Administrative Stuff

Grades and comments for Assignment 1 by tonight

Presentations to me by Thursday

Measuring User Experience

Irene Rae

Computer Sciences

CS-570 Introduction to Human-computer Interaction

Last Week

Different approaches to prototyping

The mechanics of prototyping

Experience prototyping & bodystorming

2 Weeks Ago: 5 Usability Components

Learnability How easy is it for users to accomplish basic tasks the first time they encounter the design?

Efficiency Once users have learned the design, how quickly can they perform tasks?

Memorability When users return to the design after a period of not using it, how easily can they reestablish proficiency?

Errors How many errors do users make, how severe are these errors, and how easily can they recover from the errors?

Satisfaction How pleasant is it to use the design?

Today and Thursday

What measures should you use to gather data?

How do you analyze the data?

How do you present your findings?

Measures

Usability Measures

Learnability

Efficiency

Memorability   Task
                      Performance
Errors

Satisfaction   User Experience

Common Usability Studies

According to Tullis & Albert

1. Completing a transaction

2. Comparing products

3. Evaluating frequent use of the same product

4. Evaluating navigation and/or information architecture

5. Increasing awareness

6. Problem discovery

7. Maximizing usability for a critical product

8. Creating an overall positive user experience

9. Evaluation the impact of subtle changes

10. Comparing alternative designs

Common Metrics

According to Tullis & Albert

1. Task Success
2. Task Time
3. Errors
4. Efficiency
5. Learnability
6. Issue-based metrics
7. Self-reported metrics
8. Behavioral and physiological metrics
9. Combined and comparative metrics
10. Live website metrics
11. Card-sorting data

Usability Study Scenario	Task Success	Task Time	Errors	Efficiency	Learn-ability	Issues-Based Metrics	Self-Reported Metrics	Behavioral and Physiological Metrics	Combined and Comparative Metrics	Live Website Metrics	Card-Sorting Data
1. Completing a transaction	X			X		X	X			X	
2. Comparing products	X			X			X		X		
3. Evaluating frequent use of the same product	X	X		X	X		X				
4. Evaluating navigation and/or information architecture	X		X	X							X
5. Increasing awareness							X	X		X	
6. Problem discovery						X	X				
7. Maximizing usability for a critical product	X		X	X							
8. Creating an overall positive user experience							X	X			
9. Evaluating the impact of subtle changes										X	
10. Comparing alternative designs	X	X				X	X		X		

*Tullis & Albert

Common Metrics

1. Task Success
2. Task Time
3. Errors   } Performance Metrics
4. Efficiency
5. Learnability
6. Issue-based metrics
7. Self-reported metrics
8. Behavioral and physiological metrics
9. Combined and comparative metrics
10. Live website metrics
11. Card-sorting data

Performance Metrics

Task success

Measures how effectively users are able to complete a given set of tasks

Binary vs. levels of success

"Give-up" rate

Time-on task

Measures how much time is required to complete a task

Errors

Measures mistakes made during a task

Can point at confusing or misleading parts of an interface

Performance Metrics

Efficiency

Measures the amount of effort a user expends to complete a task

Learnability

Measures how performance changes over time

1. Task Success

Levels of success

Appropriate when there are reasonable shades of gray in task success

Example levels of success

Complete success

 With/without assistance

Partial success

 With/without assistance

Failure

 Participant gave up/thought it was complete, but it wasn't

1. Task Success

Four-point scoring method

Successfully completed the task without any difficulty
1 = no problem

Successfully completed the task with small mistakes
2 = minor problem

Successfully completed the task, but had major problems
3 = major problem

Gave the wrong answer/gave up before completing the task
4 = failure/gave up

Give-up Rate

Failure because participants gave up can be used as a measure itself

Particularly important because it indicates a significant flaw in design or misunderstanding in the user

A give-up rate of 10% is considered normal across all levels of task difficulty

2. Time on Task

Time it takes the participant to complete a specified task

Also referred to as task completion time/task time

Fast is (almost) always better, except in games, learning, etc.

2. Time on Task

Ranges

A way to categorize users into fast, medium, slow

Thresholds

A way to determine whether the user completed the task within an acceptable period of time

Distributions and outliers

A way to see unusually short or long completion times

3. Errors

Incorrect actions that may lead to task failure

Appropriate in 3 key situations (Tullis & Albert):

1. When they result in **significant loss in efficiency**

2. When they result in **significant cost**

3. When they result in **task failure**

Errors include:

Entering incorrect data, making the wrong choice in a menu, taking an incorrect sequence of actions, failing to take a key action

4. Efficiency

Amount of effort required to complete a task

Cognitive vs. **physical** effort

Cognitive (e.g., finding a link on a page)

Physical (e.g., moving the mouse, entering information)

A common measure of efficiency is a combination of task success and time

5. Learnability

The time it takes users to develop proficiency of using the product

As learning occurs, efficiency improves

Two components of the task measure

Slope: learning rate

Asymptote: point of learning

LEARNABILITY
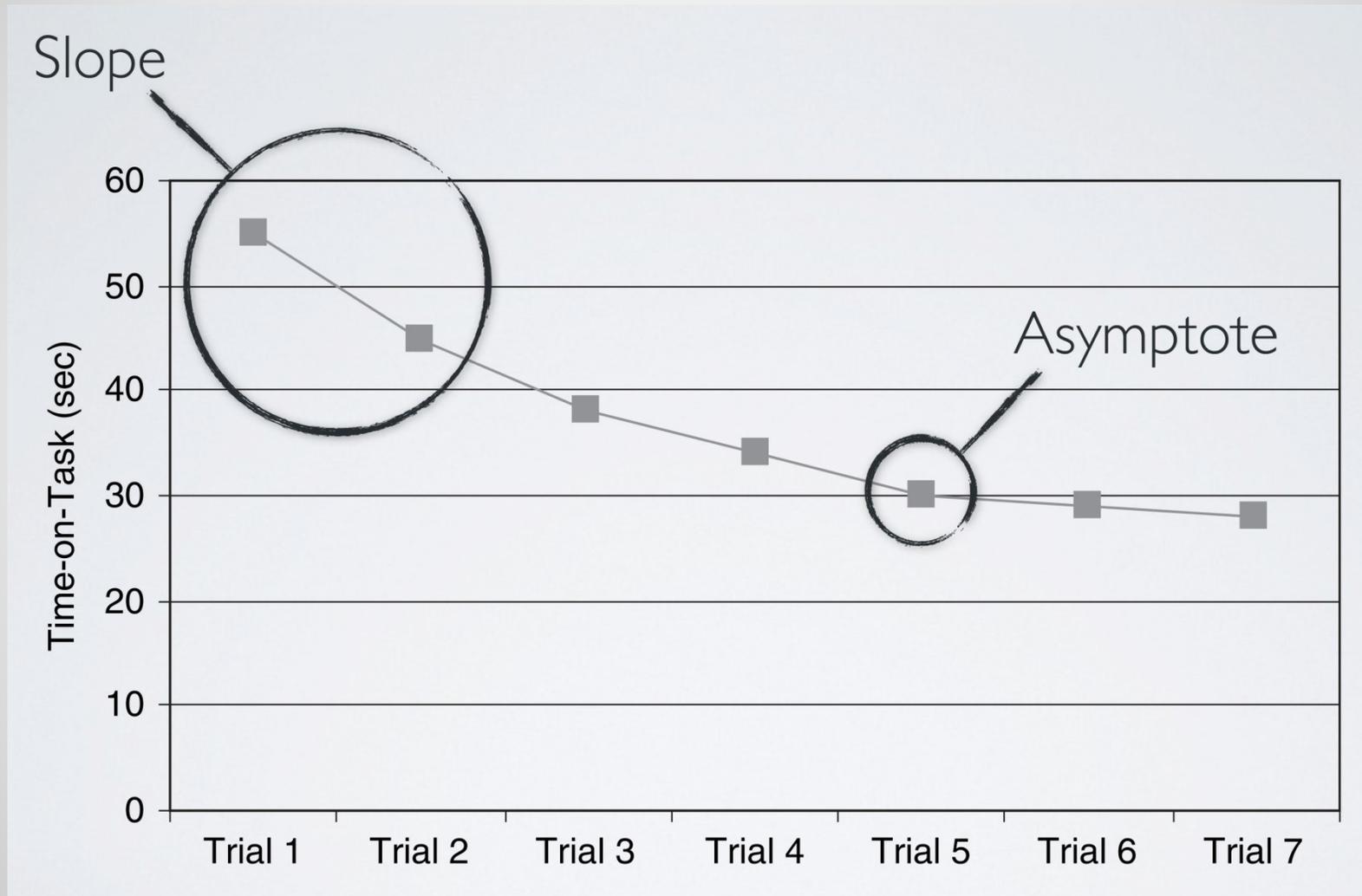
Questions?

Common Metrics

According to Tullis & Albert

1. Task Success
2. Task Time
3. Errors
4. Efficiency
5. Learnability
6. Issue-based metrics
7. **Self-reported metrics**
8. Behavioral and physiological metrics
9. Combined and comparative metrics
10. Live website metrics
11. Card-sorting data

Measuring Experience

Asking users to tell you about their experience with or perception of your system

Captured using subjective/preference measures

Data Collection

Captured using scales

Likert scales

Semantic differential scales

Rating scales

Likert Scales

Users rate their level of agreement with statements

Example:

"I found the navigation options to be confusing."

1	2	3	4	5
Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

Tradeoffs of odd and even numbers

May put an N/A option

Semantic Differential Scales

Presenting users with pairs of bipolar adjectives and asking them to rate where between the pairs the system lies

Example:

1	2	3	4	5
Easy to use				Hard to use

Keep as close to opposites as possible

Refrain from complex adjectives

Rating Scales

Similar to semantic differential scales, but more flexible and has broader use

Can be used as 5, 7, 9, or 11 point scales

Example Rating Scales

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely
	interested			afraid
	distressed			irritable
	excited			alert
	upset			ashamed
	strong			inspired
	guilty			nervous
	scared			determined
	hostile			attentive
	enthusiastic			jittery
	proud			active

Example Rating Scales
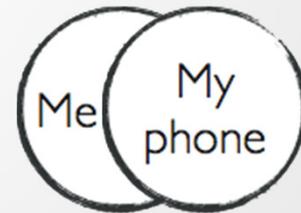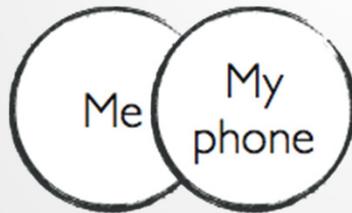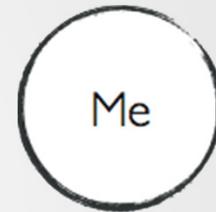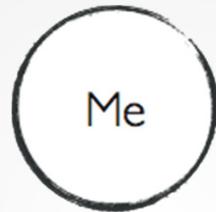
Example Rating Scales

Stress   **High Arousal**   Excitement

Low
Pleasure

High
Pleasure

Depression   **Low Arousal**   Relaxation

Hostile-
Dominant   **Dominant**   Friendly-
Dominant

Hostile

Friendly

Hostile-
Submissive   **Submissive**   Friendly-
Submissive

Collecting Data

Paper forms vs. online forms

Access-control tradeoff

Convenient online services include Qualtrics (used by UW Survey Center), instant.ly, snapsurveys.com, surveygizmo.com, surveymonkey.com, surveyshare.com, zoomerang.com

What to Measure

Ease of use

The most common task satisfaction metric

Example:

"This task was easy to complete"

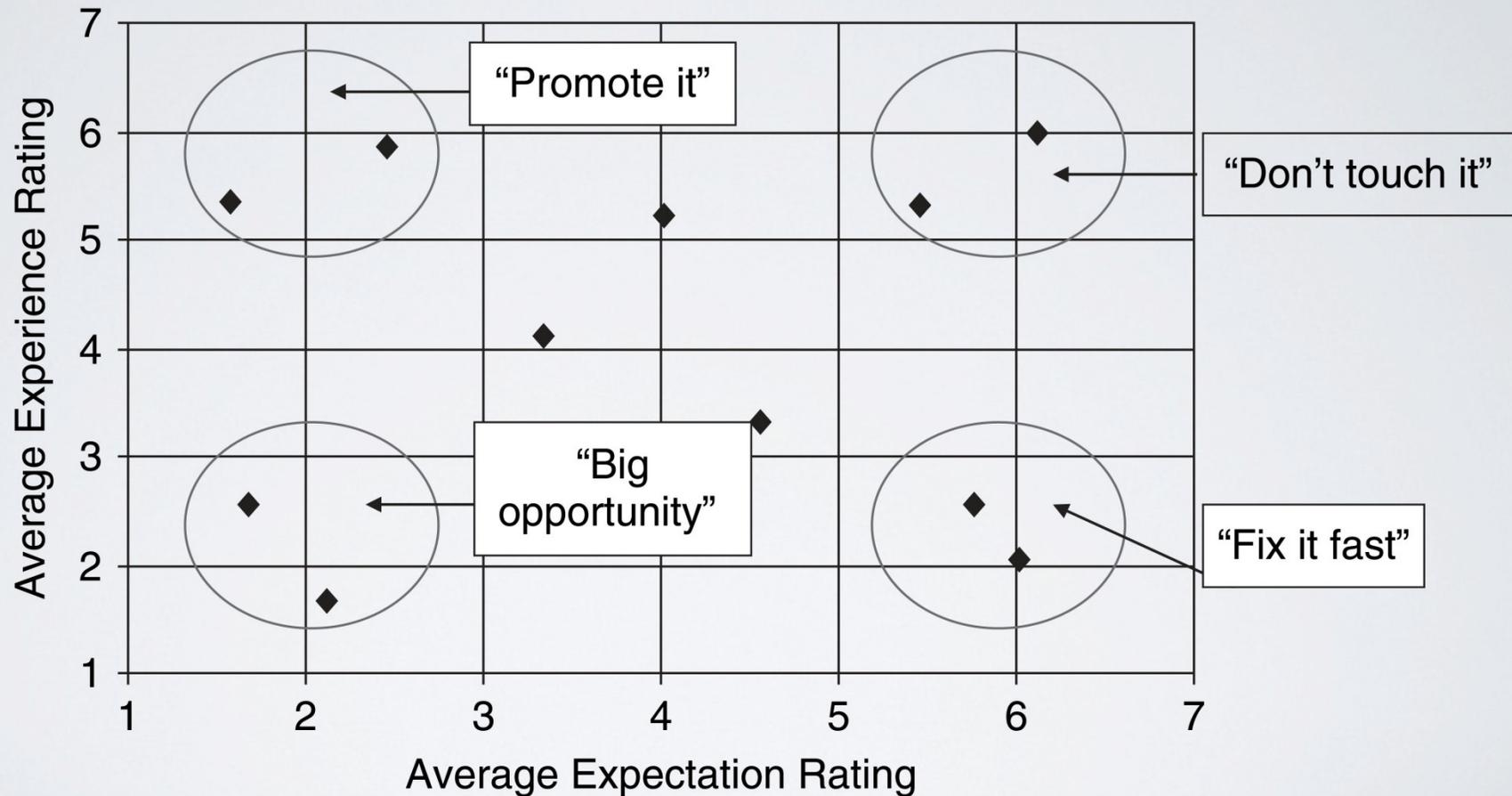| 1	2	3	4	5
Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

What to Measure

Expectation vs. experience (Albert & Dixon, 2003)

A comparison of how easy the user thought the task was going to be and how it actually was
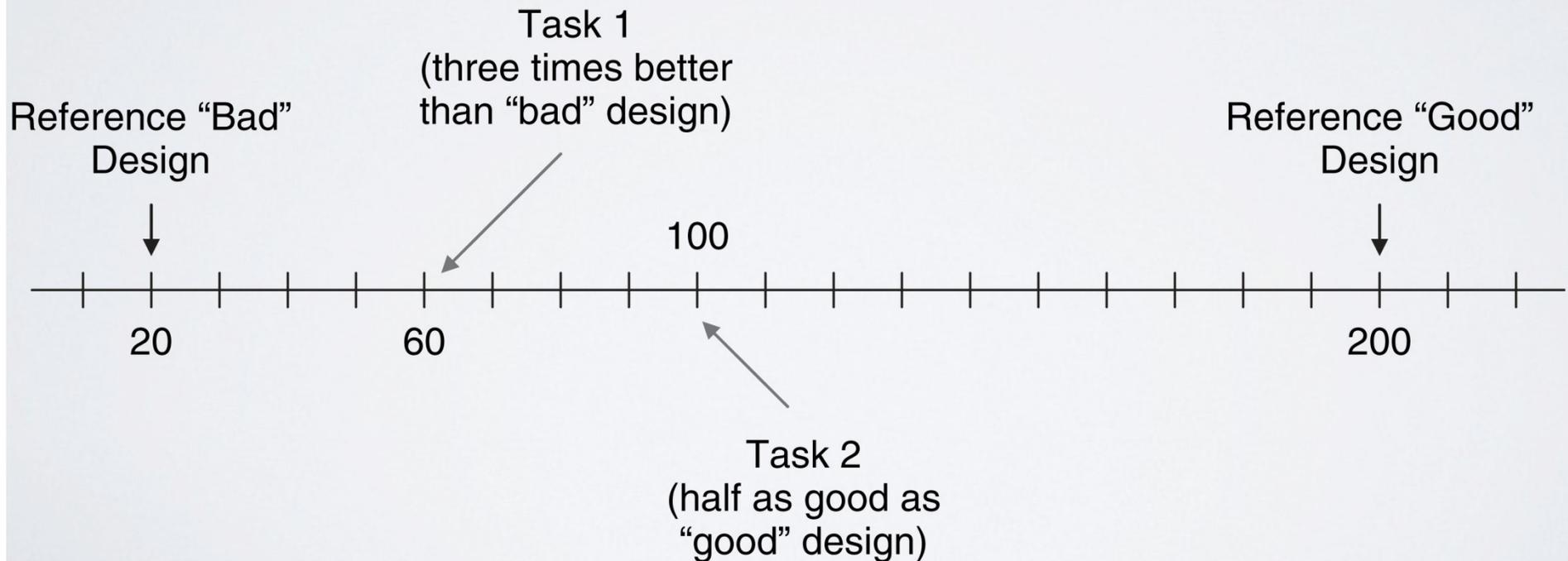
Measured through an **expectation** and an **experience** questionnaire

Expectation Measure

Usability Magnitude Estimation

Users are given reference "good" and "bad" designs and are asked to compare the tested design to the references

Standard Tools

Standard assessment tools are well-established measures of usability that are tested for reliability

System Usability Scale (SUS) (Brooke, 1996)

Computer System Usability Questionnaire (CSUQ) (Lewis, 1995)

Questionnaire for User Interface Satisfaction (QUIS) (Chin et al., 1988)

Usefulness, Satisfaction, and Ease of Use (USE) (Lund, 2001)

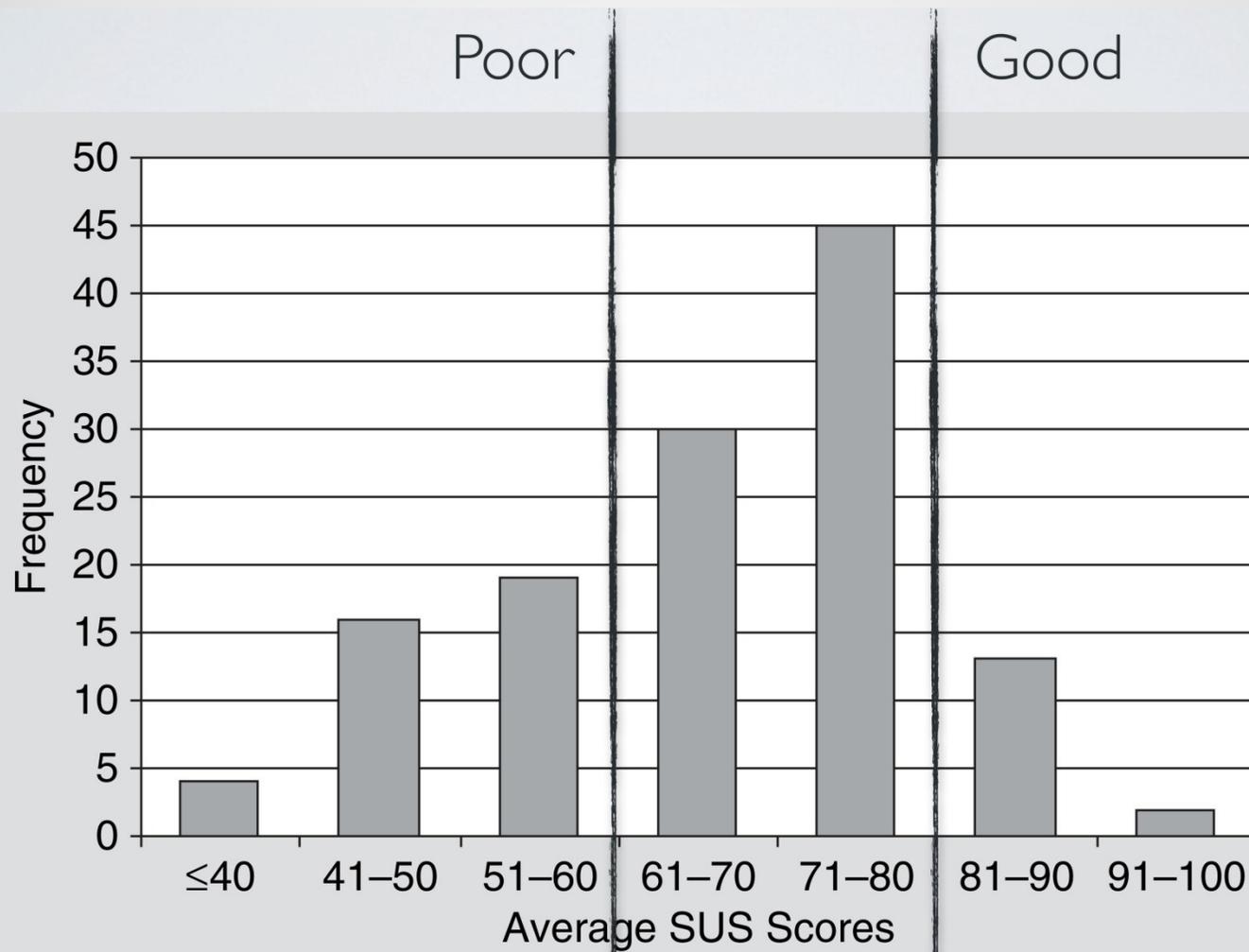Product reaction cards (Microsoft, 2002)

SUS

	Strongly disagree				Strongly agree	

1. I think that I would like to use this system frquently.
   1 2 3 4 5   → √ at 5 → 4

2. I found the system unnecessarily complex.
   1 2 3 4 5   → √ at 4 → 1

3. I thought the system was easy to use.
   1 2 3 4 5   → √ at 2 → 1

4. I think I would need the support of a technical person to be able to use this system.
   1 2 3 4 5   → √ at 1 → 4

5. I found the various functions in this system were well integrated.
   1 2 3 4 5   → √ at 2 → 1

6. I thought this system was too inconsistent.
   1 2 3 4 5   → √ at 3 → 2

7. I would imagine that most people would learn to use this system very quickly.
   1 2 3 4 5   → √ at 2 → 1

8. I found the system very cumbersome to use.
   1 2 3 4 5   → √ at 4 → 1

9. I felt very confident using the system.
   1 2 3 4 5   → √ at 5 → 4

10. I needed to learn a lot of things before I could get going with this system.
   1 2 3 4 5   → √ at 2 → 3

Total = 22   **SUS Score = 22 × 2.5 = 55**

SUS

CSUQ

		1	2	3	4	5	6	7		NA
1. Overall, I am satisfied with how easy it is to use this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
2. It was simple to use this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
3. I can effectively complete my work using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
4. I am able to complete my work quickly using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
5. I am able to efficiently complete my work using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
6. I feel comfortable using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
7. It was easy to learn to use this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
8. I believe I became productive quickly using this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
9. The system gives error messages that clearly tell me how to fix problems	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
10. Whenever I make a mistake using the system, I recover easily and quickly	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
11. The information (such as online help, on-screen messages, and other documentation) provided with this system is clear	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
12. It is easy to find the information I needed	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
13. The information provided for the system is easy to understand	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
14. The information is effective in helping me complete the tasks and scenarios	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
15. The organization of information on the system screens is clear	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
16. The interface of this system is pleasant	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
17. I like using the interface of this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
18. This system has all the functions and capabilities I expect it to have	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
19. Overall, I am satisfied with this system	strongly disagree	○	○	○	○	○	○	○	strongly agree	○
		1	2	3	4	5	6	7		NA

QUIS

OVERALL REACTION TO THE SOFTWARE		0	1	2	3	4	5	6	7	8	9		NA
1. ▣	terrible	○	○	○	○	○	○	○	○	○	○	wonderful	○
2. ▣	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
3. ▣	frustrating	○	○	○	○	○	○	○	○	○	○	satisfying	○
4. ▣	inadequate power	○	○	○	○	○	○	○	○	○	○	adequate power	○
5. ▣	dull	○	○	○	○	○	○	○	○	○	○	stimulating	○
6. ▣	rigid	○	○	○	○	○	○	○	○	○	○	flexible	○
SCREEN		0	1	2	3	4	5	6	7	8	9		NA
7. Reading characters on the screen ▣	hard	○	○	○	○	○	○	○	○	○	○	easy	○
8. Highlighting simplifies task ▣	not at all	○	○	○	○	○	○	○	○	○	○	very much	○
9. Organization of information ▣	confusing	○	○	○	○	○	○	○	○	○	○	very clear	○
10. Sequence of screens ▣	confusing	○	○	○	○	○	○	○	○	○	○	very clear	○
TERMINOLOGY AND SYSTEM INFORMATION		0	1	2	3	4	5	6	7	8	9		NA
11. Use of terms throughout system ▣	inconsistent	○	○	○	○	○	○	○	○	○	○	consistent	○
12. Terminology related to task ▣	never	○	○	○	○	○	○	○	○	○	○	always	○
13. Position of messages on screen ▣	inconsistent	○	○	○	○	○	○	○	○	○	○	consistent	○
14. Prompts for input ▣	confusing	○	○	○	○	○	○	○	○	○	○	clear	○
15. Computer informs about its progress ▣	never	○	○	○	○	○	○	○	○	○	○	always	○
16. Error messages ▣	unhelpful	○	○	○	○	○	○	○	○	○	○	helpful	○
LEARNING		0	1	2	3	4	5	6	7	8	9		NA
17. Learning to operate the system ▣	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
18. Exploring new features by trial and error ▣	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
19. Remembering names and use of commands ▣	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
20. Performing tasks is straightforward ▣	never	○	○	○	○	○	○	○	○	○	○	always	○
21. Help messages on the screen ▣	unhelpful	○	○	○	○	○	○	○	○	○	○	helpful	○
22. Supplemental reference materials ▣	confusing	○	○	○	○	○	○	○	○	○	○	clear	○
SYSTEM CAPABILITIES		0	1	2	3	4	5	6	7	8	9		NA
23. System speed ▣	too slow	○	○	○	○	○	○	○	○	○	○	fast enough	○
24. System reliability ▣	unreliable	○	○	○	○	○	○	○	○	○	○	reliable	○
25. System tends to be ▣	noisy	○	○	○	○	○	○	○	○	○	○	quiet	○
26. Correcting your mistakes ▣	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
27. Designed for all levels of users ▣	never	○	○	○	○	○	○	○	○	○	○	always	○
		0	1	2	3	4	5	6	7	8	9		NA

USE

Usefulness
- It helps me be more effective.
- It helps me be more productive.
- It is useful.
- It gives me more control over the activities in my life.
- It makes the things I want to accomplish easier to get done.
- It saves me time when I use it.
- *It meets my needs.*
- It does everything I would expect it to do.

Ease of Use
- It is easy to use.
- It is simple to use.
- It is user friendly.
- It requires the fewest steps possible to accomplish what I want to do with it.
- *It is flexible.*
- *Using it is effortless.*
- *I can use it without written instructions.*
- *I don't notice any inconsistencies as I use it.*
- *Both occasional and regular users would like it.*
- *I can recover from mistakes quickly and easily.*
- *I can use it successfully every time.*

Ease of Learning
- I learned to use it quickly.
- I easily remember how to use it.
- It is easy to learn to use it.
- *I quickly became skillful with it.*

Satisfaction
- I am satisfied with it.
- I would recommend it to a friend.
- It is fun to use.
- It works the way I want it to work.
- It is wonderful.
- I feel I need to have it.
- It is pleasant to use.

> Users rate agreement with these statements on a 7-point Likert scale, ranging from strongly disagree to strongly agree. Statements in *italics* were found to weight less heavily than the others.

Reaction Cards

The Complete Set of 118 Product Reaction Cards				
Accessible	Creative	Fast	Meaningful	Slow
Advanced	Customizable	Flexible	Motivating	Sophisticated
Annoying	Cutting edge	Fragile	Not secure	Stable
Appealing	Dated	Fresh	Not valuable	Sterile
Approachable	Desirable	Friendly	Novel	Stimulating
Attractive	Difficult	Frustrating	Old	Straightforward
Boring	Disconnected	Fun	Optimistic	Stressful
Business-like	Disruptive	Gets in the way	Ordinary	Time-consuming
Busy	Distracting	Hard to use	Organized	Time-saving
Calm	Dull	Helpful	Overbearing	Too technical
Clean	Easy to use	High quality	Overwhelming	Trustworthy
Clear	Effective	Impersonal	Patronizing	Unapproachable
Collaborative	Efficient	Impressive	Personal	Unattractive
Comfortable	Effortless	Incomprehensible	Poor quality	Uncontrollable
Compatible	Empowering	Inconsistent	Powerful	Unconventional
Compelling	Energetic	Ineffective	Predictable	Understandable
Complex	Engaging	Innovative	Professional	Undesirable
Comprehensive	Entertaining	Inspring	Relevant	Unpredictable
Confident	Enthusiastic	Integrated	Reliable	Unrefined
Confusing	Essential	Intimidating	Responsive	Usable
Connected	Exceptional	Intuitive	Rigid	Useful
Consistent	Exciting	Inviting	Satisfying	Valuable
Controllable	Expected	Irrelevant	Secure	
Convenient	Familiar	Low maintenance	Simplistic	

What Else to Measure

Specific interface attributes or elements

Visual appeal

Perceived efficiency

Usefulness

Enjoyment

Credibility

Appropriateness of terminology

Ease of navigation

Responsiveness

Questions?

Today and Thursday

What measures should you use to gather data?

How do you analyze the data?

How do you present your findings?

Next:

A crash course in experimental statistics