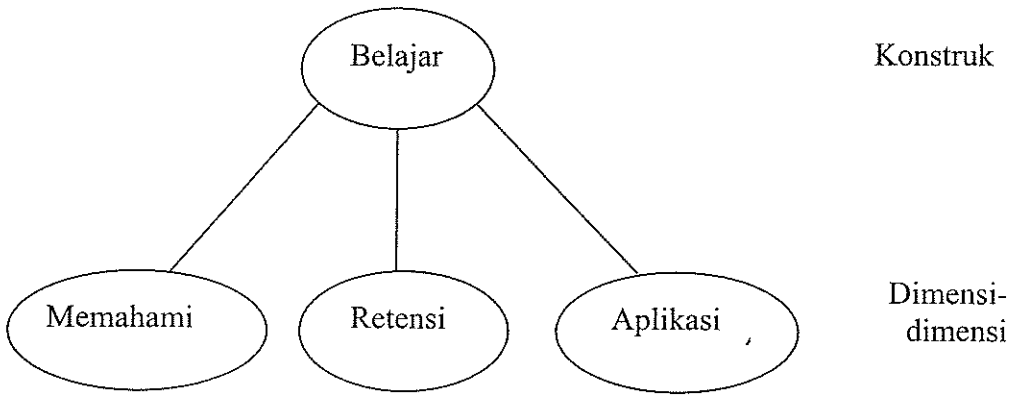
Gambar 5.1. Konstruk, dimensi-dimensi, dan elemen-elemen.

Dimensi ***(dimension)*** dari suatu konstruk adalah bagian-bagian dari properti yang menunjukkan karakteristik-karakteristik utama dari properti konstruk tersebut.

Dimensi ***(dimension)*** dari suatu konstruk adalah bagian-bagian dari properti yang menunjukkan karakteristik-karakteristik utama dari properti konstruk tersebut. Dimensi ini masih belum dapat diukur, sehingga perlu dipecah kembali menjadi elemen-elemen. **Elemen-elemen** ***(elements)*** merupakan perilaku yang dapat diobservasi dan diukur dari suatu konstruk atau dimensi.

Elemen-elemen (*elements*) merupakan perilaku yang dapat diobservasi dan diukur dari suatu konstruk atau dimensi.

Misalnya adalah konstruk belajar (*learning*). Konstruk ini dapat diukur lewat ujian yang dikerjakan. Banyak orang yang menganggap bahwa ujian saja tidak mengukur konstruk belajar. Jika demikian bagaimana konstruk belajar ini harus diukur. Sekaran (2003) memberi usulan dengan mendefinisikan konstruk belajar ini secara operasi dan memecahnya menjadi tiga dimensi yaitu 1) memahami (*understanding*), 2) retensi (*retention*) dan 3) aplikasi (*application*).



Gambar 5.2. Konstruk Belajar dan dimensi-dimensinya.

Pemecahan kedalam dimensi menunjukkan bahwa orang yang belajar harus dapat memahami apa yang diajarkan, menyimpan atau mengingat atau meretensi apa yang sudah dipahami tersebut dan dapat mengaplikasikan apa yang sudah dipahami dan diingat tersebut. Istilah-istilah memahami, retensi dan aplikasi masih merupakan hal yang abstrak walaupun sudah dapat menggambarkan konsep belajar dengan lebih detail. Oleh karena itu, ketiga dimensi ini masih perlu dipecah kembali lebih terinci supaya dapat diukur sebagai berikut ini.

1. Dimensi memahami (*understanding*) dapat diukur dengan elemen-elemen:
   a. Dapat menjawab pertanyaan-pertanyaan dengan benar.
   b. Dapat memberikan contoh-contoh yang tepat.
2. Dimensi retensi (*retention*) dapat diukur dengan elemen:
   Dapat memanggil materi dari otak setelah beberapa waktu lamanya.
3. Dimensi aplikasi (*application*) dapat diukur dengan elemen-elemen:
   a. Mampu memecahkan masalah dengan menerapkan konsep-konsep yang dipahami.
   b. Mampu mengintegrasikan dengan semua materi-materi lainnya yang relevan.

Elemen-elemen dari konstruk yang sudah dapat diukur ini kemudian dapat dijadikan sebagai item-item di kuestioner, Untuk mengukur ini diperlukan alat ukurnya yang disebut dengan skala pengukuran.

5.4. SKALA PENGUKURAN

Setelah item-item kuestioner telah didefinisikan untuk dapat diukur, maka diperlukan alat untuk mengukurnya yaitu yang disebut dengan skala. Terdapat empat macam tipe dasar dari skala yaitu nominal, ordinal, interval dan rasio.

1. **Nominal**, yaitu bernilai klasifikasi. Misalnya: laki, perempuan, untuk gender.
2. **Ordinal**, yaitu bernilai klasifikasi dan order (ada urutannya). Misalnya: penilaian (kurang, baik, sangat baik).

3. **Interval**, yaitu bernilai klasifikasi, order (ada urutannya), dan berjarak (perbedaan dua nilai berarti). Misalnya skala likert 1 sampai dengan 5, dengan jarak 1 sampai dengan 2 mempunyai jarak yang sama dengan 2 sampai dengan 3 dan seterusnya.
4. **Rasio**, yaitu bernilai klasifikasi, order, *distance* (berjarak) dan mempunyai nilai awal (*origin)*. Misalnya unit waktu sebesar 20 menit yang mempunyai nilai awal 0. Rasio dalam hal ini tidak harus dalam pembagian.

Tabel 5.1. Karakteristik-karakteristik Skala.

Skala	Karakteristik	Contoh
Nominal	- Klasifikasi	Gender berklasifikasi laki atau perempuan.
Ordinal	- Klasifikasi - Order (ada urutannya)	Kurang, baik, sangat baik.
Interval	- Klasifikasi - Order (ada urutannya) - Berjarak atau *distance* (ada perbedaan dua nilai).	Skala likert 1 sampai dengan 5, dengan jarak 1 sampai dengan 2 mempunyai jarak yang sama dengan 2 sampai dengan 3 dan seterusnya.
Rasio	- Klasifikasi - Order - Berjarak atau *distance* (ada perbedaan dua nilai). - *Origin* (mempunyai nilai awal)	Waktu 20 menit yang mempunyai nilai awal menit ke 0.

Misalnya nama variabel GROWTH dapat bernilai 1 untuk perusahaan yang bertumbuh dan 0 untuk perusahaan yang tidak bertumbuh. Nilai variabel tersebut adalah 0 dan 1 dengan tipe skala atau tipe nilai datanya adalah nominal.

Nilai data di variabel dapat juga diklasifikasikan sebagai data metrik (*metric)* dan data non-metrik (*non-metric)*. **Data metrik (*metric*)** berisi nilai kuantitatif dan yang termasuk data metrik adalah tipe data interval dan rasio. **Data non-metrik (*non-metric*)** adalah data kualitatif (Hair, 1992) yang dapat berbentuk suatu atribut, karakteristik atau kategori atau dikotomi. Yang termasuk data non-metrik adalah tipe data nominal dan ordinal.

Metoda penskalaan adalah metode pemberian nilai-nilai ke variabel.

Setelah empat macam skala dipahami (nominal, ordinal, interval dan rasio), maka perlu memberikan nilai-nilai kepada variabel sesuai dengan skalanya. Pemberian nilai-nilai ke variabel ini disebut dengan **metoda penskalaan**. Terdapat dua macam metoda penskalaan yaitu skala rating (*rating scale*) dan skala rangking (*rangking scale*).

5.4.1. Skala Rating

Skala rating (*rating scale*) digunakan untuk memberikan nilai (*rating*) ke suatu variabel.

Skala rating (*rating scale*) digunakan untuk memberikan nilai (*rating*) ke suatu variabel. Beberapa skala rating yang sering digunakan adalah sebagai berikut ini.

1. **Skala dikotomi (*dichotomous scale*).**
   Skala ini memberikan nilai dikotomi misalnya nilai Ya atau Tidak. Tipe data yang digunakan adalah nominal.

Contoh:

Apakah Anda mempunyai kartu kredit? □ Ya □ Tidak

2. **Skala kategori (category *scale*).**
   Skala ini memberikan nilai beberapa item untuk dipilih. Tipe data yang digunakan untuk skala ini adalah tipe nominal.

Contoh:

Pilih industri dari perusahaan:

_____ Pabrikan
_____ Jasa
_____ Gas dan Minyak
_____ Keuangan
_____ Lainnya

3. **Skala Likert (*Likert scale*).**
   Skala ini digunakan untuk mengukur respon subyek ke dalam 5 poin atau 7 poin skala dengan interval yang sama. Dengan demikian tipe data yang digunakan adalah tipe interval.

Contoh:

Dengan menggunakan skala Likert, apakah Anda setuju dengan pendapat ini.

	Sangat tidak Setuju	*Tidak Setuju*	*tidak tahu*	*Setuju*	*Sangat Setuju*
Kuliah di S2 menarik	*1*	*2*	*3*	*4*	*5*
Dosen memberikan wawasan	*1*	*2*	*3*	*4*	*5*
Dosen mengarahkan ke riset	*1*	*2*	*3*	*4*	*5*

4. **Skala perbedaan semantik (*semantic differential scale*).**
   Skala ini menggunakan dua buah nilai ekstrim dan subyek diminta untuk menentukan responnya diantara dua nilai tersebut di ruang yang disediakan yang disebut dengan ruang semantik. Tipe data yang digunakan adalah tipe ordinal atau interval.

Contoh:

Setuju __ __ __ __ __ *Tidak Setuju*
Pintar __ __ __ __ __ *Naif*
Besar __ __ __ __ __ *Kecil*

5. **Skala numerik (*numerical scale*).**
   Skala ini sama dengan skala perbedaan semantik hanya mengganti ruang semantik yang disediakan dengan angka-angka numerik (misalnya 1 sampai dengan 5 untuk 5 poin skala Likert atau 1 sampai dengan 7 untuk 7 poin skala Likert). Tipe data yang digunakan adalah tipe interval.

Contoh:

Setuju 1 2 3 4 5 6 7 Tidak Setuju

6. **Skala penjumlahan tetap atau konstan (*fixed or constant sum scale*).**
   Subyek diminta untuk mendistribusikan nilai responnya kedalam beberapa item yang sudah disediakan dengan jumlah yang tetap. Tipe data yang digunakan adalah tipe rasio.

Contoh:

Di dalam memilih pendidikan S2, tentukan besarnya nilai alokasi yang Anda berikan dengan total nilai 100 poin.

Fasilitas Komputer	___
Fasilitas Basis Data	___
Kenyamanan Kuliah	___
Gelar dosen tetap	___
Materi Kuliah	___
Total	*100*

7. **Skala stapel (*stapel scale*).**
   Skala ini dimaksudkan tidak hanya mengukur niattas respon dari subyek tetapi juga arah responnya. Karena nilai nol tidak disebutkan dengan eksplisit, maka tipe data yang digunakan adalah tipe interval.

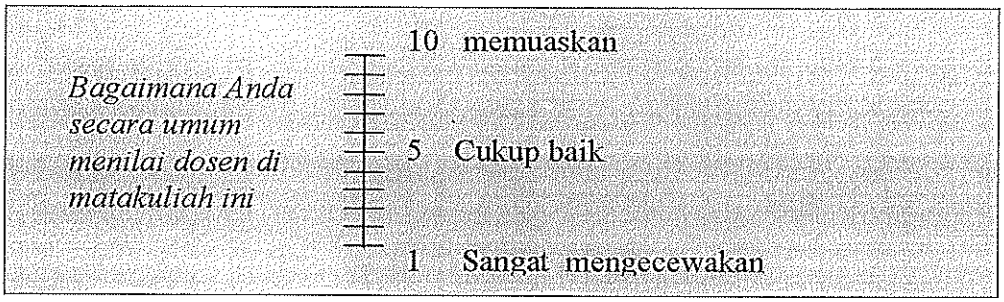
Contoh:

Tunjukkan bagaimana Anda menilai dosen yang mengajar di kelas dengan melingkari nilai jawabannya.

+3	+3	+3
+2	+2	+2
+1	+1	+1
Serius	Menarik	Pintar
-1	-1	-1
-2	-2	-2
-3	-3	-3

8. **Skala grafik (*graphic rating scale*).**
   Skala ini menggunakan grafik skala dan subyek memberi tanda pada tempat di grafik untuk responnya. Tipe data yang digunakan adalah tipe interval.

Contoh:



5.4.2. Skala Rangking

Skala rangking (*rangking scale*) membandingkan dua atau lebih obyek untuk memilih obyek yang lebih baik.

Skala rangking (*rangking scale*) membandingkan dua atau lebih obyek untuk memilih obyek yang lebih baik. Beberapa skala rangking adalah skala perbandingan-berpasangan (*paired-comparison scale*), skala rangking dipaksakan (*forced rangking scale*) dan skala komparatif (*comparative scale*).

1. **Skala perbandingan-berpasangan (*paired-comparison scale*).**
   Skala perbandingan-berpasangan digunakan untuk memilih satu dari dua obyek secara berpasangan. Jumlah pasangan yang ada adalah sebanyak (n x (n-1) / 2) dengan n adalah jumlah obyek. Misalnya jumlah obyek adalah 4, maka jumlah pasangan perbandingannya adalah (4 x (4-1) / 2) = 6. Tipe data yang digunakan adalah ordinal.

Contoh:

Diantara kandidat pasangan presiden dan wakil presiden, mana yang Anda pilih menjadi presiden perusahaan saudara:

___ *Ali*	___ *Ali*	___ *Ali*
___ *Basuki*	___ *Centil*	___ *Didik*
___ *Basuki*	___ *Basuki*	___ *Centil*
___ *Centil*	___ *Didik*	___ *Didik*

2. **Skala rangking dipaksakan (*forced rangking scale*).** Skala ini mengurutkan langsung relatif satu terhadap lainnya. Tipe data yang digunakan adalah ordinal.

Contoh:

Diantara kandidat presiden, mana yang Anda pilih menjadi presiden perusahaan saudara (beri nilai rangking 1 sampai dengan 4):

___ *Ateng*
___ *Basuki*
___ *Centil*
___ *Didik*

3. **Skala komparatif (*comparative scale*).** Skala ini membandingkan dengan standar atau *benchmark* yang lainnya. Tipe data yang digunakan adalah ordinal.

Contoh:

Dibandingkan dengan kinerja manajer periode kemarin, kinerja manajer sekarang:

Inferior		*hampir sama*		*Superior*
1	*2*	*3*	*4*	*5*

Contoh:

Contoh ini menunjukkan penggunaan skala Likert 7 poin untuk item-item konstruk kegunaan persepsian (*perceived usefulness*) di penelitian Davis (1989).

	Sangat Setuju						*Sangat Tidak Setuju*
1. Electronic mail enables me to accomplish tasks more quickly	*1*	*2*	*3*	*4*	*5*	*6*	*7*
2. Using electronic mail improves my job performance.	*1*	*2*	*3*	*4*	*5*	*6*	*7*
3. Using electronic mail increases my productivity.	*1*	*2*	*3*	*4*	*5*	*6*	*7*
4. Using electronic mail enhances my effectiveness on the job.	*1*	*2*	*3*	*4*	*5*	*6*	*7*
5. Using electronic mail makes it easier to do my job.	*1*	*2*	*3*	*4*	*5*	*6*	*7*
6. Overall, I find the electronic mail system useful in my job.	*1*	*2*	*3*	*4*	*5*	*6*	*7*

5.5. MEMBANGUN SENDIRI ITEM-ITEM KUESTIONER

Umumnya suatu penelitian menggunakan item-item konstruk sebagai kuestioner yang sudah digunakan di penelitian sebelumnya. Alasannya adalah membangun sendiri item-item konstruk sebagai kuestioner adalah hal yang tidak mudah. Membangun item-item konstruk harus didasarkan pada teori yang berhubungan dengan konstruk yang dibangun dan harus diuji terlebih dahulu validitas dan realibilitasnya. Jika terpaksa karena tidak ada item-

item konstruk di penelitian sebelumnya, maka item-item konstruk tersebut harus dibangun sendiri.

Telah dijelaskan bahwa membangun sendiri item-item kuestioner bukan hal yang mudah yang melibatkan beberapa tahapan. Tahapan pertama disebut dengan pembentukan item-item, yang tujuannya adalah membuat item-item untuk masing-masing konstruk. Tahapan kedua adalah sebelum tes (*pretest*) yang tujuannya utamanya adalah menguji validitas isi (*content validity*) supaya item-item pertanyaan tidak membingungkan dan tidak ambigius. Tahapan ketiga adalah pengujian instrumen, yaitu instrumen disebarkan ke responden-responden sebenarnya dan dari respon-respon mereka diuji validitas konstruk dan reliabilitas skalanya. Tahapan-tahapan lebih rinci adalah sebagai berikut ini.

1. Melakukan pembentukan item-item.
   Pembentukan awal dimaksudkan untuk mendapatkan nama-nama konstruk dan item-item pertanyaan awal yang membentuk konstruk-konstruk yang diinginkan. Pembentukan item-item awal ini tidak boleh sembarangan harus berdasarkan hasil-hasil penelitian sebelumnya dan sesuai dengan teori yang relevan. Oleh karena itu, item-item pertanyaan awal ini dapat diperoleh dari mengkaji penelitian-penelitian sebelumnya dan mengkaji teori-teori yang relevan.
2. Melakukan sebelum tes (*pretest*).
   Sebelum tes (*pretest*) adalah tes atau uji untuk meyakinkan bahwa item-item pertanyaan awal dapat dimengerti oleh calon partisipan atau responden survei. Jangan sampai survei sudah dikirimkan dengan biaya yang mahal tetapi responden-responden tidak memahami item-item pertanyaannya.

Sebelum tes (*pretest*) adalah tes atau uji untuk meyakinkan bahwa item-item pertanyaan awal dapat dimengerti oleh calon partisipan atau responden survei.

Disebut dengan sebelum tes (*pretest*) karena belum diuji dengan menggunakan responden-

responden sebenarnya. Pengujian sebenarnya yang menggunakan responden-responden sebenarnya dilakukan setelah instrumen selesai dibangun yang dilakukan pada tahap selanjutnya. Pada tahap ini instrumen belum selesai dibangun. Pada tahap ini, item-item perlu diuji untuk melihat kelayakan lebih lanjut dari item-item tersebut.

Validitas isi (*content validity*) menunjukkan tingkat seberapa besar item-item di instrumen mewakili konsep yang diukur.

Tujuan dari *pretest* adalah untuk menguji validitas isi (*content validity*). **Validitas isi (*content validity*)** menunjukkan tingkat seberapa besar item-item di instrumen mewakili konsep yang diukur. Jika instrumen yang digunakan secara cukup mencakup topik yang sudah didefinisikan sebagai dimensi-dimensi dan elemen-elemen yang relevan menggambarkan konsepnya, maka dapat dikatakan bahwa instrumen tersebut mempunyai validitas isi yang baik. Uji validitas isi dapat berupa **uji validitas tampang (*face validity*)** yang digunakan untuk meyakinkan bahwa item-item pertanyaan masuk akal dan benar dan biasanya dilakukan dengan meminta pendapat ahli-ahli yang bersangkutan. Uji sebelum tes juga digunakan untuk menilai validitas konstruk (*construct validity*). Hasil dari *pretest* ini adalah untuk membuang atau mengganti item-item yang kurang cocok membentuk konstruk-konstruk yang diinginkan.

Uji validitas tampang (*face validity*) untuk meyakinkan bahwa item-item pertanyaan masuk akal dan benar dan biasanya dilakukan dengan meminta pendapat ahli-ahli.

Sebelum tes (*pretest*) dapat dilakukan dengan membentuk panel pakar yang terdiri dari pakar-pakar atau hakim-hakim (*judges*). **Hakim-hakim (*judges*)** adalah pakar-pakar yang digunakan untuk mengkonfirmasi kebenaran item-item yang membentuk konstruk-konstruk. Penggunaan hakim-hakim dilakukan dengan prosedur sebagai berikut ini.

Hakim-hakim (*judges*) adalah pakar-pakar yang digunakan untuk mengkonfirmasi kebenaran item-item yang membentuk konstruk-konstruk.

a. Hakim-hakim (*judges*) diminta untuk untuk mengurutkan item-item kedalam kategori-kategori konstruk.

b. Hakim-hakim (*judges*) kemudian diberi definisi dari masing-masing konstruk. Hakim-hakim (*judges*) biasanya hanya diberi definisi dari konstruk saja tetapi tidak diberi nama atau label dari konstruknya. Penelitian Davis (1989) memberikan nama dari konstruk ke hakim-hakim (*judges*). Memberikan nama konstruk kepada hakim-hakim (*judges*) dapat menimbulkan bias yang disebut dengan **gangguan interpretasional (*interpretational confounding*)** yang terjadi karena pemberian arti yang diberikan kepada suatu konstruk oleh seorang individual yang sudah terpengaruh oleh nama yang sudah diketahuinya terlebih dahulu dibandingkan dengan arti empiris yang seharusnya diberikan kepada konstruk tersebut. Untuk menghindari bias ini sebaiknya hakim-hakim diminta untuk memberikan nama sendiri kepada konstruk-konstruk.

Gangguan interpretasional (*interpretational confounding*) adalah bias karena memberikan nama konstruk kepada hakim-hakim (*judges*).

Hasil dari penilaian oleh hakim-hakim dimaksudkan untuk menilai awal reliabilitas dan validitas konstruk. Disebut dengan penilaian awal reliabilitas dan validitas konstruk karena reliabilitas dan validitas konstruk juga akan diuji kembali pada tahapan berikutnya. Nilai reliabilitas dapat diukur dengan tingkat kesepakatan (***level of agreement***) oleh masing-masing pasangan hakim-hakim dalam mengkategorikan (mengumpulkan dalam konstruk-konstruk) item-item pertanyaan di kuestioner. Tingkat kesepakatan (***level of agreement***) oleh hakim-hakim diukur dengan Cohen's Kappa (Cohen 1960). Nilai Kappa yang dapat diterima adalah diatas 0,65 (Vessey 1984; Jarvenpaa 1989; Todd dan Benbasat 1989).

Indikator dari validitas konstruk dapat berupa konvergensi dan divergensi dari item-item didalam kategori-kategori. Jika suatu item secara konsisten

diletakkan didalam suatu kategori tertentu, maka dapat dianggap menunjukkan validitas konvergen (***convergent validity***) dengan konstruk yang berhubungan, dan validitas diskriminan (***discriminant validity***) dengan konstruk lainnya. Dari hasil prosedur pengurutan item-item ini, jika jumlah kategori-kategori yang dibuat oleh beberapa hakim-hakim, label-label yang diberikan kepada kategori-kategori tersebut, dan item-item didalamnya konsisten, maka skala-skala yang berbasis pada kategori-kategori ini dapat dikatakan menunjukkan validitas konvergen dan validitas diskriminan.

3. Melakukan uji instrumen.
   Instrumen yang sudah dikembangkan di tahap sebelumnya perlu diuji validitas dan reliabilitasnya. Uji instrumen ini paling tidak melewati dua tahapan, yaitu uji pilot (*pilot test*) dan uji lapangan (*field test*).
   a. Uji pilot.
      Setelah item-item kuestioner untuk masing-masing konstruk telah berhasil diidentifikasikan di tahap sebelumnya, tahap selanjutnya adalah melakukan uji pilot dari keseluruhan instrumen. Uji pilot dapat dilakukan sekali atau dua kali. Uji pilot yang pertama merupakan uji pilot awal dan uji pilot kedua merupakan uji pilot skala penuh.

      Di uji pilot awal, ukuran sampel dibuat kecil sekitar 10 sampai 30 responden. Di uji pilot awal ini, responden-responden tidak harus merupakan bagian dari populasi targetnya, tetapi dapat berupa individual-individual yang dipilih yang memahami tentang isu yang diteliti. Di uji pilot ini item-item kuestioner untuk semua skala dapat dikelompokkan menjadi satu grup yang umum.

Sasaran pertama dari uji pilot awal ini adalah untuk meyakinkan bahwa item-item kuestioner telah mencukupi, benar, dan dapat dipahami. Ini dilakukan dengan cara responden-responden diminta untuk melengkapi kuestioner jika dianggap ada yang kurang. Responden-responden juga diminta untuk memberi komentar mengenai panjangnya kuestioner, apakah terlalu panjang atau sudah cukup panjang. Kuestioner yang terlalu panjang akan tidak efektif, karena responden-responden akan malas dan merasa bosan mengisinya dengan akibat akan menjawab item-item pertanyaan dengan sembarangan. Di uji pilot awal ini, responden-responden juga diminta untuk meneliti kata-kata, kalimat-kalimat dan instruksi-instruksi di instrumen apakah sudah jelas dan dapat dipahami.

Sasaran kedua dari uji pilot awal ini adalah untuk penilaian reliabilitas awal dari skala-skala. Reliabilitas skala umumnya dinilai dengan skor Cronbach's ALPHA (Cronbach, 1970). Nilai ALPHA yang dapat diterima tergantung dari tujuan penelitiannya. Nunnaly mengusulkan bahwa di tahapan-tahapan awal dari riset, nilai reliabilitas 0,50 sampai dengan 0,60 dianggap cukup, dan untuk riset dasar (*basic research*), diargumentasikan bahwa meningkatkan reliabilitas melebihi nilai 0,80 sering dianggap terlalu tinggi (1967, hal. 226). Umumnya skor reliabilitas yang diterima di banyak penelitian berkisar antara 0,70 sampai dengan 0,80.

Tabel 5.2. Batasan Skor Reliabilitas Cronbach's ALPHA

Skor	Reliabilitas
< 0,50	Rendah
0,50 – 0,60	Cukup
0,70- 0,80	Tinggi

Jika skor reliabilitas rendah, maka skor ini perlu ditingkatkan. Salah satu cara meningkatkan skor reliabilitas adalah dengan menghapus item-item yang tidak konvergen (tidak mengarah ke hal yang sama) di dalam suatu konstruk tertentu. Untuk mengetahui item-item yang tidak konvergen, dapat dilihat nilai korelasi antar itemnya dan nlai variannya.

Item-item dengan korelasi yang rendah, yang akan meningkatkan ALPHA jika dihapus, atau menunjukkan varian yang rendah (memberi-kan kekuatan penjelasan yang rendah di model) merupakan kandidat-kandidat untuk dihapus. Sebelum suatu item dihapus, suatu pengecekan akan dilakukan untuk meyakinkan bahwa validitas isi (*content validity*) dari konstruk tidak akan dikorbankan.

Uji pilot kedua adalah uji pilot skala penuh (*full scale)* dari kuestioner. Sasaran utama dari pengujian ini masih untuk meyakinkan bahwa bermacam-macam skala telah menunjukkan tingkat reliabilitas yang semestinya seperti di uji pilot awal, tetapi menggunakan responden-responden yang latar belakangnya sama dengan populasi target dari penelitian. Jumlah responden di dalam sampel merupakan sampel yang kecil yang lebih besar dari sampel di uji pilot awal tetapi jauh lebih sedikit dari sampel akhirnya. Misalnya di penelitian Moore dan

Benbasat (1991) digunakan sampel sebanyak 75 responden-responden untuk uji pilot skala penuh ini.

b. Uji Lapangan Akhir (*Final Field Test*).

Uji lapangan (*field test*) adalah pengujian yang menggunakan data lapangan yang benar-benar digunakan di penelitian.

Disebut dengan **uji lapangan (*field test*)** karena pengujian ini menggunakan data lapangan yang benar-benar digunakan di penelitian. Data lapangan ini berupa responden-responden dari sampel akhir penelitian. Misalnya di penelitian Moore dan Benbasat (1991) digunakan 800 kuestioner-kuestioner yang didistribusikan kepada individual-individual di tujuh perusahaan dari beberapa macam industri.

Pengujian lapangan akhir ini dilakukan untuk menguji reliabilitas dan validitas konstruk. Uji reliabilitas dilakukan dengan cara yang sama seperti yang dilakukan di pengujian pilot sebelumnya, yaitu diukur dengan koefisien *Cronbach's alpha*. Di pengujian lapangan akhir ini, validitas konstruk biasanya diuji dengan analisis faktor (*factor analysis*). Selain menggunakan analisis faktor, validitas konstruk juga dapat diuji dengan validitas konvergen dan validitas diskriminan. Untuk contoh berikutnya, yaitu penelitian-penelitian Moore dan Benbasat (1991) dan Davis (1989) tidak menggunakan uji validitas konvergen dan validitas diskriminan, tetapi hanya menggunakan analisis faktor untuk menguji validitas konstruknya.

Untuk analisis faktor, supaya dapat dikatakan bahwa suatu item membentuk suatu konstruk dengan benar, maka item ini harus termuat dengan skor yang tinggi di konstruk tersebut dan tidak boleh termuat secara tinggi di konstruk yang lain. Comrey (1973) menunjukkan bahwa muatan-

muatan yang melebihi skor 0,45 dapat dianggap cukup, lebih besar dari 0,55 baik, 0,63 sangat baik, dan 0,71 memuaskan (*excellent*).

Tabel 5.3. Batasan Skor Muatan untuk Analisis Faktor (Comrey, 1973).

Skor Muatan	Validitas Konstruk
<= 0,45	Rendah
> 0,45	Cukup
> 0,55	Baik
> 0,63	Sangat baik
> 0,71	Memuaskan

5.6. MENGGUNAKAN ITEM-ITEM KUESTIONER DARI PENELITIAN-PENELITIAN SEBELUMNYA

Karena membangun sendiri item-item kuestioner bukan hal yang mudah yang melibatkan waktu yang cukup lama dan belum tentu hasil dari kuestioner dapat berhasil sesuai dengan yang diinginkan. Jika kuestioner yang dibangun gagal dalam arti kuestioner yang dibuat tidak valid atau tidak reliabel, maka kuestioner tersebut tidak dapat digunakan dan harus dibangun kembali. Oleh karena itu banyak peneliti yang menggunakan kuestioner yang sudah mapan dengan validitas dan reliabilitas yang tinggi dari penelitian-penelitian sebelumnya.

Pembentukan awal yang dilakukan di pembangunan sendiri kuestioner tidak perlu dilakukan jika peneliti menggunakan kuestioner yang sudah ada. Alasannya adalah kuestionernya sudah ada sehingga tidak perlu dibentuk lagi. Akan tetapi kuestioner yang sudah ada ini masih perlu diuji lebih lanjut. Kuestioner ini perlu diuji lebih lanjut karena akan digunakan di situasi (tempat dan waktu), serta responden-responden yang berbeda dengan yang digunakan di penelitian-penelitian sebelumnya asal

dari kuestioner. Apalagi jika kuestioner diterjemahkan dari satu bahasa, misalnya dari bahasa Inggris, ke bahasa yang lain, misalnya bahasa Indonesia, maka pengujian lebih lanjut perlu dilakukan.

Untuk menggunakan kuestioner dari penelitian-penelitian sebelumnya, langkah-langkah yang harus dilakukan adalah sebagai berikut ini.

1. Melakukan sebelum tes (*pretest*).
   Sebelum tes (*pretest*) perlu dilakukan karena kuestioner yang sudah ada 1) diterjemahkan ke bahasa lain, 2) diterapkan di situasi yang berbeda. Tujuan utama dari sebelum tes (*pretest*) ini adalah untuk meyakinkan bahwa item-item pertanyaan dapat dipahami, dianggap sudah benar, dan konsisten. Tujuan lainnya adalah untuk mendapatkan usulan-usulan atau rekomendasi-rekomendasi perbaikan kuestioner oleh pakar-pakar.
   a. Jika kuestioner diterjemahkan misalnya dari bahasa Inggris ke bahasa Indonesia, maka harus diyakinkan bahwa terjemahannya benar dan tidak merubah arti dari kuestioner. Untuk menguji kebenaran terjemahan ini dapat dilakukan dengan meminta ahli bahasa untuk menterjemahkan kembali hasil terjemahan dalam bahasa Indonesia ke bahasa Inggris dan kemudian mencocokkan hasil terjemahan kembali dalam bahasa Inggris dengan kuestioner aslinya yang dalam bahasa Inggris apakah terjadi perbedaan yang cukup signifikan.
   b. Kuestioner yang sudah diterjemahkan atau diambil dari penelitian-penelitian sebelumnya kemudian dapat diuji kelogisannya lewat panel pakar untuk mengkonfirmasi item-item yang membentuk konstruk-konstruk tersebut. Hasil dari *pretest* ini adalah untuk membuang atau mengganti item-item

yang kurang cocok membentuk konstruk-konstruk yang diinginkan sesuai saran dari pakar (hakim) bersangkutan.

Contoh:

Penelitian Hsu dan Chiu (2004) menggunakan teknik survei untuk mengumpulkan datanya untuk mengetahui penggunaan berlanjut dari pengisian pajak lewat elektronik (*e-filling*). Kuestioner yang digunakan tidak dibuat sendiri tetapi diambilkan dari beberapa penelitian sebelumnya. Penelitian ini melakukan uji sebelum tes (*pretest*) dengan panel pakar.

> *Pretest* dari kuestioner dilakukan oleh 10 pakar dalam penelitian sistem informasi untuk menilai konsistensi logik, kemudahan pemahaman dan urutan dari item-item dan relevansi tugas di kuestioner. Secara keseluruhan, para pakar berpendapat bahwa kuestionernya sudah jelas dan mudah untuk diselesaikan.

Contoh:

Penelitian Compeau dan Higgins (1995) yang menggunakan *pretest* untuk maksud mendapatkan umpan balik dari kuestioner.

> Uji sebelumnya (*pretest*) dilakukan kepada 40 orang termasuk akademisi dan praktisi. Hasil dari *pretest* menunjukkan bahwa masing-masing responden menjawab pertanyaan-pertanyaan dan memberikan umpan balik tentang proses dan pengukur-pengukur. Secara umum, semua responden menunjukkan bahwa kuestioner-kuestioner secara relatif jelas dan mudah untuk diselesaikan. Sejumlah usulan diberikan mengenai kata-kata dari beberapa item dan mengenai struktur dari kuestioner-kuestioner dan semua usulan sudah dimasukkan ke dalam kuestioner yang direvisi.

2. Melakukan uji instrumen.
   Setelah item-item kuestioner lolos dari uji sebelum tes (*pretest*), maka item-item tersebut perlu diuji kembali reliabilitas dan validitasnya. Uji yang dapat dilakukan adalah uji pilot (*pilot test*) dan uji lapangan (*field test*).

   a. Uji pilot.
      Di uji pilot, ukuran sampel dibuat kecil sekitar 10 sampai 30 responden. Sasaran pertama dari uji pilot awal ini adalah untuk meyakinkan bahwa item-item kuestioner telah mencukupi dan benar menurut responden. Ini dilakukan dengan cara responden-responden diminta untuk mengomentari dan memberi masukan pada kuestionernya. Sasaran kedua dari uji pilot awal ini adalah untuk penilaian reliabilitas awal dari skala-skala. Sasaran lainnya dari uji pilot adalah untuk menentukan besarnya sampel yang akan digunakan di penelitian utamanya seperti yang dilakukan oleh Higgins (1995) di contoh berikut ini.

Contoh:

Penelitian Compeau dan Higgins (1995) yang menggunakan studi pilot untuk maksud sebagai berikut ini.

> Suatu studi pilot dilakukan melibatkan 100 individual-individual. Tujuan dari studi pilot ini adalah untuk 1) mendapatkan tambahan umpan balik mengenai instrumen kuestioner, dan 2) membantu menentukan ukuran sampel untuk studi utama. Tingkat respon ditentukan dari studi pilot dibandingkan dengan jumlah sampel yang dibutuhkan untuk menganalisis model persamaan struktural. Untuk fleksibilitas maksimum, jumlah sampel yang dibutuhkan untuk analisis model persamaan struktural adalah 5 kali dari jumlah variabel di kuestioner-kuestioner. Karena jumlah variabel adalah 8, maka

jumlah minimum sampel yang diperlukan adalah 400 responden. Tingkat respon untuk studi pilot adalah 60%. Diasumsikan lebih konservatif, 50% tingkat respon digunakan, sehingga membutuhkan jumlah sampel seluruhnya adalah 1.600 responden-responden. Di penelitian ini kemudian dipilih 2.000 responden secara random untuk studi utama lapangannya.

Contoh:

Penelitian Hsu dan Chiu (2004) menggunakan studi pilot sebagai berikut ini.

Suatu studi pilot dilakukan kepada 30 orang mahasiswa tingkat master paruh waktu. Tujuan dari studi pilot ini adalah untuk mendapatkan umpan-balik tambahan tentang instrumen kuestioner-kuestioner. Responden diminta untuk menyediakan komentar-komentar pada isi dan struktur kuestioner.

b. Uji Lapangan Akhir (*Final Field Test*).
Disebut dengan uji lapangan karena pengujian ini menggunakan data lapangan yang benar-benar digunakan dipenelitian. Pengujian lapangan akhir ini difokuskan pada uji validitas konstruk. Analisis faktor biasanya digunakan sebagai penilaian untuk validitas konstruk. Analisis faktor (*factor analysis*) atau nama lengkapnya analisis komponen-komponen prinsipal (*principal components analysis*) dilakukan untuk menentukan item-item mana yang dapat masuk membentuk suatu konstruk. Lebih lanjut dan lebih detail tentang uji validitas ini dapat dilihat di Lampiran 5B.

Validitas konstruk dapat juga dinilai melalui validitas konvergen (*convergent validity*), validitas diskriminan (*discriminant validity*) dan analisis faktor. Validitas konstruk dapat juga dinilai melalui validitas konvergen (*convergent validity*), validitas diskriminan (*discriminant validity*) dan analisis faktor.

Validitas konvergen (*convergent validity*) berhubungan dengan prinsip bahwa pengukur-pengukur dari suatu konstruk seharusnya berkorelasi tinggi terhadap dirinya sendiri.

Validitas konvergen (*convergent validity*) berhubungan dengan prinsip bahwa pengukur-pengukur dari suatu konstruk seharusnya berkorelasi tinggi terhadap dirinya sendiri. Validitas konvergen (*convergent validity*) terjadi jika skor-skor yang diperoleh dari dua instrumen yang berbeda yang mengukur konstruk yang sama mempunyai korelasi yang tinggi. Validitas konvergen (*convergent validity*) dievaluasi dengan menggunakan dua kriteria berdasarkan usulan dari Fornell and Larcker (1981) sebagai berikut: (1) semua muatan-muatan faktor (*factor loadings*) seharusnya signifikan dan nilainya melebihi 0,70; dan (2) rata-rata varian yang diekstraksi (*average variance extracted* atau AVE) untuk masing-masing konstruk seharusnya melebihi varian yang diakibatkan oleh kesalahan pengukuran untuk konstruk tersebut (yaitu seharusnya melebihi nilai 0,50).

Validitas diskriminan (*discriminant validity*) berhubungan dengan prinsip bahwa pengukur-pengukur di konstruk-konstruk berbeda seharusnya tidak berkorelasi dengan tinggi. sendiri.

Validitas diskriminan (*discriminant validity*) berhubungan dengan prinsip bahwa pengukur-pengukur di konstruk-konstruk yang berbeda seharusnya tidak berkorelasi dengan tinggi. Validitas diskriminan (*discriminant validity*) terjadi jika dua instrumen yang berbeda yang mengukur dua buah konstruk yang diprediksikan tidak berkorelasi menghasilkan skor-skor yang memang tidak berkorelasi. Validitas diskriminan dapat dievaluasi lewat perbandingan dari AVE (*average variance extracted*) dengan korelasi kuadrat antara konstruk-konstruk berpasangan. Fornell dan Larcker (1981) merekomendasikan suatu pengujian validitas diskriminan, yaitu AVE seharusnya melebihi nilai korelasi kuadrat antara pasangan konstruk.

Contoh:

Penelitian Hsu dan Chiu (2004) menggunakan sampel sesungguhnya di lapangan sebanyak 100 perusahaan yang menggunakan *e-filling*. Total 800 kuestioner yang dikirimkan kepada manajer-manajer sistem informasi di 100 perusahaan-perusahaan ini. Sebanyak 157 survei dikembalikan dan yang dapat digunakan karena lengkap adalah sebanyak 149 survei.

Dengan menggunakan data lapangan ini, penelitian Hsu dan Chiu (2004) menguji reliabilitas dengan koefisien *Cronbach's alpha* dan validitas konstruk dengan muatan faktor sebagai berikut ini.

Tabel. Ringkasan dari skala-skala pengukuran.

	Pengukur Konstruk	Rata-rata	SD	Muatan Faktor	Reliabilitas Konstruk
Niat meneruskan e-jasa (*E-service Continuance Intention*) (CI)					
CI1	I intend to continue using the e-file service in the future.	5,1844	1,5241	0,978	0,9736
CI2	I will continue using the e-file service in the future.	5,1135	1,5995	0,987	
CI3	I will regularly use the e-file service in the future.	4,9787	1,5093	0,959	
Kepuasan pemkai (*User Satisfaction*) (US)					
US1	I am satisfied with the performance of e-file service.	4,8511	1,3413	0.935	0,9447
US2	I am pleased with the experience of using the e-file service.	5,0142	1,2420	0,939	
US3	My decision to use the e-file service was a wise one.	5,0213	1,2732	0,955	
US4	My feeling with using the e-file service was better than manual tax filing.	5,1348	1,3481	0,880	
Pengaruh interpersonal (*Interpersonal Influence*) (II)					
II1	My family thought I should continue using the e-file service.	4,3262	1,2733	0,805	0,8507
II2	My colleagues thought I should continue using the e-file service.	4,1631	1,2457	0,891	
II3	My friends thought I should continue using the e-file service.	4,1702	1,1524	0,892	
II4	People I knew thought that continuing to use the e-file service was a good idea.	4,5319	1,3232	0,746	
Pengaruh eksternal (*External Influence*) (EI)					

EI1	I read/saw news reports that using the e-file service was a good way of filling tax.	5,0213	1,2333	0,785	0,7892
EI2	The popular press depicted a positive sentiment for using the e-file service	4,6879	1,2769	0,776	
EI3	Mass media reports convinced me to reuse the e-file service.	4,6879	1,3946	0,780	
EI4	Expert opinions depicted a positive sentiment for using the e-file service	4,2270	1,5277	0,796	
Kegunaan persepsian (*Perceived Usefulness*) (PU)					
PU1	Using the e-file service can improve my tax filing performance.	5,4255	1,3796	0,929	
PU2	Using the e-file service can increase my tax filing effectiveness.	5,3830	1,5244	0,940	0,9319
PU3	Using the e-file service can simplify the tax filing process.	5,2908	1,3760	0,949	
Risiko persepsian (*Perceived Risk*) (PR)					
PR1	There was possibility of the e-file service malfunctioning and not performing as it was designed and therefore failing to deliver the desired benefits	5,0567	1,4870	0,772	0,7929
PR2	There might be potential monetary outlay associated with using the e-file service	4,7660	1,6020	0,859	
PR3	There was possibility of losing time when making a unsuccessful tax filing process by wasting time refilling tax.	4,8298	1,5582	0,902	
Kepermainan persepsian (*Perceived Playfulness*) (PP)					
PP1	I found using the e-file service interesting.	4,7234	1,4447	0,872	0,9478
PP2	I found using the e-file service enjoyable.	4,2766	1,4741	0,955	
PP3	I found using the e-file service exciting.	4,3262	1,5047	0,960	
PP4	I found using the e-file service fun.	4,3262	1,5095	0,931	
Kontrolabilitas persepsian (*Perceived Controllability*) (PC)					
PC1	Whether or not I file tax through the e-file service is entirely up to me.	6,1418	1,1186	0,723	0,5723
PC2	How much personal control do you feel you have over filing tax through the e-file service? (very little control/complete control)	6,2624	1,1190	0,822	
PC3	How much do you feel that whether you file tax through the e-file service is beyond your control? (not at all/very much so)	5,5957	1,5118	0,679	
Keyakinan-sendiri internet (*Internet Self-Efficacy*) (ISE)					
ISE1	I feel confident completing the task of filing income tax through the WWW.	8,5957	1,7070	0,737	0,9673

ISE2	I feel confident visiting the Web site for filing income tax by entering its address (URL) in the browser	8,8369	1,6674	0,784
ISE3	I feel confident navigating the e-file Web site by following hyperlinks.	8,8652	1,7857	0,707
ISE4	I feel confident finding information about the IRS Web site by using a search	8,8298	1,6731	0,850
ISE5	I feel confident finding information about the IRS Web site in a Web directory or portal site.	8,5816	1,9055	0,771
ISE6	I feel confident downloading software and data on the Web, e.g., certificate and software for filing income tax.	8,5532	1,7826	0,832
ISE7	I feel confident installing an application or software, e.g., software for filing income tax.	8,3901	2,0346	0,768
ISE8	I feel confident receiving an e-mail message containing the certificate data from the certificate authority.	8,5248	1,9294	0,849
ISE9	I feel confident filling out a Web form to apply for a secretary key.	8,3333	2,1303	0,789
ISE10	I feel confident filing out and submitting the forms for filing income tax.	8,7447	1,7172	0,836

Sumber: Hsu dan Chiu (2004)

Keterangan: Item-item ini memang dibiarkan masih dalam bahasa Inggris dan belum diterjemahkan kedalam bahasa Indonesia supaya didapatkan kata-kata aslinya dari sumber penelitiannya. Alasannya adalah menterjemahkan item-item di daftar pertanyaan harus sangat hati-hati karena translasi yang salah dan kurang mengena akan mengakibatkan hilangnya validitas dari item-item tersebut.

Validitas konvergen (*convergent validity*) dievaluasi dengan menggunakan dua kriteria berdasarkan usulan dari Fornell and Larcker (1981) sebagai berikut: (1) semua muatan-muatan faktor (*factor loadings*) seharusnya signifikan dan nilainya melebihi 0,70; dan (2) rata-rata varian yang diekstraksi (*average variance extracted* atau AVE) untuk masing-masing konstruk seharusnya melebihi varian yang diakibatkan oleh kesalahan pengukuran untuk konstruk tersebut (yaitu seharusnya melebihi nilai 0,50). Seperti tampak di Tabel sebelumnya, hampir semua item-item mempunyai muatan yang lebih besar dari 0,70 yang menunjukkan konvergensi item yang dapat

diterima. Salah satu pengecualian adalah pada item ketiga dari skala kontrolabilitas persepsian (*perceived controllability*) yang nilai muatannya adalah 0,679 sedikit dibawah 0,70. Nilai AVE berkisar antara 0,55 sampai dengan 0,95 (lihat Tabel berikut), yang lebih besar dari varian yang disebabkan oleh kesalahan pengukuran. Dengan demikian dua kriteria untuk memenuhi validitas konvergen terpenuhi.

Validitas diskriminan dapat dievaluasi lewat perbandingan dari AVE dengan korelasi kuadrat antara konstruk-konstruk berpasangan. Tabel berikut ini menunjukkan bahwa nilai AVE terkecil adalah sebesar 0,55 sementara korelasi kuadrat antara dua konstruk berpasangan terbesar adalah 0,53. Dengan demikian, pengujian validitas diskriminan juga dapat diterima.

Tabel. Validitas diskriminan.

Konstruk	1	2	3	4	5	6	7	8	9
1. Continuance Intention	0,95								
2. User Satisfaction	0,425	0,95							
3. Interpersonal Influence	0,309	0,361	0,68						
4. External Influence	0,047	0,076	0,338	0,60					
5. Perceived Usefulness	0,263	0,261	0,706	0,273	0,89				
6. Perceived Risk	0,244	0,183	0,594	0,292	0,662	0,77			
7. Perceived Playful	0,291	0,242	0,567	0,191	0,576	0,728	0,88		
8. Internet Self-efficacy	0,332	0,279	0,674	0,334	0,758	0,711	0,629	0,67	
9.Perceived Controllability	0,390	0,263	0,657	0,364	0,635	0,602	0,546	0,732	0,55

Sumber: Hsu dan Chiu (2004.

Keterangan:

1. Korelasi-korelasi semuanya signifikan pada tingkat 0,01.
2. Elemen-elemen diagonal adalah *Average Variance Extracted* (AVE) dan di luar diagonal adalah korelasi antara dua konstruk.

Contoh:

Misalnya adalah penelitian Hsieh et al. (2005) yang memfokuskan pada perbedaan perilaku sebelum adopsi dari penerapan akses internet gratis kepada masyarakat dan faktor-faktor yang menyebabkannya. Penelitian ini melakukan pengujian reliabilitas menggunakan Cronbach's alpha dan

validitas diskriminan dengan menggunakan AVE.

Tabel berikut ini menyajikan statistik deskriptif, reliabilitas komposit, dan rata-rata varian diekstraksi dari konstruk-konstruk. Kemudian, validitas diskriminan (*discriminant validity*) dinilai dengan meneliti apakah korelasi kuadrat antara sepasang variabel-variabel laten lebih kecil dibandingkan dengan rata-rata estimasi varian diekstraksi untuk masing-masing variabel. Hasilnya menunjukkan validitas diskriminan yang baik.

Tabel. Statistik deskriptif dan reliabilitas konstruk-konstruk.

	Grup *Underprivileged*			Grup *Privileged*		
Konstruk*	Rata-rata (Deviasi Standar)	Reliabi-litas Komposit	AVE†	Rata-rata (Deviasi Standar)	Reliabi-litas Komposit	AVE†
Attitude (3)	5.76(1.74)	0.98	0.96	4.86(1.87)	0.98	0.95
Utilitarian Outcomes (4)	5.37(1.83)	0.98	0.94	4.09(2.18)	0.98	0.97
Hedonic Outcomes (3)	5.69(1.85)	0.98	0.97	4.27(2.26)	0.98	0.97
Social Outcomes (3)	3.53(2.06)	0.95	0.87	2.68(1.72)	0.94	0.85
Subjective Norms (2)	3.83(2.16)	0.97	0.94	2.74(2.02)	0.97	0.94
Family's, Relatives', Friends' dan Peers' Influence (4)	4.01(2.18)	0.97	0.91	2.84(2.02)	0.98	0.95
Government Institutions' Influence (2)	4.23(2.27)	0.96	0.93	3.81(2.12)	0.97	0.95
PEOU (3)	5.77(1.59)	0.94	0.83	6.05(1.49)	0.95	0.86
Self-Efficacy (3)	5.89(1.68)	0.96	0.88	5.76(1.73)	0.97	0.92
Requisite Knowledge (4)	6.08(1.52)	0.96	0.86	6.28(1.24)	0.92	0.74
Perceived Ease of Use (4)	5.49(1.80)	0.94	0.80	4.96(1.87)	0.93	0.77
Time (3)	2.74(1.94)	0.93	0.83	3.18(1.95)	0.94	0.83
Accessibility (4)	2.80(1.48)	0.80	0.53	2.88(1.30)	0.71	0.55
Behavioral Intention (3)	4.91(2.37)	0.98	0.94	3.31(2.55)	0.97	0.96

*Nilai dalam kurung di konstruk adalah jumlah item.

†Rata-rata varian diekstraksi.

Contoh:

Penelitian Chau dan Hu (2002) menginvestigasi keputusan penerimaan teknologi *telemedicine* oleh individual dokter untuk pelayanan-pelayanan kesehatan di Hongkong.

Pengukuran Konstruk-Konstruk

Item-item pertanyaan yang digunakan untuk mengoperasikan konstruk-konstruk yang digunakan di penelitian ini adalah sebagai berikut ini.

1. Item-item untuk membentuk konstruk kegunaan persepsian (*perceived usefulness*) diadaptasi dari penelitian Davis (1989).

2. Item-item untuk membentuk konstruk kemudahan penggunaan persepsian (*perceived ease of use*) diadaptasi dari penelitian Davis (1989).
3. Item-item untuk membentuk konstruk kompatibilitas (*compatibility*) diadaptasi dari penelitian Taylor dan Todd (1995).
4. Item-item untuk membentuk konstruk pengaruh kolega (*peer influence*) diadaptasi dari penelitian Taylor dan Todd (1995).
5. Item-item untuk membentuk konstruk kontrol teknologi persepsian (*perceived technology control*) diadaptasi dari penelitian Taylor dan Todd (1995).
6. Item-item untuk membentuk konstruk sikap (*attitude*) diadaptasi dari penelitian Taylor dan Todd (1995).
7. Item-item untuk membentuk konstruk niat perilaku (*behavioral intention*) diadaptasi dari penelitian Davis (1998), dan Taylor dan Todd (1995).

Semua item-item tersebut diukur dengan menggunakan skala Likert 7 poin mulai dari "sangat-setuju" sampai dengan "sangat tidak-setuju." Untuk meyakinkan konsistensi jawaban dari responden-responden di daftar pertanyaan, separuh dari item-item yang ada di daftar pertanyaan dinegatipkan dan semua item-item tersebut diatur urutannya secara acak. Maksud diatur secara acak ini untuk mencegah responden memahami bahwa beberapa item mengarah ke konstruk yang sama, sehingga responden akan menjawab dengan monoton di nilai-nilai bawah saja yang disebut dengan efek lantai (*floor effect*) atau di atas saja yang disebut dengan efek langit-langit (*ceiling effect*).

Validitas Konstruk

Penelitian ini menggunakan teknik faktor analisis (*factor analysis*) untuk menguji validitas dari konstruk-konstruknya (disebut juga dengan *construct validity*). Lebih spesifik lagi,

faktor analisis yang digunakan adalah *iterated principal axis analysis* dengan *promax rotation* bukannya *principal components factor analysis* dengan suatu *varimax rotation*. Alasannya adalah karena kemungkinan faktor-faktornya akan saling berkorelasi. Secara umum, validitas konstruk (*construct validity*) dianggap memenuhi syarat atau konstruk dianggap valid jika item-item yang dimuat didalam konstruk nilainya tinggi (disebut dengan *factor loading*).

Tabel. Validitas konstruk.

Item	Muatan Faktor *(Factor Loading)*
***Perceived usefulness* (PU)**	
PU1	0,87
PU2	0,82
PU3	0,64
PU4	0,80
Compatibility (COM)	
C0M1	0,77
COM2	0,71
COM3	0,91
Perceived ease of use (PEOU)	
PE0U1	0,75
PE0U2	0,53
PE0U3	0,66
PE0U4	0,77
Peer influence (PIN)	
PINl1	0,50
PIN2	0,77
PIN3	0,99
Behavioral intention (Bl)	
BI1	0,77
BI2	0,63
BI3	0,80

Attitude (ATT)	
ATT1	0,79
ATT2	0,61
ATT3	0,57
Perceived technology control (PTC)	0,61
PTC1	0,43
PTC2	0,46
PTC3	0,46
PTC4	

Item-item di semua konstruk tampaknya mempunyai nilai *factor loading* yang cukup tinggi, kecuali beberapa item di konstruk *perceived technology control* (PTC) yang mempunyai nilai-nilai *factor loading* dibawah nilai 0,5. Ini menunjukkan bahwa konstruk ini kurang valid atau dengan kata lain item-item di konstruk (terutama item PTC2, PTC3, dan PTC4) bukan item yang valid untuk membentuk konstruk *perceived technology control* (PTC) dan seharusnya tidak dimuat (*load*) di konstruk ini.

Reliabilitas

Konstruk-konstruk perlu juga diuji reliabilitasnya (*reliability*). Uji yang dilakukan adalah dengan nilai *Cronbach alpha*. Tabel berikut ini menunjukkan hasil dari uji reliabilitas.

Tabel 6. Reliabilitas konstruk-konstruk.

Konstruk	Rata-rata	Deviasi Standar	Reliabilitas (*Alpha*)
Attitude (ATT)			
ATT1	2,83	1,03	
ATT2	2,96	1,16	0,69
ATT3	2,59	1,15	

***Perceived technology control* (PTC)**			
PTC1	2,85	1,08	
PTC2	3,94	1,34	0,55
PTC3	2,88	1,39	
PTC4	3,85	1,31	
***Perceived usefulness* (PU)**			
PU1	2,80	1,17	
PU2	3,04	1,20	
PU3	3,30	1,22	0,86
PU4	2,93	1,16	
Perceived ease of use (PEOU)			
PE0U1	3,02	1,30	
PE0U2	3,46	1,16	0,77
PE0U3	3,11	1,60	
PE0U4	3,21	1,18	
Compatibility (COM)			
C0M1	3,46	1,30	
COM2	3,23	1,24	0,83
COM3	3,28	1,16	
Peer influence (PIN)			
PIN11	3,66	1,07	
PIN2	3,30	1,03	0,75
PIN3	3,36	1,07	
Behavioral intention (BI)			
BI1	3,47	1,28	
BI2	2,98	1,26	0,86
BI3	3,23	1,26	

Hampir semua konstruk mempunyai reliabilitas yang cukup tinggi, hanya konstruk *perceived technology control* (PTC) saja yang mempunyai reliabilitas yang rendah, yaitu dengan nilai *Cronbach alpha* sebesar 0,55 yang sedikit melebihi batas ambang 0,5.

Contoh:

Data di penelitian Compeau dan Higgins (1995) dianalisis menggunakan PLS (*Partial Least Square*). Analisis PLS melibatkan dua tahapan. Tahapan pertama adalah menilai model pengukurannya, termasuk reliabilitas dan validitas diskriminan dari pengukur-pengukur. Tahap kedua adalah penilaian dari model strukturalnya. Penelitian ini menunjukkan bahwa dari uji validitas dan reliabilitasnya diketahui bahwa konstruk-konstruk di model perlu direvisi.

> Untuk penilaian model pengukuran (*measurement model*), individual muatan-muatan item dan konsistensi internal yang nilai lebih besar dari 0,70 dianggap baik (Fornell dan Larcker, 1981). Untuk validitas diskriminan, item-item seharusnya memuat lebih tinggi untuk konstruk-nya sendiri dibandingkan dengan konstruk-konstruk yang lain di model, dan rata-rata varian antara konstruk-konstruk dan pengukuran-pengukuran mereka seharusnya lebih besar dibandingkan dengan varian-varian antara konstruk-konstruk mereka sendiri.

Pengukuran Model

Pengukur-pengukur dari empat konstruk yaitu konstruk dukungan (*support),* konstruk keyakinan-sendiri (*self-efficacy*), konstruk perasaan (*affect*), dan konstruk penggunaan (*use*) memenuhi syarat reliabilitas dan validitas diskriminan. Dengan demikian, tidak ada perubahan apapun yang dilakukan untuk keempat konstruk-konstruk ini. Akan tetapi, konstruk-konstruk yang lainnya mengalami permasalahan-permasalahan dalam pengukurannya. Permasalahan-permasalahan yang terjadi dan perbaikannya dijelaskan berikut ini.

1. Tiga item di konstruk dorongan oleh orang-orang lain (*encouragement by others*) tidak berkorelasi tinggi dengan item-item pengukur yang lain. Dorongan menggunakan dari keluarga, dari kolega, dan dari bawahan-bawahan

berkorelasi rendah dengan item dorongan menggunakan dari teman-teman kerja dan manajer-manajer. Dengan demikian tiga item ini dikeluarkan dari model untuk pengujian selanjutnya.

2. Permasalahan yang sama terjadi pada pengukur-pengukur di pemakaian oleh orang-orang lain (*others' use*). Pemakaian sesungguhnya oleh keluarga, oleh teman-teman, dan oleh bawahan-bawahan tidak memuat tinggi ke konstruk. Dengan demikian, ketiga item ini juga dikeluarkan dari model.
3. Pemeriksaan terhadap muatan-muatan untuk ekspektasi-ekspektasi hasil (*outcome expectations*) menunjukkan bahwa kemungkinan terjadi dimensi-dimensi yang redunden di konstruk ini. Di dimensi ini, tampaknya mewakili skala-skala hasil-hasil yang hubungan dengan pekerjaan (*job related outcomes*) dan yang lainnya berhubungan dengan hasil-hasil personal (*personal outcomes*). Hasil-hasil yang berhubungan dengan pekerjaan (*job related outcomes*) termasuk item-item semacam "*If I use a computer, I will increase the quality of output of my job*," sedang hasil-hasil personal (*personal outcomes*) termasuk "*If I use a computer, I will increase my sense of accomplishment.*" Untuk model yang direvisi, konstruk ekspektasi-ekspektasi hasil-hasil (*outcomes expectations*) dipecah menjadi dua konstruk, yaitu ekspektasi-ekspektasi hasil berhubungan-kinerja (*performance-related outcome expectations*) dan ekspektasi-ekspektasi hasil berhubungan-personal (*personal-related outcome expectations*).
4. Muatan-muatan item individual di konstruk kecemasan (*anxiety*) menunjukkan hasil yang lemah, menunjukkan suatu masalah di pengukur konstruk kecemasan ini. Pemeriksaan kembali pengukur-pengukur dan dari hasil analisis faktor eksploratori menunjukkan sejumlah dimensi-dimensi. Ray dan Minch (1980) juga menemukan bahwa konstruk ini merupakan konstruk multi-dimensi. Di penelitian ini, disamping dimensi kecemasan (*anxiety*) itu sendiri, terdapat pula dimensi-dimensi

bagian yaitu suatu keinginan untuk belajar lebih tentang komputer-komputer (*a desire to learn more about computers*), kepercayaan-kepercayaan tentang belajar untuk menggunakan komputer-komputer (*beliefs about learning to use computers*), dan kepercayaan-kepercayaan tentang kemanfaatan komputer-komputer di bisnis dan edukasi (*beliefs abot the appropriateness of computers in business and education*). Tiga dimensi bagian ini kemungkinan mempengaruhi pem-bentukan konstruk kecemasan (*anxiety*) yang tidak membentuk kecemasan (*anxiety*) itu sendiri. Dengan demikian hanya dimensi utama saja yang akan digunakan dan ke tiga dimensi bagian tidak digunakan. Untuk dimensi utamanya, empat item ditemukan muat di dimensi ini yang diperoleh dari hasil analisis faktor. Item-item ini sudah benar menunjukkan perasaan kecemasan menggunakan komputer, tidak hanya sekedar percaya (*belief*) bahwa menggunakan komputer akan menimbulkan kecemasan.

Pengujian Model Pengukuran

Koefisien-koefisien reliabilitas dan validitas diskriminan tampak di Tabel berikut ini.

Tabel. Reliabilitas dan validitas diskriminan.

Konstruk	RKI[+]	1	2	3	4	5	6	7	8	9
1. Dorongan (*encouragement*)	0,87	0,80								
2. Pemakaian oleh orang-orang lain (*others' use*)	0,80	0,52*	0,72							
3. Dukungan (*support*)	0,91	0,24*	0,18*	0,79						
4. Keyakinan-sendiri (*self-efficacy*)	0,95	0,20*	0,18*	-0,10**	0,81					
5. Ekspektasi-ekspektasi hasil - Kinerja	0,87	0,27*	0,22*	-0,09**	0,32*	0,72				
6. Ekspektasi-ekspektasi hasil – Personal	0,87	0,19*	0,11**	-0-,12*	0,17*	0,49*	0,76			
7. Perasaan (*affect*)	0,87	0,20*	0,15*	-0,13*	0,49*	0,48*	0,32*	0,75		
8. Kecemasan (*anxiety*)	0,87	-0,11**	-0,07	-0,00	-0,50*	-0,23*	-0,05	-0,51*	0,79	
9. Penggunaan (*use*)	0,82	0,17*	0,24*	-0,05	0,45*	0,41*	0,24*	0,47*	-0,37*	0,73

Keterangan: Elemen-elemen diagonal adalah akar kuadrat dari varian bersama antara konstruk-konstruk dengan pengukur-pengukurnya. Elemen-elemen di luar diagonal adalah korelasi-korelasi kuadrat antara konstruk-konstruk. Untuk validitas diskriminan, elemen-elemen diagonal harus lebih besar dari elemen-elemen di luar diagonalnya.
* $p<0,01$, ** $p<0,05$
+ reliabiltas konsistensi internal (*internal consistency reliability*).

Untuk reliabilitas, semua konstruk mempunyai nilai koefisien reliabilitas berkisar dari 0,81 sampai dengan 0,95 yang lebih besar dari batas minimal yang disyaratkan, yaitu 0,70. Hasil ini menunjukkan bahwa semua konstruk dapat diandalkan.

Untuk validitas diskriminan, varian yang dibagikan antara masing-masing konstruk dan pengukur-pengukurnya (nilai elemen di diagonal di Tabel) melebihi nilai varian yang dibagikan antara konstruk-konstruk (nilai elemen di luar diagonal). Dengan demikian, model pengukuran valid.