

An Experiment in Experiment Participation: the relevance of studies in human psychology and physiology to web design.

Michael Trotter, u7134416

Abstract

This report details my participation in two studies conducted by the ANU School of Psychology and aims to discuss their relevance to web development and the design of user-evaluation experiments within the field of human-computer interaction (HCI). The details of the experiments are discussed, along with a description of my experience and my comments on the strengths and weaknesses of each study design. Finally, their contribution to understanding not only HCI but better experimental design is considered.

Introduction

In September 2022 I enrolled in two experiments being conducted by the ANU School of Psychology, titled 'Comparing Measurements for the Deployment of Spatial Attention' and 'How does Social Anxiety Influence Decision-Making?'. This report looks at the details of these experiments, my observations made during the experimental activities, and their relevance in the design of experiments, including user-evaluation, in the investigation of HCI for web development and design.

This report finds that their contributions are potentially two-fold in that they add to a body of work that aids the development of systems that create positive HCI experiences, while also informing better experimental design for HCI evaluation.

Experiment 1 - Comparing Measurements for the Deployment of Spatial Attention

Experiment summary

The purpose of this experiment was to evaluate a measure of spatial attention termed *exploratory breadth*, against a current widely used measure known as *attentional breadth*. Exploratory breadth is the spatial extent over which eye movements are made across a scene in response to visual stimuli, while attentional breadth is the measure of the extent of spatial attention from a fixed point in a scene.

My contribution to this study consisted of undertaking two tasks. The first was to count backwards in fixed increments while viewing images of natural environments. I was then shown pairs of similar images and asked to select which of the two were from the set I had been shown. For the second task I was shown arrangements of T or H characters that were themselves possibly arranged in a T or a H. I was asked to count backwards again while I identified which character was displayed at either scale (individual character or character arrangements). While undertaking both tasks my eye movements were tracked across the screen to measure my spatial attention and calculate the two metrics to be compared.

My experience

Both tasks required a high degree of concentration and were similarly taxing in their need to focus on two simultaneous tasks (the backwards count and the visual elements). I've recently gone through the training to get a pilot's licence and I found that the high mental load during these exercises was very similar to the early stages of learning to fly where a conscious effort is still required to manage each different task that must be performed simultaneously.

I was not provided with feedback on my performance, though subjectively it felt as though I improved throughout the experiment as I became more familiar with the expectations. With the character identification task, the pace of response was much higher and there were moments I perceived as a flow state (Csikszentmihalyi, 1990). I found a degree of satisfaction in the feeling of performing optimally and with focus entirely on the task.

Strengths and weaknesses

This experiment was well conducted and well organised. Clear explanations of the tasks were provided in what appeared to be a consistent and repeatable format. The use of physical guides for the subjects and well considered lighting suggests that the eye-tracking results are likely reliable. However, it was not apparent whether these experiments would meet the stated outcomes:

“...to determine which measurement is a more useful representation of the spatial deployment of attention.” (Wyche, 2022).

Based on my participation and reading of the provided documents it was not clear how ‘usefulness’ would be quantified. It is possible that performance metrics for each of the tasks will be compared with the spatial attention measurements to identify patterns of cross-correlation, but this was not explicitly stated.

Experiment 2 – How does Social Anxiety Influence Decision-Making?

Experiment summary

The aim of this experiment was to understand how social anxiety influences decision making. The participant information sheet and documentation on SONA stated that participants would complete questionnaires and undertake computer-based behavioral tasks, and that “you may also be required to give a speech in front of a camera which will be live-streamed to several professionals” (Tan, 2022).

On attending the lab to participate in this study I was informed up-front by the researcher that there had been “some technical difficulties” and that I would no longer be required to give a speech as a part of the study. I then completed two questionnaires requesting demographic information and details about my psychological well-being, as well as a computer-based task. Both questionnaires included questions designed to test my attentiveness and truthfulness such as “Select ‘somewhat positive’ for this question” or “How often do you eat [made up food]?”. The computer-based task involved randomly sampling from one of two possible digital jars with differing ratios of blue/red marbles and nominating when you were certain from which distribution the samples were drawn.

After finishing, I was told that there was never any expectation I present a speech, as was originally stated. The purpose of this statement was to elicit feelings of anxiety and 50% of the study participants were only told that the speech was not required after completing the tasks.

My experience

Participation in this experiment was without stress for myself, though my experience may have been different had I been allocated to the other cohort.

The computer-based task was straight-forward, though there were levels of ambiguity as I don’t believe establishing complete confidence was possible with the number of allowed samples. As a result, I perceived the definition of ‘certainty’ under this constraint to be subjective, making it somewhat difficult to find a consistent approach to the task. After completing the experiment, I found that this task is well studied and is known as the ‘Beads Task’ (Phillips & Edwards, 1966), and has previously been used to examine the effects of anxiety disorders on decision making (Jacoby *et al*, 2014). The ambiguity that I noted in the task is an inherent part of measuring the subject’s response to uncertainty.

Strengths and weaknesses

The strength in this experiment was in the study design. Although a double-blind study is not possible with this scenario, the use of single-blind randomised control means that with large enough sample sizes many inherent biases can be excluded from the results.

However, the strength of the results rests on the ‘theatrics’ of the delivery. To induce an anxiety response in half of the cohort the participants need to believe that they will be delivering a speech. As I was not assigned to this cohort, I can’t comment on how well this part of the study was executed. However, I was not given any reason to doubt the statement that I would be making a speech, or the story about technical

difficulties, so it is likely that the integrity of the results will have been maintained through consistent, believable, and scripted delivery or other similar strategies to control for this.

Comparison of the experiments

Similarities and differences

The most fundamental difference between these two experiments is their manipulation of the conditions of the experiment. The first seeks to better measure our physiological and cognitive functions as they would take place under normal though challenging conditions, while the second seeks to measure changes in cognitive processes brought about by induced differences in psychological state. There is a similarity between them though in that both experiments, at least for some participants, apply stressors to better measure the response variables. In the case of experiment 1 this was in the form of cognitive workload, while experiment 2 involved inducing an anxious state.

My experiences between experiments though were very different. As I was not exposed to the stressor for experiment 2 owing to my position in the control group, I found that experiment 1 was quite taxing while experiment 2 was relatively easy.

There were also environmental differences between experiences. Experiment 1 required the measurement of physiological responses and so necessitated an unnatural, though not uncomfortable, position with the physical frame for eye-tracking and the experimenter to be present. This contrasted with experiment 2 where I was simply required to use the computer while left alone in the room. Although these environmental conditions were not problematic in any way, there is evidence that things like researcher presence does impact upon experimental results (Webster, 1997; Friesen *et al*, 2020).

Both experiments were well organized, well considered, and well delivered. It is possible that additional resourcing may have improved study outcomes, though the degree to which this might be true is uncertain. An example of this might have been additional researcher presence and a different location for experiment 2 in order to improve the believability of the requirement to deliver a speech.

My feelings on participation

There were engaging elements of both experiment's tasks. As already mentioned, I derived a sense of satisfaction from the experiment 1 tasks mostly due to their fast pace and the challenge of managing cognitive load. While the ambiguity of experiment 2's Bead Task made it more difficult for me to engage meaningfully, it was not unenjoyable.

The topics of both experiments were of interest to me, and I will be following up to try and obtain the results when they are published. I would be interested in further details on experiment 1 as some of the specifics of how the experiment results would be applied to the stated outcomes were not included in the participant information sheet.

Relevance of user-participation experiments to web design and development

The experiments described here are relevant to the field of HCI and have implications for user participation in web design. Both aim to tell us something about the way in which a user is likely to draw information from a computer-based interaction (experiment 1) and how we respond to information provided to us through an interaction (experiment 2). Experiment 1 aims to define a useful measure for the placement of user spatial attention, how we visually process information and how information is attenuated through media. Being able to accurately quantify these factors has implications for the design of web interfaces. This experiment also has relevance in the field of bio-inspired computing where deep learning computer vision models have analogous concepts of spatial attention that have been influenced by studies in human psychology. Experiment 2 aims to define how user emotional and psychological state effects how we process information. Understanding these processes has implications for web interface design and accessibility, and in understanding how to guide user interactions.

Studies in HCI that utilize user-evaluation and usability assessments, such as the eReader evaluation conducted by Rampaul & Gedeon (2015), are useful in that they aim to closely mimic the types of interactions users are likely to have with technology in their every-day life. While these studies are critical, the relevance of their results are often tied to specific cultural, social, and technological contexts (Tractinsky, 1997). Rampaul & Gedeon note that at the time of publication in 2014 32% of Americans owned an eReader, a figure that declined to 19% in 2015 and has seen a general downwards trend since then (PRC, 2015). This places their results in a very different technological context today where many more users have extensive histories of interaction with tablets and mobile devices, and where fewer own eReaders.

The strength of the experiments discussed here is different in that they aim to tell us something about the fundamental nature of human interaction and psychology. There is a two-fold benefit to this work in terms of contributing to HCI evaluation. Firstly, these results can be translated directly to inform user interaction design – How do we present web content to users in a way that maximises its uptake? What are the likely impacts of design choices for a web interface designed to be used in environments of high psychological stress? (eg. emergency department equipment, websites providing mental health services). Secondly, these results can also feed into better user evaluation designs – how might psychological state impact perceived usability? Does usability have a correlation with more targeted deployment of spatial attention?

While neither experiment I attended included user-evaluation, it is clear that many of the aspects of experimental design discussed here are of direct relevance in the design of user-evaluation studies for HCI. It is also apparent that the topics being investigated are of direct relevance to HCI design.

References

- Csikszentmihalyi, Mihaly. 1990. *FLOW: The Psychology of Optimal Experience*. Harper & Row.
https://mktgsensei.com/AMAE/Consumer%20Behavior/flow_the_psychology_of_optimal_experience.pdf (Accessed: October 30 2022).
- Friesen, K, Zhang, Z, Monaghan, P, Oliver, G & Roper, J. 2020. *All eyes on you: how researcher presence changes the way you walk*. Nature. Science Reports 10. Article Num. 17159.
- Jacoby, R, Abramowitz, J, Buck, B & Fabricant, L. 2014. *How is the Beads Task related to intolerance of uncertainty in anxiety disorders?* Journal of Anxiety Disorders. Vol 28. Pp. 495-503.
- Pew Research Centre. 2015. The Demographics of Device Ownership.
<https://www.pewresearch.org/internet/2015/10/29/the-demographics-of-device-ownership/> (Accessed: October 30 2022).
- Phillips, L & Edwards, W. 1966. *Conservatism in a Simple Probability Inference Task*. Journal of Experimental Psychology. Vol. 72. No. 3. Pp. 346-354.
- Rampaul, U & Gedeon, T. 2015. *Evaluation of e-Readers: A Preliminary Analysis*. Australian Journal of Intelligent Information Processing Systems. Vol. 14. No. 3.
- Tan, N. 2022. *How does Social Anxiety Influence Decision-Making? – Participant information sheet*. Not published, included with this submission.
- Tractinsky, N. 1997. *Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues*. Proceedings of the ACM SIGCHI Conference on Human factors in computing systems. Pp. 115-122.
- Webster, C. 1997. *Effects of researcher presence and appeal on response quality in hand-delivered, self-administered surveys*. Journal of Business Research. Vol. 38. Iss. 2. Pp. 105-114.
- Wyche, N. 2022. *Comparing measurements for the deployment of spatial attention – Participant information sheet*. Not published, included with this submission.