

Evaluation of technologies for collaborative interaction



| **tedco**

Summary & Learning outcome

The students learn how to evaluate the user experiences with the system. Students list the methods of evaluating user experience, and propose which method works best for their own system. Students also reflect on why studying user experiences is important, and in which phases of the design process this can play a meaningful role.

- Students are able to explain the difference between formative and summative evaluation
- Students are able to explain the difference between analytical evaluation and user testing
- Students are able to come up with examples of methods to evaluate collaborative technologies that fall within either of the four types

Recommended readings

- Ramage, Magnus (2010). Evaluating collaborative technologies: a simple method. In: Donelan, Helen; Kear, Karen and Ramage, Magnus eds. Online Communication and Collaboration: A Reader. Abingdon: Routledge, pp. 73–77.
- Sharp, Preece, & Rogers. 2019. Interaction Design: Beyond Human-Computer Interaction. **Chapter 14**. John Wiley & Sons. ISBN: 111954730X, 9781119547303

Content

- Why, when, and how to evaluate
- Evaluation approaches
 - Formative versus summative
 - Analytical versus user testing
- An evaluation framework
- Summary
- References

Why, when and how to evaluate

Why evaluate?

- Check that **users** can use the technology for collaborative interaction
 - Note that 'designers are not users' (Jacob Nielsen)
- Get feedback about early design ideas
 - Evaluation does not have to wait until the final product
- Find out which problems occur in the design
 - Allowing you to fix major problems before final design

When to evaluate?

- It depends
 - If the technology is new, evaluate after initial sketches
 - If the technology exists, this might be the first step
- Evaluation during design is called *formative evaluation*
- Comparison with benchmark or competitive technology for collaborative interaction is called *summative evaluation*

Formative Evaluation:
during development



Summative Evaluation:
after development

Formative Evaluation:
during development

Formative evaluation: Strengthens or improves the product design

Often formative evaluations are done at an earlier stage than summative evaluations

Summative Evaluation:
after development

Formative Evaluation:
during development

Formative evaluation: Strengthens or improves the product design

Often formative evaluations are done at an earlier stage than summative evaluations

Summative evaluation examines and summarizes the effects or outcomes of a product design.

Summative Evaluation:
after development

How to evaluate?

- Analytical evaluation
 - Involve an expert-role
 - One who is practiced in usability methods
- Usability testing
 - Testing the technology rather than the user
 - Use of the technology often recorded
 - Sometimes a specific hypothesis is tested

Formative Evaluation:
during development

Formative evaluation: Strengthens or improves the product design

Often formative evaluations are done at an earlier stage than summative evaluations

Analytical Evaluation:
without users

User Testing:
with users

Summative evaluation examines and summarizes the effects or outcomes of a product design.

Summative Evaluation:
after development

Formative Evaluation:
during development

Formative evaluation: Strengthens or improves the product design

Often formative evaluations are done at an earlier stage than summative evaluations

Analytical evaluation: Involves only experts, these experts report on possible problems. This can be seen as a prediction of what will happen when real users use the product.

Analytical Evaluation:
without users

User Testing:
with users

Summative evaluation examines and summarizes the effects or outcomes of a product design.

Summative Evaluation:
after development

Formative Evaluation:
during development

Formative evaluation: Strengthens or improves the product design

Often formative evaluations are done at an earlier stage than summative evaluations

Analytical evaluation: Involves only experts, these experts report on possible problems. This can be seen as a prediction of what will happen when real users use the product.

Analytical Evaluation:
without users

User Testing:
with users

User Testing: Involves real users, shows what really happens in the lab or in a more informal setting. The users involved in user testing do not evaluate the product, they only interact with it so that the team/evaluation expert understands their reactions.

Summative evaluation examines and summarizes the effects or outcomes of a product design.

Summative Evaluation:
after development

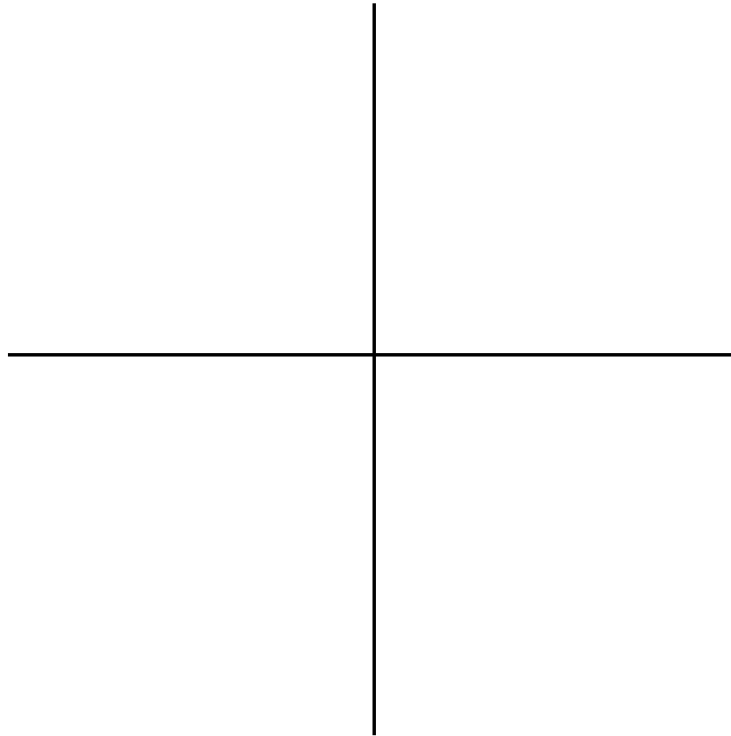
Evaluation approaches

Formative Evaluation:
during development

Analytical Evaluation:
without users

User Testing:
with users

Summative Evaluation:
after development



Formative Evaluation:
during development

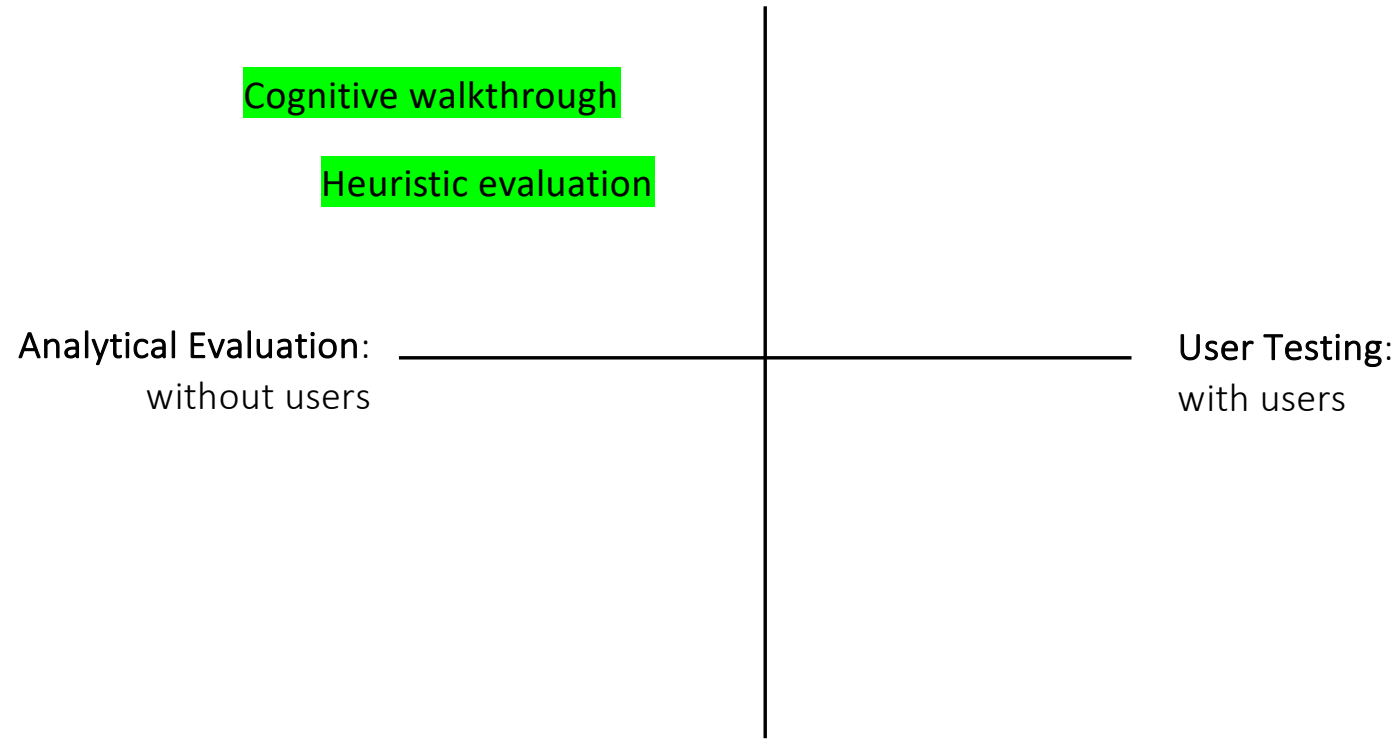
Cognitive walkthrough

Heuristic evaluation

Analytical Evaluation:
without users

User Testing:
with users

Summative Evaluation:
after development



Cognitive walkthrough

Steps "the happy path"
1
2
3

Before you can start a cognitive walkthrough, you need a complete, written list of the actions needed to complete the task with the interface — the **'happy path'**

Sometimes, creating the happy path is all you need to do to realize there is a problem with the interface

If your happy path has many actions, there's no need to continue with the review: you've found a serious problem already

Once you have the happy path, you're ready to start the walkthrough

Cognitive walkthrough

Steps "the happy path"	<i>Will the user realistically be trying to do this action?</i>	<i>Is the control for the action visible?</i>	<i>Is there a strong link between the control and the action?</i>	<i>Is feedback appropriate?</i>	Comments/issues
1					
2					
3					

One or more expert evaluators perform the evaluation by walking through the steps, placing it in the context of a **certain scenario**

They ask themselves **four questions**

Cognitive walkthrough

Steps "the happy path"	<i>Will the user realistically be trying to do this action?</i>	<i>Is the control for the action visible?</i>	<i>Is there a strong link between the control and the action?</i>	<i>Is feedback appropriate?</i>	Comments/issues
1					
2					
3					

One or more expert evaluators perform the evaluation by walking through the steps, placing it in the context of a **certain scenario**

They ask themselves **four questions**

The evaluator attempts to come up with a "success story" for each step in the process. If they cannot come up with one, they instead create a "failure story" and assess why the user might not accomplish the task based on the interface design. These insights are then used to improve the design

Cognitive walkthrough

- Advantages:
 - Can find problems before an expensive user test
 - Can find problems that would not be found when only testing with a small number of users
- Disadvantages:
 - The value of the data is limited by the skills of the evaluators
 - Analysis focuses on the words and graphics used on the screen (mostly suited for interfaces)
 - Lack of info on frequency and severity of problems
 - Labor intensive

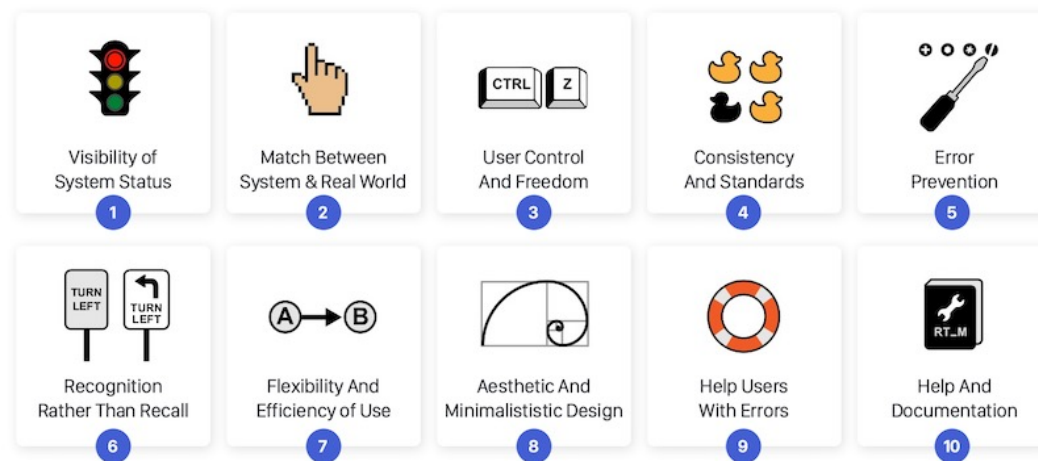
Heuristic evaluation

- First developed by Jakob Nielsen for desktop applications.
 - A cheap and easy evaluation method, and the most popular of the usability analytical methods
- Small set of evaluators (3-5) examines a technology and judges its compliance with recognized usability principles
 - Each individual evaluator inspects the interface alone
- After all evaluations have been completed the evaluators meet, exchange their conclusions, and aggregate their findings

Heuristic evaluation

- The most famous heuristics are the 10 usability principles developed by Jakob Nielsen

<https://www.interaction-design.org/literature/article/user-interface-design-guidelines-10-rules-of-thumb>



Formative Evaluation:
during development

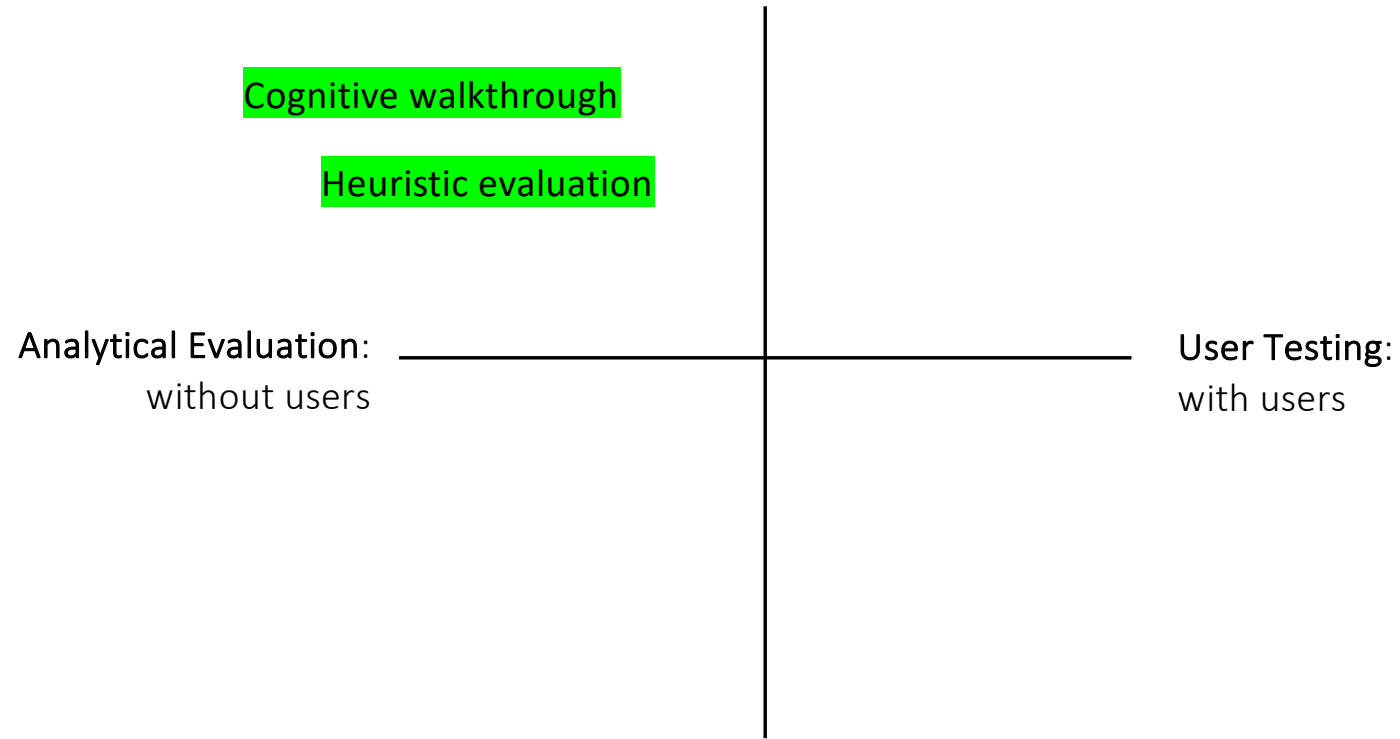
Cognitive walkthrough

Heuristic evaluation

Analytical Evaluation:
without users

User Testing:
with users

Summative Evaluation:
after development



Formative Evaluation:
during development

Cognitive walkthrough

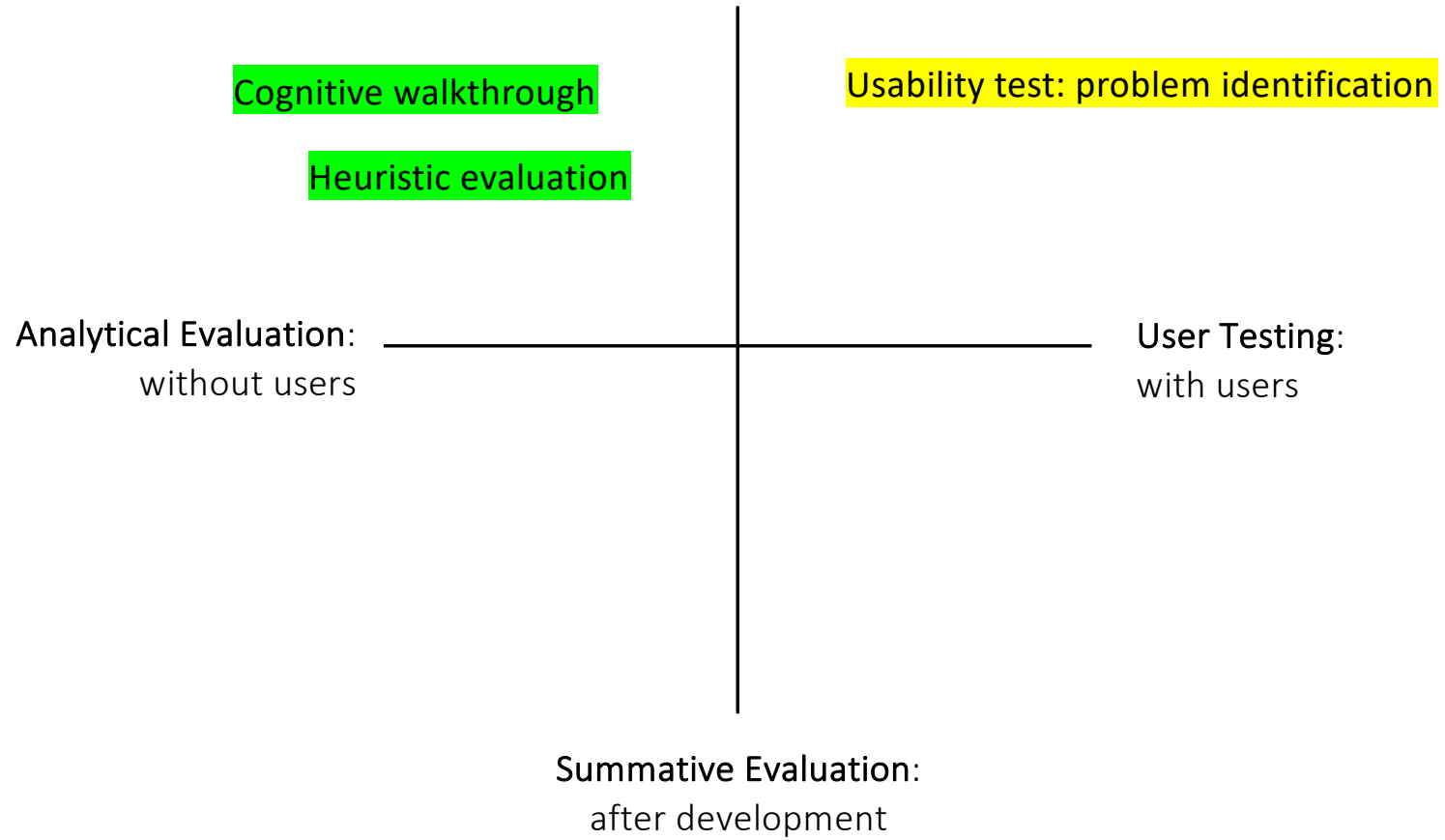
Heuristic evaluation

Usability test: problem identification

Analytical Evaluation:
without users

User Testing:
with users

Summative Evaluation:
after development



Usability test: problem identification

- Can be done on a (paper) prototype, a partly developed product, or an almost finished product
- Usually follows the Thinking Aloud protocol
- Goal: To understand how the user approaches the interface and what considerations the user keeps in mind when using the interface -> detect usability problems

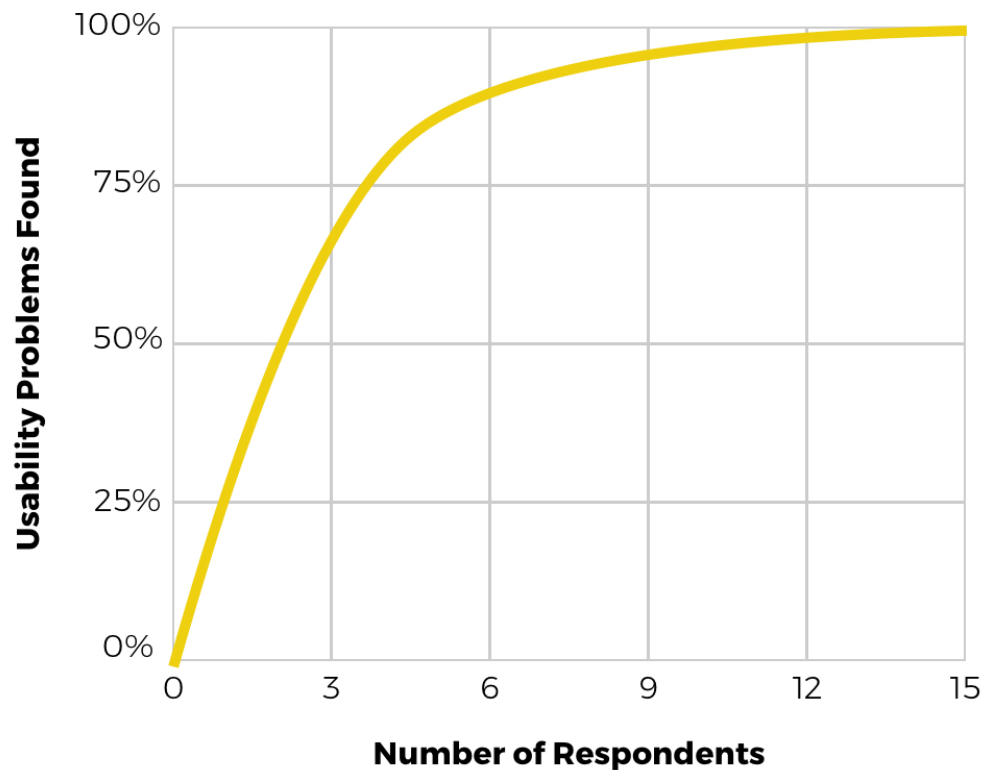
Usability test: typical misconception

- Many people think you do a user test on a finished product to prove/show that the product is good
- As you now know: User/Usability testing is not typically done at the end
- A usability test is often meant to detect problems, not to show that there are no problems -> a summative test can be done at the end if clear requirements have been defined beforehand (e.g., how long it should take for the user to perform a task)

Usability test: users

- Most often advised: 3-5 users
- The reason for this suggested number is that one user might find only a part of your product's problems
- The second user will find some additional problems, which might overlap with those of the first user or could also be new ones
- When do you stop finding new problems?

Usability test: users



What this graph tells us:

1. Testing with zero users gives zero insights
2. Testing with a single user already gives about 31% of all usability problems
3. 5 users give you insights on 88% of the problems
4. With 15 respondents, you will find 100% of the problems

Formative Evaluation:
during development

Cognitive walkthrough

Heuristic evaluation

Usability test: problem identification

Analytical Evaluation:
without users

Guidelines checking

User Testing:
with users

Summative Evaluation:
after development

Formative Evaluation:
during development

Cognitive walkthrough

Heuristic evaluation

Usability test: problem identification

Analytical Evaluation:
without users

Guidelines checking

User Testing:
with users

Usability metrics

Questionnaires (UEQ)

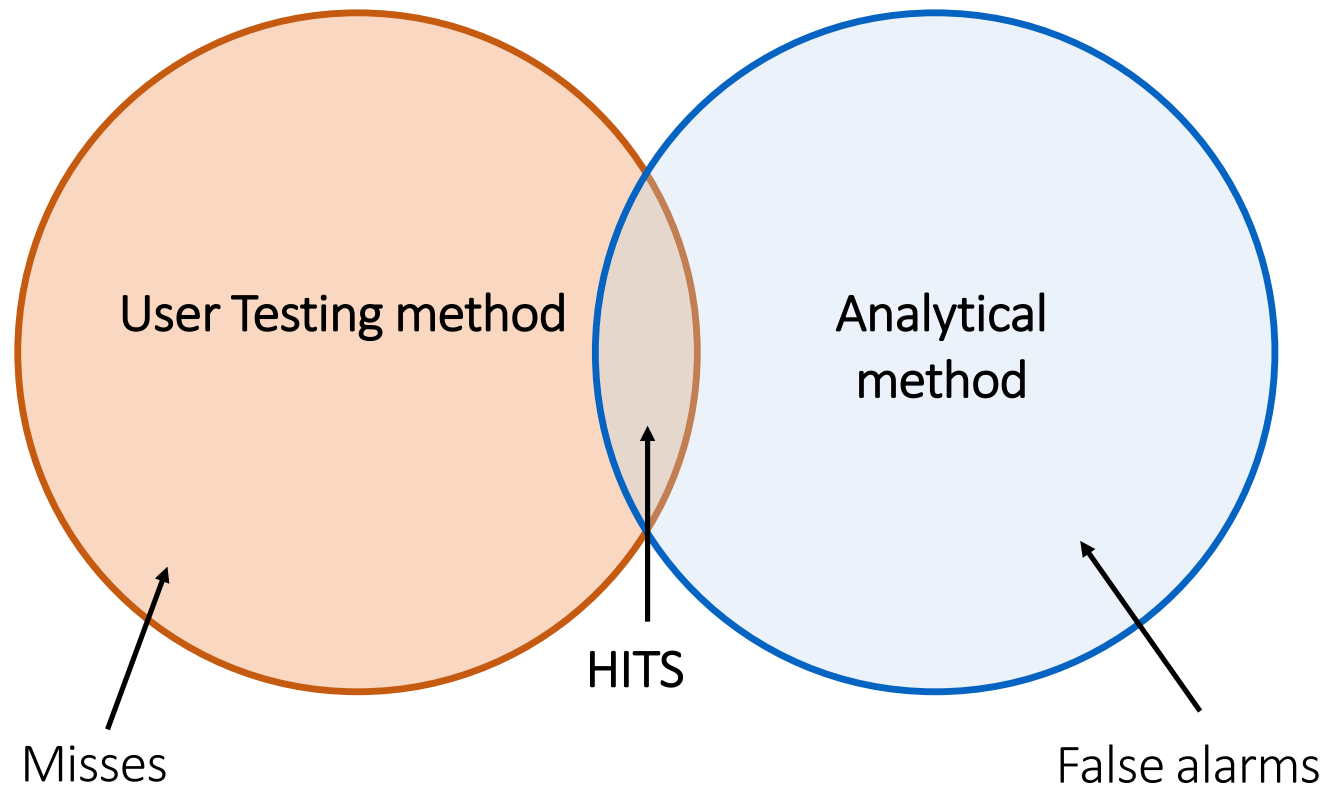
User Studies

Summative Evaluation:
after development

Summative methods

- Guidelines checking
 - The team makes a list of guidelines. An evaluator checks whether the product implements all guidelines and reports back to the team
- Usability testing metrics
 - If the team has set up metrics for the final product (e.g., a user should be able to perform a certain task within x seconds), a test can be done to see whether users are indeed able to do so
- Questionnaires
 - Questionnaires such as the User Experience Questionnaire (<https://www.ueq-online.org/>) could be used to assess UX metrics

User testing vs. Analytical methods



Examples with technologies for collaborative interaction

Example: Discord

- Discord is a communications app in which voice, video, and text chat are shared with others
- If we want to detect potential problems with the interface, we can do a user study
 - Ask users to open the app and do several tasks that are specific for the functionality of Discord
 - While doing the tasks, they could think out loud, allowing us to easily find out what barriers they face in terms of usability
- Outcomes of the user study will inspire further development, fixing some issues that harm its usability

Example: Discord (continued)

- Now imagine the app were to be used for tracking communication in a school where:
 - Each class has their own channel
 - Each channel has its own subgroups for different assignments
 - Each assignment is performed by several groups of students
- The users of the app are groups of groups, where each individual student could be in multiple groups
- The user study could take this into account by:
 - Including tasks in the think-out-loud protocol that span groups
 - Include users from multiple groups to combine experiences

Example: Miro

- Miro is an online collaborative whiteboard that enables distributed teams to work together and brainstorm with digital sticky notes, plan or manage agile workflows
- Imagine we have just begun developing the platform and we want to do a heuristic evaluation
 - We would ask experts to judge the app on usability principles
 - As the app represents a real-life whiteboard with sticky notes, we would check the extent to which it matches a physical board
 - We would also want to check the amount of control users have
- Together this will inform us about the expected usability

Summary

- A number of different evaluation approaches were introduced that can be used to evaluate the design and use of technologies for collaborative interaction
- The approaches can be classified as formative or summative, and analytical or user testing
- Example evaluation approaches of existing apps were shown

References

- Collazos, C. A., Guerrero, L. A., Pino, J. A., Renzi, S., Klobas, J., Ortega, M., Redondo, M. A., & Bravo, C. (2007). Evaluating Collaborative Learning Processes using System-based Measurement. *Educational Technology & Society*, 10 (3), 257-274
- Damianos, L., Hirschman, L., Kozierok, R., Kurtz, J., Greenberg, A., Walls, K., ... & Scholtz, J. (1999). Evaluation for collaborative systems. *ACM Computing Surveys*, 31(2)
- Mandryk, R. L., & Inkpen, K. M. (2004, November). Physiological indicators for the evaluation of co-located collaborative play. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 102-111).
- Malone, T. W., and Crowston, K. (1994). The interdisciplinary study of coordination. *ACM Computing Surveys*, 26(1), 87-119.
- Neale, D. C., Carroll, J., & Rosson, M. B. (2004). Evaluating computer-supported cooperative work: Models and frameworks. 112-121. Paper presented at Computer Supported Cooperative Work - Conference Proceedings, CSCW 2004, Chicago, IL, United States.
- Olsson, T., Jarusriboonchai, P., Woźniak, P. et al. Technologies for Enhancing Collocated Social Interaction: Review of Design Solutions and Approaches. *Comput Supported Coop Work* 29, 29–83 (2020).
- Ramage, Magnus (2010). Evaluating collaborative technologies: a simple method. In: Donelan, Helen; Kear, Karen and Ramage, Magnus eds. *Online Communication and Collaboration: A Reader*. Abingdon: Routledge, pp. 73–77.
- Reddy, M. C., Shabot, M. M., & Bradner E. (2008). Evaluating collaborative features of critical care systems: A methodological study of information technology in surgical intensive care units, *Journal of Biomedical Informatics*, 41 (3), 479-487
- Sharp, Preece, & Rogers. 2019. *Interaction Design: Beyond Human-Computer Interaction*. Chapter 14. John Wiley & Sons. ISBN: 111954730X, 9781119547303
- Sharp, Preece, & Rogers. 2019. *Interaction Design: Beyond Human-Computer Interaction*. Chapter 15. John Wiley & Sons. ISBN: 111954730X, 9781119547303
- Sharp, Preece, & Rogers. 2019. *Interaction Design: Beyond Human-Computer Interaction*. Chapter 16. John Wiley & Sons. ISBN: 111954730X, 9781119547303



Thanks for listening

Disclaimer: The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein