**Project details (continued)**

Numerous theories within the field of psychology point to the idea that for a given individual, there is an optimal level of difficulty for a given task. Moreover, these theories state that as the difficulty of the task moves away from this optimal level, the individual’s levels of enjoyment, performance, and rate of learning will suffer. However, there is relatively little literature on finding this optimal level or a proxy for it.

In the following study, I will attempt to find a rough approximation for this optimal level of difficulty. Since difficulty is an abstract concept which cannot always be measured directly, I will use the related quantity “percent correct” as a proxy for the relationship between the difficulty of a task and the level of skill which an individual possesses for that task. I hypothesize that the majority of individuals will prefer a percent correct between 70 and 90 for all tasks.

It is possible that a participant’s sensitivity to incorrect trials may play a role in how difficult the participant wants the tasks to be. A possible measure of this sensitivity to negative results on a task is given by the Neuroticism (N) personality dimension from the Big Five personality inventory. High-N individuals tend to have stronger negative emotional reactions, thus it seems likely that high-N individuals will give more mental weight to incorrect trials. I hypothesize that high-N individuals will seek to relieve themselves of this exposure to negative stimuli, and thus that there will be a positive correlation between the a participant’s N and his or her preferred percent correct on all tasks.

The purpose of this study is three-fold. The first purpose is to establish the percentage correct on a computer-based novel task which corresponds to the preferred difficulty level of a participant. The second purpose is to validate an algorithm for predicting the difficulty of future trials of a task based on a participant’s outcomes on recent trials of that task. The third is to try to establish a relationship between a participant’s scored N level and that participant’s preferred percent correct at optimal difficulty.

In order to generate data for this study, participants will complete 25 trials on each of three tasks. For the first 15 trials of each task, the difficulty level may vary from task to task as controlled by an algorithm. For the next 10 trials, the difficulty level will be fixed at a level decided on by the participant. The first task is a visual discrimination task, the second is a verbal discrimination task, and the third is a kinesthetic action task. Screen captures of these tasks can be found in Appendices A, B and C, respectively.

The visual discrimination task consists of estimating the relative sizes of circles which circumscribe triangles. The triangles are drawn but the circles are not. Small differences in the sizes of the circles present a harder task, while large differences provide an easier one. The triangles are shown on the screen and the participant is asked to determine which of the undrawn circles is larger, or if they are the same. The difference in radii between the circles can vary from one pixel to 50 pixels.

The verbal discrimination task consists of counting the number of words in a list which rhyme with a target word in a small amount of time. The more words in the list, the harder the task. List words fall into four categories. As an example, consider the target word “our”. The four categories are: rhyming lookalikes (“sour”); rhyming non-lookalikes (“tower”); non-rhyming lookalikes (“pour”); and non-rhyming non-lookalikes (“grower”). The participant has four seconds to examine the list. When these four seconds are up, the list disappears and the participant is directed to make a choice as to how many of the words in the list rhymed with the target word. The number of words in the list can vary from one to ten.

The kinesthetic action task consists of navigating a ball through a procedurally generated maze to a target square. The participant uses the up, down, left and right arrow keys to move the ball. If the ball touches the wall, it will reappear at the start point. The larger the ball, the harder the task. The radius of the ball can vary from one pixel to 65 pixels.

Total time per participant will be approximately 10 minutes. Task types will be completely counterbalanced such that each participant has an equal chance of receiving the task blocks in any given order.

QIVs for this study will be “game experience” (GE), “task type” (TT), and “neuroticism” (N). GE has two levels: “low”, and “high”. “High” GE means a participant plays video games for five or more hours per week on average. “Low” GE means a participant plays fewer than five hours of video games per week on average. TT refers to the type of task in a given trial. The three levels for TT are “visual”, “verbal” and “kinesthetic” (please see Appendices A, B and C for examples). N is scored on a scale from 1 to 5 (please see Appendix G for a description of the questions and scoring mechanism of the Neuroticism scale). GE and N are between-subjects variables, and TT is a within-subjects variable. DVs will be “chosen difficulty level” and “was participant successful?”. For each item in each task, there will be one recorded value for each of these DVs. In addition, three pieces of demographic data will be collected for each participant: age, gender, and household income.

The tasks will be administered via a program which can be accessed anonymously on the website www.ndnuvideogamestudy.com (please see Appendix I for a screen capture of the website). When a participant completes all trials, data will be transmitted to a server. The server will store data in a database which contains no identifying information.

20 participants will be needed for the pilot study, and 100 participants will be needed for the full study. These 120 participants will be recruited through two methods: flyers posted around the NDNU campus, and online posts on <http://www.reddit.com/r/samplesize>, [www.facebook.com](http://www.facebook.com), and on <http://sfbay.craigslist.org/etc/>. The recruitment flyer can be found in Appendix D. The online post can be found in Appendix H.

Before starting the tasks, participants will be prompted for their ages. Individuals below the age of 18 will not be allowed to continue. After the age verification step, participants will be shown a consent form (please see Appendix E), and will have to click on a button labelled “I accept” to continue. On this page will be a suggestion to print the consent form. Once consent is given, the participant will be asked for his or her gender and household income. Next, the participant will be guided through a series of eight to twelve questions which constitute a Neuroticism personality inventory (please see Appendix G). Once all of this is completed, the participant will be given the three tasks in random order. After the tasks are finished, the program will display a debriefing form (please see Appendix F), along with a suggestion to print this form. Before the program closes, participants will be able to add comments to their submissions.

Once data are collected, they will be analyzed. Averages will be taken of the percentage correct for each participant and task. Grand means will be established for each task type, each game experience level, and for the entire set. MANOVA and ANOVA tests will be run to see if there are significant differences in percentage correct at preferred difficulty for experience level, task type, and neuroticism score. I expect relatively small variation in values for game experience and task type, but I expect a significant main effect of neuroticism score on percent correct. If the p-values are large, I will conclude that percentage correct is a portable measurement which can be applied to find an optimal difficulty. If the p-values are small and significant, I will conclude that there are factors which affect the optimal percentage correct for an individual.

The second analysis will involve comparing a prediction from a curve-fitting procedure using data from the first 15 elements of a task to the actual outcome from the next 10 elements of a task. If the predictions line up with the actual outcomes, I will conclude that prediction by curve fitting is a viable method for establishing optimal difficulty. If not, I will discard the method and try to find something more appropriate for future research.

A demographic analysis will also be run to see how the sample in this data set compares to the general population. Given the subject matter and the proposed modes of gathering subjects, I expect the sample will skew toward being male, young, and high-income.