

Final Project Proposal – Video games as tasks for cognitive science research

Topic:

Throughout this semester, I have been working with the Video Games datasheet from the Corgis database. This dataset appealed to me because I hope to study human cognition using video games as cognitive tasks in my future research. I believe that if we want to understand the bridge between subjective and objective experience, we need better ways of probing the mind and that video games are the best method available. The research team Thompson et al. shares this belief. In their 2013 paper titled “Video Game Telemetry as a Critical Tool in the Study of Complex Skill Learning” published in PLoS ONE the group made the bold claim that real-time strategy games can serve cognitive science in the same way that *Drosophila* has served biology. They go on to list 7 reasons why video games make for good experimental tasks.

1. The environments are dynamic as opposed to the traditional psychology task (Stroop test) that are static. – More ecologically valid.
2. Games are psychologically engaging, leading to motivated participants. – Greater Presence
3. Telemetry methods enable the accurate and precise collection of motor performance and attention allocation data. – Higher resolution data
4. Many performance measures are collected non-invasively. – Less likely to produce expectancy or framing effects.
5. Massive amounts of performance data. – Central limit theorem
6. A variety of data points. – Facilitates between and within-subject investigations.
7. Large range of expertise. – Individual strategy and optimization comparisons.

I would add to this list that:

8. That video games fulfill the ethological criteria of being naturalistic, meaning it isn't a task mimicking human behavior, it is human behavior that can be ethically observed.
9. Video games are physically stationary and mentally active, allowing a variety of physiological equipment to be incorporated into this research.
10. They facilitate repeatable investigations with a temporal component. – Many instances of the same scenario can be played through, and timestamps of importance can be recorded and compared alongside physiological and neurological measures, like respiration rate or pupil dilation and EEG or fMRI.

Altogether, these points create the argument that video games should be considered for the next generation of cognitive investigations. However, with any novel approach, a course forward has not been set. Therefore, my final project will be a data-driven selection of game titles that are suitable for use in behavioral research.

Dataset:

The [dataset](#) itself was created by Austin Cory Bart in May of 2017, but the data was originally collected by Dr. Joe Cox. This CSV file contains information about the sales and playtime of over a thousand video games released between 2004 and 2010. This dataset provides 36 data points (columns) about the 1213 games (rows). Some data points that I found interesting in the context of my project were:

- Genres – A string type variable containing the list of genres the title belongs to (titles falling under multiple genres were coded as such)
- Playstyles Polled – An integer variable showing the number of players that completed the game in any way.
- Average Completionist Times – A float variable displaying the average time that players reported completing everything in the game, in hours.
- Playstyles Rushed – A float variable representing the fastest time that players reported completing the game in any way, in hours. This is the maximum over all the other categories, meaning that this is the fastest single reported time across all of the categories.

Unfortunately, all the “How Long to Beat” data was collected via self-report, which is not ideal, but likely fine for the purposes of this type of investigation since this is merely for identifying possible candidates for a class project. A second concern about this dataset is that many of the games are outdated, but that may work in favor of research since most participants would at this point have no prior exposure to the game titles. Overall this dataset provides path-independent variables about playstyle and popularity for a large range of semi-immersive game titles.

Research Questions:

Currently, the study of psychophysiology falls short in two major ways. First, many of the results from primary literature studies are failing to be replicated by other research teams, which has been dubbed the *replication crisis*. Secondly, many of the tasks used in this field of research lack *ecological validity*, the measure of how test performance predicts behaviors in real-world settings (<https://www.britannica.com/science/ecological-validity>). Shoring up these shortcomings means better diagnostic tools for mental disorders, generalizability of results that may possibly be incorporated into automated technologies, and most importantly to me, providing insights into the nature of being a subjective experience navigating an objective world.

Thus the questions that can be asked are:

1. How can video game data be used to address the replication crisis?
2. Are there game titles that could be considered ecologically valid tasks for psychophysiological studies?

A combinatory solution to these two questions is to find relatively short, highly repeatable, and popular game titles that simultaneously probe cognitive processes like executive functions (planning/decision-making), memory, and spatial navigation. This solution addresses the replication crisis by facilitating multiple trials in highly controllable and consistent task environments, as well as, increasing the ecological validity by being a function-led dynamic cognitive task. Thus the overall focus of this project will be finding a subset of suitable contenders for this type of investigation.

Visualizations/software:

The visualization software I intend to use is Tableau. I feel that using the Gephi program will not produce any meaningful visualization for my project and that most of what I would like to do with an R/Python environment can be done much more efficiently in the Tableau program. This program seems to be best suited for my analysis because of its interactivity and user-friendly nature. For my visualizations, I envision using stack bar graphs, line graphs, and pie charts to display the process of eliminating game

titles, along a variety of exclusion criteria, such as genre types, a playtime range, console type, and title ratings. The aim of these exclusion criteria is to find a game that strikes a balance between experimental practicality and informative promise. My gut tells me that driving games developed for stand-alone consoles that take between 15 and 25 hours to finish, which are also rated E for everyone are the best candidates for this work.