



Methods Review

From: An Overview of GUR Methods by Michael Medlock - *Games User Research*. Oxford University Press. – Presented by Lennart Nacke, <http://gameux.chicourse.com/>

GUR METHODS

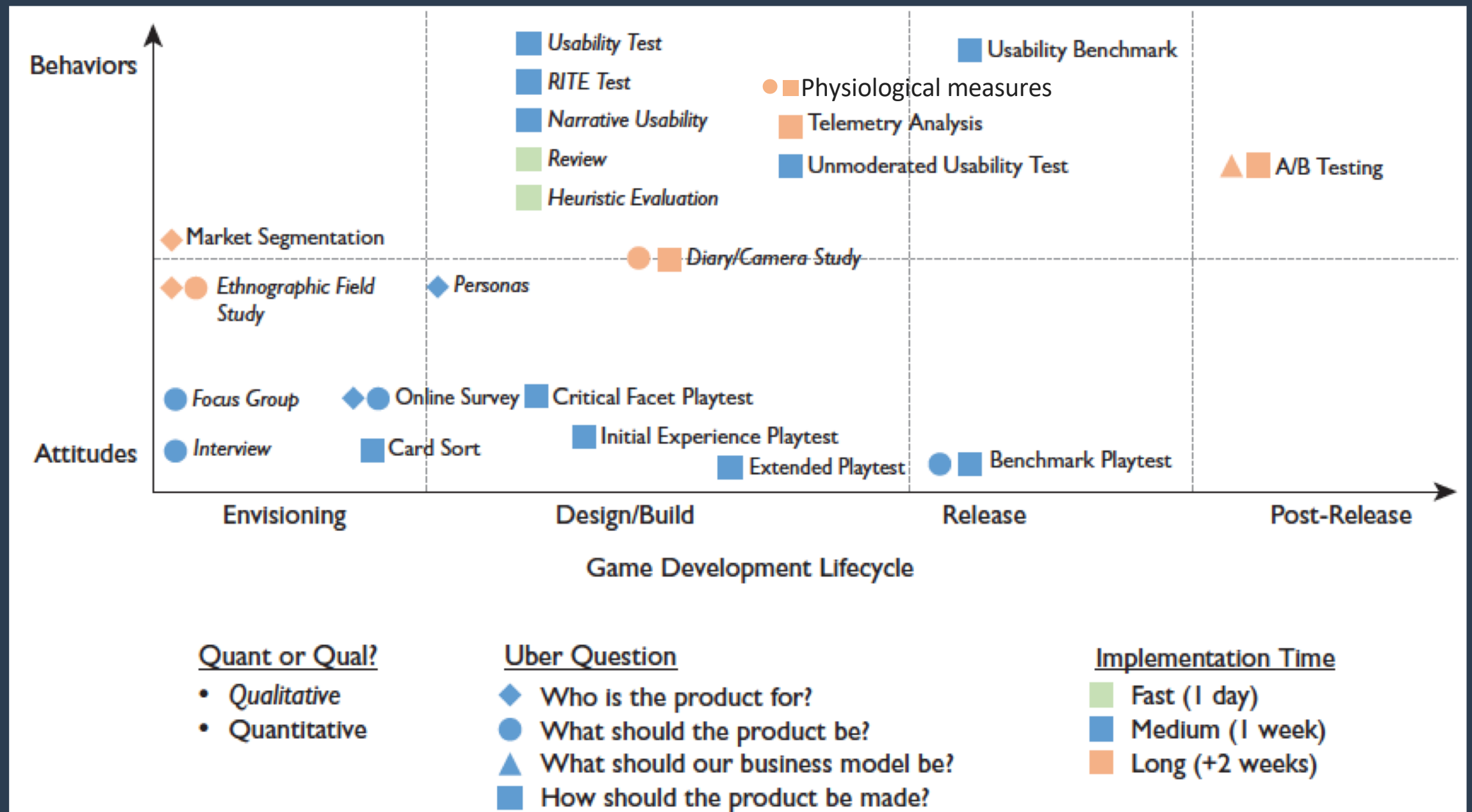
1. A/B Testing
2. Benchmark Playtest
3. Card Sort
4. Critical Facet Playtest
5. Diary Study
6. Ethnographic Study
7. Extended Playtest
8. Focus Groups
9. Heuristic Evaluation
10. Interview

11. Initial Experience Playtests
12. Market Segmentation
13. Narrative Usability
14. Online Survey
15. Personas
16. Review
17. Rite Test
18. Telemetry
19. Usability Tests
(Unmoderated, Benchmark, Regular)
20. Observation
21. Physiological Sensors

GAMES USER RESEARCH METHODS

- Who is the product for?
 - Who are the users?
 - Who should the users be?
- What should the product be?
 - What do our users do with our game?
 - What do our users wish they could do?
- What should our business model be?
- How should the product be made?

METHODS



An overview of GUR methods. Michael C. Medlock, Microsoft Game Studios. Chapter 7. Games User Research.

OBSERVATION



BEHAVIOURAL OBSERVATION

Especially for a game's designers watching somebody play the game is of high value. Seeing how players deal with the game's challenges, where they get stuck or frustrated in the game level, can lead to profound insights into the gameplay experience for game designers.

But: try to be wary of the Hawthorne effect.

Only tells them **what** is happening when players play their game, it does not answer **why** something is happening to a player or **how** the player felt when it was happening

USABILITY TEST

A 1:1 observational technique in which participants attempt to perform a variety of tasks with a prototype or game in development. Observers note what each user does and says and performance data are recorded. Test with about 5 users for each prototype.

Pro

- Good at identifying issues that keep users from playing a game well.
- Good at getting insight into why users are not able to accomplish tasks or meet the usability goals of a game.

Con

- Poor at understanding or quantifying attitudes. For example, if the game or game facet is 'liked' or 'disliked'.

REVIEW

One or more evaluators examine a game or prototype and judge if there will be issues based on their experience watching other users use similar games.

Pro

- Fast to implement. Can catch known problematic issues quickly.
- Often a good starting point.

Con

- Does not give actual behavioural or attitudinal data from participants.
- Will not catch all the important issues. Only as good as the reviewer used.

A/B TESTING

A controlled experiment in which two or more alternatives for a design are randomly assigned to users of the product 'in the wild'. Then the behaviour of interest is measured via telemetry to see which design performed better.

Pro

- Definitive answers to usage questions around the designs in question.
- Quantifies the impact of a design compared to another design.
- One of the few true experiments.

Con

- Difficult and expensive to set up.
- All alternatives have to be designed, coded and working.
- Does not tell the researcher why the winning design 'won'.

BENCHMARK PLAYTEST

A standardized attitudinal test run with a larger sample size (e.g., 35+ participants). Used to compare against other standardized attitudinal tests run in the exact same way on other products.

Pro

- Standardized way to compare how much things are liked or disliked.
- Allows for meaningful comparisons if a game or game facet is 'liked' or 'disliked'.
- It gives meaning to measures taken in the future.

Con

- Since it is self-report data it does not accurately tell the researcher why participants felt the way they did.
- Since the benchmark happens at the end of development, it does not help the game it is used for.
- Instead it helps other games that come after the benchmark as a reference point.

CARD SORT

A group of participants organize topics into categories that make sense to them and label these groups. Alternatively a group of participants place topics into groups that have already been created for them to see if the groupings make sense to them, or if there is similarity between individuals in understanding of the groups.

Pro

- Gives insight into how users think about how the information in a space relates to each other.
- Gives an excellent starting point for organizing menus or other hierarchical structures.

Con

- Does not give definitive answers to questions of grouping or hierarchy.
- Often final navigational behaviour is different from what users claim their navigational behaviour will be.
- Is a starting point, but does not completely finalize decisions around information architecture.

CRITICAL FACET PLAYTEST

A survey technique which assesses attitudes and perceptions about very specific core experiences (e.g., the aiming model in vertical levels, the camera for 3D platformers, the steering control in turns for a racing game).

Pro

- Gives quantified insight into user's attitudes about specific core game experiences.

Con

- Does not indicate if these core experiences are ultimately important to the overall enjoyment of the game.
- Since it is self-report data, it does not accurately tell the researcher why participants felt the way they did.

DIARY/CAMERA STUDY

A qualitative technique in which participants are provided with the materials and structure to record daily events, tasks and perceptions around a game in order to gain insight into their behaviour and needs over time.

Pro

- A relatively cheap qualitative way to track some behaviours and attitudes over time.

Con

- Not good for quantification.
- Can have some loss due to reliance on participants to consistently fill out the diary and take pictures.

ETHNOGRAPHIC FIELD STUDY

A holistic qualitative observational study of users in the context of their actual environment over a period of time.

Pro

- Deep qualitative observational insight into user behaviour.
- Gives a good idea of why the participants do what they do.

Con

- Does not allow for quantification and generalization of the behaviours and attitudes to others.

EXTENDED PLAYTEST

A survey technique which assesses the attitudes and perceptions users have of a game in development over an extended period of time. Often run across 2 days and 16 hours of testing.

Pro

- Gives quantified insight into user's attitudes about the game experience over time.
- Can be especially good at highlighting differences between 'levels' of the game.

Con

- Since it is self-report data it does not accurately tell the researcher why participants felt the way they did.

FOCUS GROUPS

Qualitative technique in which a group of people are asked about their perceptions, opinions, beliefs, and attitudes about a game or game experience. Questions are asked in an interactive group setting where participants are free to talk with other group members.

Pro

- Good for generating new ideas and getting participants to generate ideas by interacting with one another.

Con

- Poor at getting accurate behavioural or attitudinal data on how users either use or feel about games (since their interaction with other participants influences their answers).

HEURISTIC EVALUATION

One or more evaluators examine a game or prototype and judge its compliance with recognized game usability principles (the 'heuristics').

Pro

- Fast to implement.
- Can catch known problematic issues quickly.
- Often a good starting point.

Con

- Does not give actual behavioural or attitudinal data from participants.
- Will not catch all the important issues.
- Often only as good as the reviewer or the heuristics.

INTERVIEW

A 1:1 technique in which an interviewer asks participants questions about their perceptions, opinions, beliefs, and attitudes about a game or game experience.

Pro

- Good at understanding what each individual believes about their game experience.
- Can help augment the understanding of why users do what they do.

Con

- Poor at getting accurate behavioural data on how users use a game.
- Does not quantify a user's attitudes about a game experience.

INITIAL EXPERIENCE PLAYTEST

A survey technique which assesses the attitudes and perceptions users have of a game in development over the first 2 hours of use.

Pro

- Gives quantified insight into user's attitudes about the initial game experience.

Con

- Since it is self-report data it does not accurately tell the researcher why participants felt the way they did.

MARKET SEGMENTATION

A survey technique usually combined with cluster analysis used to divide the market for a product into groups of customers with identifiable needs and characteristics.

Pro

- Can be good for quantifying how portions of a market self-report as feeling or behaving.
- A good starting point combined with ethnographic work for understanding users.

Con

- If the questions used are poor, or not focused on the correct things then the needs and characteristics that segments are divided into can be meaningless or misleading.
- Relies on self-report, so can be incorrect about actual behaviour (especially purchasing behaviour).

NARRATIVE USABILITY

A 1:1 observational technique in which participants attempt to understand the narrative, from beginning to end, with a prototype of the narrative before it is implemented in the game.

Pro

- It can help identify areas of confusion, misaligned expectations, genre interactions, and gaps in logic allowing the designer to either correct or embrace these beliefs.
- It does so early in development before changes are costly.

Con

- It does not tell the researcher if the users will “like” the narrative.
- Or otherwise provide meaningful data on how the user feels about the narrative.

ONLINE SURVEY

A flexible survey technique which allows researchers to assess the attitudes and perceptions users have on a wide variety of topics. Can be qualitative or quantitative.

Pro

- Can give quantified insight into user's attitudes about many things. Very flexible.

Con

- Since it is self-report data it does not accurately tell the researcher why participants felt the way they did.
- Also the stimulus is rarely controlled.

PERSONAS

Fictional characters created to represent the different user types that might use a game in a similar way.

Pro

- Gives a team someone concrete to focus on when making a product.
- Can build team empathy for users and help with some fast decision making.

Con

- Can become over generalized or overused when the underlying data for the persona do not support the decisions being made based on it.

RITE TEST

A 1:1 observational technique in which participants attempt to perform a variety of tasks with a prototype or game in development, while observers note what each user does and says. Performance data are recorded. After each participant the game or prototype may be changed to see if the change solves the issue previously observed.

Pro

- Good at identifying issues that keep users from playing a game well.
- Good at getting insight into why users are not able to accomplish tasks or meet the usability goals of a game.
- Good at seeing if a change actually solves an issue previously observed.
- Excellent at getting teams in a 'fix' mentality.

Con

- Requires more buy-in from team to set up.
- Poor at understanding or quantifying attitudes.
- For example, if the game or game facet is 'liked' or 'disliked'.

TELEMETRY ANALYSIS

Automatic collection of behavioural data from users of the product. This can be on any behaviour that the system can 'count' (e.g., deaths, collisions, levels completed, time taken to do thing).

Pro

- Excellent at quantifying behavioural data. When done well gives an unparalleled assessment of 'what' the participant did.

Con

- Takes a long time to set up. If the wrong kind of data is collected it can be unhelpful. Can be very hard to analyse. Does not tell the researcher why the user engaged in the behaviour.

UNMODERATED USABILITY TEST

A 1:1 observational technique in which participants attempt to perform a variety of tasks with a prototype or game in development while an automated system notes what each user does and says and performance data are recorded.

Pro

- Good at identifying issues that keep users from playing a game well.
- Good at running something fast, at scale.

Con

- Poor at understanding or quantifying attitudes. For example, if the game or game facet is 'liked' or 'disliked'.
- Also because the sessions are unmoderated it is hard/impossible to follow up with questions to understand why participants did what they did in the moment.

USABILITY BENCHMARK

A standardized 1:1 observational test run with a larger sample size (e.g., 20+ participants). Focused on tasks and task completion. Used to compare against other standardized observational tests run in the exact same way.

Pro

- Standardized way to compare how users perform with products.
- Allows for meaningful comparisons of observable behaviour on a game or game facet.
- It gives meaning to behavioural measures taken in the future.

Con

- Since it is often focused on task or task completion it does not holistically quantify attitudes well.
- Also, since the benchmark happens at the end of development, it does not help the game it is used for.
- Instead it helps other games that come after it as a reference point.

TAKEAWAYS

1. Ensure that your company can grow its UX maturity level over time
2. Use the method that provides **most insights** to your problem
3. **Combine** your methods to make up for shortcomings of methods with strengths of other user research methods