Insignificant Effects of Coloring and Labeling on Memory Span and Serial Recall

Noora N. Saleh*1, Michael D. Huffer2

1Department of Mathematics, College of Education, University of Al-Qadisiyah, Al Di-waniyah, Iraq.
2Department of Mathematics, College of Sciences and Humanities, Ball state University, Indiana, USA.

Abstract

Human perception involves many cognitive processes, such as memory, attention, and critical thinking. An important cognitive process is memory, which is usually connected with the storing and retrieval of information. Different colors and labeling have diverse physiological effects on humans. Our investigation aimed to determine if a change in color or labeling would have a significant effect on memory span and serial recall. However, our results do not support that coloring and labeling have significant impacts on a subject’s memory.

Keywords: Memory span; Serial recall; Two-way factorial design; General linear model (GLM); Log-transformation.

Introduction

Human memory is influenced by numerous factors. This has raised some interesting and challenging questions in cognitive psychology. Two interesting aspects of the human cognitive process are those of serial recall and memory span [1]. Serial recall is the human ability to recall information in the specified order in which it was presented. Memory span is the length of time a subject can remember the order without an error.

A previous study indicated that color is believed to be the most important visual experience to human beings [2]. Color is an important aspect in the educational setting and is widely used in many marketing ploys. The human mind has used color as a crucial tool of survival to identify pleasurable or dangerous situations. Color can influence the level of attention and gives rise to emotional arousal.

*Email: noora.saleh@qu.edu.iq
which contributes to controlling processes that will later enhance memory performance [3]. The results from previous studies indicated that color and labeling could affect the human attention level. This effect on attention level could determine a subject’s ability to recall a given sequence [4]. In our experiment, we aimed to test the effects that changes in coloring and labeling of a simple game would have on a subject’s memory span and serial recall. We collected data from a group of Ball State University students, where our experimental set up included a two-way factorial design with color and labeling being our factors of interest. During our experiment, we chose to test the following hypotheses.

Null Hypotheses:
The mean lengths of the sequences are equal for the different coloring options in the game.
The mean lengths of the sequences are equal for the different labeling options in the game.
The mean lengths of the sequences are equal for the interaction between the different coloring and labeling options in the game.

Alternative Hypotheses:
The mean lengths of the sequences are not equal for the different coloring options in the game.
The mean lengths of the sequences are not equal for the different labeling options in the game.
The mean lengths of the sequences are not equal for the interaction between the different coloring and labeling options in the game.

The options of the game are discussed with more details within the methods section. Some humans have the ability to recall large lists of digits by associating a unique color to each respective digit. This association is called photism, which allows a human to have extraordinary memory recall [5]. Based on previous studies, we anticipated that there would be significant evidence from the data to support a difference in sequence length for our two factors of interest.

Materials: Using the Memorathon game, each participant was randomly assigned one of the six pre-determined treatment combinations for their experimental trial. The first factor in the treatment combination consisted of eight colored buttons (1) or eight gray buttons (2). The second factor was whether the buttons were labeled sequentially (1), randomly (2), or had no labeling (3). Each trial was set at pre-determined levels for the other three factors in the game. There were a total of eight buttons for each experimental trial and eight distinct sounds for their respective buttons, while the sequence was presented at a medium speed. The participants used a mouse on a MacBook Pro laptop computer while wearing over the ear headphones to complete their experimental trial.

Methods
Subjects: The participants in this study were recruited from the Tally Food Court in the Student Center. We approached forty-two participants for our experiment, who were present over a three-day period. The sample included a total of thirty-eight Ball State University Students, two Ball State University Staff Members, and two High School students. Eleven of the participants were males and thirty-one were females. Seven of the participants were international students. The participants had an age range from 16 to 46, where most of them were between the ages of 18 and 23. The majority of the participants majored in psychology, speech pathology, or business.

Procedure: Upon recruiting participants for our experiment, we collected information on age, sex, major, and if they were international students or not. We then instructed each participant that they would sit down, go through a practice game, and then we would change the settings for their experimental trial. The practice game would allow each participant to get accustomed to the layout and objective of the game. The practice trial consisted of no sound, three colored buttons, and sequential labeling, while the sequence was presented at a slow speed. Each practice trial did not last longer than two minutes. After the participants completed their practice game, we changed the settings and informed them to play until they failed the game. Each participant completed only one experimental trial.

Results
We were interested in whether coloring, labeling, or their interaction has an effect on a subject’s memory span and serial recall. For our two-way factorial design, our statistical model was

\[
y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha \beta)_{ij} + e_{ijk} \quad i=1,2 \quad j=1,2,3 \\
\quad k=1,2,\ldots,n.
\]
Where our response factor (sequence length) was modeled by our two factor levels, namely coloring (i = 1, 2) and labeling (j = 1, 2, 3). There were six different interaction combinations; (Colored, Sequential), (Colored, Random), (Colored, Off), (Gray, Sequential), (Gray, Random), and (Gray, Off). Analyses on our two-way factorial design were conducted using a general linear model (GLM). By initially looking at the main effects and interaction plots (Figure-1) from our data set, we anticipated that labeling and the interaction between coloring and labeling possibly have effects on the mean sequence length.

![Figure 1](image)

**Figure 1**-Main Effects and Interaction Plots of Color/Labeling for mean sequence length.

The GLM outputs for the mean sequence length across our two factors for the six treatment conditions are presented in Table-1.

**Table 1**-General Linear Model (GLM) of mean sequence length vs. Color and Label.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Type</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>fixed</td>
<td>2</td>
<td>1, 2</td>
</tr>
<tr>
<td>Label</td>
<td>fixed</td>
<td>3</td>
<td>1, 2, 3</td>
</tr>
</tbody>
</table>

Analysis of Variance for LengthSeqComp, using Adjusted SS for Tests

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Seq SS</th>
<th>Adj SS</th>
<th>Adj MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Label</td>
<td>2</td>
<td>2.29</td>
<td>2.29</td>
<td>1.14</td>
<td>0.10</td>
<td>0.902</td>
</tr>
<tr>
<td>Color*Label</td>
<td>2</td>
<td>1.71</td>
<td>1.71</td>
<td>0.86</td>
<td>0.08</td>
<td>0.925</td>
</tr>
<tr>
<td>Error</td>
<td>34</td>
<td>396.57</td>
<td>396.57</td>
<td>11.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>400.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By analyzing the output and setting our level of significance to 0.05, we were able to determine that we do not have evidence to support a significant difference in the mean sequence length for our factor of color group versus gray group (p = 1.00). The analysis also indicates not evidence of significant differences in the mean sequence length for the sequential labeling, random labeling, and no labeling factors (p = 0.902). Also, the interaction between the two factors did not produce any significant
differences in the mean sequence length values (p = .925). By analyzing the residual outputs in Figure-2, we were concerned with the normality and the constant variance assumptions of our model.

Figure 2-Residual plots for mean sequence length.

The normality plot in the top left of Figure-2 suggests that the residuals at the upper and lower ends fall below the normal fit. In the residual versus fits plot, we can see that the variance fluctuates erratically. After analyzing the individual variance plots for each factor, we decided to use a log transformation on our data.

Using the same statistical model for our original data, we obtained the GLM output for our log-transformed data, as shown in Table-2.

**Table 2-General Linear Model (GLM) of log sequence length vs. Color and Label**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Type</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>fixed</td>
<td>2</td>
<td>1, 2</td>
</tr>
<tr>
<td>Label</td>
<td>fixed</td>
<td>3</td>
<td>1, 2, 3</td>
</tr>
</tbody>
</table>

Analysis of Variance for ln(length), using Adjusted SS for Tests

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Seq SS</th>
<th>Adj SS</th>
<th>Adj MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>1</td>
<td>0.0039</td>
<td>0.0039</td>
<td>0.0039</td>
<td>0.93</td>
<td>0.00</td>
</tr>
<tr>
<td>Label</td>
<td>2</td>
<td>0.0696</td>
<td>0.0696</td>
<td>0.0348</td>
<td>0.27</td>
<td>0.025</td>
</tr>
<tr>
<td>Color*Label</td>
<td>2</td>
<td>0.0022</td>
<td>0.0022</td>
<td>0.0011</td>
<td>0.01</td>
<td>0.991</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>4.5921</td>
<td>4.5921</td>
<td>0.1276</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>4.6678</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Sq = 0.357154  R-Sq(adj) = 0.00%

By analyzing the output in Table-2 and keeping our level of significance the same, we were able to determine that we do not have evidence to support a significant difference in the log mean sequence length for our color versus gray factor (p = 0.863). The analysis also indicates that we do not have evidence from our data to support significant differences in the mean sequence length for the sequential labeling, random labeling, and no labeling factor (p = 0.763). The interaction between the two factors also did not produce significant differences in the mean sequence lengths (p = 0.991). By comparing the log-transformed data to the original data, there was an increase in the significance of our two factors of interest for the difference in the mean sequence length. Even with this increase in significance, we were unable to reject our original hypotheses. It is important to note that the adequacy of our model for our transformed data presented in Figure-3 has improved.
We found that the residuals for our log-transformed data have a better fit than the residual plots shown in Figure-2. The residuals also have a more constant variance across the fitted values. There still is a fluctuation in the variance but most of the residuals seem to fall within the horizontal bands. Using the GLM outputs in tables 1 and 2, we do not have a significant evidence to support that memory span and serial recall are affected by changes in color, labeling, or both.

**Discussion**

We originally anticipated that the coloring, labeling, and their interaction would have a larger impact on a subject’s ability to recall a given sequence. Most of the subjects completed their trial with a final sequence length of 8 units, regardless of their treatment combinations. It is important to note that the best performance in our experiment had a final sequence of 19; this trial was completed under the gray factor level and the no labeling factor level. This subject had a completed sequence twice as large as the average for the other forty-one subjects. Based on previous studies, we did not anticipate that the highest sequence length would result from a gray factor level experiment trial. The lack of effect from color or labeling on sequence length from our analyses could be attributed to the variability among subjects. Since each subject was only allowed to participate in the trial once, we could not truly capture the significance in coloring or labeling. Creating an experiment where each subject receives each of the six treatment combinations could eliminate the variability among subjects. This type of experiment would also allow for a smaller number of subjects to complete the analysis of variance. After the data collection, we perceived that the practice game was necessary since some of the subjects were not previously informed about this type of games. However, with the addition of the practice game, we perceived that some of the subjects might have already started learning it prior to the experimental trial. To deal with this learning factor, it might have been appropriate to present a picture of the game prior to the trial to explain the objective.

**References**