Project exam in Cognitive Psychology PSY1002

Autumn 2007

674107

Course responsible:
Kjellrun Englund

Stroop Effect

Dual processing causing selective attention.

674107

November 26, 2007

Abstract

This document reports on the findings of an experiment conducted to explore the Stroop effect. The experiment is based on the classical, incongruent words and colors, experiment. The findings from the experiment proved to be unreliable results not directly in compliance with the Stroop effect theory. Mostly caused by the lack of participants.

Introduction

The Stroop effect is a selective attention effect often exemplified through the reading of incongruent words and colors. Stroops theory was that selective attention would appear as easily through visual-, as well as, for audible processing. To test his hypothesis, Stroop introduced several visual experiments (MacLeod, 1991). One in which incongruent words and colors are shown, with the task to read the word or the color. This introduces a dual processing task on the participant, in which he or she must process two types of information at the same time. Causing one type to interfere with the other. The hypothesis is that words are more intuitively "built-in" - automatized, leading to a faster processing of words than of colors, resulting in faster reading of words than of the colors (Sternberg, 2007).

Method

Participants in the experiment are 6 students, all from NTNU, taking the course Cognitive Psychology. The background for the participants is; age span 22 ± 2, with technical and general education. The participants should be appropriate for the experiment and assumptions about basic fluent english knowledge and color knowledge should be valid.

Material used is the program CogLab2.0 which provides a sub program for the Stroop Effect experiment. Each participant was given/had an installation of this on their personal computers.

The experiment was held in a group room on NTNU shared by others non-participants. As well, one participant was located on another room. Unfortunately this does not provide an ideal laboratory or sterile environment for the experiment.

The actual experiment consists of a classical Stroop Effect test. Displaying words spelling a color in a font of a different or same color. The set of colors used is limited to red, green and blue. The participants task is to identify the *font color* as quickly as possible using the keyboard. This is repeated at least 45 times to ensure stability in the results. In addition the test data consists of 15 tests in which the colors of both spelling and font are the same . The results from each individual test are then saved by the participant and was distributed immediately after the experiment.

Results

As already mentioned, the system used for the experiment provided results for each individual experiment. For this experiment the measurements were reaction times on two different conditions. One in which font and spelling are the *same*, and one in which font and spelling are *different*. The scale used was milliseconds. For each individual result the measured reaction times were averaged using traditional arithmetic average. The average data from all participants were then combined. Estimations based on the data were; total average for both conditions, standard deviation using MS Excels (*stddev()*), standard error ($\frac{\sigma}{\sqrt{n}}$). Then a Student T-test was performed on the data with the implied hypothesis that the results of the two conditions were the same. The results can be found in figure 1 including the standard error. The Student T-test resulted in an 18.38% probability of the hypothesis which, according to a standard 5% confidence interval, indicates to *keep* the hypothesis.

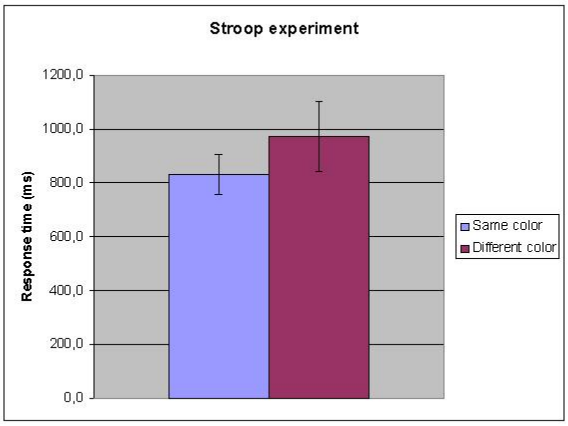

Figure 1: Calculated results from experiment.

Discussion

The hypothesis or expected results from the CogLab2.0 experiment was to have significant differences between same color and different color. With response time for same color significantly smaller than for different colors (Francis et al., 2007).

The results, however, shows that there are no significant differences between the two conditions. This is indicated by the overlapping standard error bars in figure 1 and is further supported by the Student T-test which indicates a fairly large possibility for the hypothesis (18.38%). This was not expected.

One may argue that the total average response time in itself is somewhat according to our hypothesis. And further that the population the standard errors are based on is very small (6) thus giving unreliable results. This is however not necessarily valid arguments in themselves. The fact that the population in the experiment is extremely small will give unreliable statistical analysis. Student T-test is based on a normally distributed population and one can say the standard deviation analysis also relies on Gaussian distribution. This leads to the question whether the population is normally distributed. A population of 6 more or less equalized persons is per definition not considered as a normally distributed population - thus stating the fact that the statistical analysis of our results can not be trusted in relation to scientific hypothesis. When the Stroop Effect is tested in a scientific environment, usually, the population is much bigger (MacLeod, 1991), and according to the statistical rule-of-thumb, greater than 30.

The environment in which the experiment was carried out was not ideal and could also have inflicted the results.

An intuitive explanation for the results could have been that the participants lacked knowledge on english words and/or colors. This is however rejected based on the fact that every participant is a student on a higher level and none suffered from color-blindness or alike.

As a conclusion of our results. We have found that the results from the experiment was *not* according to our hypothesis and thus not according to the theory on Stroop Effect. We can not conclude that the Stroop Effect is not valid since our population was too small to give any actual valid results.

Bibliography

Francis, G., Neath, I., and VanHorn, D. R. (2007). *CogLab2.0 on a cd*. Thomson Wadsworth.

MacLeod, C. M. (1991). Half a century of research on the stroop effect: An integrative review. *Psychological Bulletin*, 109:163–203.

Sternberg, R. J. (2007). *Cognitive Psychology*. Thomson Wadsworth.

Visual Search

Exploring object features in visual attention

674107

November 26, 2007

Abstract

This document reports the findings of an experiment conducted with the goal of exploring processing required for visual search under different conditions. The required degree of processing in relies upon the features of the target object and the distractors. Stating different combinations of these produces visual sets of objects with a target that presumably presents different complexity and requires different degree of processing. Visual attention theory states that an object with distinct features requires the least amount of processing regardless of the distractors. While for a non-distinctive object the degree of processing will depend upon the distractors.

Our findings from the analysis of our results is in direct compliance with the hypothesis. However we have found that we had too few participants resulting in an unreliable statistical analysis and thus affecting the overall findings from the experiment.

Introduction

The experiment explores attention theory in visual search. Especially focusing on feature-integration in visual attention theory (Treisman and Gelade, 1980). When we visually search for something specifically, the degree of processing required depends on several conditions and affects the complexity of searching. The most distinct and classical considerations are on the difference between object conjunction and object feature. The proposed hypothesis is that a feature object requires less attention processing to find, regardless of the numbers of distractors. When object and distractors have the same feature (conjunction) the attention processing is more complex and the numbers of distractors will affect the complexity (Treisman and Gelade, 1980).

Method

The 6 participants in the experiment are all in the age span of 22$\pm$2 years, students at NTNU and taking the course Cognitive Psychology. None of the participants have any form of color blindness.

Material used in the experiment is the software provided in the CogLab2.0 software (Francis et al., 2007). The participants used their own personal computers for the execution of the software.

The execution of the experiment was held in a group room at NTNU which is shared by both participants and none participants, unfortunately, not providing an ideal environment. Not all participants were located in the same environment and there were

CogLab2.0 provides a specific experiment for investigating visual search. The method used is to measure the response time, under different conditions, in milliseconds. The conditions used in the experiment are based on classical theory (Treisman and Gelade, 1980) and includes *feature search* and *conjunctive search*. The two different conditions are implemented using two distinctly different shapes; circles and squares, and two distinctly different colors; blue and green. Alternating the shapes and colors to produce either feature or conjunctive conditions. Feature condition is accomplished using *blue squares* as distractors and a *green circle* as target. Conjunctive condition is accomplished using *both green circles* and *blue circles* as distractors and, again, a *green circle* as target. The number of distractors are alternated between; *4, 6* and *64*, during the experiment. The experiment alternated between target present and not in both conditions serving two purposes - prevent "cheating" and have more comparable datasets.

Results

The results given by the software used in the experiment is per participant. The data consists of averages of the response time measured on the different conditions. This gives us essentially four datasets for each participant; conjunctive absent/present and feature absent/present. Averaging the four data sets, using standard arithmetical average, gives us the data represented in figure 1. Further on statistical analysis was performed on the data, standard deviation, standard error and two different Student T-tests. The Student T-tests were performed with the intention of providing statistical evidence of the results. The first hypothesis was if feature present was the same as conjunctive present. Tested on each of the three "distractor-sets". The result of this test was a probability for the hypothesis. The probability calculated in each case was far less than 5% which suggest to discard the hypothesis. The other hypothesis tested was on feature present with different number of distractors, 4 and 64, to see if the data was the same. The result of this was a probability of 21.6% which indicates to keep the hypothesis.

Discussion

The results from the experiments are in compliance with the hypothesis stated earlier. From the hypothesis and the theory on feature-integration (Treisman and Gelade, 1980), we could expect to have high response time on conjunctive search with an increase related to increased number of distractors. We also expect a higher response time when the target is absent. Rational is that one can stop when the

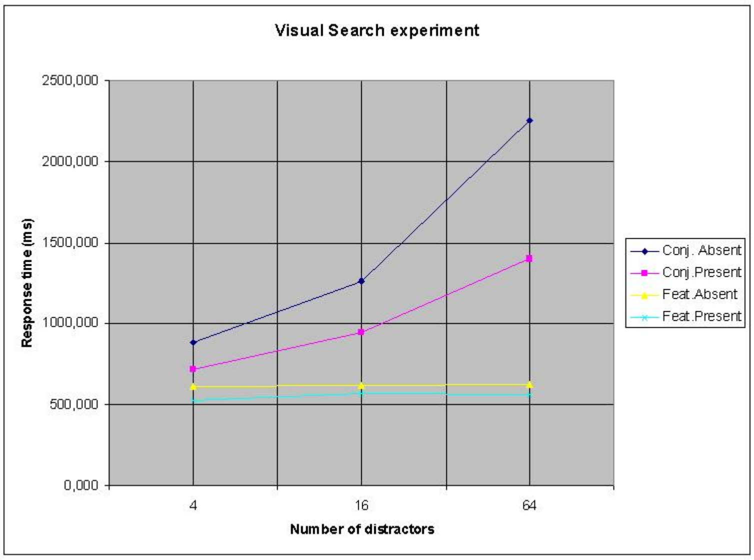

Figure 1: Calculated results from experiment.

target is found, usually, that is occurring before the whole space is searched. From the results we see this connection clearly on figure 1 with the blue (diamond) and pink (square) graphs.

For feature search we expect, from the hypothesis, that the response time should be unrelated with the number of distractors and be more or less the same. The results supports the hypothesis and the theory (Treisman and Gelade, 1980) well. This is represented in figure 1, yellow (triangle) and light-blue (cross) graphs. Which both are unaffected by the increased number of distractors.

The results could be a coincidence, since the fact that the population is extremely small will affect the viability of the statistical analysis. Since this experiment is more a proof of concept rather than scientific research, we conclude that the experiment and hence the results are correct and that they support the overall theory and hypothesis (Treisman and Gelade, 1980), (Sternberg, 2007).

Bibliography

Francis, G., Neath, I., and VanHorn, D. R. (2007). *CogLab2.0 on a cd*. Thomson Wadsworth.

Sternberg, R. J. (2007). *Cognitive Psychology*. Thomson Wadsworth.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136.

Memory Span

Exploring limitations of short term memory

674107

November 26, 2007

Abstract

These are the findings of an experiment conducted in the search for limitations of the short term memory. The basis for the experiment is that the short term memory holds approximately seven items. In the exploration of this hypothesis we use different approaches for the items, each with it's own expected process-complexity. Using these different items in a classical approach to the experiment enables the analysis of limitations of the short term memory. The findings in the report are in compliance with the expected results based on the theory of memory span. However we found the datasets to be insufficient in terms of reliability for the statistical analysis used in the report.

Introduction

The experiment explores the capacity of the short term memory. In the field of psychology there are different views on whether the short term memory actually exists, is part of the long term memory or is in collaboration with the long term memory (Sternberg, 2007). The most widely accepted theory is that there is a short term memory or working memory. The report will base itself on this theory. Short term memory is able to hold a range of items for a short period of time. The range is variable, subjective and can be extended through training. A common theory, based on empirical investigations, state that the range of items is expected to be 7 ± 2 items (Miller, 1956).

In our investigations we use different approaches to measure the limitations of the memory span. The conditions used in the experiment consists of numbers, letters and words with different phonology. The hypothesis is that digits, distinct letters and short words will provide less distracting factors than similar letters and long words. Thus enabling a larger memory span and culminating to approximately seven items, with decreasing number of items for the more process-demanding items. This approach to measurement is more or less an adaptation of the classical experiment used by Hayes (Hayes, 1952).

Method

Participants involved in the execution of the experiment were all in the age span of 22 ± 2 years, students at NTNU taking the course Cognitive Psychology and assumed fluent basic english reading skills. In this experiment it was used 6 participants.

Materials used in the experiment was essentially a computer software provided in CogLab2.0 (Greg Francis and VanHorn, 2007). The software is specifically made for this experiment. Participants provided their personal computers for running the software.

Environment for the experiment was an office at NTNU shared by participants and none participants - not providing an ideal laboratory setting.

The procedure for the experiment is to measure the maximum numbers of items recalled from a list displayed for a short period of time - 1 second. The lists are grouped into five categories, or conditions; numbers, letters different sound, letters similar sound, short words and long words. Each list has a maximum of ten items. The number of items in the list increases for each correct recall attempt performed by the participant - resulting in a performance measure for the short term memory span.

Each participant was solely responsible in conducting the experiment and distributing the results produced by the CogLab2.0 software.

Results

Each participant provided their raw data from the CogLab2.0 software after completing the experiment. The data collected was essentially the *final list length* for each of the five conditions.

Statistical analysis of the data included; standard arithmetical average on each condition, standard deviation, standard error ($\frac{\sigma}{\sqrt{n}}$) and Student T-test. In addition, a grouping and averaging of the results, in compliance with the hypothesis was performed. One group consisted of; Numbers, letters different and letters similar. The other group consisted of; short words and long words. The results are presented in figure 1.

Analysis of the two groups resulted in averages of list lengths for the two groups. For the number and letter group, the average length was 6. The word group resulted in an average length of 4.

The Student T-test was performed on the two groups, already mentioned. The hypothesis in question was whether the two groups were statistically the same. Using a 95% confidence interval for evaluating the result. Performing the Student T-test resulted in a 0.8% probability for the hypothesis. This clearly states that the two groups are not the same.

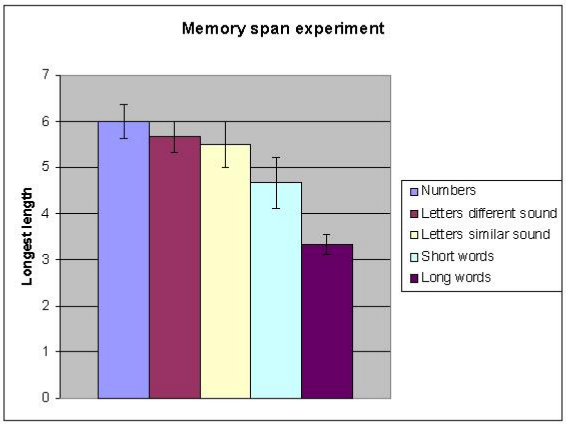

Figure 1: Calculated results from experiment.

Discussion

The results produced by the experiment are in compliance with the hypothesis stated earlier. We had expected to find that numbers would be easier to remember and long words to be the "hardest" to remember - setting different limitations on the short term memory span. Figure 1 presents this descending memory span according to the complexity of each condition.

We notice that the maximum number of items held in the short term memory is 6 and the lowest at around 3. This is not in compliance with the hypothesis on 7 ± 2 items (Miller, 1956). This can be the result of the increased complexity the long words provide, resulting in higher required processing. However, the results are based on a very limited populations, including only six participants. The fact that this will influence the statistical analysis can not be avoided.

To conclude. In our limited experiment we have found that the results were in compliance with the stated hypothesis (Hayes, 1952), but not proving the theory on seven items in the short term memory (Miller, 1956). Taking the limited datasets used into account we can not validly conclude anything from our results.

Bibliography

Greg Francis, I. N. and VanHorn, D. R. (2007). *CogLab2.0 on a cd*. Thomson Wadsworth.

Hayes, J. (1952). Memory span for several vocabularies as a function of vocabulary size. *Quarterly Progress Report*.

Miller, G. A. (1956). The magical number seven, plur or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63:81–97.

Sternberg, R. J. (2007). *Cognitive Psychology*. Thomson Wadsworth.

Serial Position

Free recall of items from a sequence

674107

November 26, 2007

Abstract

This document reports on the findings from an experiment conducted in search of compliance with the serial positioning effect. The serial positioning effect states that one remembers the last and first few items in a sequence better than the middle items upon recall. In the experiment we used a sequence of random letters shown in a limited amount of time and free recall as the recall constraint. Our findings from the experiment were not consistent with the hypothesis. In addition the datasets produced by the experiment was found to be insufficient with respect to the statistical theory used for analysis.

Introduction

The theory on serial position is essentially an hypothesis which states that an items position in a list is crucial for its ability to be remembered. Given a list of items to remember, the items in the middle are likely to be forgotten. The items at the end of the list are most likely to be remembered referred to as the *recency effect*. Items at the start of the list are second most likely to be remembered referred to as the *primacy effect*. Based on a free recall of the list. (Murdock, 1962).

Method

To explore our hypothesis we conducted an experiment. The experiment consisted of 6 participants, all in the age span of 22 ± 2 years, students at NTNU and taking the course Cognitive Psychology. None of the participants have any form of dyslexia or similar.

CogLab2.0 (Francis et al., 2007) provided software for the execution and data collection for the experiment. Participants used their personal computers for execution of the experiment. The environment in which the experiment was performed was an office at NTNU. The office was a shared space with both participants and non-participants. This did not provide an ideal laboratory-like environment.

The experiment consists of measuring the correctness of an item in a list and the associated position for that item. The list consists of 10 random letters which was each displayed one letter at a time for 1 second. After been displayed the list, the participant should then freely recall letters that he or she remembered. The position in the list were irrelevant for the recall of the letter. This was repeated 15 times, using different lists.

When the experiment was performed, each participant was responsible for saving and distributing the results produced by CogLab2.0.

Results

Each participant provided results for their experiment. This consisted essentially of the percentage of correctness related to the position in the list.

Statistical analysis that were performed on the data set included; standard arithmetical average, standard deviation using MS Excels built-in function (std-dev()), standard error ($\frac{\sigma}{\sqrt{n}}$) and Student T-tests. The average calculations is visualized in figure 1 with linear lines drawn between the data to better illustrate the trend. Additionally the standard error for each of the discrete positions are visualized with ± error bars.

The Student T-tests consisted of three tests in compliance with the hypothesis (Murdock, 1962). The hypothesis states three distinct groups in a list; first few items, middle items and last few items. Intuitively the list of 10 items were divided into three groups with the middle group having 4 items. The Student T-tests tested whether the data in the; first few and the middle group were the same, last few and middle were the same and first few and last few were the same. This resulted in a probability of; 0.91%, 35.24% and 4.22% respectively. Using a 95% confidence interval results in discarding the first and last hypothesis while not discarding the second.

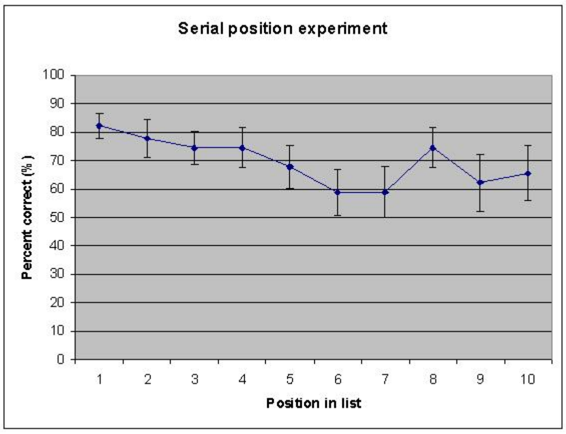

Figure 1: Calculated results from experiment.

Discussion

From the result and the statistical analysis we can explore the relationship towards the earlier stated hypothesis and the overall theory on serial position and especially the primacy and recency effects (Sternberg, 2007).

In figure 1 we see that the first items are remembered good, with an decreasing degree when nearing the middle of the list. This is somewhat in compliance with the theory on the primacy effect. The occurance of this is further supported by the first Student T-test which results in a very low probability that the first few and middle items are processed equally.

On the other hand, for the last few items we see that they are more or less remembered equally as the middle items. This is not expected according to our hypothesis and is opposite of what the recency effect suggests. The phenomenon is further supported by the Student T-test which clearly indicates that the last items and middle items are significantly processed equally.

We also checked whether the first few and last few items were processed equally. In the last Student T-test we get a probability which is fairly close to the limitations of our confidence interval, but still below. This is supporting the theory that the primacy and recency effects are different. In light of the earlier analysis of the results, however, we can not state anything concrete on this test.

The experiment was conducted with a population consisting of far to few participants (6) to be recognized as statistical reliable. Further, the trials were repeated only 15 times for each participant. Statistical analysis used on the results relies upon Gaussian distributed data, which this clearly is not. This culminates into the conclusion that we can not rely on the analysis of the results. In addition to this, it can occur confusion on what the term "few" actually is. In our analysis we have used mathematical segregation of the positions. One can argue that few is not $\frac{1}{3}$ of the items. Taken this into account would have altered slightly on our analysis, nevertheless, the conclusion would have been the same.

Bibliography

Francis, G., Neath, I., and VanHorn, D. R. (2007). *CogLab2.0 on a cd.* Thomson Wadsworth.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64:482–488.

Sternberg, R. J. (2007). *Cognitive Psychology.* Thomson Wadsworth.

Word Superiority

Context dependant detection of letters

674107

November 26, 2007

Abstract

The word superiority effect is examined using classical approaches for investigations. Testing how well participants detect letters presented under different context, enables for investigation of the word superiority theory (Reicher, 1969). Analysis of results produced by the experiment leads to a non-decisive conclusion, discussing the reliability of the datasets and the selected population.

Introduction

The human perception of letters is assumed to have be somewhat context dependent. Word superiority theory states that letters are easier recognized when presented in the context of a known word than when presented in isolation. (Reicher, 1969). The superiority of letters in context is not intuitive since the information level is intuitively larger for a word than for a single letter. However the effect is empirically investigated and found to match the theory. In the conducted experiment we revisit the classical experiment of word superiority for investigation of the theory.

Method

The experiment used 6 participants, all students from NTNU with no known reading problems and with normal vision. In addition it is assumed that the participants have good English reading skills.

Materials used in the experiment is essentially software provided in the CogLab2.0 package (Francis et al., 2007) which is specifically made for the experiment.

The execution of the experiment was conducted in an office, provided by NTNU, shared by both participants and non-participants. Not providing an ideal, laboratory-like, environment.

The experiment consists of displaying either a word or an isolated letter very briefly ($40ms$). Then a marker indicating which position the subject is to report on is displayed. Then, two choices are presented to the subject. The letters presented are on the form so that they both will produce a word if the displayed item is a word. Making it somewhat more difficult for the participant to rationalize his/her decision.

The trials are repeated 96 times with a combination of 15 four letter words and 16 isolated letters. Assembly providing sufficient stability for the results.

When participants finish the experiment CogLab2.0 provides results of the trials which the participant distributes.

Results

The datasets obtained by each participant in the experiment were averages of the percent of correct detections in the two conditions; Letter in a word and letter in isolation.

In order to get the overall result for the experiment several mathematical operations were performed on the datasets. Including; standard arithmetic average, standard deviation, standard error ($\frac{\sigma}{\sqrt{n}}$) and Student T-test. Figure 1 presents the arithmetical average for the two conditions and the associated standard error derived by the standard deviation.

A Student T-test was performed on the dataset with the aim of discovering similarities between the two conditions. The stated hypothesis which was tested upon is whether results for letter in word and letter in isolation were similar. The Student T-test resulted in a 12.91% probability in favour of the hypothesis. Using a 95% confidence interval this leads to keep the hypothesis implying that the results from each condition is similar.

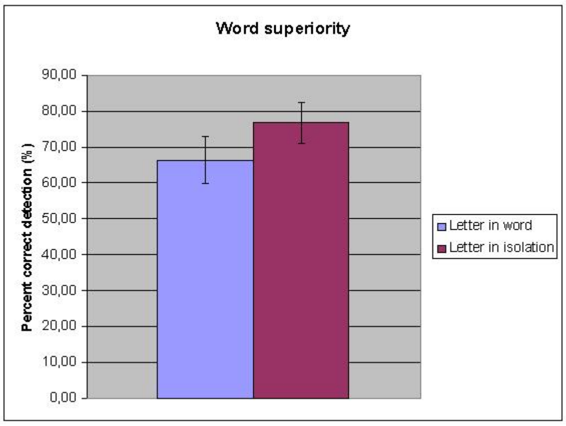

Figure 1: Calculated results from experiment.

Discussion

The results obtained from the experiment is clearly not in compliance with the stated hypothesis mentioned earlier. From figure 1 we see that the participants, in general, detected letters in isolation better than letters presented in the context of a lexical word. This is the opposite of what is stated in the word superiority effect (Reicher, 1969), and is hence, not expected. In addition, we see from the performed Student T-test that the two conditions in the datasets have a fairly high probability of being similar. This is also not expected and directly incongruent with the hypothesis, which states that the two conditions should be significantly different.

We must consider the possibility that participant truly did not see nor perceive the letter or the word when it is displayed. This would result in pure guessing for the targeted letter. From a purely statistical point of view this would have resulted in 50% detection regardless of the condition. From our results in figure 1 we see that the detection is fairly higher than 50% even when the standard error is taken into account. This leads to the conclusion that participants actually perceived the words and letters displayed.

Some research indicates that the use of visual masks can bias the detection of the letter. (Jordan and de Bruijn, 1993). Considering the fact that the experiment used off-the-shelf materials based on well known approaches to the experiment (Francis et al., 2007), will provide a somewhat solid base for the experiment. We will not go further than to mention this bias, considering the limited scope of this report.

Most notably for the results are the lack of participants which can not be considered representative for the overall population which the theory is based on. Using only 6 participants in the experiment will have impact on the reliability of the performed statistical analysis. Considering that some of the statistical analysis base themselves upon Gaussian distributed datasets.

To conclude we can not derive any reliable conclusions from this experiment. Mostly caused by the unreliable datasets and lack of representable population.

Bibliography

Francis, G., Neath, I., and VanHorn, D. R. (2007). *CogLab2.0 on a cd*. Thomson Wadsworth.

Jordan, T. R. and de Bruijn, O. (1993). Word superiority over isolated letters: The neglected role of flanking mask contours. *Journal of Experimental Psychology*, 19(3):549–563.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimuls material. *Journal of Experimental Psychology*, 81:275–280.