

Quan Nguyen Hien

STAT 515 - Project 1

Professor Bruce Craig

1) Problem Description:

The study seeks to assess the ability of blind users to comprehend images on the web using a Tactical Image Enhancer by determining the most suitable image formats for easy comprehension. Fourteen common web images will serve as the basis for this experiment, which will be further transformed with two different types of image: 1) Symbol and 2) Picture. Furthermore, the images will also undergo further simplification into three levels: 1) Edge Detection, 2) Segmentation, and 3) Original picture as the reference. In total we will have 84 different images. This experiment will encompass a total of six treatment levels aimed at comparability. Participants from the Indiana School for the Blind, trained in using the Tactical Image Enhancer, will be recruited to examine these prepared images. The primary variable of interest is the time each participant spends on correctly identifying each image, measured in seconds. It is anticipated that each participant can handle up to 20 images before fatigue impacts their performance. Further details on experiment design and rationale will be presented in the subsequent section.

2) Proposed Design of the Experiment

In designing the experiment, all combinations of the two treatment factors, **Image Type** and **Simplification Level**, will be considered, resulting in six distinct combinations or treatment levels. Each participant will be presented with 18 different images, with each combination of **Image Type** and **Simplification Level** appearing three times in randomized order to mitigate order-related variability. For clarity, Symbol and Picture forms will be denoted as I1 and I2, respectively, for Image Type, while Original, Edge Detection, and Segmentation will be denoted as S1, S2, and S3, respectively, for Simplification Level. An example of the randomized ordering of images for Participant 1 could be as follows: The following could be an example with randomized ordering of images given to Participant 1 with any image will have the associated factor in a randomized order (e.g.: I2 S3, I1 S3, I1 S2, I2 S1, I2 S2, I1 S1, I2 S1, I1 S3, I2 S3, I1 S2, I2 S2, I2 S1, I1 S1, I2 S3, I1 S3, I1 S2, I2 S1, I2 S2, I1 S1). This ensures that each treatment combination is replicated three times for each participant. It is recommended to employ a total of 84 different images, each repeated three times (resulting in 252 observations), distributed evenly among 14 participants, with

each receiving 18 images. The time taken by each participant to correctly identify each image will be measured for each observation. Since it is anticipated that each image will take a participant only 1 or 2 minutes to identify, each session will not exceed 40 minutes per participant, reducing the likelihood of performance being affected by fatigue.

Additionally, the proposed model incorporates 2 Treatment factors **Image Type** and **Simplification Level**; block Factor **Pictures** with 14 different levels which represent 14 different common images used, block factor **Students** with 14 different participants who are recruited for this experiment. There is a possibility of having an interaction effect between the combination of **Image Type** and **Simplification Level** along with block factor **Pictures** should be included in the final model since we use 84 different common pictures in this experiment which is expected to have subject to subject variability.

$$TimeSpent_{ijkl} = \mu + ImageType_i + Simplification_j + (ImageType * Simplification)_{ij} + Pictures_k + (Pictures * ImageType * Simplification)_{ijk} + Students_l + \epsilon_{ijkl}$$

We also have the following sample ANOVA table:

<i>Sources</i>	<i>Degree of Freedom</i>
Image Type	1
Simplification Level	2
Image Type * Simplification Level	2
Pictures	13
Pictures * Image Type * Simplification Level	65
Students	13
Error	155
Total	251

3) *Power Calculation and Further Analysis plan*

In this part, I will discuss the power analysis and further plan for analysis of this experiment with the main purpose of comparisons among treatment groups. With $\alpha = 0.05$, based on the following analysis, we have that with replications of 3 times per block factor,

an appropriate difference of 3 (seconds), and assumed variance of 10 seconds, we would have the power of 93.04%. Therefore we have the evidence that we can use the proposed experiment that gives a power of 93.04% with 3 replications per block.

Obs	a	b	d	alpha	var	n	dfe	nc	fcut	beta	pow
1	6	14	3	0.05	10	2	155	12.6	2.27252	0.22506	0.77494
2	6	14	3	0.05	10	3	155	18.9	2.27252	0.06952	0.93048
3	6	14	3	0.05	10	4	155	25.2	2.27252	0.01785	0.98215
4	6	14	3	0.05	10	5	155	31.5	2.27252	0.00399	0.99601

In terms of the proposed plan for analysis, I will divide it into 2 different parts: Hypothesis Testing and Post-hoc analysis. Regarding Hypothesis Testing, based on model results, we can test the main effects of **Image Type** and **Simplification Level** to see if they have any significant evidence in terms of time taken for image identification. We also examine the interaction effect between those two main effects **Image Type** and **Simplification Level** to see whether an interaction effect is appropriate in this model. On the other hand, we can also examine the interaction effect between the two main effects and the block effect **Pictures**. If this interaction effect is not significant, we will re-run the model with the exclusion of the interaction effect. For the post-hoc analysis, we can conduct post-hoc comparisons to identify specific pairwise differences between levels of **Image Type** and **Simplification Level**. The following would serve as some proposed ideas of comparisons:

- + Difference between reference group (original image) and other simplification level (edge detection or segmentation). This type of multiple comparisons could be conducted using Dunnett's Test. We can have the following sample form for one comparison.

$(\mu_{11..} + \mu_{21..} - (\mu_{12..} + \mu_{22..}))/2$: with $\mu_{11..}$ represents the mean time taken for all observations that have Symbol form and Original Image; $\mu_{21..}$ represents the mean time taken for all observations that have Picture form and Original Image; $\mu_{12..}$ represents the mean time taken for all observations that have Symbol form and Edge Detection simplification; $\mu_{22..}$ represents the mean time taken for all observations that have Picture form and Edge Detection simplification.

- + Difference among each different combination of Treatment factors . This type of comparison could be conducted using Tukey's HSD test.

4) *Conclusion*

In summary, the experiment aims to provide insights into factors influencing blind comprehension of tactical images with the Tactical Image Enhancer. The proposed design is anticipated to yield a statistical power of at least 93.04%, involving the recruitment of 14 distinct participants and the replication of 84 different images three times, resulting in a total of 252 observations. Additional considerations regarding the experimental procedure and analysis are open to discussion.