Exercise 2: Identifying problems

Scenario: Your friend Frejman approaches you and describes a playtest he recently conducted for a single-player mobile puzzle game targeted at young children. The goal of the test was to evaluate first-time user experience and player understanding.

Challenge: Frejman has some notes on his playtest. Frejman is looking for advice on how he can improve his future test designs.

Frejman's notes on the playtest:
- There were 12 participants between the ages of 25 and 40
- Participants played for 10 minutes each, some finished the tutorial and some didn't
- Participants were recruited online via social media
- All participants had at least moderate prior gaming experience
- Tests were conducted in a variety of settings, half were in the lab, two were outside on the street, and four were in participants' houses
- Some of the sessions had multiple participants back-to-back, for these the participants were seated together and played one after the other
- Participants were interviewed right after gameplay with the following set of questions:
    - Why did you like this game?
    - Was the tutorial easy?
    - Do you feel like you understand how to play?
    - Do you like puzzle games?
    - Would you play this game again? Why or why not?

Identify flaws with the design and/or execution of this test and explain why they are problematic.

Exercise: Discuss at your table 10 minutes then group then class discussion 10 minutes (20m)

Exercise 2: Bonus Challenge: More UX test problems

Scenario: Another friend of yours is telling you about a UX test she knows some of the analytics people conducted at the company Immersion. She is worried there might be some issues with bias and validity in the conclusions.

Challenge: In the following example, locate **at least five** problems with experimental bias or validity. Which group can find the most?

A group of volunteers, all hard-core fans of the zombie-horror FPS game "Undead Overkill 2019", are recruited for a user test with the company Immersion, which developed the game and aims for a global release next week. The game is two-player coop against AI agents. Test takes place in the offices of Immersion and is an observation test. Test rooms are painted lime green. Chairs are comfortable. Participants can see each others screens. There are no pre-test questions asked of the participants. There are 20 volunteer participants, all adults aged 18-23 and all male. Two of them are university professors.

The purpose of the user test is to evaluate two specific levels (levels 1 and 2) of the game and the overall UX measured across five pre-determined constructs that the lead user analyst – a quantitative specialist - thought were cool. These are "kills", "deaths", "fun", "distance travelled" and "most damage done with one shot". The constructs are self-report, with one question for each. Response options are on a 7-point likert scale, e.g.:

1. *How many kills did you score during the session? 0 – none .... 6 – more than anyone else!*
2. *How many times did you die during the session? 0 – heroes never die! .... 6 – a lot*
3. *Was the game fun? 0 – yes! .... 6 – most fun I ever had!*
4. *How many meters do you think you covered in the game? 0 – none ... 6 – 5000 meters*
5. *How much damage did your best shot do? 0 – no damage, I hid nothing .... 6 – 20000 points*

The analyst suggests take the mean value of each scale and then the mean across the five questions for an accurate measure of the UX of playing the level. The analyst is planning on calculating the UX score for the groups playing each level to see which one is best.

18 participants tried level 1, 2 tried level 2. Participants play in groups of 2, sequentially. Two of the participants are friends and ask to play together. Halfway through the experiment, one of the observers starts helps one of the participants. One of the participants had problems completing the self-report questionnaire due to a phobia of questionnaire completion.

The room the UX tests were carried out in have, due to an unfortunate coincidence, a direct view of an active slaughterhouse. During one of the tests in the 18 participant group, one of the level 1 developers comes running in and spends five minutes accusing the active group of 2 players of being stupid and not playing the game right. Another of the groups has problems playing the level due to bugs. One player uses a cheat code to score a large number of kills, another is obviously suffering from the cold. One finds bunnies very cuddly. Of the 18 participant group, 7 teams play on a Monday morning, the remaining two groups on a Thursday afternoon. It is nice weather both days. The level 1 participants play for 20 minutes, the level 2 participants for 30 minutes. One observer in the 18 participant group is wearing a humorous Christmas sweater. The level 1 groups get water to drink, the level 2 groups strawberry cordial.

Exercise: Discuss at your table if you have time left from the previous exercise.