Kwantlen Psychology Student Journal
Issue 3, July 2021
Kwantlen Polytechnic University

Effectiveness of 3-D Compared to 2-D Signage on Recycling Behaviour

Mario Jr. Anuales[1] and Kevin Hamilton

Abstract

Using 3-D objects as examples, rather than 2-D icons on signs, to help people learn recycling categories has shown mixed results in observational studies, so an online experimental study was conducted to attempt to clarify the findings. The main hypothesis was that participants would perform faster and more accurately if they learned the recycling categories through images of 3-D objects rather than by 2-D icons. Furthermore, several exploratory hypotheses were suggested: Participants given both types of signage—3-D + 2-D—would perform better than the 3-D and 2-D conditions on their own, and subjective workload and user engagement would predict differences in performance between conditions. An ANOVA found no differences between any of the three conditions in terms of accuracy of sorting performance, subjective workload, or user engagement. However, the 3-D + 2-D condition demonstrated a significant, small-to-medium sized increase in sorting speed when compared to the other two conditions, suggesting that combined 3-D + 2-D signage speeds up decision making without negatively impacting accuracy. One possible explanation is that redundancy of information in the combined condition reduced uncertainty and led to increased speed. However, replication of this result is required because of some limitations inherent to the current study.

Keywords: recycling, signage, sorting task, 3-D vs 2-D signage

[1] Mario.Anuales@email.kpu.ca; Written for a Psychology Honours Thesis. See Appendix H for Acknowledgements

Effectiveness of 3-D Compared to 2-D Signage on Recycling Behaviour

Globally, an estimated 8.3 billion metric tons of plastic was produced between 1950 and 2015, of which only 2.5 billion metric tons had been recycled (Wang, 2019). Despite Canadians becoming more supportive of environmentally conscious behaviours in recent years, there is little evidence of a matching increase in actual sustainable behaviour. This was supported by a survey indicating that 72% of Canadians reported a gap between their intentions and actual behaviours involving the environment (Huddart et al., 2009). Although recyclers and non-recyclers have similar attitudes and motivations towards the environment, a meta-analysis on the determinants of recycling behaviour found that consumer knowledge is one of the best predictors of recycling (Hornik et al., 1995). Specifically, Hornik et al. suggested that a significant barrier to recycling may be a simple lack of knowledge or understanding about how to do it properly. Therefore, it is critical to understand how to effectively communicate information about proper recycling.

Typically, recycling instructions utilize two-dimensional (2-D) images or drawings on signage to indicate what items to place in which recycling bin. Another possible method of presenting this information is to use actual physical examples of the items to be recycled. These items can be displayed in clear containers known as three dimensional (3-D) displays, located near the appropriate recycling bin. However, research into how effective these 3-D displays are compared to 2-D signage is sparse and inconclusive.

There is evidence from a number of studies in different fields that using 3-D rather than 2-D images is more effective at conveying information. For example, a meta-analysis focused on math education found that using physical objects versus abstract 2-D math symbols showed a small to moderate effect in favor of learning with physical objects (Carbonneau et al., 2013). Other studies have found positive effects when using 3-D pictures instead of 2-D pictures in hazard recognition training, where hazards needed to be identified and categorized as a way of improving safety. For example, training with 3-D images has been found to improve hazard recognition in off-road driving (Merritt & Cuqlock-Knopp, 1991) as well as in underground mining (Barrett et al., 1998; Barrett & Kowalski, 1995), and in manual tree falling (Hamilton et al., 2013). Inference based on these training studies suggests that 3-D displays may also be useful at improving categorization and sorting behaviour, such as we find in recycling. The current study was designed to address this issue.

Studies to date examining 3-D signage and recycling behaviour have yielded mixed results. In one observational recycling study conducted at a University of British Columbia (UBC) restaurant, 3-D displays using actual items as examples were able to significantly increase sorting accuracy for napkins and pizza plates, compared to 2-D displays and text (Foster, 2016). However, another study at UBC, conducted in several student residence buildings, found that 3-D displays slightly reduced contamination in the organics bins at only one of the residences (Fu et al., 2016).

Unpublished observational research from UBC has also found inconclusive results on the effectiveness of 3-D signage. One of the authors of this research suggested that there may be “simply too many items in [the] market-place” (I. Zelenika, personal communication, 2019) and that, as a result, people are simply confused about how to categorize them for recycling. Although it has not yet been examined, if Zelenika’s contention is true, one might expect this confusion to be reflected in high measures of mental workload experienced by sorters during recycling. The present study also examined this issue.

While these types of observational studies from UBC may be more ecologically sound than laboratory studies, observational studies are notorious for being affected by extraneous uncontrolled variables that potentially confound the results. This lack of control for potential confounds may help clarify the mixed findings from the UBC studies. Additionally, the Hamilton et al. (2013) study that used 3-D images in forestry hazard recognition suggested that greater motivation or engagement when using 3-D compared to 2-D training images could have introduced a confounding effect leading to better hazard recognition in the 3-D condition. Therefore, the present study included a measure of engagement as well.

Thus, because of the better control afforded by laboratory compared to observational studies, a highly controlled experimental study was designed to examine the efficacy of 3-D compared to 2-D signage in promoting recycling behaviour. In addition to looking at the efficacy of 3-D vs 2-D signage on recycling speed and accuracy, the present study also examined ratings of personal engagement and subjective workload as variables that might be related to the efficacy of recycling behaviour. The computer-based study looked at three signage conditions: 3-D, 2-D, and 3-D + 2-D. The rationale for adding the combined condition was that, to date, there is no evidence looking at the effectiveness of the two forms of signage combined compared to 2-D or 3-D alone. It could be argued that redundancy of information is a way of helping to decrease uncertainty about sorting categories.

Hypotheses

The primary hypothesis was that

1) Sorting performance would be better when using 3-D as compared to 2-D signage. Performance was examined by measuring accuracy and speed of sorting.

Thus, our exploratory hypothesis concerning the combined signage was that

2) 3-D + 2-D signage would lead to better recycling performance than either 2-D or 3-D alone.

   We were also interested in looking at type of signage and sorting in relation to two other possible effects. The first concerned mental workload, and the second was engagement or motivation. In this regard, two further exploratory hypotheses were that

3) Mental workload would be higher using 2-D compared to 3-D signage and that using 2-D + 3-D it would be lower than with 2-D or 3-D alone; and
4) Engagement would be lower using 2-D compared to 3-D signage and that using 2-D + 3-D it would be lower still than with 2-D or 3-D alone.

Method

Participants

One-hundred forty-eight participants were obtained from the psychology research subject pool of a medium-sized Canadian university. All participants were university students with no evidence of visual or motor impairment. Most participants were young, female, and Canadian residents. The originally intended sample size of 158 participants was calculated through G*Power analysis (Faul et al., 2007) using the following parameters: ANOVA fixed effects interaction, special, main effects and interactions; $f = 0.25$ (medium sized effect); $\alpha = 0.05$; power = 0.80; numerator $df = 2$; 3 groups. Participants were compensated for taking part through bonus course credit (0.5%).

Design

There was only one IV, namely, the signage condition. The signage condition had three levels: 2-D, 3-D, 2-D + 3-D. The primary dependent variables were average response time (measured in seconds) and accuracy (number of correct sorts, out of a total of 48). Exploratory dependent variables were subjective workload and user engagement.

A short demographics survey was included, which asked about gender, age, and highest level of education completed. These variables were used to describe the study participants and were not used, at this time, to analyze results of the study.

In addition to the demographics survey, a questionnaire on recycling attitudes (adapted from Werner et al., 1995) was included just before the conclusion of the study. This questionnaire provided further participant information; however, this data was also not examined at this time.

Materials

Overall, the study utilized a consent form, a demographics survey, a training and practice task, a main sorting task (either 3-D, 2-D, or 3-D + 2-D for each participant), a subjective workload measure (National Aeronautics and Space Administration [NASA], 1986), a participant engagement measure (adapted from O'Brien & Toms, 2010), a recycling attitudes measure (Werner et al., 1995), and finally a debrief of the study. Further details about the signage, items sorted, the measures used, and the experimental procedures are provided below.

Signage

There were three types of signage used: 2-D, 3-D, and 2-D + 3-D. The 2-D signage was based on the existing recycling signage at Kwantlen Polytechnic University (KPU), with slight modifications. These were four signs, one for each of the sorting categories: organics, paper, recycling, and waste. Each sign featured a large text label of the category, a colour-coded background, and four iconic graphics representing the example items that belonged to the category (see Appendix A).

The 3-D signage consisted of four photographs, again corresponding to each of the four sorting categories. Each photograph contained the four example items used for the category and had the category label, in colour-coded text, above the photograph (see Appendix B). The 2-D + 3-D condition simply displayed both the 2-D and 3-D signage.

A total of 48 recyclables—12 in each of the four categories (organics, paper, recycling, and waste)—were shown to participants to sort. These recyclables are generally what is found in waste bins on the KPU Surrey campus and were selected with input from Sustainable Kwantlen Student Association and waste audits from Environmental Protection Technology classes at KPU. Items in the organics category were defined as any food items and objects made from biodegradable materials (e.g., napkins and takeout containers made from biodegradable paper). Paper items were defined as any item made of paper that was recyclable yet not biodegradable (e.g., paper coffee cups). Recycling items were defined as anything made of recyclable plastic or glass. Waste items were defined as anything made of nonrecyclable material. Items deemed recyclable were according to KPU's on-campus categories as of early 2019, rather than being based on local domestic recycling rules. See Appendix C for a complete list of all recyclables that were used in the study, organized by category.

Measures

The engagement measure was adapted from an existing survey intended to rate shopper engagement while completing an online shopping task (O'Brien & Toms, 2010; see Appendix D). The original engagement scale examined engagement along four dimensions: Focused Attention, Perceived Usability, Aesthetic Appeal, and Reward. In the process of adapting this measure to the sorting task, some wording was slightly changed, and irrelevant items removed, while preserving the same four dimensions of engagement. Since one of the questions depended on the display shown to the participant, three versions of the survey were made, one for each condition. The reliability of the adapted surveys, measured via Cronbach's alpha, was .84 for the 2-D, .84 for the 3-D, and .81 2-D + 3-D. The overall mean score was 2.64 and the standard deviation was 0.52.

The NASA Task Load Index (NASA-TLX; NASA, 1986; see Appendix E), subjective workload measure has been widely used since the 1980s and is considered an industry standard for subjective workload measurement. The NASA-TLX asks participants to rate their task experience using six 21-point scales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. The overall mean was 8.21, and the standard deviation was 4.15.

The recycling attitudes measure used was adapted from the Recycling Attitudes Questionnaire (RAQ; Werner et al., 1995). The RAQ is a rating scale that asks participants about their recycling attitudes, behaviours, and self-concept as a recycler.

Procedure

The sorting task was administered in the form of an online study created using Qualtrics (www.qualtrics.com). Training and practice tasks were included to introduce and familiarize participants with the signage used in the task and to help minimize learning effects within conditions. See Appendix F for a screenshot of a recycling trial.

After agreeing to take part in this research, the online study randomly assigned each participant to one of the three signage conditions and then sent them to a training/practice session where they were familiarized with the task. The training task consisted of five-steps: an introduction, followed by four practice trials. The introduction featured and explained the signs and sorting examples that the participant would see during the main task; it also instructed participants to ignore local recycling rules and to only use the signs and/or examples to guide their sorting decisions. The four practice trials each featured a different item to sort, one for each category indicated by the signage. Sorting was accomplished by clicking on the appropriate sign for the item presented. The four items in these practice trials were given in a fixed order and did not vary between signage conditions. The practice trials provided feedback by either telling the participant their choice was correct or by repeating the presentation with the incorrect choice removed. After completing the four practice trials, the participant was informed-that the main sorting task was about to begin.

In the main sorting task, each item to be sorted was presented on a separate page. At the top of the page, participants saw 2-D, 3-D, or 2-D + 3-D signage, corresponding to the condition they were assigned. At the bottom of the page was an image of the item to be sorted, and four buttons that each corresponded to one of the four signage categories. After the participant selected one of the buttons to make a sorting decision, they could move onto the next page/item until they sorted all 48. Participants did not receive feedback on their sorting decisions in the main task.

Response time was recorded on a per-page/per-item basis, measuring how many average seconds it took until the participant made a sorting choice and clicked the submit button. Qualtrics also recorded which choice was selected and, during later data analysis, the number of correct choices was tallied into a score.

After sorting all 48 items, the task ended and participants completed the NASA-TLX, the engagement measure, and the RAQ. Upon completing these, participants received a debriefing and were thanked. The entire survey took roughly twenty to thirty minutes to complete.

Results

The signage analysis consisted of two 3-level (Signage: 2-D signs, 3-D boxes, 2-D + 3-D) one-way analyses of variance (ANOVA) with speed and accuracy as the main dependent variables. Speed (response time; RT) was recorded as the average number of seconds it took to make a sorting decision and click the submit button. Accuracy was measured by scoring the participant's sorting decisions, where each correct sort was worth one point, out of a total of forty eight. Subjective workload and user engagement were also analyzed with one-way ANOVAs. For all analyses, alpha level was set at .05. Effect sizes for significant differences were reported in terms of eta-squared for the ANOVA and Cohen's *d* for specific comparisons between conditions.

Time and Accuracy ANOVA

Levene's test was conducted and returned non-significant results for both time and accuracy, $p = .57$ and $p = .48$ respectively, indicating no violation of the assumption of homogeneity. Visual inspection of the histograms for the time and accuracy variables indicated three outlier scores: two for time and one for accuracy. Further analysis confirmed these outliers had *z*-scores either above 3.29 or below -3.29. The pattern of ANOVA results, however, was not affected by the presence of the outliers, so the results are presented with their data included.

The result of the ANOVA on the mean RT data was significant, $F(2, 139) = 5.33$, $p = .006$, with a small to medium effect size, $\eta^2 = .071$, at an observed power of .83. Follow up post-hoc testing using the method of Least Significant Difference (LSD) indicated two significant differences. Both of these comparisons involved the 3-D + 2-D combined condition. Specifically, the 3-D + 2-D combined condition was significantly faster than either the 2-D or 3-D condition. The combined 3-D + 2-D condition had a small effect size when compared with 2-D, $p = .006$, $d = -1.97$, and with 3-D, $p = .004$, $d = -2.05$.

With regard to accuracy, the ANOVA examining the number of errors was non-significant, $F(2, 139) = .83$, $p = .44$, indicating no differences between signage condition for sorting accuracy. Means, standard deviations, and 95% confidence intervals for the average time can be seen in Table G1.

User Engagement and Subjective Workload

Levene's test indicated non-significant results for user engagement, $p = .79$, and subjective workload, $p = .69$, preserving the assumption of homogeneity. The ANOVA examining user engagement was non-significant, $F(2, 139) = .91$, $p = .41$, as was the ANOVA examining subjective workload, $F(2, 139) = 1.17$, $p = .32$. Thus, there were no differences between signage conditions' average scores on engagement or subjective workload (see Table G2).

Discussion

The results of this study did not support the hypothesis that presenting information using 3-D images would increase accuracy and decrease sorting time compared to using 2-D images. This contradicts the research literature that found positive effects when using 3-D examples during math instruction or hazard recognition training in mining, off-road driving, and forestry (Barret & Kowalski, 1995; Barrett et al., 1988; Carbonneau et al., 2013; Hamilton et al., 2013).

Of particular note is recognizing that what was referred to in this study as 3-D in fact was not true 3-D. The photographs of the example items were taken by a simple cellphone camera, whereas the 3-D stimuli in the earlier 3-D studies were created using actual 3-D photography and required special viewing glasses to see the 3-D images properly. Since this was required to be an online study as per COVID-19 restrictions, using 3-D photographs as stimuli was not a viable option. This difference might account for the discrepancy between the results of this study and previous findings.

Furthermore, the lack of significant differences between the 2-D and 3-D signage might be explained in that the 2-D condition had some 3-D visual cues while the 3-D condition lacked salient 3-D visual cues, rendering both of them as effectively 2.5-D stimuli. The 2-D condition contained monocular cues of depth perception such as occlusion and height in visual field (Cutting & Vishton, 1995).

Another explanation for why the current study does not support the earlier 3-D studies may be that the task of recycling behaviour qualitatively differs from learning math and identifying hazards in off-road driving, mining, and forestry. In recycling, a person with a waste object needs to compare this object to examples on the signage and find the closest match; on the other hand, math instruction involves using the example to visualize an abstract math concept (e.g., multiplication tables illustrated by rows of blocks), and hazard recognition involves recognizing objects/situations in the environment as being similar to the examples in safety training (e.g., identifying a rotting branch on a tree by remembering the example shown in training).

When compared to past research on the effectiveness of 3-D, the results of the present study indicate that 3-D examples do not have a significant effect on recycling performance in an experimental context. These results contrast with some of the findings of the previous 3-D display studies previously mentioned that were able to find significant results in observational settings (Foster, 2016; Fu et al., 2016; I. Zelenika, personal communication, 2019), but not all of those studies found this effect. However, more research into this topic is required before definite conclusions can be drawn.

While the results of this study do not directly support the view that 3-D images lead to better sorting outcomes, the effects observed do partially support the first exploratory hypothesis, which looked at 3-D and 2-D signage together. Here, the combined 3-D + 2-D condition had faster performance than either the 3-D or 2-D conditions separately. However, the 3-D only and 2-D only conditions did not differ from each other in sorting speed.

Because speed and accuracy are interrelated (speed/accuracy trade-off: Pachella & Pew, 1968) and participants showed no differences in accuracy between conditions, the RT differences observed with the 3-D + 2-D condition can be unambiguously interpreted as indicating valid differences in sorting performance.

One possible explanation for the difference in performance seen with 3-D + 2-D signage is the redundancy of information present in the combined condition, since it uses both types of signage presented together. This redundancy of information may reduce a participant's uncertainty during sorting decisions. Redundancy is sometimes used in signs to reduce the chance of miscomprehension or uncertainty. For instance, the principle of redundancy reducing uncertainty is used in the design of many road signs, where for example the instruction to STOP is conveyed by text, colour and shape of the sign (Shinar et al., 2003).

Lastly, the results do not support hypotheses 3 or 4, as there were no significant differences between signage conditions on user engagement or subjective workload. This does not support the suggestions to study these variables that were made by previous studies (Hamilton et al., 2013; I. Zelenika, personal communication, 2019).

Limitations

A key limitation of this study pertained to the research being conducted online rather than in a more ecologically valid setting. In reality, people can receive recycling information from a variety of sources such as signage above the bin or nearby, informational/sensory cues from the recyclables (such as identifying an item's material as paper or plastic, or reading recycling information on the packaging), and more. In this study, the participant was instructed to use only the signage provided to help make their sorting decision. While this helps internal validity by controlling for extraneous variables that might be present in an observational study, it also does not replicate the same context that someone experiences when recycling.

For example, in the real world, if someone wanted to properly recycle a used coffee cup, they would have to remove the plastic lid, remove the cardboard sleeve, and wash the cup until it was clean before sorting them; these items would then be placed in the waste bin (plastic lid), paper bin (cardboard sleeve), and recycling bin (clean, empty cup), respectively. In the current study, participants did not have to do these extra steps and the items to be sorted were presented individually. Moreover, there were very few "dirty" items used in the study.

There was a slight difference in the information presented in the 2-D signage compared to the 3-D signage. Due to the limitations of making icons using simple vector graphics, it was difficult to clearly represent certain items individually, such as a coffee cup lid, and these items were subsequently paired with other items to make the representation clearer. In order to make it clear which of the items was being represented, a large “X” was placed on the non-relevant item. e.g. to represent a coffee cup lid, the lid was shown on top of the coffee cup but the cup itself was crossed out by an “X.” These “X” symbols were not used in the 3-D condition. See Appendices B and C to compare these representations.

There were also some limitations with the Qualtrics’ survey building program used to construct this study. The page timer only worked by measuring time until the first click, the last click, and page submission. Time until page submission was used to calculate time in this study. This means there may have been extra time added if participants stayed on the page for a while between making their sorting decision and deciding to submit and move to the next page. Additionally, the pages containing the sorting trials had many large images on them and the participants may have only been able to view about one third of the page at time if they were using a smartphone to complete the survey; in this case, the participants would then have to spend some time scrolling down past the signage images in order to make a sorting choice, which might also have added extra time. Lastly, there were some other differences between this sorting task and reality that may have affected the results. The participant likely had to sort far more recyclables at one time in this task (48 items) than they would normally, and this could have caused annoyance or boredom if they found the task repetitive. Similarly, another differing factor is focus; the goal of recycling quickly and accurately is the main focus during the task in this experiment, and participants might not devote that much conscious effort to recycling accuracy in their everyday life. Furthermore, while the participants were instructed to ignore their local recycling rules and follow the signs, they may have experienced confusion if they relied on their previous recycling experience or attempted to search for recycling information online while completing the task.

Future Research

Since there seems to be a research gap regarding studies that look at combining 2-D signage with 3-D examples to present recycling information, more studies, especially experimental ones, looking at this area are needed. Conducting this study in a virtual reality setting can retain the control of an experimental setting while increasing ecological validity. Replications of this study would help determine if the demonstrated effect on sorting time is robust and consistent. Furthermore, since redundancy of information may possibly explain the effects seen in this study, recycling studies that specifically examine signage employing redundancy of information might help clarify these findings. Lastly, Zelenika (personal communication, 2019) also mentioned that an alternative to signs was having a designated person near recycling bins and directing people on how to sort, which showed effectiveness at large events. Studies like these can help clarify which design methods are the most effective at encouraging recycling and thus prevent pollution caused by excessive waste production.

References

Barrett, E. A., Wiehagen, W. J., & Peters, R. H. (1988). *Application of stereoscopic (3-D) slides to roof and rib hazard recognition training* (Vol. 9210). US Department of the Interior, Bureau of Mines.

Barrett, E. A., & Kowalski, K. M. (1995). *Effective hazard recognition training using a latent-image, three-dimensional slide simulation exercise.* US Department of the Interior, Bureau of Mines.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, *105*(2), 380–400. https://doi.org/10.1037/a0031084

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion* (pp. 69–117). Academic Press. https://doi.org/10.1016/B978-012240530-3/50005-5

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Foster, K. (2016, April 19). The effectiveness of 3D display cases in the AMS Nest. *UBC Open Collections.* https://doi.org/10.14288/1.0343156

Fu, A., Siu, A., & Misra, R. (2016). The effectiveness of 3D displays on sorting behaviour in post- secondary institutions. *UBC Open Collections.* https://doi.org/10.14288/1.0343157

Hamilton, K., Saffari, N., & Colman, J. (2013). Training to recognize worksite hazards using 3D and degraded imagery. *44th Annual Conference of the Association of Canadian Ergonomists*.

Hornik, J., Cherian, J., Madansky, M., & Narayana, C. (1995). Determinants of recycling behavior: A synthesis of research results. *The Journal of Socio-Economics*, *24*(1), 105–127. https://doi.org/10.1016/1053-5357(95)90032-2

Huddart, E., Beckley, T. M., McFarlane, B. L., & Nadeau, S. (2009). Why We Don't "Walk the Talk": Understanding the Environmental Values/Behaviour Gap in

Canada. *Human Ecology Review*, *16*(2), 151–160. https://www.jstor.org/stable/24707539

Merritt, J. O., & Cuqlock-Knopp, V. G. (1991). Perceptual training with cues for hazard detection in off-road driving. *Proceedings of the SPIE, Stereoscopic Deisplays and Applications II*, 133–138. https://doi.org/10.1117/12.46302

National Aeronautics and Space Administration. (1986). *NASA Task Load Index.* Retrieved from https://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf

Pachella, R. G., & Pew, R. W. (1968). Speed-accuracy tradeoff in reaction time: Effect of discrete criterion times. *Journal of Experimental Psychology*, *76*(1), 19–24. https://doi.org/10.1037/h0021275

Shinar, D., Dewar, R. E., Summala, H., & Zakowska, L. (2003). Traffic sign symbol comprehension: A cross-cultural study. *Ergonomics*, *46*(15), 1549–1565. https://doi.org/10.1080/00140130320001216155

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, *61*(1), 50–69. https://doi.org/10.1002/asi.21229

Wang, T. (2019, September 2). *Global plastic waste*. Retrieved December 12, 2019, from https://www.statista.com/topics/5401/global-plastic-waste/

Werner, C. M., Turner, J., Shipman, K., Twitchell, F. S., Dickson, B. R., Bruschke, G. V., & von Bismarck, W. B. (1995). *Recycling Attitudes Questionnaire* [Database record]. APA PsycTests. http://doi.org/10.1037/t26411-000

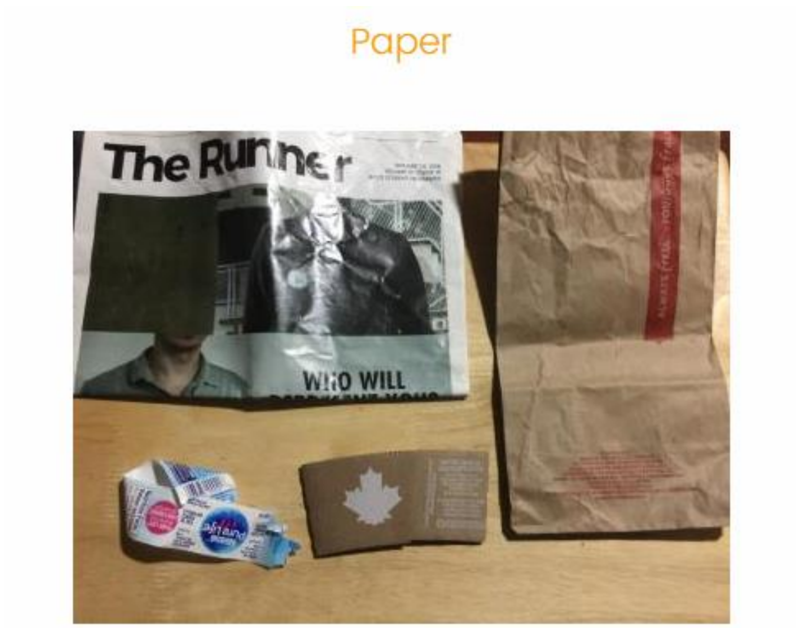
Appendix A

2-D Recycling Signs

Appendix B

Examples (3-D condition)

Pictured above, clockwise starting from top left: an empty white cardboard takeout container with dark brown paint stains, a red apple, a paper muffin cup with muffin crumbs stuck to it, and brown paper napkins with dark brown paint stains. All items are on a small wooden table.



Pictured above, clockwise starting from top left: a newspaper, a small brown paper bag, a brown drink sleeve made from cardboard, a water bottle label made of laminated paper. All items are on a small wooden table.

Recycling

Pictured above, from left to right: empty juice box with the attached straw removed, empty plastic water bottle with label and cap removed, empty energy drink can, empty coffee cup with lid and drink sleeve removed. All items are on a small wooden table.

Waste

Pictured above, clockwise starting from leftmost item: empty small potato chip bag, plastic straw in sealed paper wrapper, coffee cup lid made of plastic, plastic water bottle cap. All items are on a small wooden table.

Appendix C

List of Recyclables Used

Compost	Paper	Recycling	Garbage
Grassroots To-go container	TimHortons Sleeve	TimHortons Coffee Cup (clean)	TimHortons Coffee Cup Lid
Donut	Starbucks Sleeve	Starbucks Cup (clean)	Starbuck Coffee Cup Lid
Soiled paper plate	GrassRoots Sleeve	GrassRoots Tea Cup (clean)	GrassRoots Tea Cup Lid
Soiled napkins	Starbucks frappucino glossy paper label	TimHortons IceCap Cup (clean)	TimHortons Plastic straw
Leftover chips	Paper label (beer bottle)	IceCap Cup Plastic lid (clean)	Plastic utensils
Unfinished Burger	Glossy paper label (pop bottle)	Plastic pop bottle (empty)	Plastic pop bottle cap
Leftover muffin	Glossy paper label (water bottle)	Plastic water bottle (empty)	Plastic water bottle cap
Soiled paper muffin cup	TimHortons Paper bag	Juice box (empty)	Chip bag
Apple	Newspaper	Beer Bottle (clean)	Juice box plastic straw
Oranges	Magazine	Starbuck frappucino glass bottle (empty)	Starbucks frappucino Bottle cap
Sandwich leftovers	Flyer	Soda can	Beer Bottle cap
Salad leftovers	Lined paper	Energy drink can	Candy/gum wrapper

Appendix D

Engagement Scale

1. I lost myself in this sorting experience
2. I was so involved in my sorting task that I lost track of time
3. I blocked out things around me when I was sorting.
4. When I was sorting, I lost track of the world around me.
5. The time I spent sorting just slipped away.
6. I was absorbed in my sorting task.
7. During this sorting experience I let myself go.
8. I was really drawn into my sorting task.
9. I felt involved in this sorting task.
10. This sorting experience was fun.
11. I felt interested in my sorting task.
12. This sorting task was worthwhile.
13. I consider my sorting experience a success.
14. This sorting experience did not work out the way I had planned.
15. My sorting experience was rewarding.
16. The signs/3-D display was aesthetically appealing.
17. The signs/3-D display appealed to my visual senses.
18. I felt frustrated while doing this sorting task.*
19. I found this sorting task confusing. *
20. I felt annoyed while doing this sorting task. *
21. I felt discouraged during this sorting task. *
22. This sorting task was mentally taxing. *
23. This sorting experience was demanding. *
24. I felt in control of my sorting experience.
25. I could not sort some of the things I needed to do on this task.*

Scale administered on 5-point scale with “strongly disagree” and “strongly agree” at the respective endpoints. Items with asterisks are reverse coded.

Questions 16 and 17 will change wording to match the relevant condition. e.g. “The sign was aesthetically appealing” for the 2-D condition, and “The sign and 3-D display was aesthetically appealing” for the 2-D + 3-D condition.

Appendix E

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

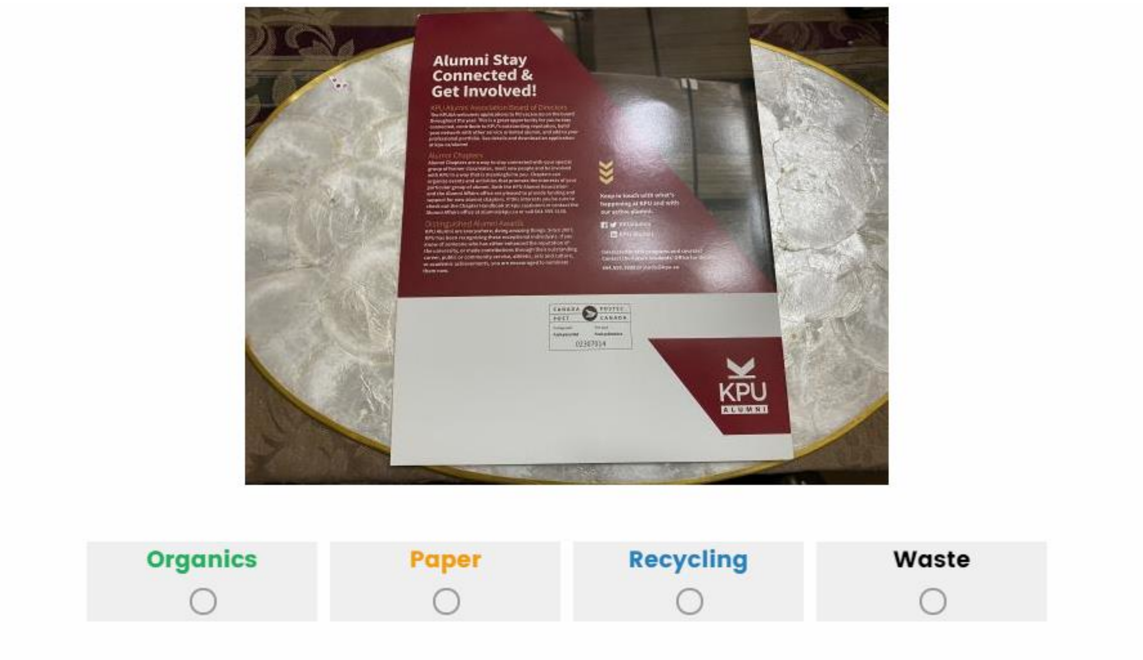
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Appendix F

Screenshot of recycling trial



Appendix G

Table G1

Means, Standard Deviations, and 95% CIs for Average Time and Total Score According to the Three Levels of Signage Condition

Dependent Variable	Signage Condition	*M*	*SD*	*95% CI*
Average Time	2-D	9.13	4.14	[7.91, 10.34]
	3-D	9.20	2.56	[8.45, 9.94]
	3-D + 2-D	7.16	3.46	[6.13, 8.18]
	Total	8.52	3.54	[7.93, 9.10]
Total Score	2-D	37.98	5.14	[36.47, 39.49]
	3-D	39.14	6.11	[37.39, 40.90]
	3-D + 2-D	39.54	6.98	[37.47, 41.62]
	Total	38.89	6.11	[37.87, 39.90]

Note. CI = confidence interval. Time is measured in seconds.

Table G2

Means, Standard Deviations, and 95% CIs for Ratings on the User Engagement Scale and NASA-TLX (Subjective Workload) According to the Three Levels of Signage Condition

Dependent Variable	Signage Condition	*M*	*SD*	*95% CI*
User Engagement	2-D	2.57	.524	[2.41, 2.72]
	3-D	2.64	.532	[2.49, 2.79]
	3-D + 2-D	2.71	.490	[2.57, 2.86]
	Total	2.64	.516	[2.55, 2.72]
NASA-TLX	2-D	7.93	4.08	[6.74, 9.13]
	3-D	7.76	4.22	[6.53, 8.98]
	3-D + 2-D	8.97	4.14	[7.74, 10.20]
	Total	8.21	4.15	[7.52, 8.90]

Note. CI = confidence interval. NASA-TLX = National Aeronautics and Space Administration Task Load Index. User engagement ratings were from 1 – 5, while TLX ratings were from 1 – 21.

Appendix H

Acknowledgements from Mario Jr. Anuales

I would like to thank my supervisor, Dr. Kevin Hamilton: For agreeing to take me on as an Honours student in the first place, for all of the excellent feedback and guidance he has provided, and for his patience and understanding.

I would also like to thank the many people that helped me throughout this process, without whom this work would not have been possible.

I would like to start with my family for the invaluable support they have given me.

Additionally, I would like to thank the following people for providing help during the conceptualization of this study: Dr. Daniel Bernstein for teaching the lecture portion of the 2020 Kwantlen Psychology Honours program, as well as my fellow Honours cohort members, they provided valuable criticism and suggestions; Ms. Lisa Hubick, a KPU librarian, who assisted me during my literature review; Mr. Zebulon Fastabend, a KPU student and community member, who provided a practical perspective by discussing recycling signage best practices in his workplace with me; Dr. Jiaying Zhao and Dr. Ivana Zelenika from the UBC Behavioural Sustainability lab who corresponded with me and provided critical information on my research topic; and Dr. Farhad Dastur of the KPU Psychology department for his suggestion regarding the appropriate types of images to use in the signage.

Furthermore, I would also like to thank all of the people that provided support during the implementation of this study: Ms. Erin Pedersen, the head of the Sustainable Kwantlen Student Association (KSA), who consulted with me on recycling at KPU, and provided me with recycling audits from previous years as well as copies of the recycling signage currently used on campus that would become the basis for the signage used in my own study; Dr. Jocelyn Lymburner of the KPU Psychology department who provided access to her Qualtrics account; Mr. Anthony Gentilezo, my cousin and a graphic designer, who created some of the icons used on the signage in my study; and finally, I would like to thank Ms. Ivy Ng and the staff at the KPU Psychology labs for their assistance in getting my study on SONA and helping me test my study.

Lastly, I would like to thank KPU's Office of Research Services for providing funding for this project through the Student Led Research Grant (SLRG).