Investigating Image Comprehension with a Tactile Image Enhancer: A 2x3 Factorial in a Randomized Complete Block

Madison Dunn
Stat 582

Introduction:

The goal of this study is to investigate a blind person's ability to comprehend Web images using a Tactile Image Enhancer. Web images will be categorized into two types: pictures and symbols, each undergoing one of three simplification processes: staying in original form, edge detection, or segmentation. By examining these different image types and simplification levels, the study aims to find which combinations of these variables will be most effective for a blind person to view with a Tactile Image Enhancer.

Methods:

Participants:

Twenty participants will be recruited from the Indiana School for the Blind.

Procedure:

Before beginning the experiment, subjects will undergo a brief orientation to ensure they are comfortable with the Tactile Image Enhancer. Participants have the opportunity to ask questions and practice using the Tactile Image Enhancer. The experiment itself will involve each subject sitting at a table while the researcher presents them with an image and the subjects are asked to identify it via the Tactile Image Enhancer. The time it takes the subject to correctly identify the image will be recorded in seconds. There will be no time limit for identifying the image. In the event the subject is unable to identify the image, the researcher will move on to the next image, and an 'X' will be marked for this trial, and this trial will be removed from any calculations. The subject will view nine images, then take a five-minute break to help reduce fatigue, and then view another nine images for a total of eighteen images. The subject is then free to leave. This procedure will be repeated for all twenty subjects.

Experimental design:

This study will utilize a 2x3 factorial design with randomized complete blocks. The first factor is image type with two levels: picture and symbol. The second factor is simplification level with three levels: original, edge detection, and segmentation. These two factors will result in six treatment combinations, shown in **Figure 1.** There are 14 Web images and the six combinations have been applied to each, so there are 84 possible images to be used in this study. These image combinations are summarized in **Figure 2**.

Figure 1. Six treatments for study

Treatments	
1. Photo-Original	2. Symbol-Original
3. Photo-Edge Detection	4. Symbol-Edge Detection
5. Photo-Segmentation	6. Symbol-Segmentation

The study will be blocked on subject. Each subject will receive every treatment. This will help reduce between-subject variability. Image combinations are shown in **Figure 2.** Each participant will receive three images from each of the six treatment rows, resulting in 18 images. The order of the 18 images will be entirely randomized but balanced to ensure the participant gets exactly three images from each treatment. For example, a subject may receive image 1 with picture and edge detection applied, then receive image 12 with symbol and segmentation applied. Images may be repeated. These results will be averaged to give each subject one final result in seconds for each of the six treatments. **Figure 3** shows an example of what the data will look like once collected. This will be repeated for each subject. The underlined portion within the table will be where the actual data will be displayed. The math shown in this figure is just for demonstration purposes.

Figure 2. Image combinations:

Treatment	Image				
1	1	2	3	…	14
2	1	2	3	…	14
3	1	2	3	…	14
4	1	2	3	…	14
5	1	2	3	…	14
6	1	2	3	…	14

Figure 3. An example of data collection for one subject

	Subjects (blocks)	1 Image type	
		Photo	Symbol
Simplification Level	Original	$\frac{(Trial\ 1 + \ trial\ 2\ + \ trial\ 3)}{3}$ = Average time (s)	$\frac{(Trial\ 1 + \ trial\ 2\ + \ trial\ 3)}{3}$ = Average time (s)
	Edge Detection	$\frac{(Trial\ 1 + \ trial\ 2\ + \ trial\ 3)}{3}$ = Average time (s)	$\frac{(Trial\ 1 + \ trial\ 2\ + \ trial\ 3)}{3}$ = Average time in seconds
	Segmentation	$\frac{(Trial\ 1 + \ trial\ 2\ + \ trial\ 3)}{3}$ = Average time in seconds	$\frac{(Trial\ 1 + \ trial\ 2\ + \ trial\ 3)}{3}$ = Average time in seconds

Analysis:

The analysis will involve an examination of the main effects of image type and simplification level, as well as their interaction effect, on comprehension time. Blocking will be a part of the model as well. The linear model will be as follows:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij}\,\delta_k + \epsilon_{ijk} \qquad \begin{cases} i = 1, 2 \\ j = 1, 2, 3 \\ k = 1, 2, \ldots, 20 \end{cases}$$

- τ_i = image type effect
- β_j = simplification level effect
- $(\tau\beta)_{ij}$ = interaction effect
- δ_k = block effect
- ϵ_{ijk} = error component

An ANOVA table will be calculated, including the sum of squares, degrees of freedom, mean squares, and F-statistics. The ANOVA table is presented in **Figure 4**. Only the degrees of freedom have been calculated. The sum of squares, mean square, F-values, and P-values will be calculated once data collection is complete. Formulas and notation have been included in the table for convenience.

Figure 4. Analysis of Variance Table

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	P-Value
Image Type (A)	$a - 1 = \underline{1}$	SS_A	MS_A	$\frac{MS_A}{MS_E}$	*P-value*
Simplification Level (B)	$b - 1 = \underline{2}$	SS_B	MS_B	$\frac{MS_B}{MS_E}$	*P-value*
Image Type x Simplification Level (AB)	$(a-1)(b-1) = \underline{2}$	SS_{AB}	MS_{AB}	$\frac{MS_{AB}}{MS_E}$	*P-value*
Blocks (subjects)	$n - 1 = \underline{19}$	SS_{block}	MS_{Block}	$\frac{MS_{Block}}{MS_E}$	*P-value*
Error	$(ab - 1)(n-1) = \underline{95}$	SS_E	MS_E		
Total	$abn - 1 = \underline{119}$	SS_{Total}			

From this ANOVA table, main effects and interaction effects will be tested. This will help to answer whether image comprehension time depends on image type, simplification level, or an interaction of both. The model fit will also be assessed via a goodness of fit test with a likelihood ratio test, looking at the deviance residuals. Predictions can be made via the model as well.

Analysis will also be done within blocks. Multiple comparisons will be completed with Tukey's HSD tests. Within each block, fifteen possible combinations can be looked at. For example, treatment 1 (Photo-Original) vs treatment 2 (Photo-Segmentation) for subject 1. Through these tests, conclusions can be made about which treatment groups have greater average comprehension time compared to others.

Conclusion:

In this study, the comprehension abilities of blind individuals using a Tactile Image Enhancer to interpret images on the web will be investigated. To analyze the data, a 2x3 factorial design with randomized complete blocks will be employed, with blocking by subject to reduce variability. The time taken for participants to correctly identify each image will be measured, allowing comprehension across different image types and simplification levels to be assessed.

Specifically, the Tukey tests completed in the analysis portion will help determine which image type and simplification level is most effective for blinder users to identify images with the Tactile Image Enhancer. The model will help predict comprehension time in seconds based on which levels of the two variables are used.

There are several advantages to using this experimental design. The design controls for variability among participants through blocking, which in turn should help increase the power of the study. There will be increased precision in terms of treatment effects. Balancing and randomization are completed to help reduce possible bias. Lastly, each subject viewing multiple images and having an average comprehension time per treatment should help the study be more representative of overall comprehension time, not just the ability to recognize one specific image.

Throughout the study, several potential issues will appear. The design attempts to reduce possible order effects and fatigue by randomizing the order of image presentation and incorporating a break between trials. However, these factors may still be influenced by participants' individual abilities and might not be completely removed. Additionally, other factors such as age and cognitive ability could act as confounders, possibly affecting how long it takes a participant to identify an image or their ability to use the Tactile Image Enhancer. Future studies could address these confounders by blocking participants based on age or gender and similar analyses from this study could be done.