# Human-Computer Interaction (HCI)

# DECO2500/7250

Dr Chelsea Dobbins

deco2500@itee.uq.edu.au

Room 78-305

# 08

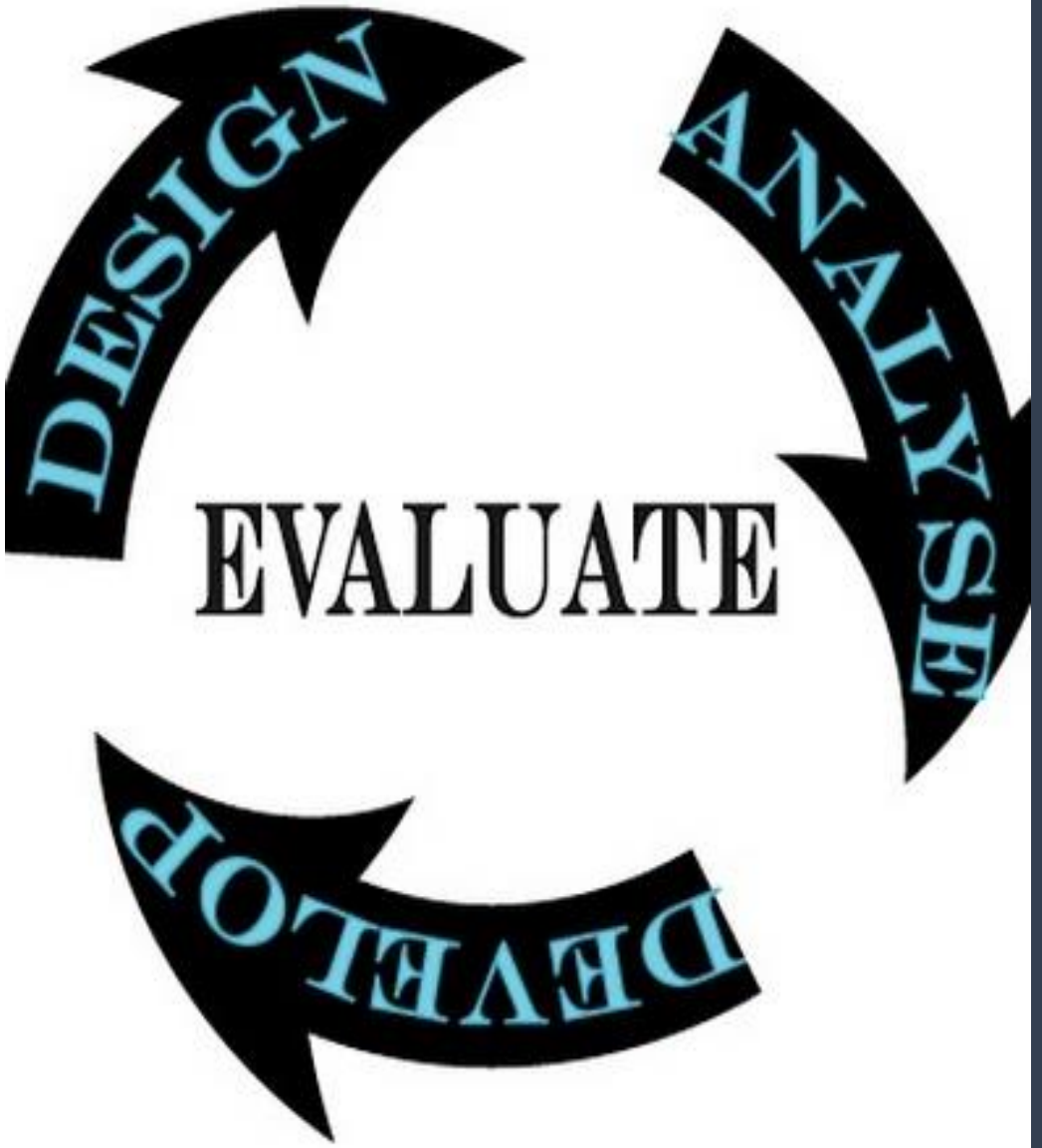
# User-Based Evaluations and Data Analysis

# In this session…

- User–Based Evaluations
  - Data analysis
  - System Usability Scale (SUS)
  - Technology Acceptance Model (TAM)
  - Time on Task
  - Think Aloud

# User–Based Evaluations

- Tests the usability and functionality of a system
- Occurs in collaboration with users
- Considered at all stages in the design lifecycle
- Many evaluation methods exist
  - System Usability Scale
  - Think Aloud
  - Time on Task



# Data Analysis

- One of the most important components
- Weak analysis produces inaccurate results
- Inaccurate results hinder the authenticity of the research and make the findings unusable
- Vital to choose your data analysis methods carefully to ensure that your findings are *insightful* and *actionable*
- When to use each method depends on the research questions

# Data-Gathering Techniques

| Technique | Good for | Kind of data | Advantages | Disadvantages |
|---|---|---|---|---|
| **Interviews** | Exploring issues | Some quantitative but mostly qualitative | Interviewer can guide interviewee if necessary. Encourages contact between developers and users | Time-consuming. Artificial environment may intimidate interviewee |
| **Focus groups** | Collecting multiple viewpoints | Some quantitative but mostly qualitative | Highlights areas of consensus and conflict. Encourages contact between developers and users | Possibility of dominant characters |
| **Questionnaires** | Answering specific questions | Quantitative and qualitative | Can reach many people with low resource | The design is crucial. Response rates may be low. Unless carefully designed, the responses may not provide suitable data |
| **Direct observation in the field** | Understanding context of user activity | Mostly qualitative | Observing gives insights that other techniques don't give | Very time-consuming. Huge amounts of data are produced |
| **Direct observation in a controlled environment** | Capturing the detail of what individuals do | Quantitative and qualitative | Can focus on the details of a task without interruption | Results may have limited use in the normal environment because the conditions were artificial |
| **Indirect observation** | Observing users without disturbing their activity; data captured automatically | Quantitative (logging) and qualitative (diary) | User doesn't get distracted by the data gathering; automatic recording means that it can extend over long periods of time | A large amount of quantitative data needs tool support to analyze (logging); participants' memories may exaggerate (diary) |

# Quantitative and Qualitative

| | Usual raw data | Sample **qualitative** data | Sample **quantitative** data | Initial processing |
|---|---|---|---|---|
| Interviews | Audio recordings. | Responses to open questions. | Age, job role, years of experience. | Transcription of recordings. |
| | Interviewer notes. | Video pictures. | Responses to closed questions | Expansion of notes |
| | Video recordings | Respondent's opinions | | |
| Questionnaires | Written responses. | Responses to open questions. | Age, job role, years of experience. | Clean up data. |
| | Online database | Responses in 'further comments' fields. | Responses to closed questions | Filter into different data sets |
| | | Respondent's opinions | | |
| Observation | Observer's notes. | Records of behavior. | Demographics of participants. | Expansion of notes. |
| | Photographs. | Description of a task as it is undertaken. | Time spent on a task. | Transcription of recordings. |
| | Audio and video recordings. | Copies of informal procedures | The number of people involved in an activity | Synchronization between data recordings |
| | Data logs. | | | |
| | Think-aloud | | | |

# Simple Quantitative Analysis

## Quantitative Data Analysis Methods

### Descriptive Analysis

The first level of analysis, this helps researchers find absolute numbers to summarize individual variables and find patterns.

A few examples are...

- **Mean:** numerical average
- **Median:** midpoint
- **Mode:** most common value
- **Percentage:** ratio as a fraction of 100
- **Frequency:** number of occurrences
- **Range:** highest and lowest values

### Inferential Analysis

These complex analyses show the relationships between multiple variables to generalize results and make predictions.

A few examples are...

- **Correlation:** describes the relationship between 2 variables
- **Regression:** shows or predicts the relationship between 2 variables
- **Analysis of variance:** tests the extent to which 2+ groups differ

- Standard deviation or variance (square root of) average deviation from the mean

https://humansofdata.atlan.com/2018/09/qualitative-quantitative-data-analysis-methods/

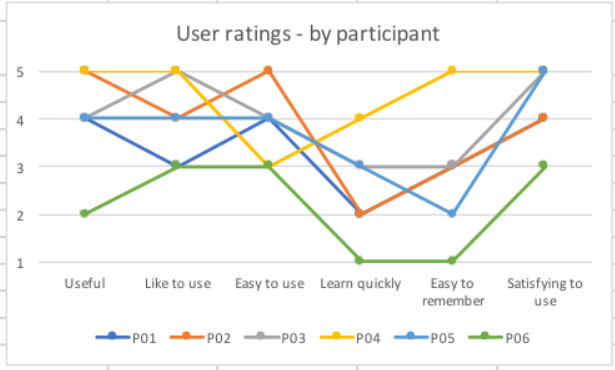
# Simple Quantitative Analysis

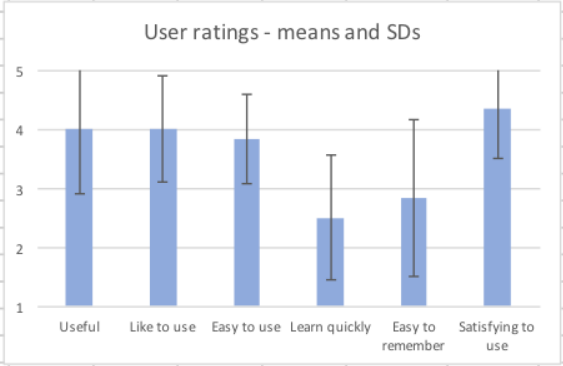
Questionnaire with 5-point Likert items on it (strongly disagree to strongly agree)

| | | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
|---|---|---|---|---|---|---|
| | The app is useful | 1 | 2 | 3 | 4 | 5 |
| | I like to use the app | 1 | 2 | 3 | 4 | 5 |
| | It's easy to use the app | 1 | 2 | 3 | 4 | 5 |
| | I learned the app quickly | 1 | 2 | 3 | 4 | 5 |
| | It's easy to remember how to use the app | 1 | 2 | 3 | 4 | 5 |
| | The app is satisfying to use | 1 | 2 | 3 | 4 | 5 |

Collation of results from six participants (P01–P06).

| Participants | Useful | Like to use | Easy to use | Learn quickly | Easy to remember | Satisfying to use |
|---|---|---|---|---|---|---|
| P01 | 4 | 3 | 4 | 2 | 3 | 4 |
| P02 | 5 | 4 | 5 | 2 | 3 | 4 |
| P03 | 4 | 5 | 4 | 3 | 3 | 5 |
| P04 | 5 | 5 | 3 | 4 | 5 | 5 |
| P05 | 4 | 4 | 4 | 3 | 2 | 5 |
| P06 | 2 | 3 | 3 | 1 | 1 | 3 |
| | | | | | | |
| Average | 4.0 | 4.0 | 3.8 | 2.5 | 2.8 | 4.3 |
| Standard deviation | 1.1 | 0.9 | 0.8 | 1.0 | 1.3 | 0.8 |

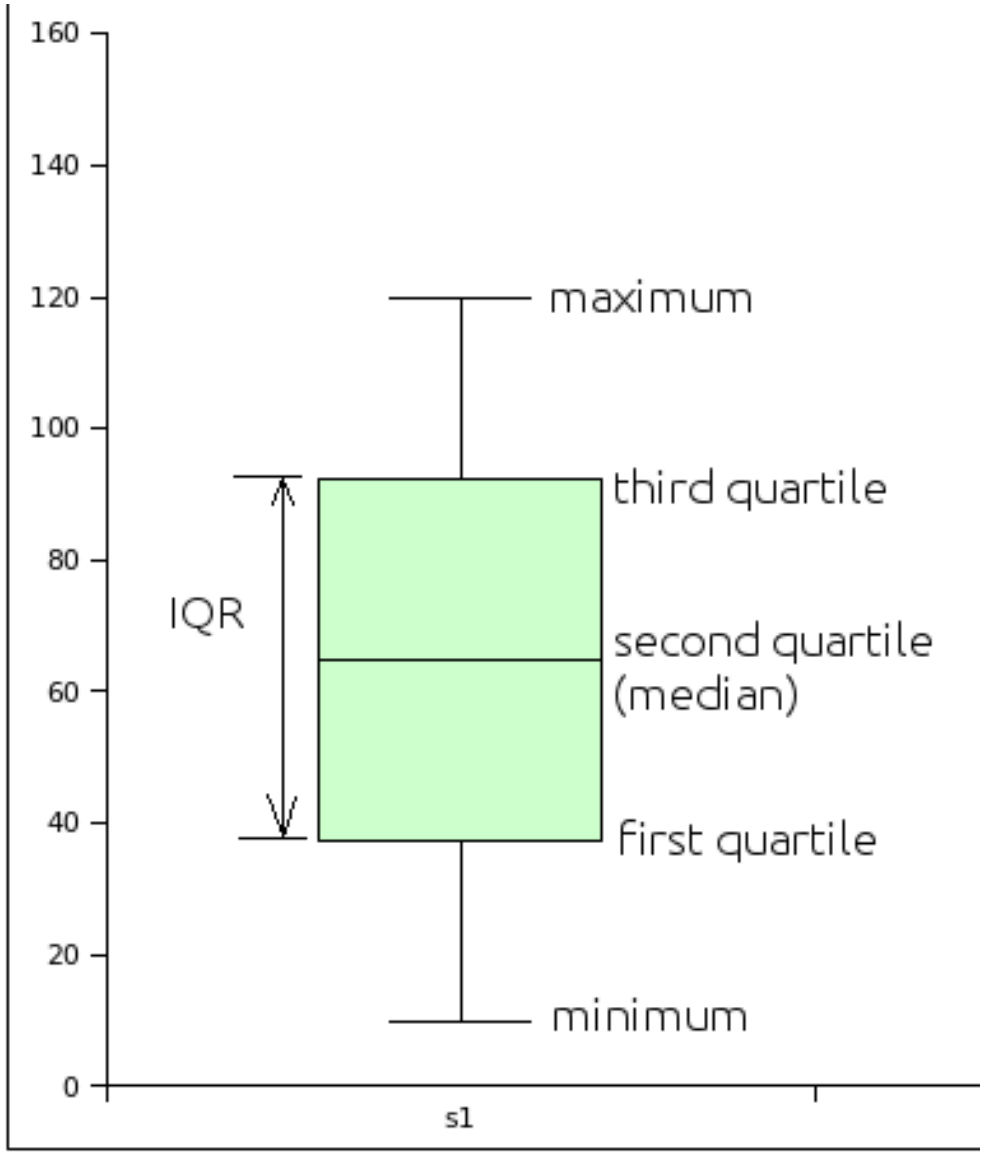


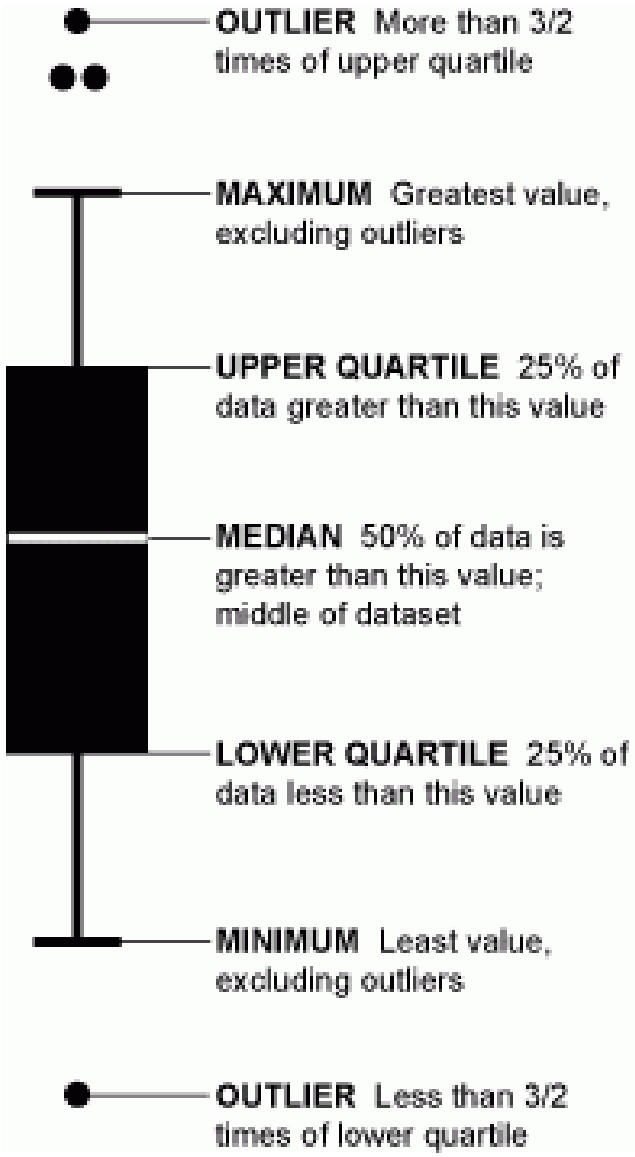




Some ways of viewing data:
- Raw numbers (1–5 ratings)
- Trends (means, SDs)
- Participant patterns

Choose a graph more suitable for your data

# Simple Quantitative Analysis



# Qualitative Analysis

- Many complex and socially based phenomena in HCI cannot be easily quantified
- Qualitative research aims to:
  - Understand the qualities of a particular technology
  - How people use it in their lives
  - How they think about it
  - How they feel about it
  - Understand why
  - Explain data and patterns


A. Adams, P. Lunt, and P. Cairns, "A Qualitative Approach to HCI Research," in *Research Methods for Human-Computer Interaction*, 2008, pp. 138–157

# Qualitative Analysis

## Qualitative Data Preparation and Analysis

**Get familiar with the data**

Start by reading the data several times to get familiar with it and start looking for basic observations or patterns. This also includes transcribing the data.

**Revisit research objectives**

Revisit the research objective and identify the questions that can be answered through the collected data.

**Develop a framework**

Identify broad ideas, concepts, behaviors, or phrases and assigns codes to them. This is helpful for structuring and labeling the data.

**Identify patterns and connections**

Start identifying themes, looking for the most common responses to questions, identifying data or patterns that can answer research questions, and finding areas that can be explored further.

https://humansofdata.atlan.com/2018/09/qualitative-quantitative-data-analysis-methods/

# Qualitative Analysis

## Content Analysis

- Analyse documented information
- Usually used to analyse responses from interviewees

## Narrative Analysis

- Analyse content such as from interviews, observations from the field, or surveys.
- Focuses on using the stories and experiences shared by people to answer the research questions

## Discourse Analysis

- Used to analyse interactions with people
- Analyses the social context in which the communication between the researcher and the respondent occurred.
- Looks at the respondent's day-to-day environment and uses that information during analysis

## Grounded Theory

- Uses the data to explain why a certain phenomenon happened.
- Studies a variety of similar cases in different settings and using the data to derive causal explanation

## Thematic Analysis

- Used to deduce the meaning behind the words people use
- Discovers repeating themes in text that reveal key insights into data
- Outcome is a code frame that captures themes in terms of codes, also called categories

## Sentiment Analysis

- The process of detecting positive or negative sentiment in text
- Focuses on the polarity of a text (*positive, negative, neutral*)
- Can detect specific feelings and emotions (*angry, happy, sad*, etc), urgency (*urgent, not urgent*) and even intentions (*interested v. not interested*)

# System Usability Scale (SUS)

- Measures perceptions of usability – Brooke (1986)
- A "quick and dirty" survey scale to quickly and easily assess the usability of a product or service
- Generally technology agnostic and non-proprietary


Image source: https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/

# Benefits of SUS

- Cheap and
- Quick
- Very easy to administer
- An industry standard
- Small sample size acceptable
- Valid

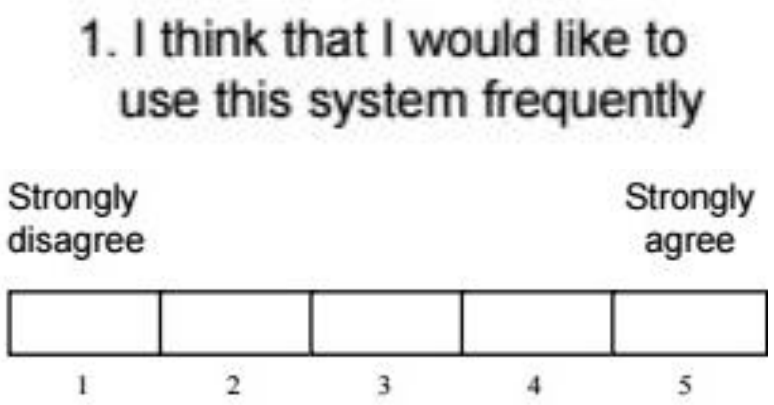
# Drawbacks When Using SUS

- Complex scoring system
- Results not a percentage
- Results need to be normalized
- Not a diagnostic tool
- Might tell you THAT there is a problem, but not much about the problem, why it is there, or what to do about it
- Not to be used in isolation

# Measuring Perceptions of Usability

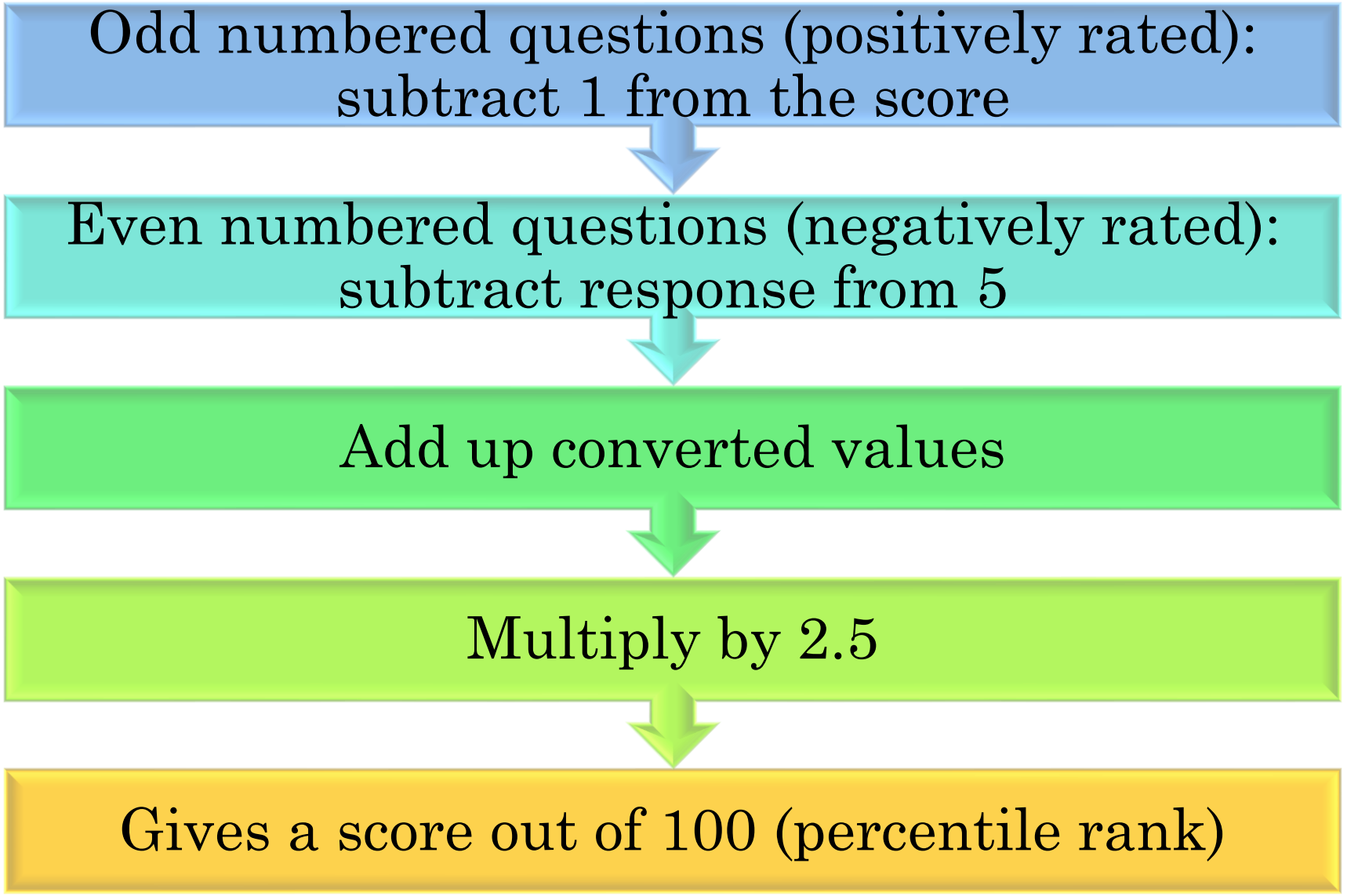
| Question |
| --- |
| 1. I think that I would like to use this system frequently. |
| 2. I found the system unnecessarily complex. |
| 3. I thought the system was easy to use. |
| 4. I think that I would need the support of a technical person to be able to use this system. |
| 5. I found the various functions in this system were well integrated. |
| 6. I thought there was too much inconsistency in this system. |
| 7. I would imagine that most people would learn to use this system very quickly. |
| 8. I found the system very cumbersome to use. |
| 9. I felt very confident using the system. |
| 10. I needed to learn a lot of things before I could get going with this system. |

# Measuring Perceptions of Usability




| Question |
| --- |
| 1. I think that I would like to use this system frequently. |
| 2. I found the system unnecessarily complex. |
| 3. I thought the system was easy to use. |
| 4. I think that I would need the support of a technical person to be able to use this system. |
| 5. I found the various functions in this system were well integrated. |
| 6. I thought there was too much inconsistency in this system. |
| 7. I would imagine that most people would learn to use this system very quickly. |
| 8. I found the system very cumbersome to use. |
| 9. I felt very confident using the system. |
| 10. I needed to learn a lot of things before I could get going with this system. |

# Analysing SUS



# Interpreting SUS

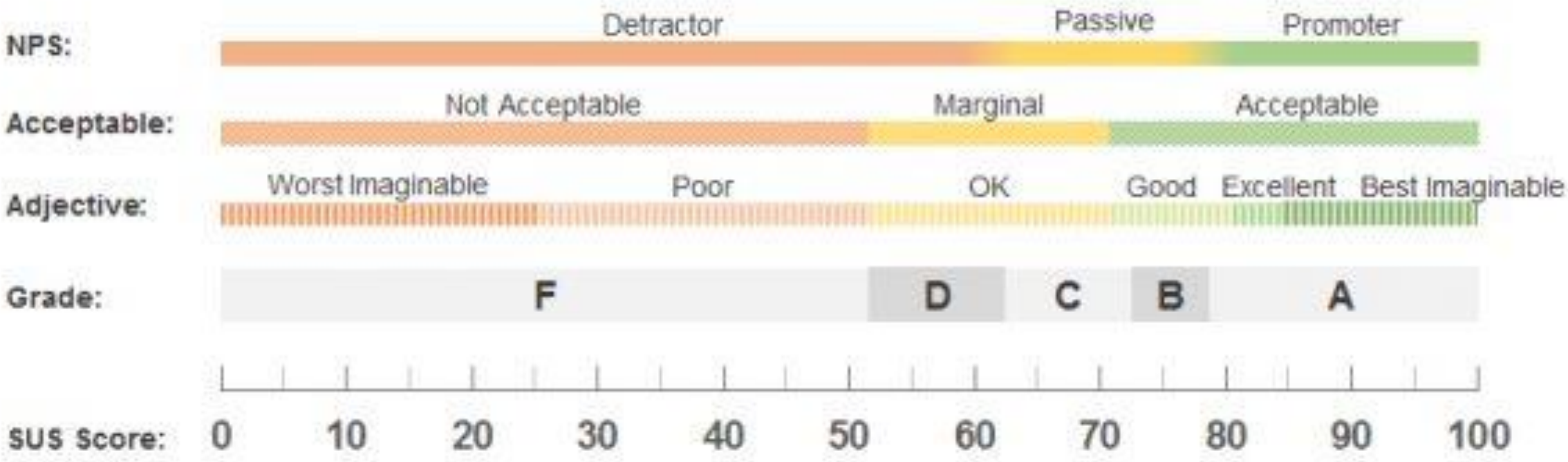



Image source: https://measuringu.com/interpret-sus-score/

# Interpreting SUS

- >= 80.3 (A rated)
- The average score (at the 50th percentile) is 68 (C rated)
- =< 51 (F rated)

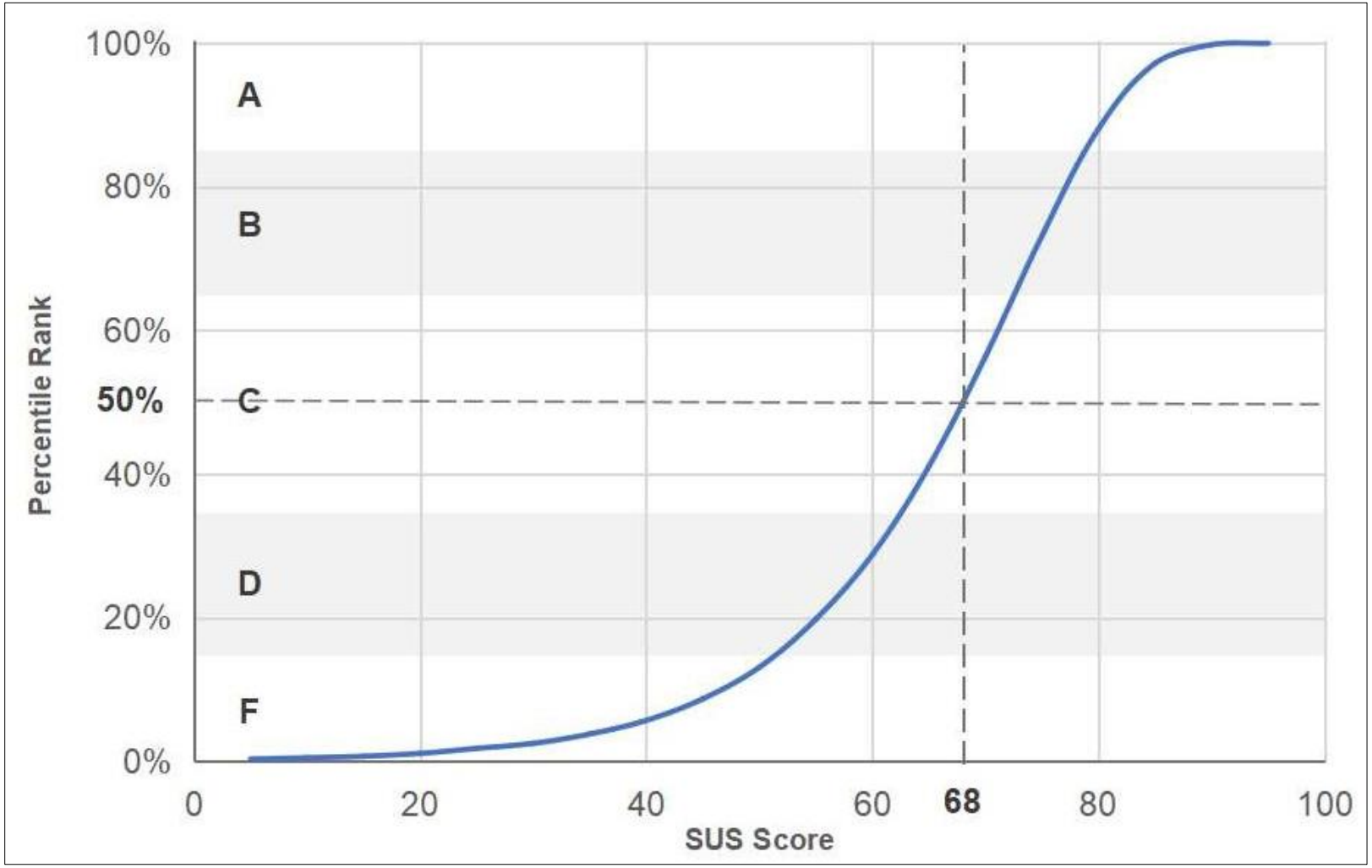
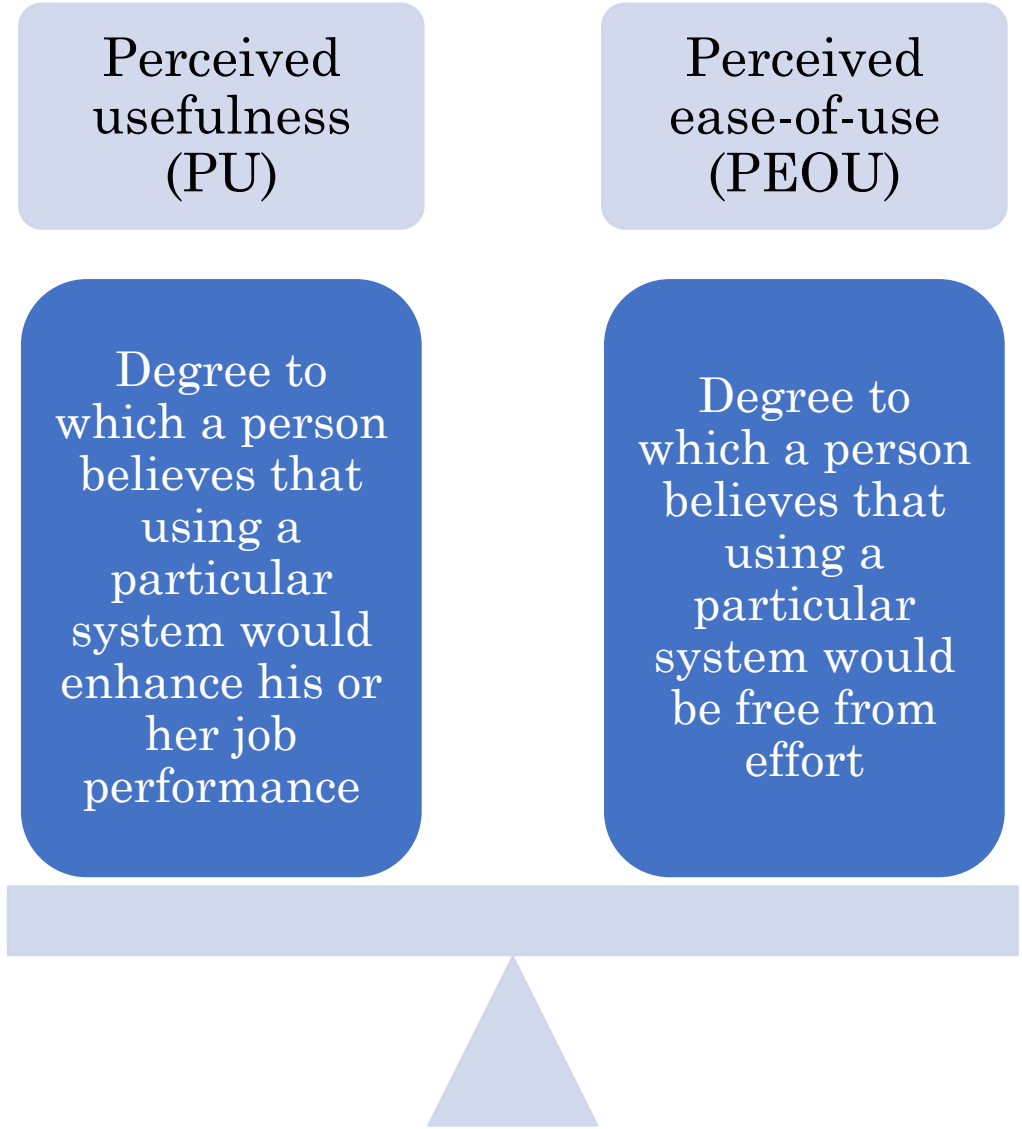



Image source: https://measuringu.com/interpret-sus-score/

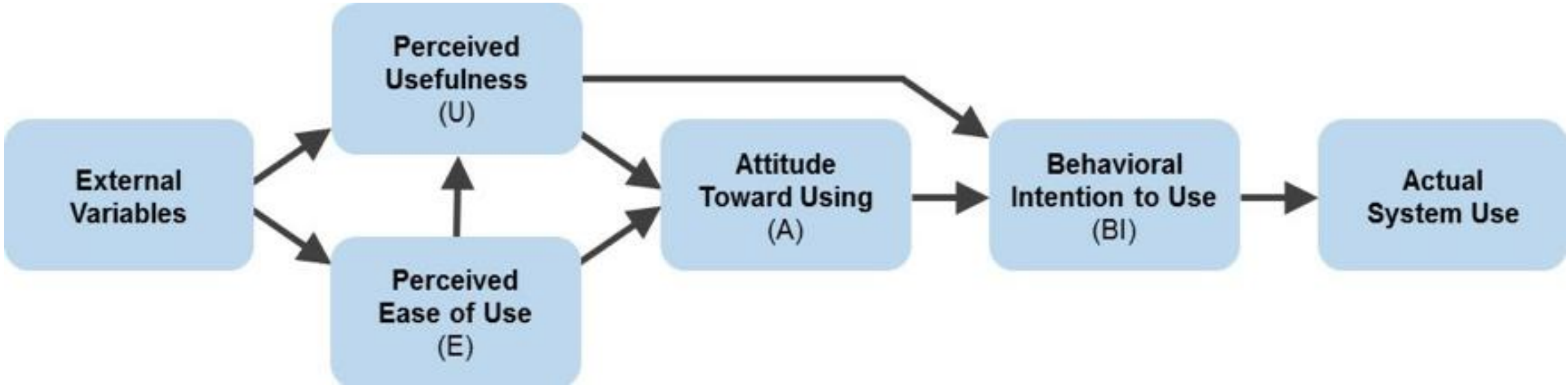
# Technology Acceptance Model (TAM)

- Information systems theory that illustrates how users come to accept and use a technology (Davis, 1989)




Davis, F. D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly, 13 (3): 319–340, doi:10.2307/249008, JSTOR 249008

# Technology Acceptance Model (TAM)



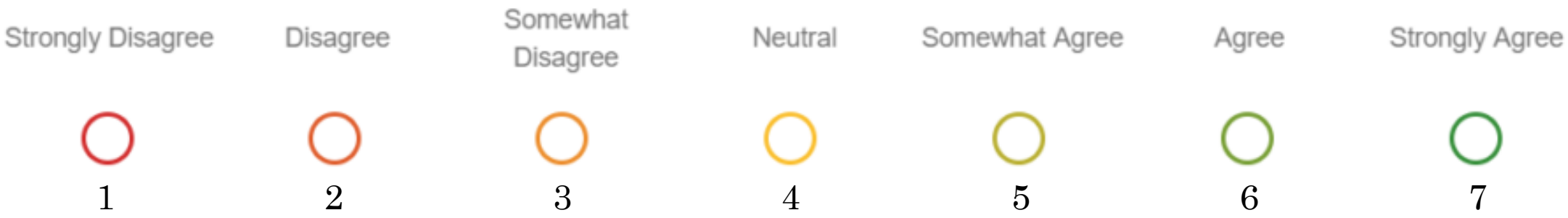

Technology Acceptance Model (TAM) from Davis, 1989.

# Technology Acceptance Model (TAM)

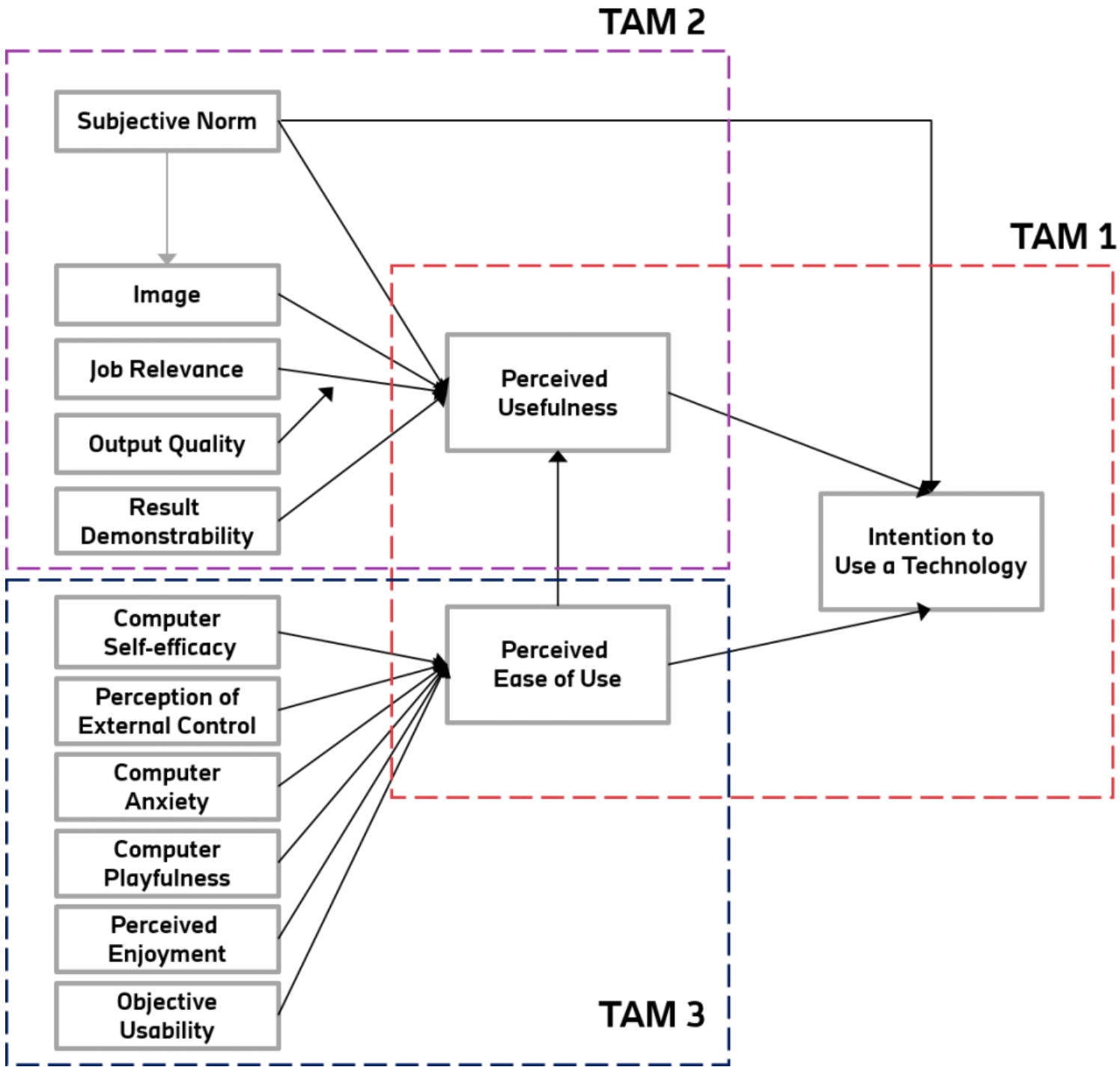
| Dimension | Question |
|---|---|
| **PU1** | I can accomplish my [description of task] more quickly using [name of system] |
| **PU2** | I can accomplish my [description of task] more easily using [name of system] |
| **PU3** | [Name of system] enhances my effectiveness in utilizing [type of service] |
| **PU4** | [Name of system] enhances my efficiency in utilizing [type of service] |
| **PU5** | [Name of system] enables me to make better decisions in utilizing [type of service] |
| **PU6** | Overall, I find [name of system] useful |
| **PEOU1** | Learning to use [name of system] is easy for me |
| **PEOU2** | It is easy to use [name of system] to accomplish my [task] |
| **PEOU3** | Overall, I believe [name of system] is easy to use |
| **ATT1** | In my opinion, it is desirable to use [name of system] |
| **ATT2** | I think it is good for me to use [name of system] |
| **ATT3** | Overall, my attitude towards [name of system] is favourable |
| **ITO1** | I will use [name of system] on a regular basis in the future |
| **ITO2** | I will frequently use [name of system] in the future |
| **ITO3** | I will strongly recommend others to use [name of system] |

# Technology Acceptance Model (TAM)

- Items are scored on a 7-point Likert scale
- Statistics are then generated for each dimension
  - Average
  - Standard Deviation



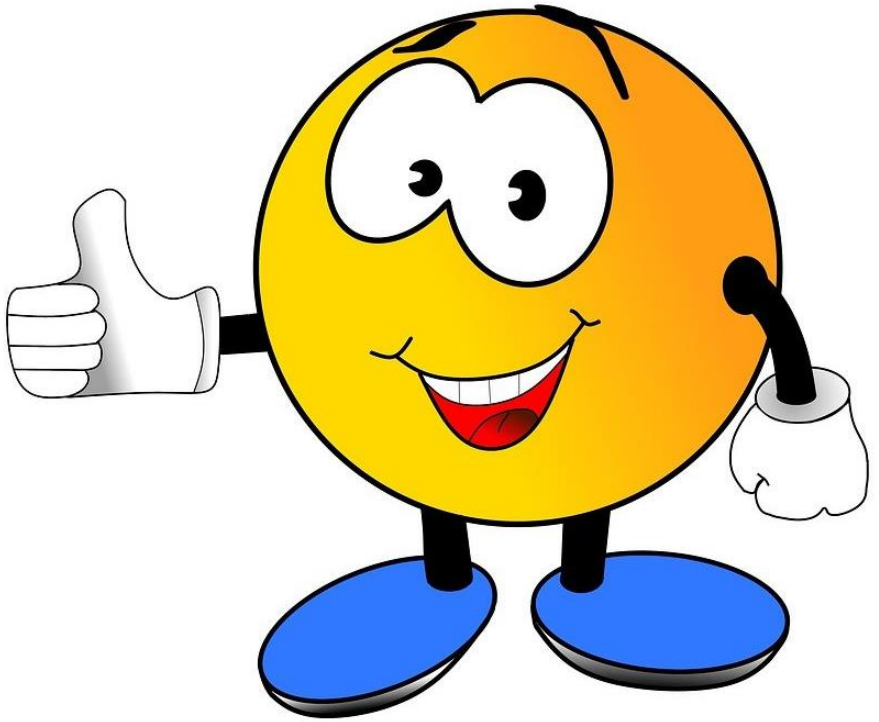
# TAM 2 and TAM 3




https://acceptancelab.com/technology-acceptance-model-tam

# Benefits of TAM

- TAM 1
  - Easy to understand
  - Has demonstrated a high level of predictiveness in many contexts
- TAM 2 and 3
  - Takes external and social influences into consideration as well
  - Both models have been successfully applied to a wide variety of innovations



# Drawbacks of TAM

- TAM 1
  - Originally developed for the adoption of IT at the workplace
  - Neglects the diverse needs of users, including subjective norms or social impact
  - The central constructs Perceived Usefulness (PU) and Perceived ease-of-use (PEOU) provide no information about how to make technology more useful and easier to use
- TAM 2 and 3
  - Very complex due to the multitude of factors incorporated

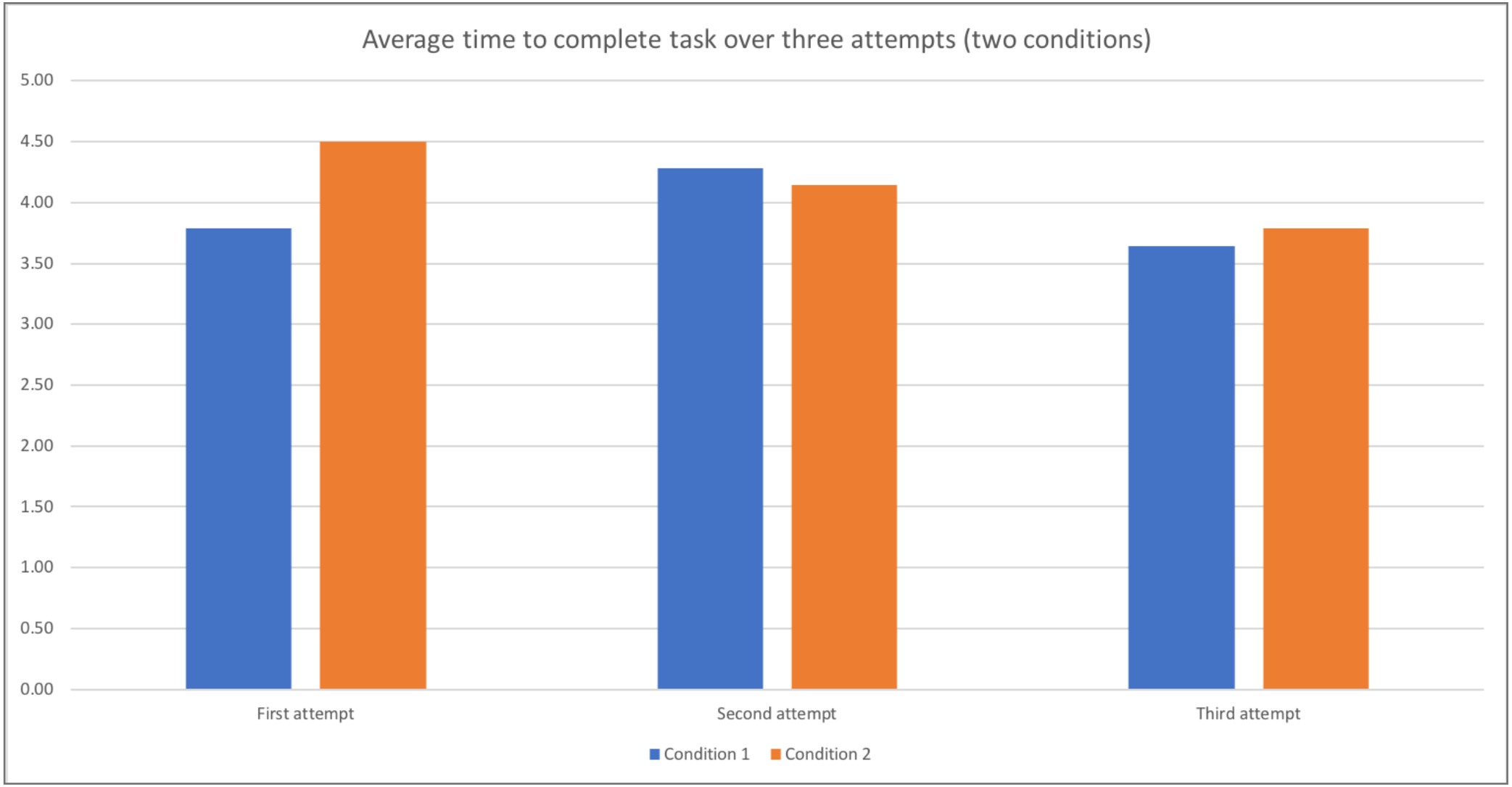
# Time on Task

- Easy and understandable making it used often
- Used to identify usability problems
- Often combined with a Think Aloud protocol
- Used in both formative and summative evaluation

# Time on Task

- Define your protocol so that it is **repeatable**
- Three "events" to measure:
  - Average task completion time
  - Mean time to failure
  - Average time on task

# Time on Task – Analysis



# Time on Task – Analysis

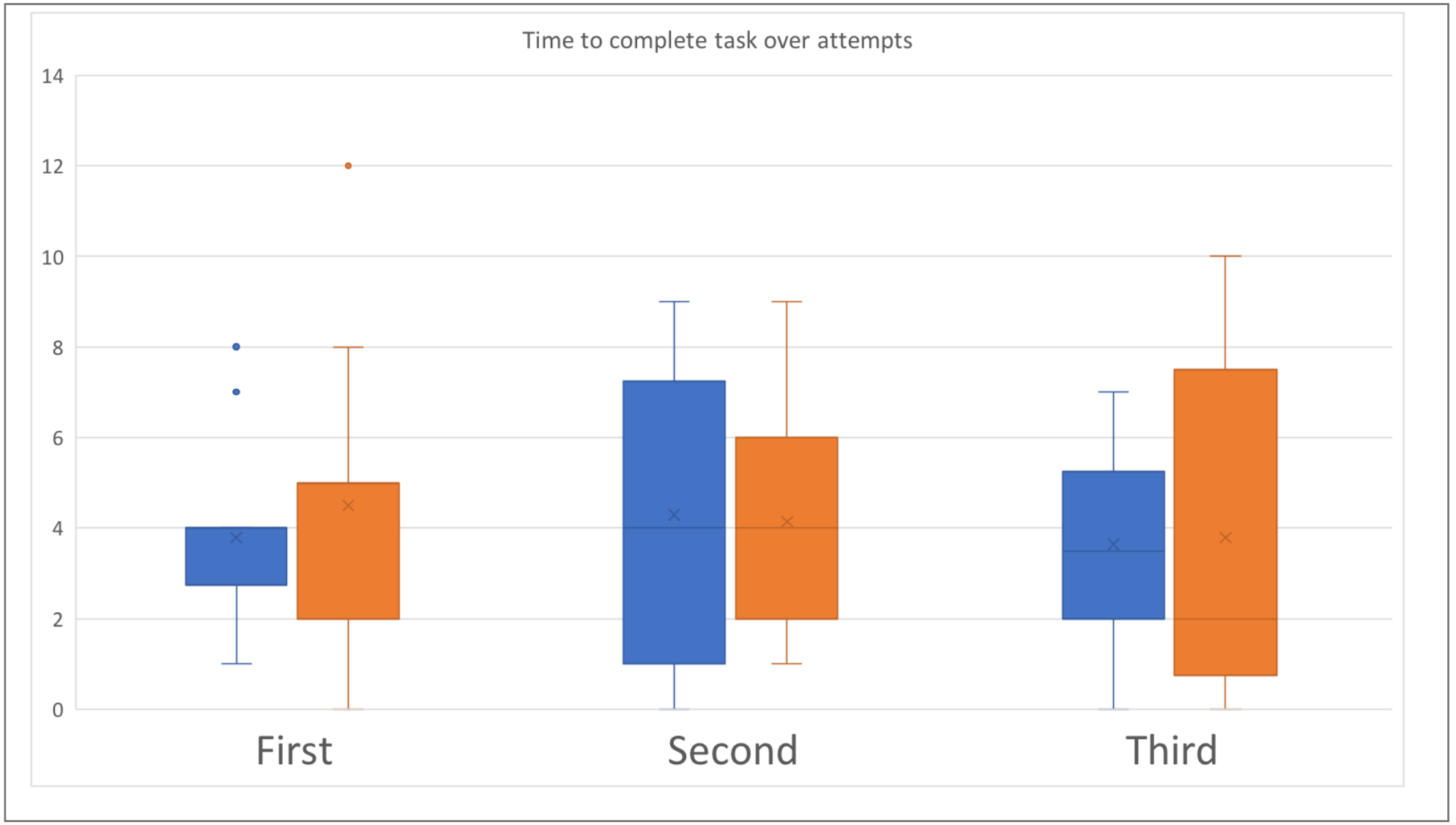



Image source: https://flowingdata.com/2008/02/15/how-to-read-and-use-a-box-and-whisker-plot/

# Think Aloud

- Users verbalize their thoughts, feelings, and opinions while interacting with a system
- Useful for capturing a wide range of cognitive activities
- Two variations exist:
  - Specific task
  - Open-ended
- One usability expert and a minimum of 4 users should be observed

Image source: https://www.pinterest.com.au/pin/453245149972280324/

# Think Aloud

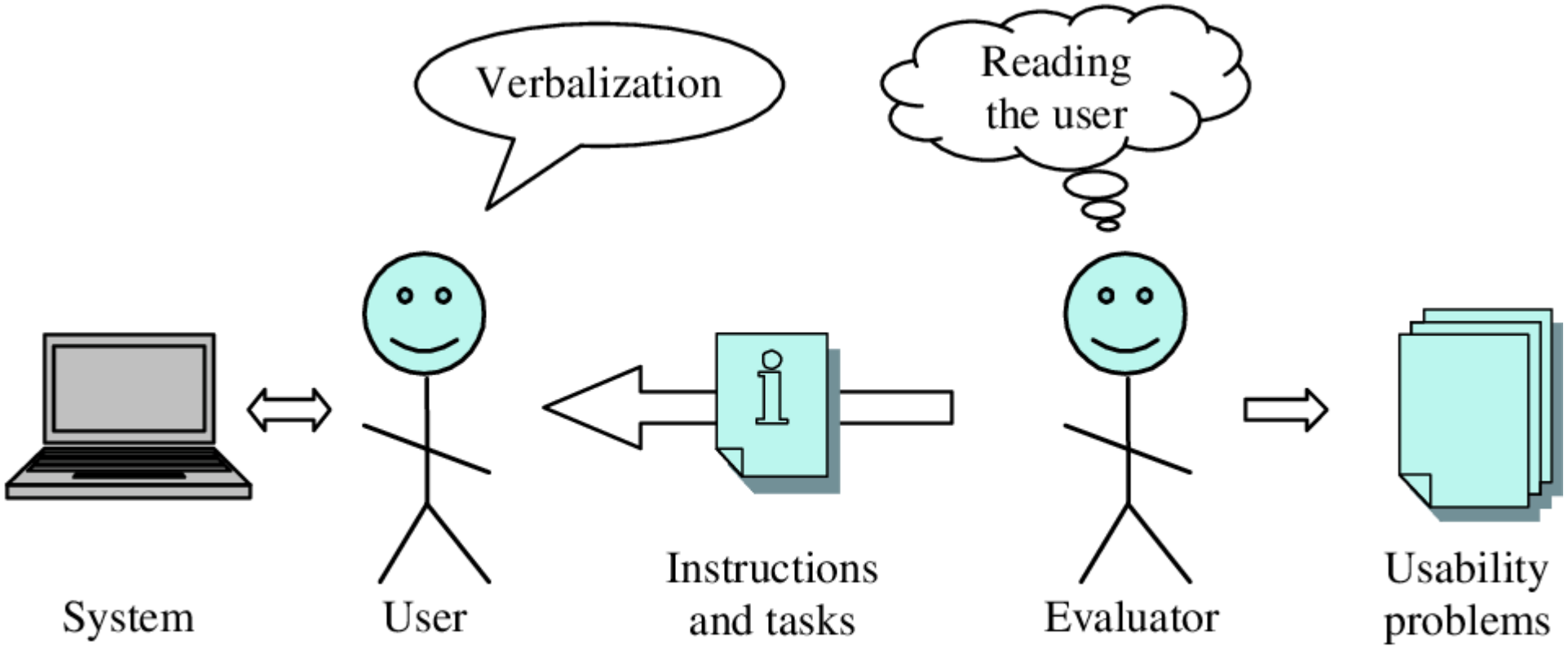



**Figure 1. Reference model of TA.**

Clemmensen, Torkil, Morten Hertzum, Kasper Hornbæk, Qingxin Shi and Pradeep G. Yammiyavar. "Cultural Cognition in the Thinking-Aloud Method for Usability Evaluation." ICIS (2008).

# Think Aloud Benefits

- Useful for understanding how the user approaches the interface
- Rapid, high-quality, qualitative feedback
- Broad range of detailed data
- Errors can be clarified
- Flexible
- Meaningful dialogue
- Versatile
- Usually the majority of major issues can be found

# Think Aloud Drawbacks

- Small sample size, so can be difficult to know the relative importance of problems identified
- Talking aloud changes the time spent on tasks
- Can be complex to undertake both as an experimenter and participant

# Presenting Findings

- Only make claims that your data can support
- Dependant on the audience, purpose, data gathering and analysis undertaken
- Graphical representations are always powerful
- Other techniques:
  - Rigorous notations (e.g. UML) may be too rigid
  - Storyboards and personas to create scenarios and use cases

# Summary

- There are many complex and socially-based experiences in HCI that cannot be easily quantified or experimentally manipulated
- There are many evaluations methods for usability, each uses a different approach and theory
- Need to understand the theory behind an approach to understand what specifically the questionnaire is measuring
- Data analysis depends on the data gathered

# Looking ahead…

- In our next session, we will look at **Evaluating Usability: "Expert" or "Non-User" Evaluations**