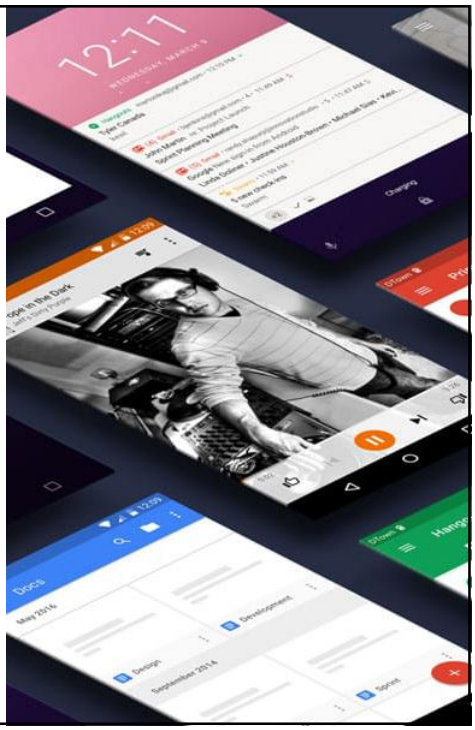
TI5033

DESAIN & EVALUASI ANTARMUKA

Restyandito, Ph.D



1

TI5033

Performance Based Metrics

Measuring the User Experience (Tullis & Albert, 2023)
chapter 4

2

Pengantar

Lima pengujian dasar dari *performance metrics yaitu*:

1. Task success
2. Time on task
3. Errors
4. Efficiency
5. Learnability

3/23

3

Task Success

Task Scenario

Membuat Task Scenario

- Buat task yang realistis

user goal : browse produk yang ditawarkan dan membeli suatu barang
poor task : belilah sepasang sepatu tenis warna biru dengan merk Nike
better task: belilah sepasang sepatu yang harganya di bawah Rp. 500.000,-

Asking a participant to do something that he wouldn't normally do will make him try to complete the task without really engaging with the interface

4/23

4

Task Success
Task Scenario

Membuat Task Scenario

- Buat task yang dapat dilakukan *(actionable)*

user goal : mencari film dan jam tayang
poor task : Anda ingin menonton film pada hari Minggu siang. Pergilah ke www.cgv.id dan katakan pada saya apa yang akan anda 'click' berikutnya
better task: Gunakan situs www.cgv.id untuk mencari informasi mengenai film ingin anda tonton pada hari Minggu siang

It's best to ask the user to do the action, rather than asking them how she would do it

5/23

5

Task Success
Task Scenario

Membuat Task Scenario

- Hindari memberikan petunjuk dan menjelaskan langkah yang harus dilakukan

user goal : melihat nilai
poor task : Anda ingin melihat hasil ujian TTS. Pergilah ke situs e-class, sign in, dan katakan pada saya menu apa yang akan anda pilih untuk mendapatkan transkrip nilai anda.
better task: Carilah nilai hasil ujian TTS anda di situs e-class

Avoiding clues does not mean being vague !!!

6/23

6

Task Success

Binary Success

- Untuk mengukur tingkat keberhasilan, maka hasil atau tujuan dari setiap task harus jelas
- **Binary Success** adalah cara paling sederhana untuk mengukur tingkat keberhasilan (pass / fail)

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10	Average
Participant 1	1	1	1	0	1	1	1	1	0	1	80%
Participant 2	1	0	1	0	1	0	1	0	0	1	50%
Participant 3	1	1	0	0	0	0	1	0	0	0	30%
Participant 4	1	0	0	0	1	0	1	1	0	0	40%
Participant 5	0	0	1	0	0	1	0	0	0	0	20%
Participant 6	1	1	1	1	1	0	1	1	1	1	90%
Participant 7	0	1	1	0	0	1	1	1	0	1	60%
Participant 8	0	0	0	0	1	0	0	0	0	1	20%
Participant 9	1	0	0	0	0	1	1	1	0	1	50%
Participant 10	1	1	0	1	1	1	1	1	0	1	80%
Average	70%	50%	50%	20%	60%	50%	80%	60%	10%	70%	52.0%

Table 4.1 Task success data for 10 participants and 10 tasks.

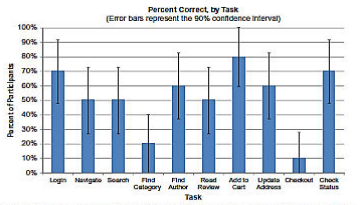


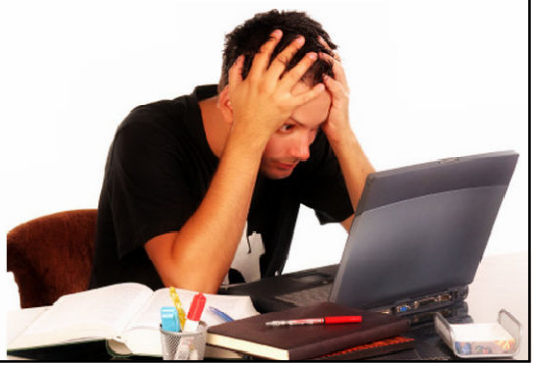
Figure 4.1 Task success rates for the data in Table 4.1, including a 90% confidence interval for each task.

7

Task Success

Binary Success

- Beberapa jenis kegagalan:
  - → responden menyerah
  - → dihentikan oleh moderator
  - → melebihi batas waktu yang ditentukan
  - → hasil / jawaban yang diberikan responden salah

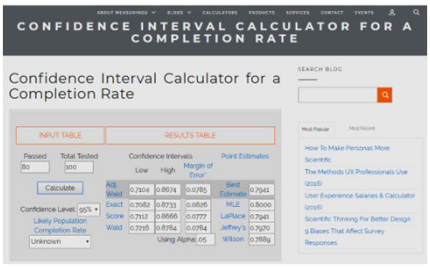


8

Task Success

Binary Success

- Menghitung success rate: → Success / number of participants
- Menghitung confidence interval:
  menggunakan binomial confidence interval
  → Wald Method, Exact Method
  → Confidence Interval Calculator (Sauro & Lewis, 2005)
  http://www.measuringu.com/wald



9/23

9

Task Success

Level of Success

- Jika keberhasilan tidak dapat ditentukan hanya menggunakan kategori sukses/tidak.
  Mis : Task → mencari kamera yang paling murah dengan resolusi ≥ 10 MP dan lensa zoom ≥ 5 optical zoom
- Cara 1 (6 tingkat):
  Complete Success ([1] without assistance , [2] with assistance)
  Partial Success ([3] without assistance , [4] with assistance)
  Failure ([5] wrong result , [6] user gave up)
- Cara 2 (4 tingkat):
  1. **No problem** (menyelesaikan task tanpa masalah dan efisien)
  2. **Minor problem** (menyelesaikan task dengan sedikit masalah)
  3. **Major problem** (menyelesaikan task dengan susah payah)
  4. **Failure/gave up** (gagal menyelesaikan task)

10/23

10

Task Success

Level of Success

Bantuan dapat berupa :

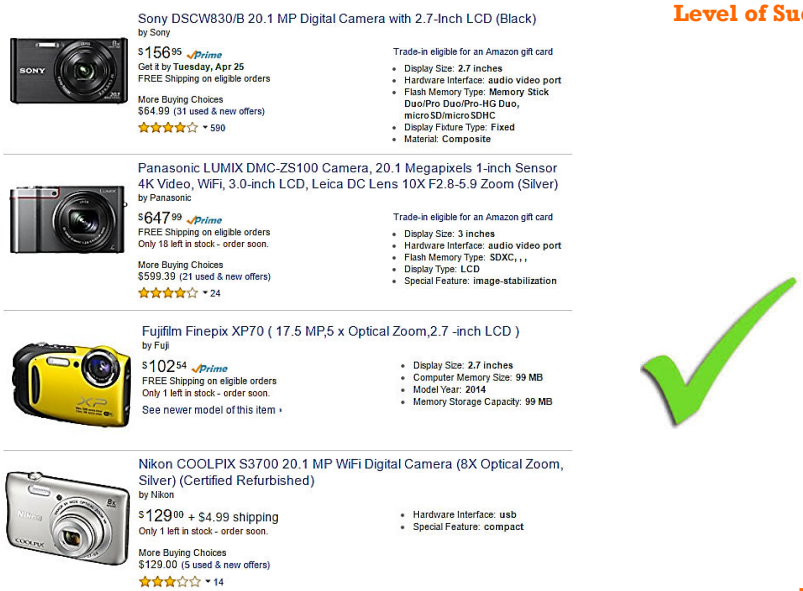
- Moderator mengajak peserta kembali ke halaman beranda atau mereset ulang ke awal sebelum melakukan tugas (*pretask*) → orientasi ulang peserta (dari kebingungan).
- Moderator mengajukan pertanyaan yang mengarahkan → mendorong pengguna berpikir tentang perilaku atau pilihannya dengan cara yang berbeda.
- Moderator menjawab pertanyaan atau memberikan informasi yang membantu peserta menyelesaikan tugas.
- Peserta mencari bantuan dari sumber luar.

11/23

11

Task Success

Level of Success



12/23

12

Task Success

Level of Success

- Contoh:

Berikut adalah hasil pengujian suatu website terhadap 4 orang responden yang diminta menyelesaikan 6 buah task.

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
User 1	F	F	S	F	F	S
User 2	F	F	P	F	P	F
User 3	S	F	S	S	P	S
User 4	S	F	S	F	P	S
	Note: S = success, F = failure, P = partial success					

success rate = (9 + (4*0.5)) / 24 = 46%

Menurut Jakob Nielsen, kebanyakan situs memperoleh nilai kurang dari 50%
(https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/)

13/23

13

Task Success

Level of Success

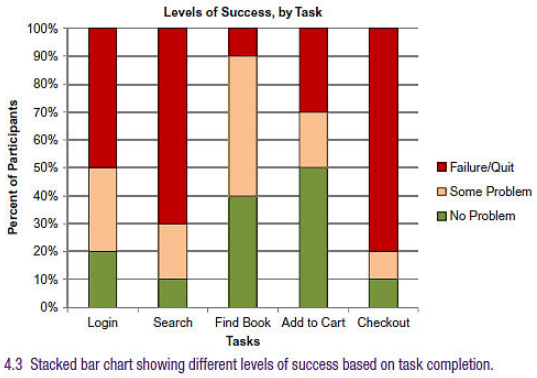


Figure 4.3 Stacked bar chart showing different levels of success based on task completion.

14/23

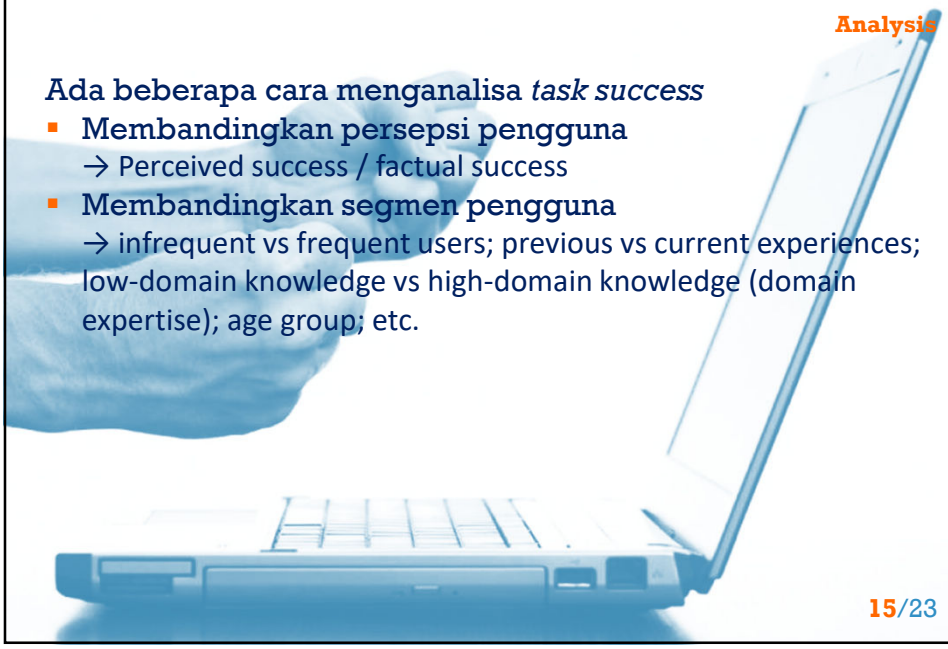
14

Task Success

Analysis

Ada beberapa cara menganalisa *task success*

- Membandingkan persepsi pengguna
  → Perceived success / factual success
- Membandingkan segmen pengguna
  → infrequent vs frequent users; previous vs current experiences; low-domain knowledge vs high-domain knowledge (domain expertise); age group; etc.



15/23

15

Time on Task

Collecting and Measuring Time on Task

- Time on task merupakan salah satu cara untuk mengukur efisiensi suatu produk.
- Beberapa cara mencatat waktu adalah dengan menggunakan:
  - stop watch (ada baiknya dilakukan oleh 2 atau lebih orang)
  - aplikasi
  - video recording

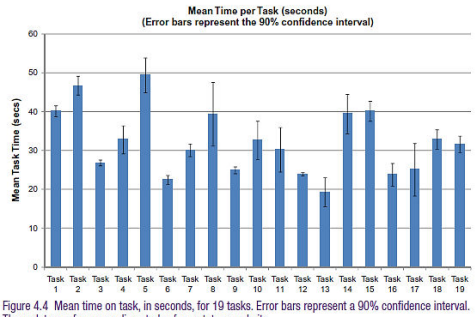


Figure 4.4 Mean time on task, in seconds, for 19 tasks. Error bars represent a 90% confidence interval. These data are from an online study of a prototype website.

16

Time on Task

Issues in Measuring Success

- Beberapa hal yang harus diperhatikan pada saat melakukan pengukuran time on task.
  → participant diberitahu bahwa waktu dihitung ?
  → apakah perlu menentukan range waktu atau menggunakan threshold ?
  → apakah terdapat outliers ?
  → semua atau yang sukses saja ?

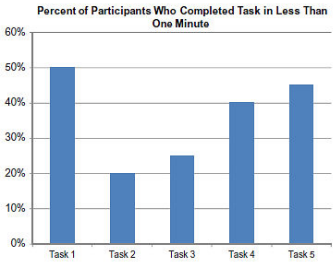


Figure 4.5 An example showing the percentage of users who completed each task in less than 1 minute.

17/23

17

Error

Collecting and Measuring Error

Kapan melakukan pengukuran:

- Jika kesalahan mengakibatkan ketidakefisienan yang besar
- Jika kesalahan mengakibatkan biaya yang besar
- Jika kesalahan mengakibatkan kegagalan task

Apa saja yang disebut kesalahan (error):

- Memasukkan data yang salah
- Memilih menu/daftar yang salah
- Melakukan task dengan urutan aksi yang tidak benar
- Gagal melakukan hal yang penting / utama

18/23

18

Error

Collecting and Measuring Error

- Mengukur kesalahan
  - single opportunity error (0= no error, 1=error)
  - multiple opportunity error (0= no error, n=n error)
- Cara menganalisa kesalahan
  - frekuensi terjadinya kesalahan untuk tiap task
  - jumlah kesalahan yang dibuat oleh setiap participants
  - menentukan threshold
  - dampak dari suatu error (severity)

19/23

19

Efficiency

Collecting and Measuring Learnability

- Jumlah mouse click / keystroke.
- Jumlah site visit

→ lostness $= L = \sqrt{[\left(\frac{N}{S} - 1\right)^2 + (\frac{R}{N} - 1)^2]}$

N: jumlah halaman unik yang dikunjungi selama tes
S: jumlah total semua halaman yang dikunjungi selama tes
R: jumlah minimum halaman untuk menyelesaikan suatu task

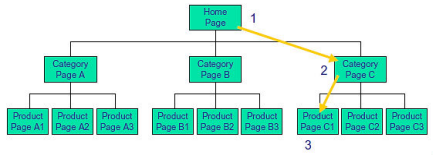


Figure 4.7 Optimum number of steps (three) to accomplish a task that involves finding a target item on Product Page C1 starting from the home page.

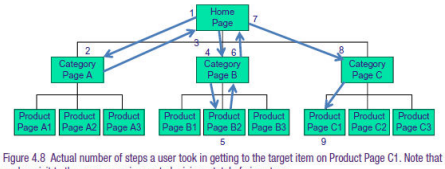


Figure 4.8 Actual number of steps a user took in getting to the target item on Product Page C1. Note that each revisit to the same page is counted, giving a total of nine steps.

→ lostness $= L = \sqrt{[\left(\frac{6}{9} - 1\right)^2 + (\frac{3}{6} - 1)^2]} = 0.60$

20/23

20

Efficiency

Collecting and Measuring Learnability

- Kombinasi task success dan waktu.

	Task Completion Rate	Task Time (min)	Efficiency (%)
Task 1	65%	1.5	43
Task 2	67%	1.4	48
Task 3	40%	2.1	19
Task 4	74%	1.7	44
Task 5	85%	1.2	71
Task 6	90%	1.4	64
Task 7	49%	2.1	23
Task 8	33%	1.3	25

Table 4.3 The efficiency measure is simply the ratio of task completion to task time in minutes[a].
[a]Of course, higher values of efficiency are better. In this example, users appear to have been more efficient in performing tasks 5 and 6 than the other tasks.

$$Efficiency = \frac{completion\ rate}{average\ task\ time}$$

$$Efficiency = \frac{number\ of\ success\ tasks}{total\ time\ on\ all\ tasks}$$

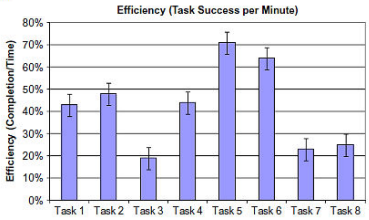


Figure 4.9 An example showing efficiency as a function of completion rate/time.

21/23

21

Learnability

Collecting and Measuring Efficiency

- Membandingkan performance metrics (time on task , number of steps , numbers of error, etc) pada beberapa percobaan.

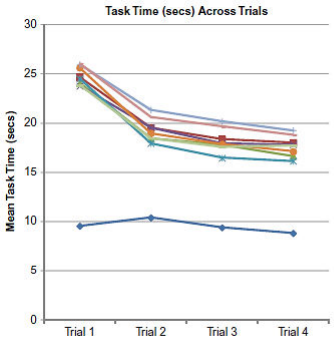


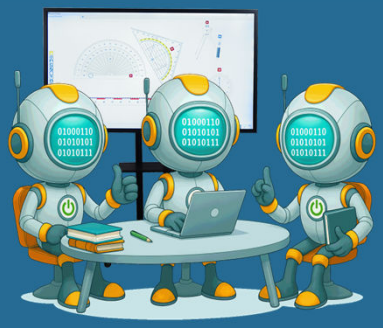
Figure 4.12 Looking at the learnability of different types of on-screen keyboards.

22/23

22

Class Activity

1. Pilihlah sebuah aplikasi mobile yang menurut anda belum/tidak dikenal oleh teman anda.
2. Buatlah 3 task skenario
3. Mintalah 5 orang teman anda mengerjakan task tersebut
4. Analisalah dan buatlah kesimpulan berdasarkan hasil uji usabilitas yang dilakukan



23

24