Watching People Making Decisions: A Gogglebox on Online Consumer Interaction

Chris Barry
National University of Ireland Galway
Galway, Ireland

chris.barry@nuigalway.ie

Mairéad Hogan
National University of Ireland Galway
Galway, Ireland

mairead.hogan@ nuigalway.ie

Abstract

This paper presents a research study, using eye tracking technology, to measure participant cognitive load when encountering micro-decision. It elaborates and improves on a pilot study that was used to test the experiment design. Prior research that led to a taxonomy of decision constructs faced in online transactional processes is discussed. The main findings relate to participants' subjective cognitive load and task error rates.

Keywords: Eye tracking, decision constructs, cognitive load, NASA-TLX, micro-decision.

1. Introduction

This research is part of an ongoing effort to shine a light on more subtle aspects of ethics in information systems design. The research began as anecdotal, personal observations about some questionable customer service practices in the low cost carrier (LCC) sector in Ireland. Several studies, mostly qualitative, were conducted to understand and to solicit user views on design features creating distance between the consumer and the firm [7], [3], [4]. The work broadened outwards to examine how firms were presenting choices to users. The specific context of this research is the Business-to-Consumer (B2C) transactional process from consumer commitment to an online purchase to payment conclusion. This process has become crowded with an increasing number of micro-decisions, such as the purchase of additional insurance or faster delivery. These decision points, are increasingly ambiguous and time-consuming [3].

This study describes an experiment that measures the cognitive load users experience when making online decisions. Eye tracking equipment was used to collect physiological data on user interactions; a self-assessment survey was used to collate perceptions about cognitive load; and immediately afterwards, users were prompted to express feelings and opinions. The main focus of this paper is the analysis of the subjective cognitive load and the error rate in the interactions, the remaining data will be the subject of later reporting. The findings present some fascinating insights into how subtle differences in the design and framing of decisions can lead to significantly different perceptions and outcomes.

2. Analyzing Online Decision Constructs

2.1. The Influence on User Choice

The classical view of decision-making in economics is that individuals will behave rationally by objectively weighing and ranking alternative options, according to their preferences, and choosing appropriately. The model assumes the phrasing of alternatives, and the order they are presented in, makes no difference to how individuals choose. This theory has been significantly contradicted or inherently questioned by many authors [51], [33], [47], [35], [8], [32], [34], [50], [6]. So, quite to the contrary, individuals or consumers have been shown to be: influenced by whether information has been positively or negatively framed; persuaded by the context and

the presentation of choice; impacted by influential labels; affected in their choice by default values; influenced in their choice by opt-in or opt-out decision formats. Additionally, firms can advantage themselves when questions are framed and defaults are presented in particular ways.

2.2. A Taxonomy of Online Transactional Decision Constructs

What is clear is a variety of elements can make micro-decisions unnecessarily complicated and potentially subject to error [2]. The complexity is contributed to by: question framing; default values; levels of persuasion; whether decisions are optional or necessary; how users expect decisions to be presented; and the unconventionality of some decision constructs. With this in mind the authors set about identifying an exhaustive list of decision constructs [28] and produced a taxonomy identifying seven decision constructs used in the B2C transactional process (see Table 1). At the macro level decisions are either essential or optional. Essential decisions are those that must be made, such as choosing a shoe size or a delivery method. Optional decisions were identified as being either opt-in, opt-out or must-opt. A must-opt decision is neither an opt-in nor an opt-out, it occurs when a user cannot continue through the transactional process without explicitly choosing to accept or decline an option. All decisions were also classified as being either pre-selected or un-selected. A study was conducted [5] to confirm that each construct identified in the taxonomy is used in practice. The study concluded that although most constructs are not problematic, the opt-out decision construct was often presented in a way that users could inadvertently make unwanted choices. Variants of the opt-out decision construct are the subject of this study.

Table 1. Taxonomy of transactional decision constructs.

Decision Construct	Default Value	Normal Presentation	Framing
Un-selected opt-in	Don't receive the option	Un-selected	Acceptance
Pre-selected opt-in	Don't receive the option	Selected	Rejection
Un-selected opt-out	Receive the option	Un-selected	Rejection
Pre-selected opt-out	Receive the option	Selected	Acceptance
Must-opt	Cannot proceed	Multiple option variants, un-selected	Normally acceptance
Un-selected essential decision	Cannot proceed	Multiple option variants, un-selected	Normally acceptance
Pre-selected essential decision	Variant selected	Multiple decision variants, one selected	Normally acceptance

2.3. Cognitive Load

According to Grimes and Valacich [23] cognitive load, or mental workload, can be defined as: "the mental effort and working memory required to complete a task". Considerable research in the area of cognitive load in computer-based learning applications has shown it impacts negatively on learning [39], [9], [32], [30]. While less research on cognitive load in e-commerce transactions has been conducted, higher cognitive loads have been shown to negatively affect both time to complete tasks and user satisfaction in e-commerce applications [48]. Additionally, higher mental workload corresponds to lower perceived usability for webpages [36].

Cognitive load can be measured in multiple ways. The main approaches include: subjective measures; direct objective (or physiological) measures; and indirect objective measures (for example, electroencephalography (EEG) or cardiovascular metrics) [38]. The subjective measures generally use Likert scales for self-reporting of stress or other indicators of mental load. Some of the more commonly used measures include the Subjective Workload Assessment Technique (SWAT), the NASA-Task Load Index (NASA-TLX) and the Workload Profile (WP). Each of these measures lead to a global workload index that is sensitive to the level of difficulty in the task [46]. Think-aloud can also be used to measure cognitive load [16], albeit qualitatively, rather than quantitatively.

2.4. Subjective Measurement Scales

Measurement scales commonly used to determine cognitive load include uni-dimensional scales, such as the Modified Cooper-Harper Scale (MCH), and the Overall Workload Scale (OW), as well as multi-dimensional scales, such as NASA-TLX and SWAT [39]. Rating scales require the user to indicate the mental effort required to complete a task. Research indicates people can put a numerical value on their perceived mental effort [22], [44], resulting in their use in much research. NASA-TLX and SWAT are the most commonly used measurement scales of subjective cognitive load [18]. However, SWAT is not sensitive for low cognitive load tasks, unlike the NASA-TLX [37]. Hence, NASA-TLX is considered to be superior to SWAT in terms of sensitivity [27] and is frequently used as a benchmark when assessing other measures [1], [54], [17].

While NASA-TLX was originally developed for use in the aviation domain, its use has spread to other areas, including the medical profession, data entry and decision-making. Additionally, it has been translated into multiple different languages. Hart [24] examined 550 studies in which NASA-TLX was used and found most of these studies were concerned with some form of question relating to interface design or evaluation. Modification of the scale occurred in many of the studies, with subscales being added, deleted or modified. Modifications include either eliminating the weighting or analyzing the subscales individually, either in conjunction with, or instead of, the overall workload measure. Hart [24] concluded "NASA-TLX has achieved a certain venerability; it is being used as a benchmark against which the efficacy of other measures, theories, or models are judged."

2.5. Subjective Cognitive Load

Based on the research detailed above in Section 2.4, NASA-TLX was deemed the most appropriate measurement scale to use in this study. NASA-TLX was the culmination of a multi-year research programme that resulted in a multi-dimensional rating scale, and derives an estimate of workload that is both reliable and sensitive [25]. The programme determined the contributing factors to an individual's subjective perception of physical and mental workload. These were narrowed down to 6 factors: mental demand; physical demand; temporal demand; performance; effort; and frustration level. The definitions for these can be seen in Table 2.

Table 2. Rating scale definitions for NASA-TLX.

Rating Scale Definitions	
Scale	**Definition**
Mental demand	The level of mental and perceptual activity required for the task
Physical demand	The level of physical activity required for the task
Temporal demand	The level of time pressure felt
Performance	The level of success in reaching the goals of the task
Effort	The level of work, both mental and physical, required
Frustration level	The level of frustration felt during the task

According to the NASA-TLX user manual [41], the participant assigns a score on a 21-point scale ranging from 0-100 on each factor. Additionally, each of these factors are weighted by the participants according to their perception of the contribution of each factor to the workload of a given task. This weighting can be done while carrying out the task, or afterwards while replaying the task and requires the participant to weight each of the factors by indicating which one was most relevant to the task in a series of paired comparisons. However, more recent studies [53], [30], [10], [23] have used a slightly modified version of the NASA-TLX, known as NASA-Raw Task Load Index (NASA-RTLX). Rather than weighting the factors, each is assigned equal weight and the overall workload is obtained by summing the values and dividing by the number of factors used. Studies have shown [43], [39], [24], [53] this modified

version to be as effective as the original, with the added benefit of being a much simpler approach.

In addition, Hart and Staveland [25] determined the individual factors can be used independently to garner information about the various aspects of workload. Hart [24], in her review of the usage of NASA-TLX states the analysis of subscale ratings instead of, or in addition to, an overall rating demonstrates "one of the continuing strengths of the scale: the diagnostic value of the component subscales." Studies have also adapted the measure in various ways: using a 5-point scale [21], [49] rather than the original 21-point scale with values between 0 and 100; changing the wording to increase the relevance to the tasks [26], [40]; and using only some of the subscales [20], [52].

3.   Eye Tracking Research

Another important aspect to the study was to use eye tracking technology to track users' gaze when making micro-decisions. While it is possible, and indeed desirable, to listen to users describe what they see and experience, eye tracking data contribute objective measurements of the visual pattern of the interaction.

Eye tracking technology involves the projection of a light and a video camera on a person's eye to identify where they are looking on a screen [42]. The usual pattern of eye movement on webpages is much more erratic than one might anticipate. When someone does want to concentrate on an area of interest, they fix their gaze on it and it then comes into sharp focus. As their gaze remains on something for more than 3 milliseconds (depending on how you wish to specify it) this become known as a fixation, while the movements in-between fixations are known as saccades.

In HCI and web usability research, eye tracking has been extensively used [14], [12], [13], [31]. By studying what users do and do not look at, it is possible to determine where they are concentrating their attention [45]. Through the examination of eye movement patterns, conclusions may then be drawn regarding the decision-making strategies users adopt [11], [19], [31]. The potential of gathering hard, physiological data about participant behaviour in interactive decision-making was a key motivation for developing an Eye Tracking Laboratory at the authors' university.

4.   Preparation for the Study

It was anticipated that the design and construction of a study of user decision-making in online transactional processes was going to be a substantial undertaking and, using research technology new to the authors, a risky process. Thus, it was decided to run a pilot study to validate the research design for a more extensive eye tracking study. The purpose was two-fold: to learn from the process of constructing an eye tracking experiment; and to fine-tune the research instruments [29].

This pilot study, using eye tracking and the qualitative Cued Retrospective Think Aloud (RTA) technique, examined potentially problematic decision constructs [29]. It explored the impact of decision constructs on users' decision-making and their cognitive processes during interactions. The main lessons ascertained were to ensure participants: are fully briefed before commencing the test; perform the interaction as instructed working with neither haste nor labouring the tasks; and are de-briefed after the test to ascertain insights into their behaviour. The key contributions of the study were the identification of improvements to be made to the research design, robust experiment administration and the refinement of research instruments.

5.   The Research Study

The research study was successfully concluded in May 2018. Data was collected from 114 participants, 456 experiment trials, 2736 Interest Area data sets and 23 Cued RTA sessions. The study was made up of three dimensions: an eye tracking experiment; self-assessment evaluations (subjective cognitive load); and Cued RTA sessions.

5.1. Eye Tracking Experiment

Significant effort was spent planning the eye tracking study. The pilot study had highlighted the risks associated with the use of the eye tracking technology. While initially it was intended to study each of the decision constructs in the taxonomy, it soon became clear that the scope of the study would be unachievable and would require hundreds of test trials. Instead opt-out decision constructs, the most problematic in the taxonomy, were selected. Indeed their problematic nature has been recognized by the European Union, who prohibit their use in distance selling [15]. Each of the variants of opt-out decisions were examined: Un-selected Rejection Framing; Pre-selected Acceptance Framing; Pre-selected Rejection Framing; and Pre-selected Neutral Framing (see Table 3).

Table 3. Opt-out decision constructs.

Construct Name	Construct Type	Default Value	Framing
Un-Selected Rejection	USR	Un-selected	Rejection
Pre-Selected Acceptance	PSA	Pre-selected	Acceptance
Pre-Selected Rejection	PSR	Pre-selected	Rejection
Pre-Selected Neutral	PSN	Pre-selected	Neutral

Four screens of consumer decisions, (each of the opt-outs in Table 3 and illustrated in Table 4) were randomized and presented to participants. The core webpage screen was a breakdown insurance product to which participants were asked if they require an enhanced monthly-costed, add-on feature. Each screen had a single decision point with a checkbox beside it. Participants were instructed to make a decision to buy or not to buy the add-on.

Table 4. Decision constructs presented to participants.

Construct Type	Decision Construct
PSA	☑ Rescue Plus includes free car hire and travel expenses. I want to purchase Rescue Plus.
PSN	☑ Rescue Plus.
PSR	☑ Rescue Plus includes free car hire and travel expenses. If you would rather not purchase Rescue Plus, please untick this box.
USR	☐ Rescue Plus includes free car hire and travel expenses. If you would rather not purchase Rescue Plus, please tick this box.

5.2. Self-assessment Evaluation

After each decision, participants were taken through three screens with scales on them. They were asked to rank each decision in terms of their performance, the mental demand and the level of frustration experienced. Collectively these constitute the subjective cognitive load relevant to this study. The scales are explained as follows:

  i.   Mental Demand: How much mental and perceptual activity did you spend for this task? 1= Low, 5 = High

  ii.   Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you during the task? 1= Low, 5 = High

  iii.   Performance: How successful do you think you were in accomplishing the task goals? 1= Good, 5 = Poor

The pilot study had highlighted the importance of avoiding any perception that participants were under time pressure. Thus, they were instructed to work at their normal pace; not to feel under pressure to complete quickly or to over analyse it; and to look at and read whatever information they normally would to make the decision.

5.3. Cued RTA Sessions

About one in four participants took part in a Cued RTA session immediately after their session. They were shown an animated playback of their interaction showing eye movements for each of the four decision screens. The lead researcher prompted the participant to articulate their thought processes and feelings they had during the interaction. Each of the sessions, taking from 5 to seven minutes, were documented by a scribe and a sound recording. The analysis of these sessions is the focus of a further publication.

6. Exploratory Analysis

Once the data was gathered, exploratory analysis was conducted. The main focus of this analysis was subjective cognitive load while conducting the tasks and participant error rate, where the number of errors made by participants was examined.

The opt-out decisions presented to participants were typical of the type of micro-decisions encountered by users as they navigate the transactional process on websites. They were based on the variety of opt-out constructs identified during a desk analysis of 57 different websites [6] and are based on actual decision constructs encountered. They were all opt-out decisions, meaning the participant needed to take action to decline the purchase, with the default option being to opt-in and purchase the item. Opt-out decisions are generally recognized as being problematic, with users often inadvertently making a purchase, or opting in to a mailing list [8].

6.1. Error Rate

The first step in exploring our data was to examine the error rate for the different construct types.

Table 5. Error rate

Construct Type	Correct	Incorrect	Total
PSA	94 (82%)	20 (18%)	114
PSN	87 (76%)	27 (24%)	114
PSR	86 (75%)	28 (25%)	114
USR	72 (63%)	42 (37%)	114
Total	339 (74%)	117 (26%)	456

As can be seen in Table 5, the number of incorrect selections varied between the construct types, with PSA having the highest number of correct selections and USR having the highest number of incorrect selections. Overall, the number of incorrect selections made by participants was quite high, ranging from 18% (PSA) to 37% (USR). Chi-square test was conducted to determine whether the relationship between construct type and error rate was significant. The test indicated significance (χ^2 (3, N=456) = 11.715, p<0.01). However, the value for ϕ was 0.158, indicating the association was weak, only accounting for 2.5% of the variation.
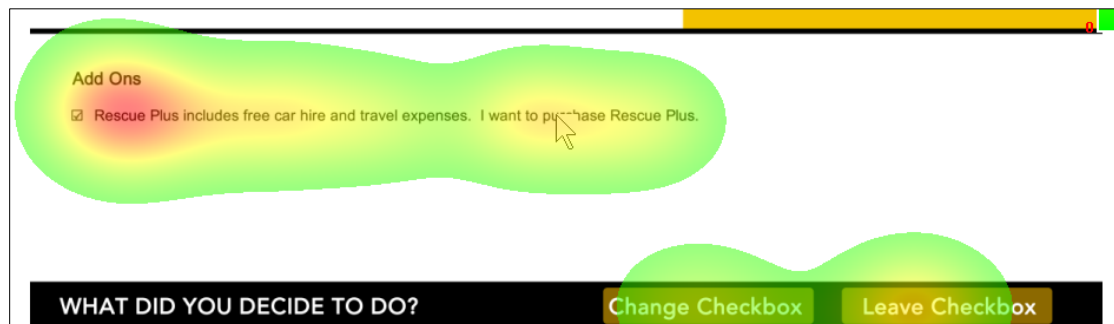
Fig. 1. Heat map of a pre-selected opt-out with acceptance framing (PSA)

The results from this study, whereby 26% of the decisions participants made during the experiment were incorrect, support previous research [8] that found opt-out constructs to be problematic. Additionally, the significant difference in error rates suggests that, while opt-outs in general are error prone, some are more problematic that others.

The heat maps in Figures 1 and 2 illustrate gaze concentration where green indicates less time and red indicates more time focusing on the text. Figure 1 is the heat map of a pre-selected opt-out with acceptance framing (PSA), the construct that had fewest errors, while Figure 2 is the heat map of an un-selected opt-out with rejection framing (USR), the construct with the most errors. As can be seen, participants spent considerably longer examining the rejection-framed text (i.e., If you would rather not purchase…) than the acceptance-framed text (i.e., I want to purchase…). Despite spending more time reading the USR construct text, participants still made twice as many errors.
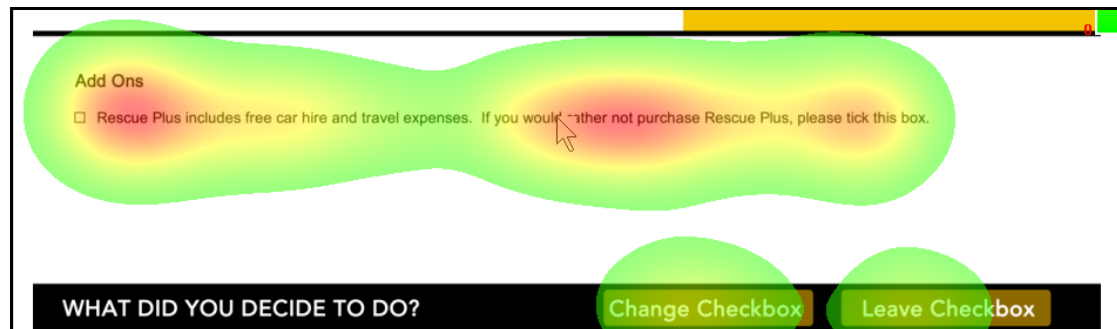

Fig. 2. Heat map of an un-selected opt-out with rejection framing (USR)

The error rate is broadly in line with the frequency of construct types found in general use by Authors [5], who found most opt-outs are PSA, with considerably fewer PSN, USR and PSR. The comparatively smaller, though still high, error rate for PSA may be explained by the fact that consumers are more used to seeing opt-outs in this format and, when seeing a pre-selected checkbox, may be more likely to assume it is an opt-out, and so requires action if the user does not wish to purchase the product. In contrast, the higher rate for PSR and PSN may be explained by the rejection or neutral framing being more rarely encountered but still having the pre-selected checkbox. The pre-selected checkbox may be suggestive of an opt-out to the participant, while the framing may confuse them due to it being unusual for an opt-out, thus resulting in a higher error rate. The USR has an unselected checkbox, which may be more suggestive of an opt-in to the participant, and when combined with the more rarely encountered rejection framing may explain the considerably higher error rate.

6.2. Cognitive Load

The cognitive load was measured using NASA-RTLX, where participants rated their interaction with each construct type on a scale of 1-5 for 3 factors: Mental Demand, Frustration and Perceived Performance. It was decided to use NASA-RTLX as it is simpler to administer and has been determined to be as effective as the original [43], [39], [24], [53]. Cognitive load was determined by summing the scores for each factor and dividing by 3.

Table 6. Mean cognitive load

Construct Type	Mean	Std. Deviation	N
PSN	2.11	0.885	114
PSN	2.13	0.938	114
USR	2.49	0.996	114
PSR	2.54	0.956	114

In order to assess if the participant's perceived cognitive load varied by construct type, a one-way, within-subjects ANOVA was conducted. As can be seen in Table 6, participants reported the lowest cognitive load for PSN and the highest for PSR.

Mauchly's test of sphericity was used to test the null hypothesis that the variances are equal. The test indicates the assumption of sphericity has not been violated ($\chi^2 (5) = 2.931$, p= 0.711) and so, no corrections were required. The test showed there was a significant effect (p < 0.01) of the construct type on the participants' subjective assessment of cognitive load. The cognitive load ranged from 2.11 to 2.54 on a scale of 1-5, with 1 being very low and 5 being very high (see table 6). The lowest cognitive load was reported for PSN, which also had the lowest error rate. The highest cognitive load was for PSR, which had the second highest error rate. Interestingly, despite the high error rate, the participants did not rate cognitive load very high for any of the construct types. This may be due to the nature of the micro-decisions required by the construct types. The micro-decisions are minor, everyday decisions, that users would encounter multiple times in the transactional process. Thus, the everyday nature of the decisions, coupled with the brevity of the text, may have made the participant feel that, even though some decisions were more difficult and complex than others, none warranted being described as imposing a major cognitive load.

6.3. Cognitive Load and Error Rate

A logistic regression was performed for each of the construct types to determine whether cognitive load, and the individual factors contributing to cognitive load, could be used to predict error rates. Firstly, the score for cognitive load was tested. For one of the constructs (PSN), the model significantly predicted the error rate. For the other three (PSA, PSR and USR), it did not.

For PSN, (omnibus Chi-square = 4.89, df = 1, p = 0.027), the model accounted for between 3% and 4.5% of the variance in error rate, with 100% of correct interactions successfully predicted. However, none of the predictions for unsuccessful interactions were accurate. Overall, 77.6% of predictions were accurate. The predictions for PSA, PSR and USR were not significant (omnibus Chi-square = 2.537, df = 1, p = 0.111), (omnibus Chi-square = 1.711, df = 1, p = 0.191) and (omnibus Chi-square = 0.147, df = 1, p = 0.7) respectively.

The measure of cognitive load was then broken down into the individual factors and the model was re-run with mental demand, frustration and perceived performance as predictor variables of error rate (see Table 7). For two of the construct types (PSA and PSN), the full model significantly predicted error rate. For the other two construct types, it did not.

For PSA, (omnibus Chi-square = 8.873, df = 3, p = 0.031), the model accounted for between 5.2% and 7.5% of the variance in error rate, with 96.6% of correct interactions successfully predicted. However, only 8.7% of the predictions for unsuccessful interactions were accurate. Overall, 72.1% of predictions were accurate. When the individual factors were assessed, only perceived performance significantly contributed to the prediction of error rate.

For PSN, (omnibus Chi-square = 15.808, df = 3, p = 0.001), the model accounted for between 9.1% and 13.9% of the variance in error rate, with 97.7% of correct interactions successfully predicted. However, only 10.8% of the predictions for unsuccessful interactions were accurate. Overall, 78.2% of predictions were accurate. When the individual factors were assessed, it was again only perceived performance which significantly contributed to the prediction of error rate.

Table 7. Cognitive load factors and error rate.

Construct Type	Correct Predicted	Incorrect Predicted	Total Predicted	p value
PSA	96.6%	8.7%	72.1%	0.031
PSN	97.7%	10.8%	78.2%	0.001
PSR	100%	0%	78.2%	0.325
USR	100%	0%	65.5%	0.612

The only construct for which cognitive load was a significant predictor of error rate was PSN, although it predicted only correct interactions. When cognitive load was broken down into its constituent factors, they significantly predicted error rates for PSA and PSN, although for each, the only factor that contributed to the prediction was perceived performance. The model did not significantly predict error rate in PSR or USR.

Where perceived performance significantly predicted error rate, it predicted correct interactions at a considerably higher level than incorrect interactions (see Table 7). This suggests that participants were less likely to believe they had performed poorly on the tasks and perceived their accuracy in making the micro-decisions to be higher than it actually was. This is consistent with Bellman et al.'s [8] findings that opt-outs are error prone. Early analysis of the Cued RTA sessions would also indicate over-confidence on the part of many participants who actually had high error rates. If users perceive these micro-decisions to be relatively easy to make, as suggested by their self-reported cognitive load; and erroneously over-estimating their performance when making these micro-decisions; they are more likely to inadvertently make a purchase or sign-up to a mailing list. This phenomenon may encourage firms to deliberately use design features to trick users into inadvertently making a purchase or signing up to a mailing list.

7.   Conclusions and Future Direction

The study set out to examine whether participants were able to correctly make decisions in respect of multiple, micro-decisions involved in online, commercial transactions. All the decisions were opt-outs, and while each is generally recognized as being problematic, a certain construct (the USR) hugely distorted expectations of error rates. The take-away finding here is that should a firm wish to nudge consumers toward a preferred outcome then the deliberate choice of framing and default values constitute a potent combination.

The self-reported cognitive load on participants, while not at the high end of the NASA-RTLX scales, was evident at modest levels amongst participants. Given that the micro-decisions were completed in just seconds, the manifestation of cognitive load was clearly evident. The analysis also clearly found that the construct type had a significant effect on the participants' subjective assessment of cognitive load. Not surprisingly the error rate was lowest for the decision that bore the least cognitive load, while the constructs with the higher cognitive load tallied with the higher error rates.

Where cognitive load was tested to see if it was a good predictor of error rate, the data was less convincing. Only for pre-selected neutrally-framed decisions was cognitive load a significant predictor of error rate. What was interesting was that perceived performance predicted correct interactions to a much greater extent than incorrect interactions. The implication is that participants had a much greater confidence that they were making correct decisions than was born out by their actual accuracy. The clear inference that can be drawn here is that such over-confidence leads to error rates where participants are more likely to inadvertently, and unknowingly, make decisions not in their interest, reinforcing the temptation for firms to use ambiguous design strategies

The study also yielded an enormous quantity of physiological participant data such as fixation count and duration, dwell time, blink count and saccade information – not reported here. The data will be parsed, for example, on the basis of construct type, interest area, trial analysis and gender. The granularity of the data is microscopic and initial probing indicates some promising results. The near future direction of analysis and publication will be fixed firstly on the physiological data, then the qualitative Cued RTAs and ultimately to effectively combine the three data dimensions. More Gogglebox episodes to follow!

References

1. Albers, M.: Tapping as a Measure of Cognitive Load and Website Usability. In: 29th ACM International Conference on Design of Communication, ACM, Pisa, Italy (2011)
2. Anaraky, R.G., Nabizadeh, T., Knijnenburg, B.P., Risius, M.: Reducing Default and Framing Effects in Privacy Decision-Making. In: SIGHCI 2018 Proceedings. San

Francisco, CA: Association for Information Systems, AIS Electronic Library (AISeL) (2018)

3. Barry, C., Hogan, M., Torres, A.: Low-Cost Carriers and High-Tech Barriers - User Views on Questionable Web Design Practices in Ireland. Irish Journal of Management. 31(1), 43-58 (2011)

4. Barry, C., Hogan, M., Torres, A.M.: Perceptions of Low Cost Carriers' Compliance with Eu Legislation on Optional Extras. In: 20th International Conference on Information Systems Development. Edinburgh, Scotland (2011)

5. Barry, C., Hogan, M., Torres, A.M.: Confirming a Taxonomy of Decision Constructs in Business-to-Consumer Commercial Transactions. In: 23rd International Conference on Information Systems Development, Varaždin, Croatia (2014)

6. Barry, C., Hogan, M., Torres, A.M.: Framing or Gaming? Constructing a Study to Explore the Impact of Option Presentation on Consumers, in Information System Development: Transforming Healthcare through Information Systems. Springer International Publishing. pp. 111-124 (2016)

7. Barry, C., Torres, A.M.: Tricks and Clicks - How Low-Cost Carriers Ply Their Trade through Self-Service Websites, in Self-Service in the Internet Age - Expectations and Experiences, Oliver, D., Romm Livermore, C., Sudweeks, F., Editors. Springer: New York. pp. 111 - 137 (2009)

8. Bellman, S., Johnson, E.J., Lohse, G.L.: On Site: To Opt-in or Opt-Out?: It Depends on the Question. Communications of the ACM. 44(2), 25-27 (2001)

9. Brunken, R., Plass, J.L., Leutner, D.: Direct Measurement of Cognitive Load in Multimedia Learning. Educational Psychologist. 38(1), 53-61 (2003)

10. Colligan, L., Potts, H.W., Finn, C.T., Sinkin, R.A.: Cognitive Workload Changes for Nurses Transitioning from a Legacy System with Paper Documentation to a Commercial Electronic Health Record. International Journal of Medical Informatics. 84(7), 469-76 (2015)

11. Day, R.F., Shyi, G.C.W., Wang, J.C.: The Effect of Flash Banners on Multiattribute Decision Making: Distractor or Source of Arousal? Psychology & Marketing. 23(5), 369-382 (2006)

12. Di Stasi, L.L., Antolí, A., Cañas, J.J.: Main Sequence: An Index for Detecting Mental Workload Variation in Complex Tasks. Applied ergonomics. 42(6), 807-813 (2011)

13. Djamasbi, S., Siegel, M., Skorinko, J., Tullis, T.: Online Viewing and Aesthetic Preferences of Generation Y and the Baby Boom Generation: Testing User Web Site Experience through Eye Tracking. International Journal of Electronic Commerce. 15(4), 121-158 (2011)

14. Djamasbi, S., Siegel, M., Tullis, T., Dai, R.: Efficiency, Trust, and Visual Appeal: Usability Testing through Eye Tracking. In: 43rd Hawaii International Conference on System Sciences (HICSS), IEEE, (2010)

15. European-Union: Directive on Consumer Rights, in 2011/83/EU, Union, E., Editor. (2011)

16. Eveland, W.P., Dunwoody, S.: Examining Information Processing on the World Wide Web Using Think Aloud Protocols. Media Psychology. 2(3), 219-244 (2000)

17. Finomore, V.S., Shaw, T.H., Warm, J.S., Matthews, G., Boles, D.B.: Viewing the Workload of Vigilance through the Lenses of the Nasa-Tlx and the Mrq. Human Factors: The Journal of the Human Factors and Ergonomics Society. 55(6), 1044-1063 (2013)

18. Galy, E., Cariou, M., Melan, C.: What Is the Relationship between Mental Workload Factors and Cognitive Load Types? International Journal of Psychophysiology. 83(3), 269-75 (2012)

19. Glöckner, A., Herbold, A.K.: An Eye-Tracking Study on Information Processing in Risky Decisions: Evidence for Compensatory Strategies Based on Automatic Processes. Journal of Behavioral Decision Making. 24(1), 71-98 (2011)

20. González Gutiérrez, J.L., Jiménez, B.M., Hernández, E.G., López López, A.: Spanish Version of the Swedish Occupational Fatigue Inventory (Sofi): Factorial Replication, Reliability and Validity. International Journal of Industrial Ergonomics. 35(8), 737-746 (2005)

21. Goodman, J., Brewster, S.A., Gray, P.: How Can We Best Use Landmarks to Support Older People in Navigation? Behaviour & Information Technology. 24(1), 3-20 (2005)
22. Gopher, D., Braune, R.: On the Psychophysics of Workload: Why Bother with Subjective Measures? Human Factors. 26(5), 519-532 (1984)
23. Grimes, G.M., Valacich, J.S.: Mind over Mouse: The Effect of Cognitive Load on Mouse Movement In: Thirty Sixth International Conference on Information Systems, Forth Worth, USA (2015)
24. Hart, S.G.: Nasa-Task Load Index (Nasa-Tlx); 20 Years Later. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, (2006)
25. Hart, S.G., Staveland, L.E.: Development of Nasa-Tlx (Task Load Index): Results of Empirical and Theoretical Research. Advances in Psychology. 52, 139-183 (1988)
26. Hayashi, T., Kishi, R.: Utilization of Nasa-Tlx for Workload Evaluation of Gaze-Writing Systems. In: IEEE International Symposium on Multimedia, (2014)
27. Hill, S.G., Ianecchia, H.P., Byers, J.C., Bittner, A.C., Zakland, A.L., Christ, R.E.: Comparison of Four Subjective Workload Rating Scales. Human Factors. 34(4), 429-439 (1992)
28. Hogan, M., Barry, C., Torres, A.M.: Theorising and Testing a Taxonomy of Decision Constructs. Journal of Customer Behaviour. 13(3), 171-185 (2014)
29. Hogan, M., Barry, C., Torres, A.M.: An Eye Tracking Pilot Study on Optionality - Some Lessons Learned. In: 17th Irish Academy of Management Annual Conference, Galway (2015)
30. Hoonakker, P., Carayon, P., Gurses, A., Brown, R., Mcguire, K., Khunlertkit, A., Walker, J.M.: Measuring Workload of Icu Nurses with a Questionnaire Survey: The Nasa Task Load Index (Tlx). IIE transactions on healthcare systems engineering. 1(2), 131-143 (2011)
31. Huang, Y.-F., Kuo, F.-Y.: An Eye-Tracking Investigation of Internet Consumers' Decision Deliberateness. Internet Research. 21(5), 541-561 (2011)
32. Johnson, E.J., Goldstein, D.G.: Do Defaults Save Lives? Science. 302, 1338-1339 (2003)
33. Kahneman, D., Miller, D.T.: Norm Theory: Comparing Reality to Its Alternatives. Psychological review. 93(2), 136 (1986)
34. Lai, Y.-L., Hui, K.-L.: Internet Opt-in and Opt-Out: Investigating the Roles of Frames, Defaults and Privacy Concerns. In: Proceedings of the 2006 ACM SIGMIS CPR conference on computer personnel research: Forty four years of computer personnel research: achievements, challenges & the future. Claremount, California, USA: ACM (2006)
35. Levin, I.P., Schneider, S.L., Gaeth, G.J.: All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. Organizational behavior and human decision processes. 76(2), 149-188 (1998)
36. Longo, L., Rusconi, F., Noce, L., Barrett, S.: The Importance of Human Mental Workload in Web Design. In: WEBIST2012 8th International Conference on Web Information Systems and Technologies, (2012)
37. Luximon, A., Goonetilleke, R.S.: Simplified Subjective Workload Assessment Technique. Ergonomics. 44(3), 229-43 (2001)
38. Martin, S.: Measuring Cognitive Load and Cognition: Metrics for Technology-Enhanced Learning. Educational Research and Evaluation. 20(7-8), 592-621 (2014)
39. Miller, S.: Literature Review - Workload Measures. National Advanced Driving Simulator (2001)
40. Mohamed, R., Raman, M., Anderson, J., Mclaughlin, K., Rostom, A., Coderre, S.: Validation of the National Aeronautics and Space Administration-Task Load Index as a Tool to Evaluate the Learning Curve for Endoscopy Training. Canadian Journal of Gastroentrology & Hepatology. 28(3), 155-160 (2014)
41. Nasa: Nasa Tlx Paper and Pen Manual. (1986); Available from: http://humansystems.arc.nasa.gov/groups/tlx/paperpencil.html
42. Nielsen, J., Pernice, K.: Eyetracking Web Usability. New Riders, Berkeley, CA (2010)

43. Nygren, T.E.: Psychometric Properties of Subjective Workload Measurement Techniques: Implications for Their Use in the Assessment of Perceived Mental Workload. Human Factors. 33(1), 17-33 (1991)
44. Paas, F.: Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. Journal of Educational Psychology. 84(4), 429-434 (1992)
45. Pernice, K., Nielsen, J.: Eyetracking Methodology: How to Conduct and Evaluate Usability Studies Using Eyetracking. Nielsen Norman Group Technical Report. (2009)
46. Rubio, S., Díaz, J.E., Martín, J.S., Puente, J.M.: Evaluation of Subjective Mental Workload: A Comparison of Swat, Nasa-Tlx, and Workload Profile Methods. Applied Psychology: An International Review. 53(1), 61-86 (2004)
47. Samuelson, W., Zeckhauser, R.: Status Quo Bias in Decision Making. Journal of risk and uncertainty. 1(1), 7-59 (1988)
48. Schmutz, P., Heinz, S., Métrailler, Y., Opwis, K.: Cognitive Load in Ecommerce Applications—Measurement and Effects on User Satisfaction. Advances in Human-Computer Interaction. 2009, 1-9 (2009)
49. Smuts, M., Scholtz, B., Calitz, A.P.: Usability Guidelines for Designing Information Visualisation Tools for Novice Users. In: Beyond development. Time for a new ICT4D paradigm? Proceedings of the 9th IDIA conference, IDIA2015, Nungwi, Zanzibar (2015)
50. Torres, A., Barry, C., Hogan, M.: Opaque Web Practices among Low-Cost Carriers. Journal of Air Transport Management. 15, 299-307 (2009)
51. Tversky, A., Kahneman, D.: The Framing of Decisions and the Psychology of Choice. Science. 211(4481), 453-458 (1981)
52. Wesiaka, G., Steinera, C.M., Moore, A., Dagger, D., Power, D., Bertholda, M., Alberta, D., Conlan, O.: Iterative Augmentation of a Medical Training Simulator: Effects of Affective Metacognitive Scaffolding1. Computers & Education 76, 13-19 (2014)
53. Wiebe, E.N., Roberts, E., Behrend, T.S.: An Examination of Two Mental Workload Measurement Approaches to Understanding Multimedia Learning. Computers in Human Behavior. 26(3), 474-481 (2010)
54. Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O.N., Atkins, M.S.: Workload Assessment of Surgeons: Correlation between Nasa Tlx and Blinks. Surgical Endoscopy. 26(10), 2746-50 (2012)