Creating a Usable Subject and Experimenter Interface for a Computerized Assessment of Cognitive Flexibility

Stephanie M. Pratt, Christian Gonzalez, & William Benson
George Mason University

Cognitive flexibility (CF) is an individual's ability to switch from one cognitive strategy to another based on cognitive task demands. The predominant method of measuring CF currently is the Wisconsin Card Sort (Grant & Berg, 1948). Our client created a paper puzzle to measure CF with less training and overhead costs for the experimenter (Figueroa & Youmans, 2011). Our asked us to help design a computerized version of the puzzle , in order to measure more variables in addition to just the overall time to complete the puzzle. This paper discusses our process of determining requirements, designing, usability testing, and completing iterative designs of the CF puzzle's graphical user interfaces.

INTRODUCTION

Cognitive flexibility (CF) is generally thought of as an individual's ability to switch from one cognitive strategy to a more optimal strategy based on changing task demands (Scott, 1962). Extensive research has linked CF to an individual's ability to adapt pre-existing knowledge, strategies or tools in a novel way to meet situational demands (Sprio, Feltovich, Jacobson & Coulson, 1991). CF has been applied to studies in educational psychology (Spiro & Jehng, 1990), clinical diagnoses of mental disorders (Dennis & Vander Wal, 2009) and studies of creativity and functional fixedness (Frensch & Funke, 1995). CF also has various neurological correlates to attentional control and executive function (Elslinger & Grattan, 1993).

Currently, the Wisconsin Card Sorting Task (WCST) developed by Grant and Berg in 1948 is the definitive measure of CF. Since its inception, the WCST has been adapted numerous times. However, the rules of the task have not changed very much over this measure's 63-year history. Participants have a deck of cards which they are asked to sort into 4 piles. Cards are to be sorted by one of three rules: by shape, shape color or number of shapes on card. Participants do not know the current rule and receive experimenter feedback of "correct" or "incorrect" based on how they are sorting. Once 10 cards are successfully sorted according to the active rule, the rule switches and participants need to start sorting by the new rule to receive positive feedback and continue with the task. This particular measure has been cited extensively not only for its use as a measure of CF, but as standard measure of executive function. However, due to the extensive level of training required to administer and score, the WCST is often quite a cumbersome and laborious measure to administer.

Our client, Ivonne J. Figueroa, and Dr. Robert J. Youmans created a new method of assessing an individual's CF (Figueroa & Youmans, 2011). They designed an entirely new method: a paper puzzle in which participants were instructed to follow 3 matching rules (shape, shape color, background color) to traverse through a grid of different shapes and color combinations (see Figure 1). They found that "trials to complete first categories" (one of several WCST dependent variables) significantly correlated with total puzzle completion time-the puzzle's only dependent measure.

However, as the study addressed, the paper puzzle proved to be limiting in some aspects. Measuring puzzle completion time alone was not sufficient information to warrant a new valid and reliable method of measuring CF. Using a paper and pencil method made it impossible to accurately measure different variables already present in the WCST.

Our clients' solution to this was to design a computer program that let participants complete the puzzle, while accurately collecting additional related variables of interest (e.g. trials to complete first category and number of preservative errors). As with any new interface, user experience should be a central concern. Our goal was to ensure that experimenters would want to use this tool, and participants would be able to effectively complete the task.

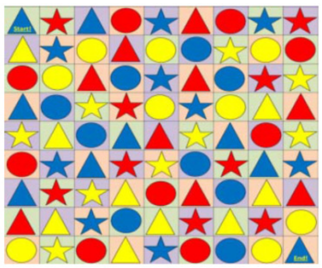

Figure 1. An Example of the Paper Puzzle

Though our client had already planned many of the new features of the program in advance, our role was to ensure the new design is usable, while adding functionality as needed. Through an iterative design process utilizing a mixed methodology, an effective and user-friendly version of Figueroa & Youman's (2011) Cognitive Flexibility Puzzle was fully realized.

The Design Process

Early in the development process, we created a list of "needs," "should haves," and "nice to haves" to ensure that our clients and our own vision were congruent. These requirements served as valuable guidelines for each step in the design process. They allowed us to have enough freedom to create a usable product, but enough structure to ensure our client's satisfaction with the end result.

One of the improvements initially suggested by our client was the addition of a "fog-of-war" feature, which would occlude all the other moves (puzzle spaces) other than the immediately available moves. This was intended to prevent participants from planning ahead in the puzzle. Since strategy switching is the primary concern in the study CF, any other cognitive processes create noise in the data.

Other "needs" included: collection of specified variables extant in the WCST, a training methodology for the participant to go through to understand how to use the puzzle, and experimenter output. These requirements served as the foundation of a successful final product.

Some of our "should haves" included features such as directional keypad puzzle navigation, grayed-out previous moves, and full screen modes. These types of additions were not completely necessary, but for adding these features added more usability and functionality.

The last set of suggestions was the "nice to haves." These suggestions included features that would add additional layers of functionality to the program, or further enhance the users' experience. These features were only to be implemented if time allowed. Some examples of these were: randomized puzzles, exporting of data into a comma separated value (.csv) format, saving data to an online database to ensure it can be accessed easily from other computers. Once these requirements were delivered to the client, they were free to add or modify any other requirements.

A calendar of deadlines was also constructed to ensure that the development of the tool stayed on track. Using a detailed project calendar allowed for little lag throughout the entire design process, and provided a clear sense for the group as to when tasks were going to be completed.

After initial planning and structuring of how our time would be spent, our team began work on the initial prototypes. The prototypes, created using Microsoft PowerPoint 2011, were of medium to high fidelity. The main purpose of these initial prototypes was to demonstrate the information architecture and expected functionality of the interface. During this phase, we also proposed the use of an independent experimenter interface to allow experimenters to manage and administer multiple series of experiments in addition to the output requested. Our client was pleased with the initial design of the prototypes, as well as the new proposed features so we moved into user testing

METHOD (STUDY 1)

Participants

Ten participants (9 male, 1 female) from George Mason University volunteered to complete Study 1. Participants were all graduate students enrolled in the Human Factors and Applied Cognition Program. Five participants completed a usability test of the puzzle and training interface. Five participants completed a usability test of the experimenter interface.

Materials

Puzzle. Originally, we had developed a functional PowerPoint prototype. However, due to a miscommunication, our clients' software developer had already submitted a Shockwave Flash (.swf) version of the puzzle prototype (see Figure 2). We decided to test this version instead of our own.

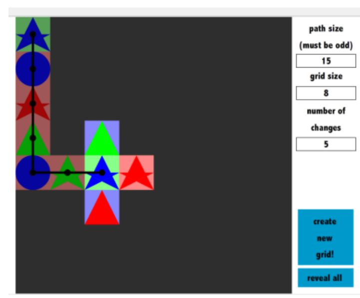

Figure 2. Puzzle Prototype Interface.

Puzzle training. A brief training PowerPoint was created to ensure participants understood the rules of the puzzle as well as how to interact with the puzzle interface.

Experimenter interface. A PowerPoint based, clickable experimenter interface was developed. This allowed us to evaluate users' navigation strategies as well as overall ease of use. For example in Figure 3, the left window ("Main Menu") provided a number of clickable options. When users clicked on "Puzzle Settings", for example, they would then be directed to the right window ("Puzzle Settings").

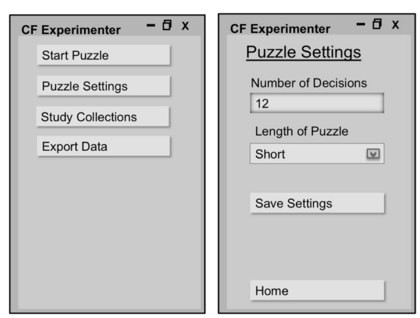

Figure 3. New Experimenter Interface Design.

Procedure

Experimenter interface. Participants were asked to complete 5 simple tasks with the interface: edit a study collection, create a new study collection, check grid size of puzzle, export data and start a puzzle. We asked users to provide a verbal protocol as they navigated the menus and give comments and feedback afterwards.

Puzzle and training. Participants were asked to follow on-screen directions, rather than complete a series of assigned tasks. Participants completed the training first, then the puzzle prototype. Upon finishing, participants verbally completed a brief usability questionnaire assess the usefulness of the training and the ease of use of the puzzle.

RESULTS (STUDY 1)

Experimenter Interface

We identified several issues with naming conventions used as well as the information architecture of the interface. Several participants were confused by the term "Study Collection", unsure if it was meant to collect data or aggregate different studies. Participants also reported feeling having separate settings for the puzzle and a study collection was redundant and confusing. Furthermore, only half of the participants used the "save settings" option indicating it was not adequately salient.

Puzzle and Training

Participants found both the puzzle and the training easy to use. No participants reported any major issues in using either interface. 4 out of the 5 participants found the training helpful and the only one who did not felt it was not helpful because it was not necessary. We also found the participants completed the puzzle in 30 seconds or less, significantly faster than the paper puzzle

DISCUSSION (STUDY 1)

We presented our findings to our client in the form of usability recommendations. Our major recommendations for the experimenter interface included: changing "study collections" to "active experiments", making distinct options for outputting collective or single participant data reports and including feedback when experimenter selects impossible settings (e.g. path size bigger than grid).

The puzzle and training required mostly minor changes. We recommended shortening the training to include just the rules and how to navigate as well as increasing the contrast for certain colors of shapes.

We also informed our client of the speed with which participants completed the puzzle and that overall; participants felt the puzzle was extremely easy. This raised concerns about the reliability and validity of the measure itself, so our client asked that we included a feature allowing for a participant to complete multiple puzzles (trials) of varying difficulties to acquire more data.

We also received input from a third party researcher utilizing the puzzle prototype to study CF and creative thinking. He suggested we add more customization to the types of shapes and colors available to the experimenter. Adding more customizability makes the product more fun, increases user freedom and may make the product more usable for children or older adults.

METHOD (STUDY 2)

Based on our findings in Study 1 as well as our client's requirements, we decided to redesign the information architecture of the experimenter interface. However, due to time constraints, a prototype for user testing was not available. Instead, we created several paper prototypes and new PowerPoint prototypes (see Figure 4) with a detailed task flow of the new experimenter interface (see Figure 5). We then performed a heuristic evaluation (Nielsen, 2005) as well as several keystroke-level models (KLM; Kierras, 2001) comparing our new design to the WCST and paper based puzzle.

RESULTS (STUDY 2)

After conducting a heuristic evaluation we found that our new design improved error prevention status by eliminating impossible settings in the experimenter interface. Puzzle settings were constrained only to possible combinations in dropdown menu. Our new design also included an experiment "recycling bin" called "Inactive Experiments", where unused experiments can be removed from user interaction without being permanently deleted. Help and documentation was made available on every screen of the interface via a dedicated help button.

Results from our KLM (Table 1) indicate that our new design requires an equivalent training time to the paper puzzle and is slightly slower than the WCST. Our new design had a significantly faster task completion time (20s) than either the WCST or the paper puzzle. However, this value could vary depending on the number of trials per experiment. Our new design automatically scores data, similar to the paper puzzle and significantly faster than the WCST.

Table 1. KLM Comparison Between Three CF Measures

	Measure		
	WCST	Paper Puzzle	New Design
Training	60s	120-180s	140s
Task Completion	5-25min	60-1000s	20s (per trial)
Scoring	5min (with training, data still needs to be input)	0 (data still needs to be input)	6s (to export .csv)

DISCUSSION (STUDY 2)

Our findings indicate that our new design is significantly faster than either of older designs. Our design automates the scoring process, eliminating opportunities for user errors while still keeping the user informed as to what data is being collected and how. Our design also allows for rapid exporting of data enabling users to quickly analyze a single participant's data or entire datasets.

Furthermore, due to the user-friendly nature of our design, extensive training is not required. In contrast, the WCST, in addition to the 5 minutes on average to score, still needs to be manually input into a data set. Also, extensive training of experimenters is required, as the scoring process is quite complex.

GENERAL DISCUSSION

After completing our analysis, we presented our conclusions to our client and software engineer. We walked both through each interface screen utilizing the task flow we created as well as detailing all our recommendations. Based on the third party researcher's suggestions, we also created a trial customization interface (Figure 6).

Currently, a new fully-functional prototype is being developed for formal usability testing on approximately thirty George Mason University undergraduate and graduate students. Future iterations will focus on making the entire interface more mobile and deployable in the form of a mobile application.

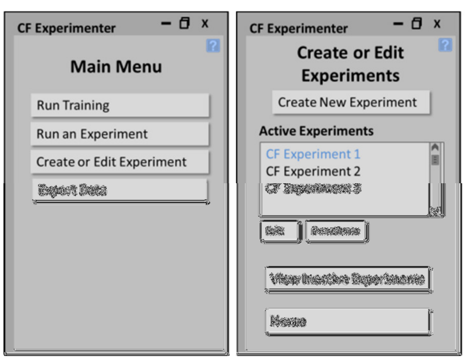

Figure 4. Iterative Design of the Experimenter Interface

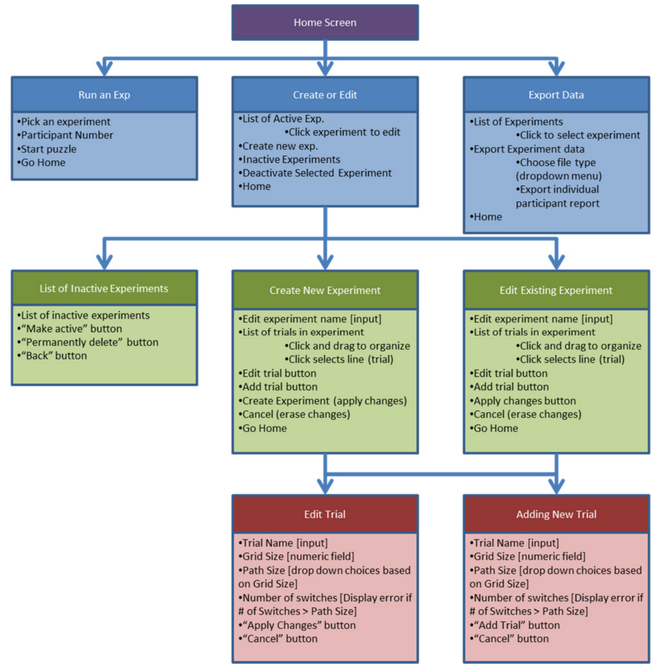

Figure 5. Experimenter Interface Task Flow

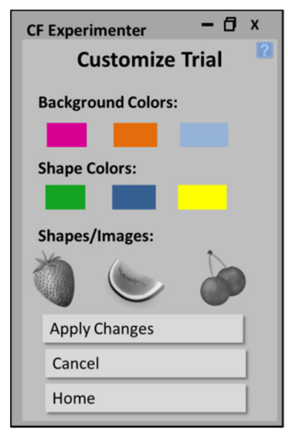

Figure 6. Experimenter Interface Trial Customization

ACKNOWLEDGEMENTS

We would like to thank Ivonne Figueroa for having us work with her on this project. We would also like to thank Dustin Rhodes for his software engineering on this effort. Finally, we'd like to thank Olga Kramarova and Alex Lim for their expertise of different portions of this project as well as Dr. Robert Youmans for his guidance and support.

REFERENCES

Dennis, J. P. & Vander Wal, J. S. (2009). The cognitive flexibility inventory: Instrument development and estimates of reliability and validity. *Cognitive Theory Research, 34,* 241-353.

Elslinger, P. J. & Grattan, L. M. (1993). Frontal lobe and frontal striatal substrates for different forms of human cognitive flexibility. *Neuropyshologia, 31* (1), 17-28.

Figueroa, I. J. & Youmans, R. J. (2011). Developing an easy-to-administer, objective, and valid assessment of cognitive flexibility. *Proceedings of The 55th Annual Human Factors and Ergonomics Society.* Las Vegas, NV. 944 – 948.

Frensch, P.A. & Funke, J. (1995). *Complex problem solving: The European perspective.* Hillsdale, NJ: Lawrence Elbram Associates, Inc.

Grant, D. A. & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl type card-sorting problem. *Journal of Experimental Psychology, 38*, 404–411.

Kierras, D. (2001). Using the keystroke-level model to estimate execution times. Unpublished report. Ann Arbor: Department of Psychology, University of Michigan

Neilson, J. (2005). *Ten Usability Heuristics*. Retrieved from: http://www.useit.com/papers/heuristic/heuristic_list.html

Scott, W. A. (1962). Cognitive complexity and cognitive flexibility. *American Sociological Association, 25,* 405-414.

Spiro, R. J., Feltovich, P. J., Jacobsen, M. J. & Coulson, R. L. (1991). Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction for Advanced Knowledge Acquisition in Ill-Structured Domains. *Educause Review*, 5, 24 -33.

Spiro, R. & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In Nix, D., Spiro, R. (Eds.), *Cognition, education and multimedia: Exploring ideas in high technology*, 163-205.