

PERKULIAHAN KE 8

Tujuan Instruksional Khusus (TIK)

- Menjelaskan metodologi dan model untuk mendukung desain sistem interaktif dan tujuan yang akan diperoleh
- Menyebutkan beberapa jenis pengambilan evaluasi
- Menyebutkan beberapa cara untuk mengevaluasi desain dan implementasi
- Menjelaskan pemilihan beberapa metode evaluasi.

Pokok Bahasan : Evaluasi pengujian dari sistem interaktif secara fungsional, Pengambilan evaluasi, Beberapa pendekatan desain evaluasi, Beberapa pendekatan untuk mengimplementasikan evaluasi.

Deskripsi Singkat : bahasan ini tentang bagaimana melakukan evaluasi terhadap tahap perancangan dan tahap implementasi.

Bahan Bacaan :

Dix, Alan et.al, *HUMAN-COMPUTER INTERACTION*, Prentice Hall, Europe, 1993, hal 363-399

Johnson, P., *HUMAN-COMPUTER INTERACTION : Psychology, Task Analysis and Software Engineering*, McGraw-Hill, England UK, 1992

TEKNIK EVALUASI

Tinjauan

- Evaluasi menguji Kegunaan (usability) dan Fungsi (functionality) dari Sistem Interaktif
- Evaluasi dapat dilakukan pada :
 - Laboratorium
 - Lapangan Pekerjaan
 - Kerja Sama dengan User
- Beberapa Pendekatan yang mengevaluasi perancangan/ desain :
 - Metode Analitik
 - Metode Review
 - Metode Berbasis Model
- Beberapa Pendekatan yang mengevaluasi implementasi :
 - Metode Experimental
 - Metode Observasi
 - Metode Query
- Metode Evaluasi Harus Dipilih Secara Cermat Dan Harus Cocok Dengan Pekerjaan

Evaluasi

- Evaluasi digunakan untuk melihat apakah hasil rancangan dengan proses uji coba system yang telah dibuat sesuai dengan permintaan pengguna (user)
- Proses ini tidak dikerjakan dalam satu fase proses perancangan tetapi melalui perancangan dengan prinsip life cycle, dengan hasil dari evaluasi dikembalikan untuk memodifikasi perancangan.

Tujuan Evaluasi

Ada 3 tujuan utama dari evaluasi :

- Melihat seberapa jauh sistem berfungsi.
Desain system memungkinkan user melakukan tugas yang dibutuhkan dengan lebih mudah. Ini tidak hanya membuat fungsionalitas yang sesuai ada di system, tetapi membuat mudah dicapai oleh user, user dapat melakukan aksi untuk melaksanakan tugas. Juga mencakup kesesuaian penggunaan system terhadap harapan user pada tugas tersebut. Evaluasi pada tahap ini meliputi pengukuran unjuk kerja dari user pada system, untuk melihat keefektifan system dalam mendukung tugas.

- Melihat efek interface bagi pengguna.
Ini mencakup pertimbangan aspek dari kemudahan system dipelajari, usability dan perilaku user. Penting juga untuk mengidentifikasikan area desain yang berlebih dari user, dengan menggunakan sejumlah informasi yang berlebih.
- Mengidentifikasi problem khusus yang terjadi pada system
Ketika menggunakan konteks yang diinginkan menyebabkan hasil yang tidak diinginkan, atau terjadi kekacauan diantara user. Ini tentunya berhubungan dengan usability dan fungsionalitas dari desain (bergantung pada sebab masalah). Tujuan ini merupakan aspek negative dari desain.

Jenis- Jenis Evaluasi

Ada 2 jenis utama dari evaluasi, yaitu yang dilaksanakan pada kondisi laboratorium dan yang dilaksanakan pada lingkungan kerja atau 'lapangan'.

- Dalam kondisi percobaan (Laboratory)
- Dalam kondisi lokasi kerja sebenarnya
- Participatory Design

Pada Kondisi Percobaan (Laboratory)

Penggunaan pengujian system ini pada ruang percobaan mempunyai beberapa kondisi diantaranya

- Laboratorium yang bagus biasanya memiliki fasilitas perekaman audio/visual yang baik, cermin dua arah, komputer beserta perlengkapannya yang mungkin tidak ada pada lokasi kerja sebenarnya.
- Operator bebas dari gangguan yang menghambat pekerjaan.
- Beberapa situasi hanya dapat dilakukan di lab., seperti sistem yang akan digunakan ditempatkan pada lokasi yang berbahaya atau lokasi yang terpencil, contoh stasiun ruang angkasa.
- Dapat memanipulasi situasi untuk memecahkan masalah dan melihat sedikit penggunaan prosedur atau membandingkan beberapa alternatif perancangan dengan situasi yang sebenarnya.
- Situasi pada laboratorium tidak dapat menggambarkan situasi ruang kerja sebenarnya dan terdapat beberapa orang yang tidak bisa bekerja pada kondisi di laboratorium.

Pada Kondisi Lokasi Kerja Sebenarnya (Lapangan)

Penggunaan pengujian system ini pada lokasi kerja sebenarnya mempunyai beberapa kondisi diantaranya :

- Tingkat gangguan yang melebihi ambang batas, tingkat-tingkat pergerakan yang besar dan interupsi yang tetap, seperti panggilan telepon menyebabkan observasi ini sulit dilakukan.
- Situasi yang lebih "terbuka" antara system dan pengguna, dimana kondisi ini tidak ditemukan pada kondisi di laboratorium
- Observasi pada lokasi kerja sebenarnya lebih baik dilakukan daripada di laboratorium dan gangguan-gangguan yang terjadi pada lokasi ini digunakan sebagai situasi yang mewakili situasi sebenarnya dan digunakan untuk proses penyimpanan dan pengambilan selama tugas.

Participatory Design

- Adalah suatu pemikiran yang melibatkan keseluruhan alur perancangan dan tidak hanya proses evaluasi saja.
 - Perancangan ini dilakukan pada ruang kerja yang melibatkan pengguna yang tidak hanya digunakan sebagai subyek percobaan tetapi juga sebagai anggota yang aktif dalam team perancangan.
 - Argumennya adalah user dianggap ahli dalam konteks dan perancangan hanya dapat efektif pada konteks tersebut jika para ahli ini memberikan kontribusi yang aktif ke perancangan.
- Mempunyai tiga karakteristik :
 - Meningkatkan lingkungan kerja dan tugas. Ini membuat perancangan dan evaluasi berorientasi kerja atau konteks dibanding berorientasi system.
 - Mempunyai sifat kerja sama (kolaborasi), yakni pengguna dilibatkan sebagai anggota team dan mempunyai kontribusi pada setiap tingkat perancangan.
 - Mempunyai pendekatan iterative, perancangan adalah suatu subyek untuk evaluasi dan revisi pada setiap tingkatan.
 - Beberapa metode untuk membantu mengkomunikasikan informasi antara pengguna dan perancang :
 - Brainstorming
 - Storyboarding
 - Workshops
 - Pensil dan kertas percobaan

Metode-metode ini tidak semuanya digunakan secara eksklusif dalam participatory design. Metode ini digunakan untuk memberikan pemahaman antara perancang dan pengguna.

Mengevaluasi Perancangan

Evaluasi terjadi setelah proses perancangan. Evaluasi pertama system idealnya dilakukan sebelum implementasi dimulai. Jika perancangan dievaluasi, kesalahan dapat dihindari karena perancangan diubah/ diperbaiki sebelumnya. Sejumlah metode dibuat untuk mengevaluasi perancangan sebelum implementasi, yaitu :

- Cognitive Walkthrough
- Heuristic Evaluation
- Review based
- Model based

Cognitive Walkthrough

Suatu usaha yang dilakukan untuk mengenalkan teori psikologi ke dalam bentuk informal dan subyektif atau dengan kata lain mempunyai tujuan untuk mengevaluasi perancangan dengan melihat seberapa besar dukungan yang diberikan ke pengguna untuk mempelajari beberapa tugas yang diberikan. Pendekatan ini dikemukakan oleh Polson, dkk. Walkthrough dilaksanakan oleh perancang atau seorang ahli dalam psikologi kognitif. Ahli bekerja melalui perancangan tugas tertentu, tahap demi tahap, mengidentifikasi masalah yang berpotensi terhadap criteria psikologi. Kemudian dibandingkan ke proses dimana perancang software akan bekerja dengan koding pada kondisi yang berbeda (menggunakan set data yang berbeda atau kondisi yang

salah, misalnya) untuk mengevaluasi unjuk kerja setiap software. Dalam pendekatan ini terdapat beberapa issue yang timbul seperti :

- Pengaruh apa yang timbul setelah tugas ini diberikan ke pengguna?
- Proses cognitive apa yang tersedia?
- Masalah pembelajaran apa yang seharusnya timbul?

Analisis difokuskan pada tujuan user dan pengetahuan. Cognitive walkthrough harus menunjukkan jika dan bagaimana interface merujuk user untuk membangkitkan tujuan yang benar dari pelaksanaan tugas yang diinginkan, dan memilih aksi yang diperlukan untuk memenuhi setiap tujuan. Untuk melakukan Cognitive Walkthrough harus mempunyai informasi yang dibutuhkan :

- Deskripsi dari suatu interface yang dibutuhkan itu sendiri
- Deskripsi dari tugas termasuk usaha yang benar untuk melakukannya dan struktur tujuan untuk mendukungnya

Dengan Informasi ini maka Evaluator dapat Melakukan langkah dari walkthrough :

- Pilih Tugas
- Deskripsikan Tujuan awal dari user
- Lakukan kegiatan/ aksi yang tepat
- Analisa proses keputusan untuk setiap kegiatan.

Cognitive Walkthrough berbasis formulir, yang disediakan untuk merujuk evaluator melalui sekumpulan pertanyaan yang berhubungan dengan tugas user dan tujuan. Cognitive Walkthrough merupakan metode analitik yang dapat digunakan untuk fase perancangan awal maupun untuk mengevaluasi system yang ada, berisi informasi yang berhubungan dengan kemampuan belajar (learnability) dari interface. Jika digunakan di awal perancangan, dapat mengidentifikasi masalah perancangan sebelum tahap prototype dan langsung difokuskan pada evaluasi selanjutnya.

Meskipun dirancang untuk digunakan perancang sendiri, tetapi perlu juga pengetahuan teori psikologi dan terminologinya agar efektif.

Contoh : Memprogram Video Dengan Remote Control. Misalkan akan memprogram video ke waktu dimulai dari jam 18.00 dan berakhir pada jam 19.15 pada channel 4 pada tanggal 4. Maka tugas tersebut adalah :

- Set waktu awal
- Set waktu akhir
- Set channel
- Set tanggal

Ini adalah tujuan dari user, memasukkan pengalaman dan pengetahuan sebelumnya. Ekspresi di atas adalah pada tingkat tinggi, tidak diberikan detail pada level interface. Dengan melihat tujuan user yang diidentifikasi, setiap aksi digambarkan dengan tugas yang lengkap dan menentukan apakah interface mendukung pembangkitan struktur tujuan yang dibutuhkan untuk aksi yang lengkap. Setiap aksi dianalisa pada formulir walkthrough. Formulir menampilkan sejumlah pertanyaan bagi evaluator sebagai pertimbangan. Contoh di atas adalah pertama yang dilakukan adalah melakukan penekanan tombol “timed recording”.

Aksi #1

Deskripsi : tekan tombol “timed recording”

.....
.....

Langkah selanjutnya adalah membandingkan struktur tujuan yang tersedia dengan struktur tujuan pada langkah awal. Hal ini dilakukan dengan empat pertanyaan:

- Akankah user gagal untuk menambah tujuan yang diperlukan ?
- Akankah user gagal untuk memindahkan tujuan yang tidak diinginkan ?
- Akankah user mengadopsi beberapa tujuan palsu berdasarkan interface ?
- Akankah user membatalkan tujuan yang masih dibutuhkan ?

Heuristic Evaluation

Diusulkan oleh Nielsen dan Molich, hampir sama dengan Cognitive Walkthrough tetapi sedikit terstruktur dan sedikit terarah. Pada pendekatan ini, sekumpulan criteria usability atau heuristic diidentifikasi dan perancangan dilaksanakan misalnya dimana criteria dilanggar. Dalam system ini terdapat beberapa criteria :

- Perilaku Sistem dapat dipastikan.
- Perilaku Sistem konsisten.
- Feedback tersedia.
- Kemampuan memori user tidak melebihi batas.
- Dialog merupakan orientasi tugas.

Tujuan dari Heuristic Evaluation adalah untuk memperbaiki perancangan secara efektif. Evaluator melakukan evaluasi melalui kinerja dari serangkaian tugas dengan perancangan dan dilihat kesesuaiannya dengan kriteria setiap tingkatan. Jika ada kesalahan terdeteksi maka perancangan dapat ditinjau ulang untuk memperbaiki masalah ini sebelum tingkat implementasi.

Pendekatan ini mudah dan relative cepat. Tidak seluruhnya subyektif, dimana criteria khusus digunakan untuk merujuk evaluasi. Selain itu juga membutuhkan level tertentu dari pengetahuan untuk mengaplikasikan heuristic. Misalnya untuk menentukan jika suatu perancangan melebihi memori user, perancang perlu mengetahui kemampuan memori manusia. Evaluasi heuristic tidak mudah dan seorang evaluator kemungkinan terlewatkan pada masalah perancangan, meskipun menggunakan criteria. Karena itu adalah latihan yang berguna untuk melaksanakan khususnya dalam kelompok evaluator kecil supaya melihat perancangan secara bebas dan menyusun hasilnya.

Review Based Evaluation

Evaluasi antara psikologi eksperimen dengan interaksi manusia dan komputer menghasilkan hasil-hasil eksperimen yang baik dan pengalaman yang nyata. Beberapa diantaranya dari domain khusus ke umum, tetapi kebanyakan berhubungan dengan isu generic dan teraplikasi pada berbagai situasi. Misalnya pada usability dari tipe menu yang berbeda, pemanggilan nama perintah dan pemilihan icon.

Dalam kenyataannya hasil eksperimen ini tidak dapat dipastikan mempertahankan keadaan yang tetap. Evaluator harus memilih data secara hati-hati, menunjuk rancangan eksperimen yang dipilih, subyek masyarakat yang digunakan, analisa pelaksanaan dan asumsi yang dibuat. Misalnya pengujian eksperimen, usability dari jenis system 'bantu' umum yang menggunakan subyek baru tidak menyediakan evaluasi yang tepat dari system 'bantu' yang dirancang bagi user ahli.

Model Based Evaluation

Pendekatan terakhir untuk mengevaluasi perancangan dengan mengkombinasi spesifikasi perancangan dan evaluasi ke dalam kerangka kerja yang sama. Contoh GOMS model, keystroke level model dan design rationale.

Mengevaluasi Implementasi

Perbedaan yang besar dengan evaluasi perancangan adalah keberadaan implementasi system yang ada dalam berbagai bentuk. Hal ini dapat dimulai dari simulasi kemampuan interaktif system, sebagai contoh Wizard of Oz, melalui fungsi prototype dasar sampai dengan system yang telah diimplementasi secara keseluruhan.

Metode Empirik : Evaluasi Eksperimen

Satu metode yang sangat kuat dalam mengevaluasi perancangan atau aspek dari perancangan adalah menggunakan eksperimen yang terkontrol. Ini menyediakan kenyataan empiris untuk mendukung klaim umum atau hipotesis.

Setiap eksperimen memiliki dasar yang sama. Evaluator memilih hipotesa untuk diuji, yang dapat ditentukan dengan mengukur beberapa atribut dari perilaku subyek. Sejumlah kondisi eksperimen ditentukan yang berbeda dalam hal nilai dari variable tetap yang terkontrol. Setiap perubahan dalam pengukuran perilaku diatributkan ke kondisi yang berbeda. Dalam bentuk dasar ada sejumlah factor penting terhadap keseluruhan kehandalan eksperimen, yang harus dipertimbangkan dalam perancangan eksperimen. Ada pun factor-faktor tersebut adalah :

- Subyek
 - Pemilihan subyek sangat penting dalam beberapa eksperimen.
 - Harus setepat mungkin dengan keinginan user.
 - Jika subyek bukan user sebenarnya maka subyek yang dipilih harus sama usia dan tingkatan pendidikan sebagai kelompok user yang dimaksud. Pengalaman menggunakan komputer secara umum dan keterhubungan system yang sedang dalam pengujian sejajar dengan pengalaman atau pengetahuan dari domain tugas.
 - Ukuran sample yang dipilih, kebanyakan ditentukan dengan pertimbangan pragmatis, kehandalan dari user terbatas. Ukurannya juga harus cukup besar untuk mewakili populasi yang diambil dalam eksperimen perancangan dan metode statistika yang dipilih.
- Variabel
 - Independent Variabel
Karakteristik suatu eksperimen yang memanipulasi untuk menghasilkan kondisi yang berbeda sebagai perbandingan.
Contoh : jenis interface, level dari help, jumlah item menu dan perancangan icon. Setiap variable ini dapat memiliki nilai yang berbeda; masing2 digunakan dalam eksperimen sebagai level variable.
 - Dependent Variabel
Variabel yang dapat diukur dalam eksperimen. Contoh pada variable independent adalah kecepatan dari pemilihan menu. Variable dependent harus diukur berdasarkan pengaruh variable independent dan tidak oleh yang lain. Pilihan umum dari variable dependent dalam eksperimen evaluasi adalah waktu untuk menyelesaikan tugas, jumlah kesalahan yang dibuat, preferensi user dan kualitas dari unjuk kerja user.

- Hypothesis
 - Merupakan prediksi yang dihasilkan dalam eksperimen.
 - Masih menggunakan variable independent dan dependent, dimana variasi di dalam independent variable akan menyebabkan perbedaan pada dependent variable.
 - Tujuan dari eksperimen adalah untuk menunjukkan bahwa perkiraan benar dengan menyangkal adanya null hypothesis, yang menyatakan bahwa tidak ada perbedaan dalam dependent variable diantara tingkatan dari independent variable.
 - Pengukuran statistika menunjukkan bahwa nilai hasil dapat dibandingkan dengan tingkat kepercayaan. Jika nilai tertentu ada, pada tingkat kepercayaan, hipotesis null tidak benar.

Perancangan Eksperimen

- Between-Groups (Randomized)
 - Masing-masing subyek diberikan kondisi yang berbeda yakni kondisi eksperimen (variable sudah dimanipulasi) dan control, yang identik dengan kondisi eksperimen kecuali untuk manipulasi ini. Control ini memastikan bahwa manipulasi bertanggungjawab untuk setiap perbedaan yang diukur.
 - Keuntungan perancangan ini adalah setiap user menghasilkan satu kondisi
 - Kerugiannya adalah dengan semakin banyak jumlah subyek yang tersedia akan menyebabkan hasilnya akan berkurang dan perbedaan antar setiap individu akan membuat bias hasil. Hal ini dapat diatasi dengan memilih dengan hati-hati subyek yang dipilih dan menjamin setiap kelompok terwakili.
- Within-Groups
 - Setiap user akan menampilkan kondisi yang berbeda
 - Jumlah user yang tersedia lebih sedikit
 - Pengaruh dari subyek lebih sedikit

Pemilihan metode eksperimen bergantung dari sumber yang tersedia, bagaimana transfer pembelajaran terjadi atau dapat dikontrol, dan perwakilan yang bagaimana dari kelompok user dipertimbangkan.

Sekali kita menentukan hipotesa yang akan diuji, variable yang akan dipelajari, subyek yang ada, dan perancangan yang sesuai, kita harus menentukan hasil analisa yang bagaimana yang akan diambil/ disimpan.

Pengukuran Statistik

Dua aturan dalam analisa dengan statistik yakni : melihat data dan menyimpan data. Menyimpan data penting jika kita akan melakukan metode analisis yang berbeda. Pemilihan analisis statistika bergantung pada tipe data dan pertanyaan yang perlu dijawab. Variabel yang digunakan terbagi dalam :

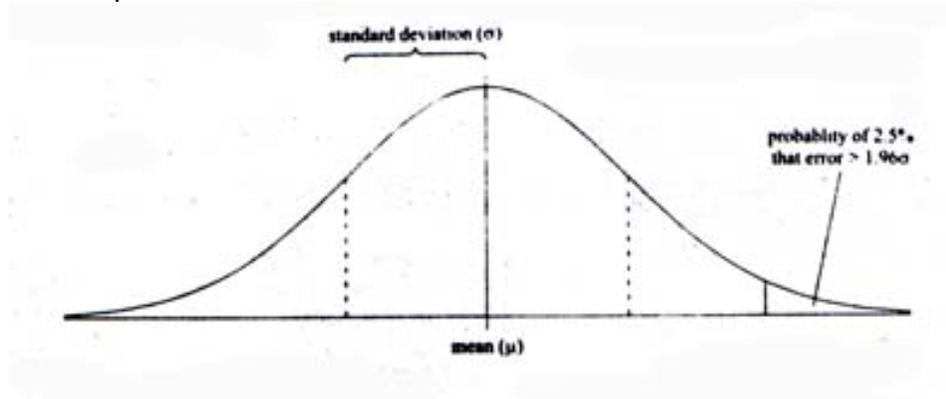
- Discrete Variables
 - Menggunakan jumlah yang terbatas dari suatu nilai atau tingkatan
Contoh: Warna layar monitor yakni : red, green atau blue.
 - Merupakan suatu independent variable
- Continous Variable

- o Menggunakan jumlah berapapun (batas atas maupun batas bawah)
Contoh : Ketinggian seseorang atau waktu yang diperlukan untuk menyelesaikan pekerjaan.
- o Merupakan suatu dependent variable

Jika variable dependent diukur dan subyeknya variasi dari eksperimen yang acak, bila variable kontinyu maka variasi acak menjadi khusus. Jika bentuk data mengikuti distribusi yang dikenal maka uji statistic khusus dan yang lebih kuat digunakan. Uji ini dikenal dengan tes parametric dan kebanyakan variasi menggunakan distribusi normal. Jika datanya tidak terdistribusi normal, maka pengujian non-parametrik yang digunakan. Uji kontijensi adalah ketika kita mengelompokkan data dengan beberapa atribut diskrit kemudian menghitung sejumlah item data dengan masing2 kombinasi atribut.

	Independent Variable	Dependent Variable	
Parametric			
	2 valued	Normal	Student's test on difference of means
	Discrete	Normal	ANOVA (Analysis Of Variance)
	Continuous	Normal	Linear (or non-linear) regression factor analys
Non-parametric			
	2 valued	Continuous	Wilcoxon (or Mann-Whitney) rank-sum test
	Discrete	Continuous	Rank-sum version of ANOVA
	Continuous	Continuous	Spearman's rank correlation
Contingency test			
	2 valued	Discrete	No special test, see next entry
	Discrete	Discrete	Contingency table and X squared test
	Continuous	Discrete	(rare) group independent variable and then as above

Contoh parametric test



Contoh non-parametric statistik

Kondisi A : 33,42,25,79,52
 Kondisi B : 87,65,92,93,91,55

Setelah diurut mulai 25,33,42,...,92,93 kemudian setelah di ranking maka 25 jadi 1, 33 jadi 2 dan seterusnya.

Kondisi A : 2,3,1,7,4
 Kondisi B : 8,6,10,11,9,5

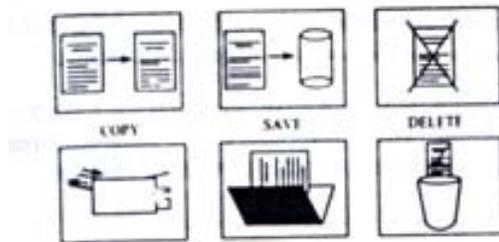
Dengan Wilcoxon test, setiap kondisi dijumlah rankingnya kemudian dikurangkan dengan nilai terkecilnya sehingga didapatkan nilai statistic U nya :

Kondisi A : $1+2+3+4+5 = 15$
 Kondisi B : $1+2+3+4+5+6 = 21$

	Rank sum	smallest	U
Kondisi A	: $(2+3+1+7+4)$	- 15	= 2
Kondisi B	: $(8+6+10+11+9+5)$	- 21	= 28

Ternyata jumlah dari kedua statistic U tersebut, $2 + 28 = 30$ merupakan perkalian dari jumlah nilai data dari masing2 kondisi 5×6 .

Contoh mengevaluasi perancangan icon



dimisalkan kita akan membuat paket pengolah kata baru yang menggunakan icon untuk presentasinya. Ada 2 jenis perancangan icon dan akan digunakan perancangan yang paling mudah bagi pengguna-user. Jenis pertama menggunakan gambar natural (berdasarkan metafora dari dokumen kertas); jenis yang lain menggunakan gambar abstrak seperti terlihat pada gambar di atas. Yang pertama harus dilakukan adalah membentuk hipotesis : tampilan yang seperti apa yang diinginkan ? Pada kasus ini kebanyakan akan mengambil gambar natural sebagai icon yang lebih mudah dalam pemanggilan karena lebih familiar bagi pengguna. Hipotesis ini mengidentifikasi variable independent untuk eksperimen, natural dan abstrak. Pada pertimbangan variable dependent, pengukuran 'lebih mudah' dilihat dari jumlah kesalahan dalam memilih dan waktu yang digunakan untuk memilih sebuah icon. Control eksperimen dilakukan dengan membuat interface yang identik dengan perancangan icon dan tugas pemilihan dibuat yang dapat diulang untuk setiap kondisi. Kemudian metode eksperimen dilakukan dengan menggunakan kelompok user yang dimaksud. Mengingat pembelajaran terlibat dalam eksperimen ini, maka perancangan within-groups digunakan dimana setiap user menggunakan kedua set icon tersebut. Kemudian hasil eksperimen secara detail dianalisis.

Teknik Observasi

- *Think aloud*

Cara yang populer dalam mengumpulkan informasi dengan melihat interaksi user. Evaluator melihat dan merekam aksi pengguna-user (dengan berbagai teknik). Konsekuensinya user ditanyai untuk mengelaborasi aksinya dengan 'thinking aloud': menjelaskan : apa yang mereka percaya terjadi, mengapa mereka mengambil tindakan, apa yang mereka coba kerjakan.

Think aloud mempunyai keuntungan yakni sederhana, membutuhkan sedikit pengetahuan untuk menampilkannya dan menyediakan pengertian yang berguna dengan interface, juga dapat digunakan untuk mengobservasi bagaimana system digunakan.

Variasi lain adalah cooperative evaluation (Monk, dkk) dimana

- Proses mempunyai hambatan yang sedikit sehingga lebih mudah dipelajari oleh evaluator.
- User mempunyai dorongan untuk menguji system
- Evaluator dapat menjelaskan kekacauan yang terjadi dan memaksimalkan pendekatan yang efektif untuk mengenali masalah

- *Analisa Protokol*

Terdapat beberapa metode untuk merekam kegiatan user :

- Paper and pencil
- Audio recording
- Video recording
- Computer logging
- User notebook

Dalam prakteknya, kebanyakan campuran dari metode perekaman digunakan.

- *Automatic protocol analysis tools*

EVA (Experimental Video Annotator)

System prototype yang berjalan pada multi-media workstation yang dihubungkan langsung ke video recorder. Evaluator dapat menemukan sekumpulan tombol yang mengindikasikan event yang berbeda. Tombol digunakan dalam sesi perekaman untuk mencatat video dengan notes. Selama sesi user bekerja pada workstation dan direkam menggunakan video dan juga logging system dan audio. Evaluator menggunakan tombol untuk mencatat event yang penting dari user dan merekam notes tambahan pada editor teks. Setelah sesi evaluator dapat menanyakannya dan menggunakannya untuk mencari informasi. Contoh : Workplace project pada Xerox PARC

- *Post-task walkthroughs*

Sering data didapat melalui observasi langsung tanpa interpretasi. Kita memiliki aksi dasar yang dapat digunakan, tetapi sedikit pengetahuan tentang kenapa. Bahkan pada saat subyek didorong untuk think aloud melalui tugas, informasi dapat saja berada pada level yang salah. Misalnya saat subyek mengatakan 'dan sekarang saya memilih menu undo', tetapi tidak mengatakan kesalahan apa yang dilakukan sehingga perintah undo diperlukan.

Walkthrough dibuat untuk mengatasi masalah tersebut dengan merefleksikan kembali ke subyek aksi mereka setelah event. Apakah ditulis atau direkam, diputar kembali ke subyek untuk memberikan komentarnya atau langsung ditanya oleh

analisis. Ini dapat langsung dilakukan jika subyek mengingat mengapa aksi tertentu dilaksanakannya atau setelah beberapa saat ketika jawaban merupakan interpretasi subyek setelah aksi. Keuntungannya adalah bahwa analisis memiliki waktu untuk menanyai pertanyaan yang sesuai dan memfokuskan pada aksi khusus.

Terdapat beberapa keadaan yang menyebabkan subyek tidak diharapkan berbicara selama observasi sebenarnya, seperti subyek sedang mengerjakan tugas yang sangat penting dan banyak. Pada keadaan ini post-task walkthroughs dapat melihat secara subyektif perilaku user.

Teknik Query

- *Interview*

Menginterview user tentang pengalaman mereka dengan system interaktif yang menyediakan informasi secara langsung dan terstruktur. Keuntungannya bahwa tingkat pertanyaan dapat disesuaikan dengan konteks dan evaluator dapat lebih mendalami pertanyaan (sering pertanyaan “mengapa ...?” atau “bagaimana jika...?” untuk mengelaborasi aspek dari respon user.

Interview efektif pada evaluasi tingkat tinggi, khususnya dalam memperoleh informasi tentang preferensi user, impresi dan perilaku. Jika digunakan bersama dengan observasi, berarti mencari klarifikasi sebuah kejadian. Agar efektif, interview perlu direncanakan, dan dibuat pertanyaan2nya.

- *Questionnaire*

Metode alternatif yang agak kurang fleksibel dibandingkan dengan interview karena pertanyaan sudah tersedia tetapi dapat meraih subyek yang banyak dan membutuhkan waktu yang tidak lama. Yang pertama harus dilakukan evaluator adalah menentukan tujuan dari questionnaire : informasi apa yang akan diambil ?

Terdapat beberapa jenis questionnaire :

- General : pertanyaan untuk mengetahui latarbelakang pengguna dan memasukkannya ke dalam subyek populasi.
- Open-ended : pertanyaan yang membutuhkan jawaban atau opini sendiri, berguna untuk mengumpulkan informasi subyektif secara umum tetapi sulit untuk dianalisis.
- Scalar : user diminta untuk memutuskan pernyataan tertentu berdasarkan skala angka, biasanya berhubungan dengan setuju atau tidak setuju terhadap pernyataan. Contoh : tidak setuju 1 2 3 4 5 setuju. Tetapi skala yang efektif adalah dari 1 hingga 5 atau 1 hingga 7.
- Multi-choice : responden diberikan pilihan respon secara eksplisit, dan hanya boleh memilih satu saja atau sebanyak yang diperlukan. Misal : software apa yang pernah anda gunakan ?
 - Pengolah kata
 - Basis data
 - Spreadsheet
 - System pakar
 - System help online
 - Kompilator
- Ranked : mengurutkan item dalam daftar dan berguna untuk mengindikasikan preferensi user. Misal :
 - Ranking lah kegunaan dari metode perintah (1 sangat berguna, 2 berikutnya, 0 jika tidak berguna)
 - Pilihan menu

- Baris perintah
- Akselerator kunci control

Tipe2 pertanyaan ini sangat berguna untuk tujuan yang berbeda-beda. Untuk mengurangi kesalahan dari responden, sebaiknya menggunakan pertanyaan tertutup seperti scalar, ranked atau multi choice.

Memilih Metode Evaluasi

Faktor yang membedakan teknik evaluasi

Ada 8 faktor yang membedakan teknik evaluasi yang berbeda yang membantu kita dalam memilih teknik yang sesuai, yaitu :

- Tingkat siklus
- Jenis evaluasi
- Tingkat objektivitas dan subyektivitas
- Jenis ukuran yang tersedia
- Informasi yang tersedia
- Kesiapan dari suatu respon
- Tingkat gangguan yang tidak secara langsung
- Sumber yang tersedia

Design VS. Implementation

Evaluasi pada tingkatan perancangan hanya membutuhkan ahlinya dan akan dianalisa sedangkan evaluasi pada tingkat implementasi membawa user sebagai subyek dari eksperimen.

Laboratory VS. Field studies

Secara ideal proses perancangan mencakup dua jenis evaluasi ini dimana pada laboratorium merupakan awal tingkatan sedangkan field studies ditempatkan pada tingkatan implementasi.

Subyektif VS. Obyektif

Pada teknik yang subyektif seperti pada cognitive walkthrough maupun think aloud mempercayakan pada evaluator atau orang yang melakukan evaluasi. Hal ini dapat menyebabkan penilaian evaluator menjadi bias sehingga diperlukan lebih dari satu orang evaluator. Sedangkan teknik obyektif seharusnya menghasilkan hasil yang berulang yang tidak tergantung pada evaluator. Eksperimen yang terkontrol merupakan contoh pengukuran obyektif. Secara ideal pengukuran secara subyektif dan obyektif seharusnya digunakan keduanya.

Qualitative VS Quantitative measures

Pengukuran menggunakan numeric lebih mudah digunakan menggunakan teknik statistik sedangkan non-numeric lebih sulit dilakukan akan tetapi menghasilkan sesuatu yang penting secara lengkap yang tidak dapat dilakukan secara numeric.

Informasi yang tersedia

Tingkatan informasi yang tersedia untuk evaluator pada setiap tingkat proses perancangan dimulai pada informasi yang rendah sampai dengan tingkatan yang tinggi.

Kesiapan suatu tindakan

Faktor lain yang membedakan teknik evaluasi adalah kesiapan mengambil tindakan yang tersedia. Seperti metode think aloud yang merekam perilaku user pada waktu interaksi. Sedangkan post task walkthrough mempercayakan pada pengumpulan kembali kegiatan user.

Intrusiveness

Merupakan teknik yang menghasilkan pengukuran segera yang dihasilkan user selama interaksi dan mempunyai resiko mempengaruhi perilaku user. Evaluator harus sensitive untuk mengurangi masalah ini tetapi tidak dapat mengubah semuanya.

Sumber-sumber

Hal terakhir yang perlu diperhatikan dalam memilih teknik evaluasi yakni : peralatan, waktu, biaya, subyek serta evaluator yang ahli.

Klasifikasi teknik evaluasi

	Cognitive Walkthrough	Heuristic Evaluation	Review Based	Model Based
Stage	Throughput	Throughput	Design	Design
Style	Laboratory	Laboratory	Laboratory	Laboratory
Objective	No	No	As source	No
Measure	Qualitative	Qualitative	As source	Qualitative
Information	Low level	High level	As source	Low level
Immediacy	N/A	N/A	As source	N/A
Intrusive	No	No	No	No
Time	Medium	Low	Low-medium	Medium
Equipment	Low	Low	Low	Low
Expertise	High	Medium	Low	High

	Experiment	Interview	Questionnaire
Stage	Throughout	Throughout	Throughout
Style	Laboratory	Lab/Field	Lab/Field
Objective	Yes	No	No
Measure	Quantitative	Quantitative/Qualitative	Quantitative/Qualitative
Information	Low/High Level	High Level	High Level
Immediacy	Yes	No	No
Intrusive	Yes	No	No
Time	High	Low	Low
Equipment	Medium	Low	Low
Expertise	Medium	Low	Low

	Think-aloud	Protocol Analysis	Post-task Walkthrough
Stage	Implementation	Implementation	Implementation

Style	Lab/Field	Lab/Field	Lab/Field
Objective	No	No	No
Measure	Quantitative	Qualitative	Qualitative
Information	High/Low Level	High/Low Level	High/Low Level
Immediacy	Yes	Yes	No
Intrusive	Yes	Yes	No
Time	High	High	Medium
Equipment	Low	High	Low
Expertise	Medium	High	Medium

Latihan

- Variabel yang dapat diukur dalam eksperimen dalam metode empiris disebut :
 - Variable independent
 - Variable discrete
 - Variable dependent
 - Variable continous
- Sistem prototipe yang berjalan pada multimedia workstation yang dihubungkan langsung ke video recorder adalah :
 - Paper and pencil
 - Experimental Video Annotator
 - User notebook
 - Computer Logging
- Yang *bukan* merupakan pendekatan perancangan evaluasi adalah :
 - Metode observasi
 - Metode review
 - Metode analitik
 - Metode model dasar
- Yang *bukan* merupakan tujuan evaluasi adalah :
 - Membantu user
 - Seberapa jauh sistem berfungsi
 - Masalah yang terjadi di sistem
 - Efek suatu interface ke user
- Metode untuk membantu mengkomunikasikan informasi antara user dan perancang, kecuali :
 - User notebook
 - Brainstorming
 - Storyboarding
 - Workshop