University of Tsukuba
School of Informatics
College of Media, Arts, Science and Technology
Undergraduate Thesis

Voluntariness based separation of attentional mechanisms in a visual search task: a possible indication for a new paradigm

Boger Matheus

Adviser: Sakai Ko

January, 2021

Abstract

Attentional mechanisms are of primary concern in psychological inquiry, and recent research [16] has given rise to new questions on what modulates the so-called "top-down" attention. In the present study, voluntariness, controlled as the presence or absence of freedom of choice, is investigated as a possible modulator of visual attention. An experiment was conducted where participants were shown four colored square cues – either all of the same color or all of different colors – and were instructed to pick a color to pay attention to in the following visual search task. In the case of all squares having the same color, no choice was possible and the trial was considered an involuntary attention trial; the other case was considered a voluntary attention one. It was predicted that voluntariness would act as an attentional effect of reward [4] or in a similar way, improving performance on voluntary attention trials. This hypothesis was confirmed to be true in regards to reaction time: voluntary trials were significantly faster by 16 ms compared to involuntary ones; but showed conflicting results in regards to hit rate: voluntary trials had a significant 9.8% decrease in accuracy compared to involuntary ones. This begs the question if this voluntariness-based attentional effect reflects a trade-off relationship in the same attentional process or is due to different voluntary and involuntary mechanisms of attention.

Contents

Figures

Tables

Chapter 1 - Introduction

1.1. Attentional studies background

Attention is commonly defined in psychology as the process by which humans, or more specifically, human brains, decide in what order to process information. Nobre has defined it as "the prioritization of processing information that is relevant to current task goals" [12]. This means that attentional processes are one of the most fundamental determiners of human behavior, and that is why research on attention is pivotal. Generally, attention is studied in relation to physical stimuli and sensory systems, and due to its ease of experimental manipulation and heavy dependence upon it by human behavior, vision is the typically preferred modality for such studies.

In order to determine which factors, be they properties of the external world or personal internal states, modulate the attentional process, researchers organize psychological experiments. Of these, one of the most common methodologies is the visual search task. In visual search tasks, participants are shown stimuli in a screen and asked to react according to the experiment's settings and the stimuli they are shown. In this way, researchers of the past were able to sketch out some properties of attention, such as its capture to stimuli that "pop-out" or its orienting separate from direct line of sight. One special paradigm of understanding of attention is the exogenous-endogenous dichotomy, also labeled as top-down and bottom-up attentional processes.

Bottom-up (or exogenous) attention is the name given to attentional processes mainly elicited by factors external to the observer, such as stimuli that stand out because of their color or shape. As such, bottom-up attention is completely involuntary and deeply embedded in basic properties of our sensorial systems and their preferences. An example of bottom-up attention at work in daily life would be the ease which one experiences when looking for red peppers amid green produce in supermarket shelves. On the other hand, top-down (or endogenous) attention is the name given to attentional processes mainly elicited by factors internal to the observer, such as the instructions one has in mind for a task or the interpretation of an arrow pointing to the right as a signal for orienting one's attention to the right. As such, top-down attention can be interpreted as an umbrella term for many voluntary and involuntary processes that occur from within, but this voluntariness-based distinction has not been in the traditional focus of research, which tended to compare top-down to bottom-up without any procedural consideration to the further breaking down of top-down attention. The most classical example is the exogenous condition on a Posner task, which relies on the interpretation of an arrow as an order for the orienting of attention. Is this interpretation voluntary or involuntary? It could be argued both ways. Nevertheless, an example of top-down attention at work in daily life would be the concentration one must maintain to find a book in a library shelf just by its numbered label.

Precisely because of this generality of the term "top-down", some researchers have pointed out that this dichotomy should be avoided, and a new way of thinking attention classification introduced [1]. Indeed, recent lines of inquiry are painting a new, finer picture of the constituent processes of "top-down" attention. Representative is the work of Theeuwes and Failing [16]. They have shown that what they call selection history plays a major role in attentional processes. This selection history does not necessarily correspond to the active engagement to task-related goals, sometimes even conflicting with them. On their terms, "the underlying notion of what is meant by selection history is that through (explicit or implicit) learning (i.e., processes shaped by past episodes of attentional selection), particular stimuli may receive 'value' that affects future selection episodes" [16]. They go on further dividing this selection history mechanism into three main categories: priming effects, statistical regularities effects, and reward effects. This last category is of special interest to the present thesis.

Theeuwes and Failing cite the work of Della Libera and Chelazzi [4] as one of the pioneering studies on the effect of reward on attentional selection. They set up an experiment where participants received a high or low monetary reward on every correct trial depending on what stimuli appeared as a target or distractor during a training phase, and found that on a later test phase this influenced the participant's performance; stimuli associated to high reward when targets increased participant's reaction time when they later appeared as distractors.

Combined with the evidence from Leotti and Delgado [9], whose work suggest that choice and reward share brain areas they activate, which might indicate an inherent reward feeling associated with freedom of choice, the present study brings focus to the relations between choice, reward and attention. Since the breaking down of top-down attention in task-oriented voluntary processes and historically-influenced involuntary ones, much focus was given on the specifics of each kind of process, but their core differences were not directly contrasted, nor was the way in which they may interact clarified. In other words, there is a gap in the literature on the effects that the presence or absence of voluntariness itself in a task may induce, and whether such effects behave in the same way reward-based effects do, due to the neurological similarity expected from choice (the utmost representation of voluntariness) and reward.

1.2. The present study

The present study consists of an experimental attempt to differentiate attentional effects elicited by the presence or absence of voluntariness. It follows from the logic that voluntariness can be made present by the inclusion of free choice to an experimental paradigm, and that its effects should be similar to reward ones. It is hypothesized that stimuli to which a participant wholly deliberately choose to pay attention are accrued with value in a manner similar to what happens in selection history effects, and should elicit a better task performance when they are task relevant. On the other hand, stimuli that are task relevant but not valuable from a voluntariness-based reward point of view are not expected to improve performance, providing a basal state for comparison. In other words, the main hypothesis of this study is that solely voluntariness-based attentional effects exist and can be observed with the presence or absence of a task relevant free choice.

The other question to pursue is the difficult visual recognition influence hypothesis. This hypothesis states that voluntariness-based attentional effects are exacerbated under situations where visual recognition is especially difficult, such as when stimuli are dimly lit or present a low contrast to the background. This second hypothesis comes from the results of a previous experiment by Suizu (unpublished), who belonged to the same research team as the author. Suizu's experiment was fundamentally the same as the one in the present study, since this study was built upon his preliminary findings. Some adjustments were made to guarantee a better controlled environment, however, and the specifics of what was changed can be found in Section 2.2. Suizu's results failed to prove a general attentional effect based on voluntariness for all tested conditions, but did find a statistically significant difference when the scope of the analysis was reduced to the darkest condition tested. This led to the hypothesis and instigated further investigation. Also, he did not control for contrast between stimuli and background luminance, which could be a confounding factor. Thus, the present study is a follow-up on Suizu's, and aims to go further the path he opened by improving the methodology used, which shall be discussed next.

Chapter 2 – Procedure

2.1. The present experiment

To test the aforementioned hypothesis of voluntariness driven attention mechanisms, an experiment was conducted where the results of trials based on either voluntary or involuntary attention could be compared. In addition, for testing of the difficult visual recognition influence hypothesis, trials were also separated according to two stimuli to background luminance contrast conditions (low and high; Michelson contrast of 0.33 and 0.6, respectively), and three stimuli luminance conditions (low, medium and high; mean physical luminance in cd/m^2 of 10.46, 19.99 and 29.41, respectively). The experiment in question was conducted two times; one time for exploratory purposes as a means to set task relevant criteria for the suitability of participants ($N = 7$, five male participants, details in Appendix), and another time for the de facto testing of the hypothesis ($N = 6$, all male participants). It was run on a computer using PsychoPy software [13][14] for stimulus presentation and reaction time recording, and took place in a dark room, where the participants were sat 57cm away from the center of the display screen, whose characteristics are summarized in Table 2.1. Participants responded to the stimuli using a regular keyboard. In both the exploratory and confirmatory runs of the experiment, all participants had normal or corrected to normal visual acuity, and one participant in the first run had self-declared a color vision disability and did not participate in the second run of the experiment. All participants belong to the School of Informatics or the Graduate School of Science and Technology of the University of Tsukuba and formally accepted to take part of the psychological experiment according to the Research Ethics Committee of the Faculty of Engineering, Information, and Systems at the same university. Participants who did not belong to the author's research team were monetarily rewarded with ¥900 per hour of collaboration in the form of Amazon Gift Cards. The experiment was conducted in Japanese and all participants were either native speakers of the language or had an intermediate to advanced proficiency in it.

TABLE 2.1. Specifications of the monitor utilized in the present experiment.

Property	**Value**
Monitor Type	EIZO ColorEdgeCG242W
Gamma	1
Size	24.1''
Resolution	1920x1200 pixels
Refresh Rate	59Hz
Maximum Luminance	160 cd/m2
White Point	D65 (x = 0.3127, y = 0.3290)

The experiment consisted of a training session and two testing sessions, with the training session being conducted two days before the first testing session, and the second testing session being conducted from two days to a week after the first. In the case of the first run of the experiment, all participants were trained in five minutes blocks of trials containing all possible conditions prepared for the experiment in a random order until they got at least 80% of trials correct for both voluntary and involuntary attention conditions or an hour passed since the beginning of training. As for the second run, participants were trained with the same restrictions as the first with the addition of a reaction time criterion: the mean of their reaction times of each trial in the block added to two times their standard deviation had to be less than 1.3s ($\mu + 2.5\sigma < 1.3$s). If by the end of an hour the participant did not achieve such criteria, he would not be allowed to advance to the testing phase of the experiment and his data would not be analyzed. There were four such participants.

The testing sessions were in principle equal in content, consisting of 12 blocks with 40 trials each and a one-minute break between every block, with the exception of the 6^{th} to 7^{th} block at the middle of the experiment where the break was prolonged to five minutes in the exploratory run and ten minutes in the confirmatory. The only difference between each session was that the first had its first six blocks' trial conditions set to have a low contrast between stimuli and background luminance with the remaining six blocks having a high contrast, while the second session had the exact reverse order. The luminance condition was kept constant during each block and was selected at random at its beginning. Inside the block, trials varied according to voluntariness and the presence or not of a target in the visual search task, such that a quarter of trials would be of voluntary condition with a target, a quarter of voluntary condition without a target, a quarter of involuntary condition with a target and a quarter of involuntary condition without a target. Furthermore, the involuntary condition with a target trials were controlled to half of them having the target match the designated color, and half having the target don't match it. Each combination of luminance and contrast conditions had two blocks, with the only difference being which color set was used for the stimuli, thus each luminance level had two color sets for control of a possible random effect due to color distribution. Each color set contained six colors regularly distributed in the CIELAB color space, and their specific values are available in the Appendix along with physical luminance measurements and background luminance levels for each contrast condition.

The trials can be separated into the task itself and a feedback part, and their overall flow is shown in Figure 2.1. The task began with the presentation of a fixation cross (0.5°x0.5° in visual angle) at the center of the screen for 500 ms, followed by the presentation of a cue for 1500 ms. The cue determined the attention condition of the trial, consisting of four squares of the same color on an involuntary trial (Figure 2.2b) and four squares of different colors on a voluntary one (Figure 2.2a). These squares were placed equidistantly with their centers on a circle of radius 8.49° from the center, and were of size 3.8°x3.8°. Their colors were chosen at random from the block's color set. After a 500 ms blank, a search array was presented for 100 ms, being made of eight colored letters of height 3.8°, disposed equidistantly on a circle of radius 12.5° from the center (Figure 2.2c). The letters were set to appear in pairs of different colors of four distinct letters, in a way that the used colors on a voluntary trial would be the four same of the cue squares and on a voluntary trial they include the squares' color along with three more randomly chosen ones from the block's color set. The target was set to be the letter H, with distractors being either U, E, P or S. In a target present trial, the other three letters to make up the array were randomly chosen, while in a target absent trial all four distractors were used. The task consisted in the participant choosing a color from the available ones in the cue and answering with the keyboard if there was a target (an H) of the same color that the participant chose in the search array, pressing "1" when there was, and "0" when there wasn't, paying respect to being as accurate and fast as possible. In this way, involuntary trials were deemed as so because since all squares had the same color, there would be no way for the participant to freely choose which color to look out for; thus, no voluntariness. On the other hand, on voluntary trials participants could freely choose one color out of four to pay attention to. This means that in this experiment freedom of choice was thought to be a proxy variable to voluntariness, and could be allowed or restrained for the creation of two voluntary and involuntary attention conditions.

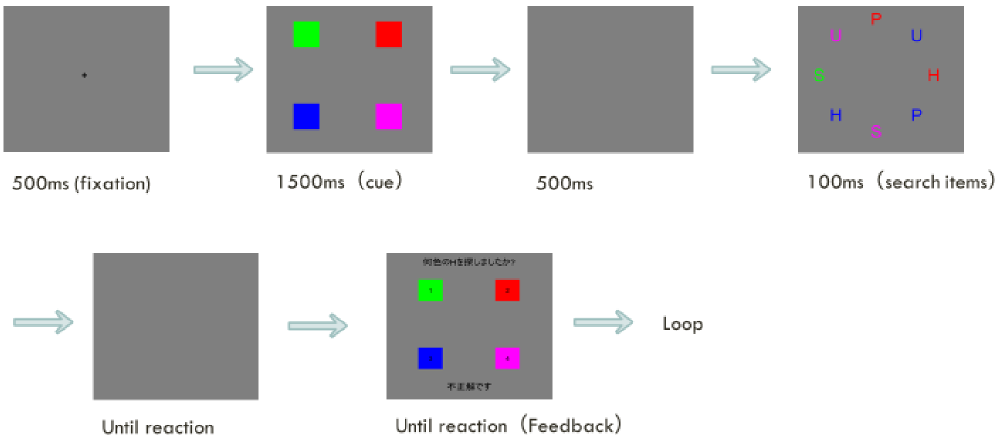

FIGURE 2.1. Schematic flow of the present experiment (colors in this figure do not correspond to the ones actually shown)

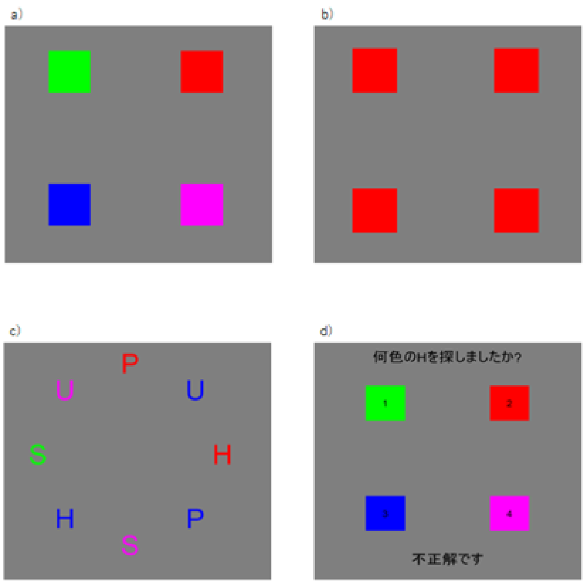

FIGURE 2.2. Example of cue, search array and feedback screens (colors in this figure do not correspond to the ones actually shown). a) Cue displayed on voluntary attention condition trials. b) Cue displayed on involuntary attention condition trials. c) Search array with target present. d) Feedback screen of an involuntary attention condition trial. The top reads “The H you were looking for was of what color?”, and the bottom reads

“Incorrect answer”, pointing that the participant incorrectly answered the last trial. There would be no such line at the bottom in a voluntary attention condition trial.

Once the search array was fully presented and then disappeared from the screen, the screen would be set to blank until the participant reacted and this reaction time was measured. Subsequently after, the feedback screen was shown (Figure 2.2d). In the top part of the feedback screen the question “The H you were looking for was of what color?” was presented, and right below it the square cues were shown in the same position as before, labeled with the numbers one to four, which corresponded to the keys “1” through “4” in the keyboard used to answer to the question. This answer was later utilized in the analysis of voluntary trials, for checking whether the participant successfully reacted to the presence or absence of the target during the task, under the assumption that the feedback was 100% correct. This assumption relies on the fact that this question’s reaction time was not measured and the participants were instructed to take their time and value accuracy over speed. This question is needed because there is no way to know which color the participant chose during the presentation of the cues if not by directly asking them. In the case of an involuntary trial, three cues were randomly chosen to be set to a different color than the original, and a feedback on the trial announced whether the participant correctly answered to the original task of finding the correspondingly colored target. Since it is possible to know which color the participant was looking for in involuntary trials, the feedback’s answer was not needed and was consequently ignored. However, the chance of each participant correctly answering the feedback question in involuntary trials was calculated and all participants had a rate higher than 95%, which signalizes the safety of assuming that they correctly answered to the feedback question in voluntary trials as well.

In conclusion, the task was to answer whether or not there was an H of the same color that the participant chose during the cue presentation, with those cues changing according to the attention condition of the trial. They could be either same-colored squares for involuntary trials or differently colored squares for voluntary trials. In addition, the background luminance and stimuli color set varied according to the luminance and contrast conditions assigned to each block, where there were three luminance levels (low, medium and high) and two contrast levels (low and high). In the end of each trial there was a feedback question in order to ascertain the hit rate of voluntary attention trials.

2.2. Differences from Suizu's experiment

As mentioned in the final parts of the first chapter, the present experiment was conducted as a follow-up of Suizu from the previous year. However, some alterations were made, either for correction or further clarification, and they will now be listed.

Cue stimuli. The present experiment shows four squares as cue for both voluntary and involuntary attention trials. In Suizu's, voluntary trials had four squares, while involuntary trials had only one square, in the center of the screen, of the same size as the ones in here.

Color values. The values for the current experiment are available in the Appendix, and they all use Adobe RGB to be displayed, and have an L value of 27, 40 or 49 (depending on luminance condition) and a C value of 30 in the L*C*h color space. On the other hand, Suizu's colors have an L value of 60, 70 or 80, each L with a corresponding C value of 35, 40 or 35, respectively, and were displayed in an sRGB environment. All C values in this paragraph are given in degrees, not radians.

Luminance levels. The current experiment has three stimuli luminance levels, each with a mean luminance in cd/m^2 of 10.46, 19.99 and 29.41, respectively. Suizu's experiment also had three luminance levels. However, they had mean luminance in cd/m^2 of 22.33, 32.20 and 44.31, respectively. In this way, the current experiment was set such that the current medium luminance condition roughly corresponds to Suizu's low luminance condition, allowing for an even darker low luminance condition in the current experiment. This allowed for a stricter test of the aforementioned difficult visual recognition influence hypothesis, where it was expected darker luminance values would show a higher voluntariness based attentional effect.

Outlier criteria. It has not been discussed yet in this thesis what was the outlier criteria used in the reaction time analysis, but this is another point of divergence between the present thesis and Suizu's. In the analyses within conducted, a reaction time of a certain trial was deemed to be an outlier and thus excluded from further analysis if it was less than 200ms long or greater than two and a half times the standard deviation of the reaction times of all trials of a participant added to their mean (mathematically, $RT > \mu + 2.5\sigma$). Suizu's criterion, however, was only that the reaction time had to be smaller than one second.

Contrast. In the present experiment, the contrast was kept constant inside a block and it had two different conditions: low (Michelson contrast of 0.33) or high (Michelson contrast of 0.6). The change in contrast was achieved via change in the luminance of background gray color. Suizu did not pay special attention to contrast and did not change the background color to match each stimulus set luminance. He also did not report background luminance in his experiment, but from the monitor specifications given, one can expect it to be around 40 cd/m^2. Therefore, the contrast in Suizu's experiment probably fluctuated from 0.28, through 0.11 to 0.05 as the luminance of the stimuli rose.

Feedback. The feedback question did not change regardless of the attention condition of the trial in the current experiment. In Suizu's, voluntary attention trials had the same feedback question as here, while involuntary attention trials asked which color was the other H, since they come in pairs. The feedback question was made the same in both types of trials for two main reasons. The first is that in involuntary trials without a target there was no way to ask which color the other H was, since there was no H to begin with. Secondly, asking which color was the other H could induce participants to slot attentional resources to that subjacent task, potentially debilitating their capacities at the main task. This could lead to systematic alterations on the results of the experiment, and was thus abolished.

ANOVA factors. One last difference between the present experiment and Suizu's is the focus given or not to color. Here, color was deemed to be sufficiently controlled and not theoretically important enough to be established as a factor of analysis (as will be shown in the next chapter). Suizu thought differently and included color as a factor in his analyses. In his work, some colors did show statistical significance in post hoc tests on a significant color and luminance interaction, but no interpretation was given to why that happened.

Chapter 3 – Results

3.1. Outlier criteria

It is common in reaction time studies to first determine if there are any outliers in the reaction time data and to properly remove them from analyses. The definition of an outlier varies from study to study, generally according only to personal preference, with no single method being held as the orthodoxically correct one [19]. In the present study, two criteria were used for outlier determination. The first consisted of considering all reaction time data less than 200ms long to be an outlier, since around 200ms is the historically recorded average of how fast a human can press a button in reaction to a clearly perceptible stimulus. Thus, any value smaller than 200ms was considered to be a recording error and was discarded. The second criterion is an intra-subject one, varying according to how fast each participant responded: any value greater than two and half times the standard deviation of a participant's reaction time data added to its mean (mathematically, $RT > \mu + 2.5\sigma$) was considered exceptionally slow for that participant and probably an indication of some lapse of attention rather than being a regular trial. These criteria were chosen for withholding proper accountability of both human and machine error and because they seemed appropriate when plotted against the full range of the raw reaction time data for each participant. Figure 3.1 shows the distribution of the raw reaction times across all participants alongside the criteria (Figure 3.1a), the reaction time distribution of all trials that were determined not to be outliers (Figure 3.1b), and finally the reaction time distribution of non-outlier trials that were correctly answered to by the participants (Figure 3.1c). Only this last set of trials was used in the subsequent analyses. Participant-specific graphs are available in the Appendix. Note that the criteria where applied participant by participant (N = 6), and not directly on the overall data, so that the criterion lines drawn in Figure 3.1 are merely representative.

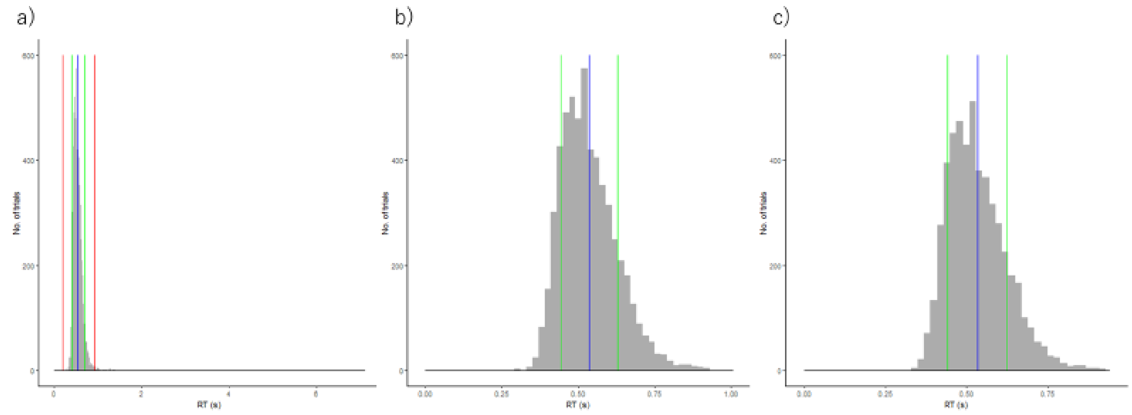

FIGURE 3.1. Reaction time distribution before and after application of outlier criteria. In all histograms, the blue line represents the mean and the green lines are the limits of the one standard deviation interval from the mean. All histograms share a bin width of 20ms, and the line along the horizontal axis stretches up to the maximum value of each distribution. a) Raw reaction time data distribution. The red lines represent where the criteria would be applied if this were a single participant data. Number of trials = 5760; mean = 545ms; standard deviation = 153ms. b) Non-outlier reaction time distribution. Number of trials = 5672; mean = 535ms; standard deviation = 93ms. c) Non-outlier and correctly answered reaction time distribution. Number of trials = 5075; mean = 534ms; standard deviation = 91ms.

3.2. Reaction time analysis

There were two main hypotheses to be tested: the voluntariness hypothesis, that voluntariness modulates attentional processes, and the difficult visual recognition influence hypothesis, that voluntariness-based effects can be especially influential under situations where visual recognition is poorer, such as low luminance and contrast trials. For clarifying the first hypothesis, a 3-way repeated measures ANOVA with factors of attention, contrast, and luminance was performed. For the second hypotheses, three 2-way repeated measures ANOVAs for low contrast, and low to medium luminance trials (e.g., factors were attention and luminance conditions for the low contrast trials) were carried out. The results of those analyses are presented in the following order: 3-way

ANOVA, low contrast trials 2-way ANOVA, medium luminance trials 2-way ANOVA, and low luminance trials 2-way ANOVA. Since the outlier and wrongly answered trials removal was not evenly distributed across conditions, the final data used in the analysis were not balanced. To balance the data, the same number of data was repeatedly bootstrapped and the median F-value was taken. Each ANOVA was done repeatedly (101 times) with random sampling from the data such that all different combinations of conditions (participant, attention, contrast, luminance) had the same number of trials each time. The number of trials sampled was always the highest possible, which means the number of trials in the condition combination with the fewest trials. The results presented here are the ones with the median F-value for the attention condition of all 101 ANOVAs, since these were deemed to be the most representative of the true result. Individual participants' reaction time data were deemed significantly not normal by the Shapiro-Wilk test of normality (see Appendix), but were considered not to be deviant enough to require special treatment (such as a bootstrap ANOVA instead of a parametric one). Furthermore, the Greenhouse-Geisser correction of sphericity was applied whenever possible, and only the corrected p-values are reported. Graphical representations of the reaction time distributions in question along with their separation by each condition of interest are displayed before the results of each ANOVA.

Overall analysis. Figure 3.2 shows the normalized distribution of the reaction time for all correct trials with categorization by attention, contrast, and luminance (Figure 3.2a, b, c, respectively).

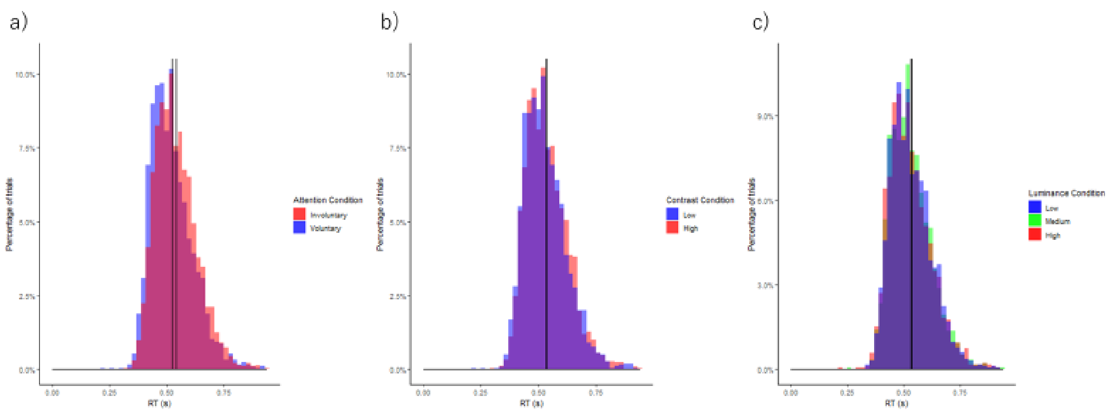

FIGURE 3.2. The normalized histograms of reaction time. The vertical black lines represent the mean of each distribution and the bin width is 20 ms. a) Distributions by attention condition. The red is of involuntary trials (N = 2672; mean = 542 ms; standard error = 2ms) and the blue is of voluntary trials (N = 2403; mean = 526 ms; standard error = 2 ms). b) Distributions by contrast condition. The blue is of low contrast trials (N = 2516; mean = 537 ms; standard error = 2 ms) and the red is of high contrast trials (N = 2559; mean = 530 ms; standard error = 2 ms). c) Distributions by luminance condition. The blue is of low luminance trials (N = 1712; mean = 532 ms; standard error = 2 ms), the green is of medium luminance trials (N = 1701; mean = 532 ms; standard error = 2 ms), and the red one is of high luminance trials (N = 1662; mean = 534 ms; standard error = 2 ms).

The reaction time distributions overlap heavily when separated by contrast or luminance condition, and only the separation by attention condition seems to have some relevance (mean difference of 16ms), although arguably limited due to a still relatively high degree of superposition. The statistical analysis results follow in Table 3.1.

TABLE 3.1. 3-way ANOVA of all trials (participant N = 6; trial N = 5075).

Factor	**F**	**p**	η^2
Attention	15.532	0.011*	0.030
Luminance	0.155	0.857	0.0004
Contrast	0.557	0.489	0.003
Attention:Luminance	0.209	0.798	0.0002
Attention:Contrast	2.366	0.185	0.0005
Luminance:Contrast	5.000	0.034*	0.012
Attention:Luminance:Contrast	0.533	0.584	0.0004

As expected, luminance and contrast did not exhibit significance, while attention did. On the other hand, luminance and contrast showed a significant interaction. This leads one to expect that, according to the difficult visual recognition hypothesis, low contrast and low luminance trials, which are the most challenging ones in matters of perception, show slower reaction times. Remember, however, that the difficult visual recognition influence hypothesis is concerned about attentional effects on challenging perceptional situations, and not about challenging perceptional situations themselves.

Thus, the hypothesis would be corroborated only by a significant interaction between all three factors, which is not the case. Still, statistical significance warrants further analyses, and the interaction plot for luminance and contrast is given in Figure 3.3.

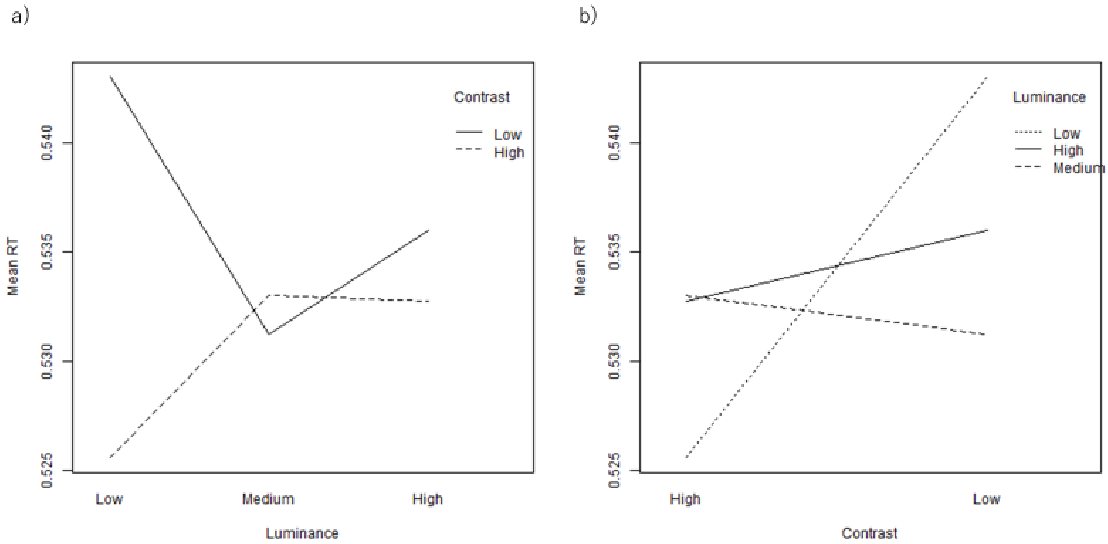

FIGURE 3.3. Reaction time interaction plot for contrast and luminance. The vertical axis represents mean reaction time in seconds. a) Mean reaction time of each contrast condition by luminance condition. b) Mean reaction time of each luminance condition by contrast condition.

As can be seen in Figure 3.3, the interaction lines of both plots are not parallel, which suggests a clear interaction. Giving special attention to Figure 3.3a, it is possible to confirm that reaction times in low contrast and low luminance trials have surprisingly high reaction times, while high contrast and low luminance trials have surprisingly low ones. This seems to be the source of the statistical significance, but it must be emphasized that this result has nothing to do with the difficult visual recognition influence hypothesis. A post hoc analysis result of contrast effects on trials of each luminance (run the same way as the previous ANOVA, but as a 1-way ANOVA with only contrast as a factor) is given in Table 3.2. Note that even the difference in contrast reaction times in low luminance trials was only marginally significant.

TABLE 3.2. Post hoc test of contrast and luminance interaction. “Luminance condition” refers to which trials were considered in each ANOVA (e.g., “Low” means that only low luminance condition trials were used in this analysis)

Factor	**Luminance Condition**	**F**	**p**	η^2
Contrast	Low	4.384	0.094	0.0417
Contrast	Medium	0.290	0.613	0.0006
Contrast	High	0.059	0.817	0.0008

2-way analyses for difficult visual recognition trials. Next, ANOVAs with relatively more difficult trials, that were medium to low luminance trials and low contrast trials, were tested. Medium luminance trials were also included because they share around the same stimuli luminance with the darkest luminance condition in Suizu’s. Figure 3.4 shows the distributions of reaction time according to attention condition in each of those types of trials.

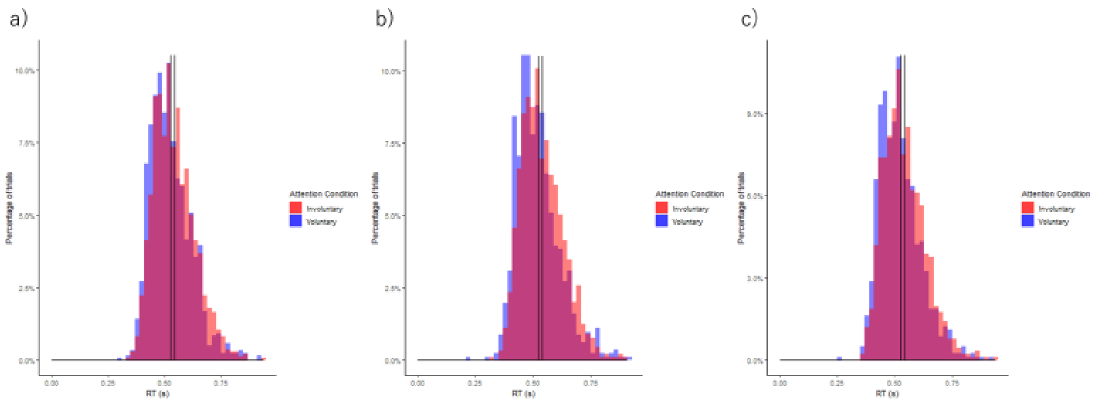

FIGURE 3.4. Normalized histograms of reaction time categorized by attention condition on difficult visual recognition trials. In all histograms, the bin width is 20 ms, involuntary and voluntary attention trials are in red, and blue, respectively, and the vertical black lines represent the corresponding mean value of the distributions. a) low contrast condition trials (involuntary distribution: N = 1333, mean = 544 ms, standard error = 2 ms;

voluntary distribution: N = 1183, mean = 529 ms, standard error = 2 ms). b) medium luminance condition trials (involuntary distribution: N = 893, mean = 541 ms, standard error = 3 ms; voluntary distribution: N = 808, mean = 522 ms, standard error = 3 ms). c) low luminance condition trials (involuntary distribution: N = 906, mean = 543 ms, standard error = 3 ms; voluntary distribution: N = 806, mean = 525 ms, standard error = 3 ms).

It is clear from Figure 3.4 that at least around the same effect present on the general analysis can be expected, but is it more accentuated in these conditions? Compared to the overall mean difference of 16ms, in these trials the mean reaction time differences between voluntary and involuntary trials were of 15ms, 19ms, and 18ms, for low contrast, medium luminance and low luminance trials, respectively, with voluntary trials having the lower mean in all cases. The results of each ANOVA are available in the following Table 3.3. It is clear that attention effects were significant, but the generalized eta squared for each attention effect, which measures their strength, is either close to the overall effect found in Table 3.1 (of 0.030), in the case of medium (0.027) and low luminance trials (0.032), or even smaller than that, in the case of low contrast trials (0.018). These results do not sustain the difficult visual recognition influence hypothesis.

TABLE 3.3. 2-way ANOVA of difficult visual recognition trials.

Factor	F	p	η^2
Low contrast trials only			
Attention	11.787	0.019*	0.018
Luminance	2.085	0.186	0.013
Attention:Luminance	0.234	0.695	0.0005
Medium luminance trials only			
Attention	13.988	0.013*	0.027
Contrast	0.013	0.915	0.00003
Attention:Contrast	0.646	0.400	0.0006
Low luminance trials only			
Attention	8.484	0.033*	0.032
Contrast	3.077	0.140	0.031
Attention:Contrast	1.393	0.291	0.002

3.3. Hit rate analysis

At first, only reaction time analyses were planned. However, it was noted that all participants consistently performed worse in matters of correct answer rates in voluntary trials compared to involuntary ones (as can be found in the Appendix). A hit rate analysis was due.

The hit rate analysis was conducted in the same way of the overall reaction time analysis, with a 3-way repeated measures ANOVA with Greenhouse-Geiser correction applied when possible of factors of attention condition, contrast condition and luminance condition. This time, though, since there was only one datum per cell (the mean hit rate of that particular condition over all participants), no random sampling to control unbalanced data was needed. The hit rate plots of each condition follow in Figure 3.5.

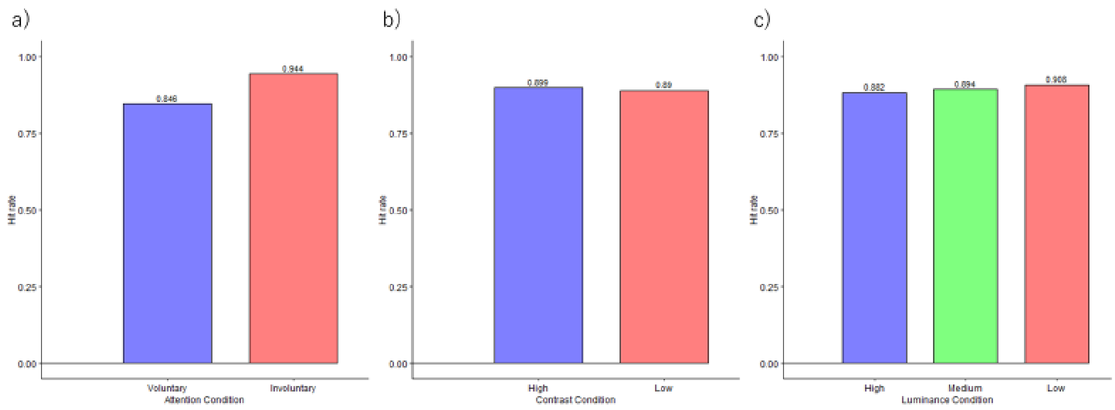

FIGURE 3.5. Mean hit rate over each condition. a) Separation by attention condition. Blue is voluntary (84.6%) and red is involuntary (94.4%). b) Separation by contrast condition. Blue is high contrast (89.9%) and red is low contrast (89.0%). c) Separation by luminance condition. Blue is high luminance (88.2%), green is medium luminance (89.4%) and red is low luminance (90.8%).

The mean difference in the rate of trials correctly answered between voluntary and involuntary conditions is of an astounding 9.8%, while between high and low contrast conditions it goes to around 1%, and as luminance decreases, the hit rate increases of about 1% per level, going from around 88% to around 90%. The small fluctuations between contrast and luminance condition were according to the expected, since it was thought that in the task in question only reaction time effects would be observed. However, that was clearly not the case of the attention condition. The ANOVA results corroborate this disparity, whose statistical significance is shown in Table 3.4. Surprisingly, the small difference found in luminance conditions were also statistically significant. A post hoc Bonferroni adjusted pairwise proportion test was run between the three luminance conditions, and only the difference between high and low luminance conditions was found to be statistically significant (high-medium: $p = 0.654$; medium-low: $p = 0.455$; high-low: $p = 0.024$). This means that a mean hit rate difference of around 2% is enough to be statistically significant, even though there is no reason to believe that this gradual fluctuation is theoretically relevant.

TABLE 3.4. Hit rate 3-way ANOVA.

Factor	F	p	η^2
Attention	60.56	0.00056*	0.599
Luminance	10.70	0.0086*	0.067
Contrast	1.10	0.342	0.011
Attention:Luminance	0.0249	0.930	0.0004
Attention:Contrast	4.33	0.091	0.019
Luminance:Contrast	0.325	0.653	0.010
Attention:Luminance:Contrast	4.028	0.075	0.0966

In summary, there was a clear gap in the hit rate observed between voluntary and involuntary conditions, with voluntary trials being more prone to be incorrectly answered. This leaves one question: if voluntary trials are found to be both more incorrectly and faster to be answered to compared to involuntary trials when participants received the exact same instructions, is this data explained by true attentional effects or just a displacement of an internal decision criterion in a trade-off function? This problem is more thoroughly treated in the next Chapter.

Chapter 4 – Discussion

4.1. Interpreting the results

In the present study, an experiment was conducted in order to test two hypotheses that relate voluntary action to attentional mechanisms. They were that voluntariness can modulate attention and that that modulation is accentuated under difficult visual recognition situations. The results supported the first hypothesis, but not the second. The reason for that conclusion lies on the fact that attentional effects were observed in the analysis involving all experimental conditions, and with only a similar (or even lower) magnitude (as measured by the generalized eta square) when the analysis was limited to trials considered especially challenging for visual recognition. However, these results do not come without their share of critic.

A first critic that could be made is the small size of the sample, especially compared to how many people were called to take part of the experiment but did not pass the participant removal criteria. In the exploratory analysis, seven people were trained and tested, of whom only three were deemed fit for the confirmatory run of the experiment. Even excluding the one participant who reported color vision disability, that would mean that just about 50% of available data was used for the final analysis. As for the new participants that were recruited especially for the confirmatory run, the same pattern was observed: the removal criteria cut down seven possible participants to only three. There is the possibility that the criteria used in this study were too strong and could have forced the selection of one specific group of people where a voluntariness effect could be observed just out of chance, and not because it is a real effect in the whole population. Setting aside questions about what common characteristic the people who passed the criteria would share, such as some sort of color picking strategy, a possible solution for this problem is conducting a statistical power analysis for theoretically determining how many participants are needed for a robust result.

Another problem is the real effect or trade-off artifact discussion briefly touched upon by the end of the last Chapter. It is known that in reaction time studies, reaction time and accuracy tend to respect a sigmoidal-like trade-off function, so that accuracy increases as the reaction time increases (or speed decreases), starting from 0% accuracy at times too fast for actual processing and steadily transitioning into 100% accuracy at a theoretical minimum reaction time, which is taken to be the typical definition of reaction time in psychological theory. Usually, participants are instructed to react as accurately and fast as possible, which results in an accuracy slightly below 100% with a faster reaction time than the theoretical one [19]. Following the same reasoning in the opposite direction leads us to interpret the presented results of a drop in both hit rate and reaction time in voluntary trials as a possible shift in the participant's trade-off characteristics. In other words, it could be that voluntary and involuntary trials share the same basic attentional process, which corresponds to a single trade-off function, and the only difference is at which point in the function each type of trial maps to, opposed to each representing a different sigmoidal for different attentional processes. Either way, it could be argued that the even if the effect were due to a simple change in criterion, it would still be a valuable and meaningful finding, given that the experimenter's instruction was the same for both attentional conditions. Maybe voluntariness-based attentional effects do not materialize as different processes with different trade-off functions, but as a slight tuning of a shared process with a shared function.

4.2. Further hypotheses – a direction for future work

As has been discussed in the previous section, new questions have risen from the present results. Of main concern is the explanation of the observed drop in both reaction time and hit rate in voluntary attention trials. The participants were instructed to react as quickly and precisely as possible, with no focus given on achieving faster or more correct performances, and yet the results in hand show the possibility of a trade-off. But is it really a change in criterion of a trade-off function what happened?

One possible further investigation is using more advanced mathematical analysis tools to compare the expected trade-off functions for voluntary and involuntary trials, or even repeat the experiment giving greater focus to speed or accuracy for an attempt to plot such a function more reliably. The introduction of analyses utilizing indexes such as BIS (balanced integration score) or LISAS (linear integrated speed-accuracy score) [10][19] which indicate the level of an observed trade-off is also a possible route.

Nevertheless, other possibilities that treat this drop in reaction time and hit rate as a real effect should also be explored. It could be that time was the problem. In the present experiment, participants had the same exposure time to the square cues in both involuntary and voluntary trials, which could have negatively affected the performance on voluntary trials since they first had to perceive which colors where available, decide on one to look for and then concentrate on it, while on involuntary trials they did not have to choose anything. This difference in number of actions could lead to a smaller activation time allotted to the color of interest in the attentional process, which would in turn lead to an overall worse performance on voluntary trials. It could also be a problem of area. On involuntary trials, participants saw four squares of the same color, compared to only one square of the chosen color on voluntary ones. Maybe this higher degree of how much of the field of vision the color of interest covered also affected activation levels bound to that color in the attentional process. In this way, changing either or both of cue exposure time and cue size according to each voluntariness-based condition and redoing the experiment could be a path for clarification.

Finally, Suizu noted in his work that maybe the cues used were not appropriate in the first place, since it is difficult to believe that people have any clear preferences over four differently colored squares. It could be that if the cues were made of more tangible categories, like fruits or faces, the selection process would be smoother and the results clearer. This line of inquiry was not taken in the present study because such cues are not as straightforwardly linkable to items in a search array as colored squares are. Not only that, they are also harder to generate. However, this suggestion still lies as a possibility for further investigation.

4.3. Conclusion

To summarize, the present experiment set out to confirm and evaluate the robustness of Suizu's findings. It was successful on verifying a statistically significant difference in reaction time of voluntary and involuntary attention trials, but failed to prove a greater effect under poorer visual conditions. Furthermore, an unexpected difference in hit rate between the two voluntariness conditions was also observed, which led to concerns about the nature of that attentional effect. It was then suggested that more experimentation should follow with a larger number of participants decided by means of statistical power analysis, or with some alterations to the present paradigm, such as size and time of presentation of the cues. Despite these difficulties, it is certainly possible to assert that in the very least this study has shown that a new taxonomical paradigm of attention based on voluntariness is not only an interesting possibility, but a promising field of inquiry with many roads in the wait for scientific exploration.

References

[1] Awh, E., Belopolsky, A.V. & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16(8), 437-443.

[2] Baayen, R.H. & Milin, Petar. (2010). Analyzing Reaction Times. *International Journal of Psychological Research*, 3(2), 12-28.

[3] Baluch, F. & Itti, L. (2011). Mechanisms of top-down attention. *Trends in Neurosciences*, 34(4).

[4] Della Libera, C. & Chelazzi, L. (2009). Learning to Attend and to Ignore is a Matter of Gains and Losses. *Psychological Science*, 20(6), 778-784.

[5] Failing, M. & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin & Review*, 25, 514-538.

[6] Gaspelin, N. & Luck, S.J. (2018). "Top-down" Does Not Mean "Voluntary". *Journal of Cognition*, 1(1):25, 1-4.

[7] Gescheider, G.A. (1997). *Psychophysics: the fundamentals* (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associative, Inc., Publishers.

[8] Khauchouf, O.T., Chen, G., Duzzi, D., Porro, C.A. & Pagnoni, G. (2017). Voluntary modulation of mental effort investment: an fMRI study. *Scientific Reports*, 7:17191.

[9] Leotti, L.A. & Delgado, M.R. (2011). The Inherent Reward of Choice. *Psychological Science*, 22(10), 1310-1318

[10] Liesefeld, H.R. & Janczyk, M. (2018). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, 51, 40-60.

[11] Liu, T., Stevens, S.T. & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. *Vision Research*, 47, 108-113.

[12] Nobre, A.C. (2018). Attention. In Wixted, J.T. & Serences, J. (Ed.). *Steven's Handbook of Experimental Psychology and Cognitive Neuroscience, Sensation, Perception, and Attention* (4th ed., pp. 241-316). John Wiley & Sons, Incorporated.

[13] Peirce, J. & MacAskill M. (2018). *Building experiments in PsychoPy*. SAGE.

[14] Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman E. & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195-203.

[15] Rens, N., Bode, S., Burianová, H. & Cunnington, R. (2017). Proactive Recruitment of Frontoparietal and Salience Networks for Voluntary Decisions. *Frontiers in Human Neuroscience*, 11:610.

[16] Theeuwes, J. & Failing, M. (2020). *Attentional Selection: Top-Down, Bottom-Up and History-Based Biases*. Cambridge University Press. Elements in Perception.

[17] Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, 1(1):29, 1-15.

[18] Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional; A Reply to Commentaries. *Journal of Cognition*, 1(1):21, 1-4.

[19] 綾部早穂、井関龍太、熊田孝恒（編）．（2019）．『心理学，認知・行動科学のための反応時間ハンドブック』．東京，勁草書房

[20] 井関龍太. (2020).「心理学者は反応時間をどう分析するか」.『基礎心理学研究』, 38(2), 253-249.

[21] 大山正、岩脇三良、宮埜壽夫．（2005）．『心理学研究法 ―データ収集・分析から論文作成まで―』．東京，サイエンス社．コンパクト新心理学ライブラリ，12.

[22] 島崎恒雄．（2015）．『心理統計学 II 分散分析とノンパラメトリック検定』．東京，培風館．現代心理学シリーズ，12

[23] 南風原朝和、市川伸一、下山晴彦．（2001）．『心理学研究法入門 調査・実験から実践まで』．東京大学出版会.

[24] 南風原朝和．（2002）．『心理統計学の基礎 統合的理解のために』．東京，有斐閣.

[25] 南風原朝和．（2014）．『続・心理統計学の基礎 統合的理解を広げ深める』．東京，有斐閣.

Appendix

A.1. Stimuli colors

The colors used to display the stimuli in the present study are as follow in Tables A.1 through A6. There are two sets of six different colors for each luminance condition. The colors were chosen to keep their saturation and hue constant through luminance levels, and to be homogenously distributed in the CIELAB color space. This was done by utilizing the L*C*h color space, where L varied according to the experimental luminance level, C was kept constant and h altered from 0º (color set 1) or 30º (color set 2) to 300º (color set 1) or 330º (color set 2) in 60º steps. The RGB values in each table are the actually presented Adobe RGB colors that correspond to the theoretically chosen L*C*h ones. Finally, 'a' and 'b' values refer to the coordinates of that color in the L*a*b color space (L being the same value for both L*C*h and L*a*b). All luminance values are given in cd/m^2.

TABLE A.1. Low luminance color set 1.

L	C	h	a	b	R	G	B	Luminance
27	30	0	30	0	27	6	13	9.96
27	30	60	15	25.98	23	9	3	9.82
27	30	120	-15	25.98	11	15	3	10.02
27	30	180	-30	0	3	18	13	10.21
27	30	240	-15	-25.98	0	17	34	10.72
27	30	300	15	-25.98	12	11	34	10.63
Mean luminance: 10.23, Lumi. Std. Deviation: 0.34, Lumi. Amplitude: 0.90								

TABLE A.2. Low luminance color set 2.

L	C	h	a	b	R	G	B	Luminance
27	30	30	25.98	15	28	7	6	10.87
27	30	90	0	30	17	13	2	10.82
27	30	150	-25.98	15	6	17	6	10.43
27	30	210	-25.98	-15	0	18	24	10.45
27	30	270	0	-30	4	14	38	10.71
27	30	330	25.98	-15	22	8	24	10.88
Mean luminance: 10.69, Lumi. Std. Deviation: 0.19, Lumi. Amplitude: 0.46								

TABLE A.3. Medium luminance color set 1.

L	C	h	a	b	R	G	B	Luminance
40	30	0	30	0	52	18	29	20.20
40	30	60	15	25.98	46	23	9	19.76
40	30	120	-15	25.98	26	33	9	19.73
40	30	180	-30	0	10	37	29	19.65
40	30	240	-15	-25.98	8	35	62	20.53
40	30	300	15	-25.98	28	25	62	20.26
Mean luminance: 20.02, Lumi. Std. Deviation: 0.33, Lumi. Amplitude: 0.88								

TABLE A.4. Medium luminance color set 2.

L	C	h	a	b	R	G	B	Luminance
40	30	30	25.98	15	53	19	16	20.10
40	30	90	0	30	36	28	7	19.64
40	30	150	-25.98	15	17	36	16	19.67
40	30	210	-25.98	-15	7	37	46	19.99
40	30	270	0	-30	15	30	68	20.28
40	30	330	25.98	-15	43	20	46	20.01
Mean luminance: 19.95, Lumi. Std. Deviation: 0.23, Lumi. Amplitude: 0.64								

TABLE A.5. High luminance color set 1.

L	C	h	a	b	R	G	B	Luminance
49	30	0	30	0	76	30	45	29.97
49	30	60	15	25.98	69	37	18	29.05
49	30	120	-15	25.98	41	50	18	28.69
49	30	180	-30	0	20	57	45	28.92
49	30	240	-15	-25.98	17	53	88	29.64
49	30	300	15	-25.98	45	40	88	29.70
Mean luminance: 29.33, Lumi. Std. Deviation: 0.46, Lumi. Amplitude: 1.27								

TABLE A.6. High luminance color set 2.

L	C	h	a	b	R	G	B	Luminance
49	30	30	25.98	15	77	32	28	29.64
49	30	90	0	30	55	44	15	29.53
49	30	150	-25.98	15	29	55	27	29.25
49	30	210	-25.98	-15	15	56	67	29.31
49	30	270	0	-30	27	47	96	29.78
49	30	330	25.98	-15	64	33	68	29.42
Mean luminance: 29.49, Lumi. Std. Deviation: 0.18, Lumi. Amplitude: 0.52								

Lastly, Table A.7 shows the L*a*b, L*C*h and Adobe RGB values, along with the measured luminance of the background gray used in each luminance and contrast level. These background colors where selected in order to yield a Michelson contrast of around 0.33 in low contrast trials and 0.6 in high contrast trials. Once again, all luminance values are given in cd/m^2.

TABLE A.7. Background gray for each luminance and contrast condition.

Luminance Condition	Contrast Condition	L	a	b	R	G	B	Luminance
Low	Low	40.20	0	0	29	29	29	20.11
Low	High	57.17	0	0	64	64	64	40.62
Medium	Low	57.17	0	0	64	64	64	40.62
Medium	High	77.14	0	0	132	132	132	79.86
High	Low	68.04	0	0	97	97	97	59.99
High	High	91.68	0	0	204	204	204	120.5

A.2. Outlier removal by participant

The following Figures A.1 through A.10 present the reaction time distribution over all trials (Figures A.1a-A.10a) of each participant whose data was used for analysis in either the exploratory run of the experiment (here referred to as Experiment 1) or the confirmatory run of the experiment (here referred to as Experiment 2), along with the distributions resultant from the removal of outliers (Figures A.1b-A.10b) and the subsequent removal of incorrectly answered trials (A.1c-A.10c). The outlier criteria proposed and applied in the study were: a) reaction time smaller than 200ms; b) reaction time larger than two times and a half the standard deviation of the participant in question added to the mean (mathematically, $RT > \mu + 2.5\sigma$). In all figures, the bin width is of 20ms, and the blue bar indicates the mean reaction time, while the green bars represent the limits of the one standard deviation interval from the mean. The red bars in Figures A.1a-A.10a represent the outlier criteria, such that all trials outside the interval denoted by those bars were removed. Additionally, in Figures A.1a-A.10a the horizontal bar along the x axis stretches up to the maximum recorded reaction time. The number of trials in each distribution are given in each Figure's caption along with the distribution's mean and standard deviation. The hit rate for each attention condition of each participant and the p-value obtained from the Shapiro-Wilk test of normality when applied to the correct trials distribution can also be found in the correspondent participant's Figure's caption.

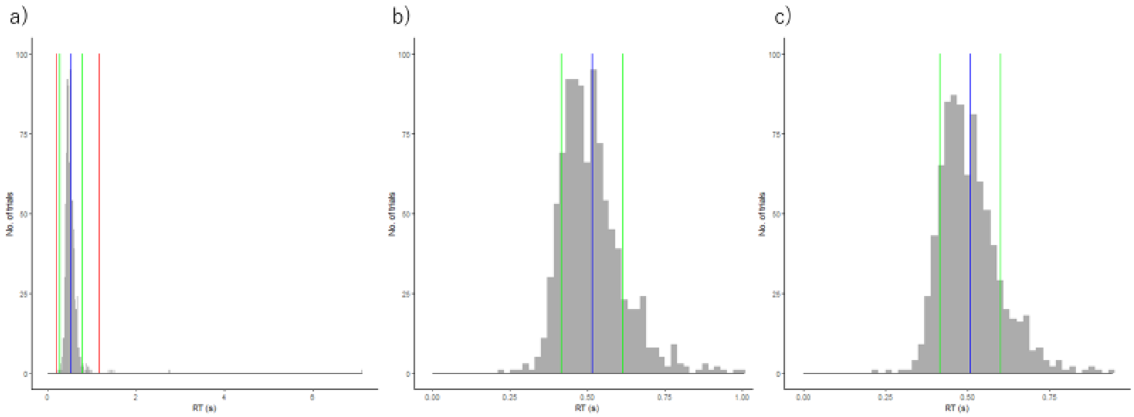

FIGURE A.1. Reaction time distribution for Participant A (male; took part in Experiments 1 and 2). a) Distribution over all trials. N = 960; mean = 527 ms; standard deviation = 253 ms. b) Distribution of non-outlier trials. N = 954; mean = 513 ms; standard deviation = 100 ms. c) Distribution of non-outlier and correctly answered trials. N = 832; mean = 508 ms; standard deviation = 91 ms. Shapiro-Wilk test of normality p-value: 2.21e-16. Hit rate for voluntary trials: 86.0%. Hit rate for involuntary trials: 92.4%.

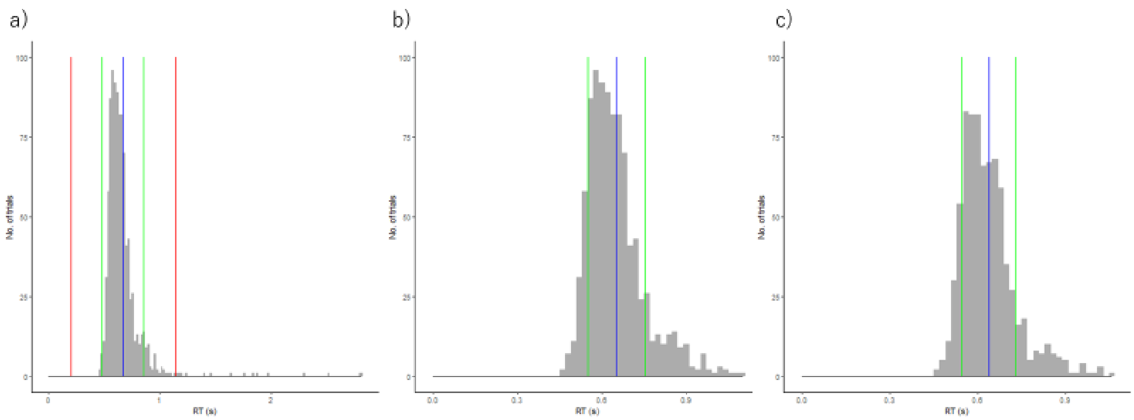

FIGURE A.2. Reaction time distribution for Participant B (female; took part in

Experiment 1). a) Distribution over all trials. N = 960; mean = 668 ms; standard deviation = 190 ms. b) Distribution of non-outlier trials. N = 944; mean = 650 ms; standard deviation = 102 ms. c) Distribution of non-outlier and correctly answered trials. N = 765; mean = 637 ms; standard deviation = 93 ms. Shapiro-Wilk test of normality p-value: 5.71e-21. Hit rate for voluntary trials: 70.8%. Hit rate for involuntary trials: 91.3%.

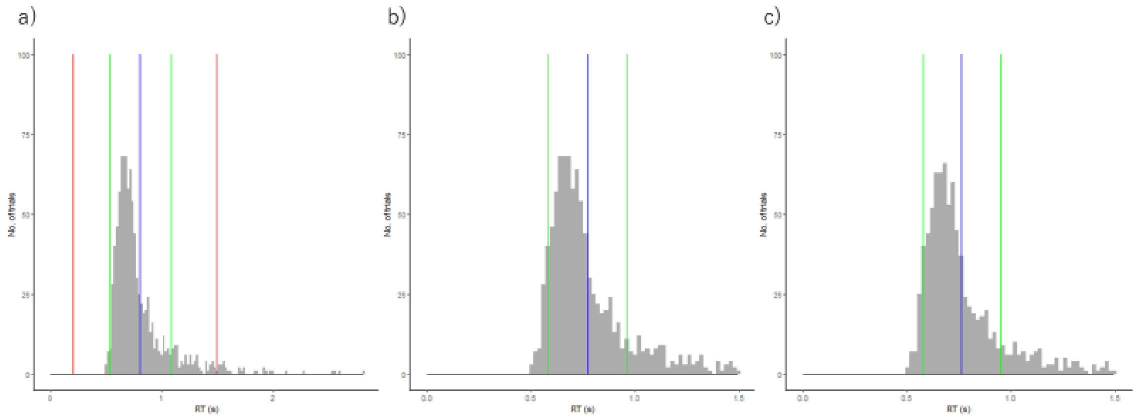

FIGURE A.3. Reaction time distribution for Participant C (male; took part in Experiment 1). a) Distribution over all trials. N = 960; mean = 806 ms; standard deviation = 276 ms. b) Distribution of non-outlier trials. N = 928; mean = 770 ms; standard deviation = 191 ms. c) Distribution of non-outlier and correctly answered trials. N = 828; mean = 761 ms; standard deviation = 186 ms. Shapiro-Wilk test of normality p-value: 5.34e-28. Hit rate for voluntary trials: 83.4%. Hit rate for involuntary trials: 95.0%.

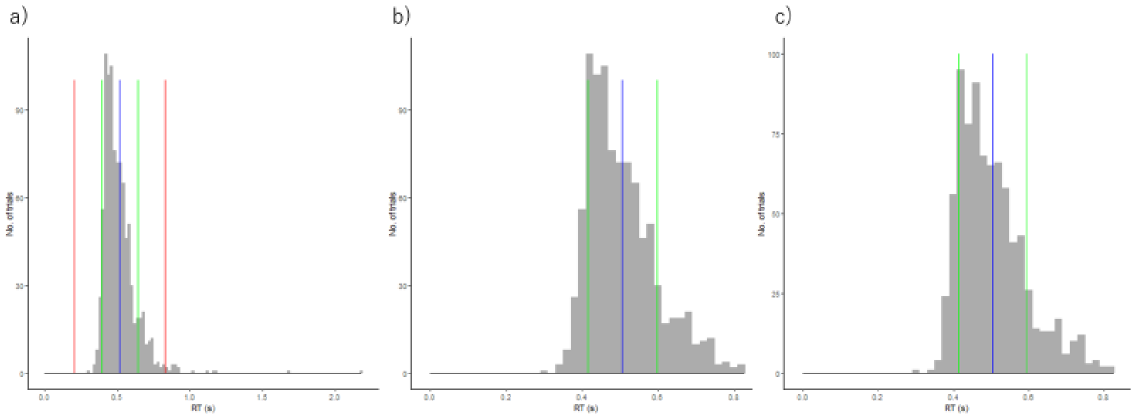

FIGURE A.4. Reaction time distribution for Participant D (female; took part in Experiment 1). a) Distribution over all trials. N = 960; mean = 515 ms; standard deviation = 125 ms. b) Distribution of non-outlier trials. N = 943; mean = 505 ms; standard deviation = 91 ms. c) Distribution of non-outlier and correctly answered trials. N = 813; mean = 504 ms; standard deviation = 90 ms. Shapiro-Wilk test of normality p-value: 1.34e-18. Hit rate for voluntary trials: 75.3%. Hit rate for involuntary trials: 97.2%.

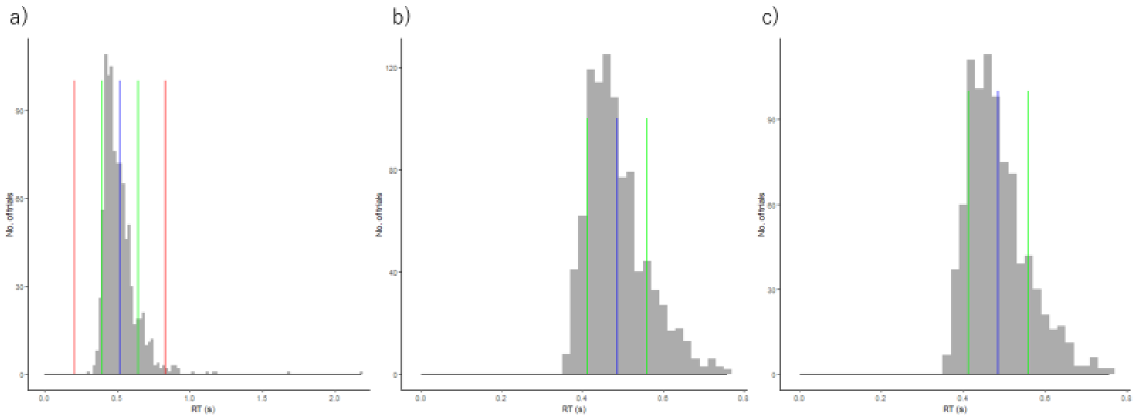

FIGURE A.5. Reaction time distribution for Participant E (male; took part in Experiments

1 and 2). a) Distribution over all trials. N = 960; mean = 494 ms; standard deviation = 106 ms. b) Distribution of non-outlier trials. N = 945; mean = 485 ms; standard deviation = 74 ms. c) Distribution of non-outlier and correctly answered trials. N = 865; mean = 484 ms; standard deviation = 73 ms. Shapiro-Wilk test of normality p-value: 2.31e-18. Hit rate for voluntary trials: 87.5%. Hit rate for involuntary trials: 95.6%.

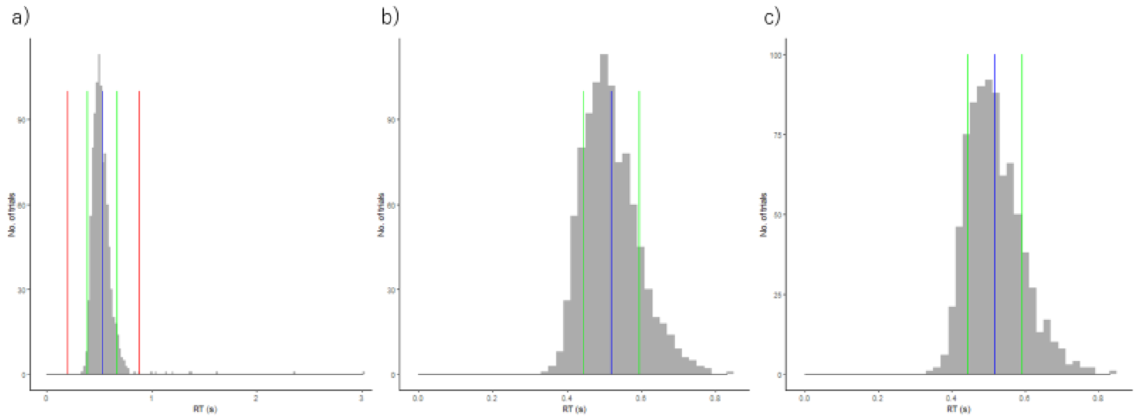

FIGURE A.6. Reaction time distribution for Participant F (male; took part in Experiments 1 and 2). a) Distribution over all trials. N = 960; mean = 529 ms; standard deviation = 141 ms. b) Distribution of non-outlier trials. N = 950; mean = 518 ms; standard deviation = 75 ms. c) Distribution of non-outlier and correctly answered trials. N = 809; mean = 517 ms; standard deviation = 74 ms. Shapiro-Wilk test of normality p-value: 1.73e-12. Hit rate for voluntary trials: 80.4%. Hit rate for involuntary trials: 89.9%.

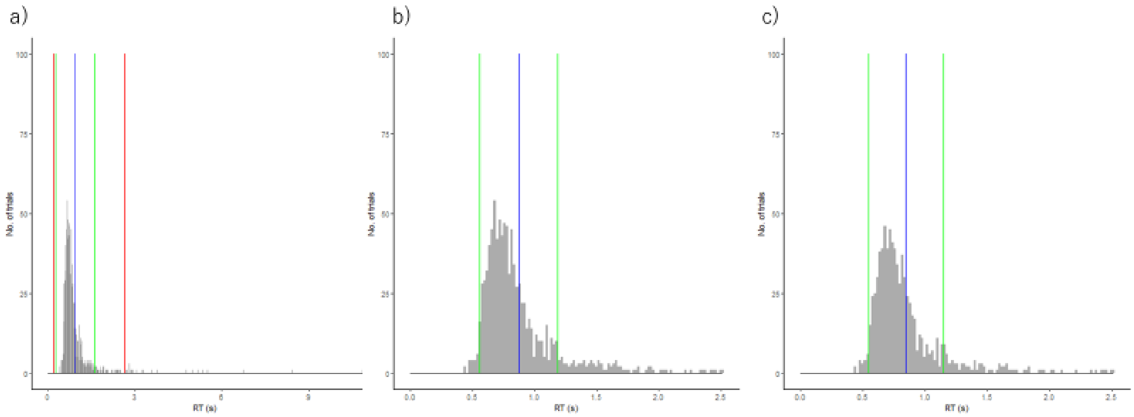

FIGURE A.7. Reaction time distribution for Participant G (male; took part in Experiment 1; self-reported a color vision disability). a) Distribution over all trials. N = 960; mean = 952 ms; standard deviation = 678 ms. b) Distribution of non-outlier trials. N = 937 ms; mean = 871 ms; standard deviation = 315 ms. c) Distribution of non-outlier and correctly answered trials. N = 781; mean = 849 ms; standard deviation = 302 ms. Shapiro-Wilk test of normality p-value: 1.41e-32. Hit rate for voluntary trials: 78.4%. Hit rate for involuntary trials: 88.4%.

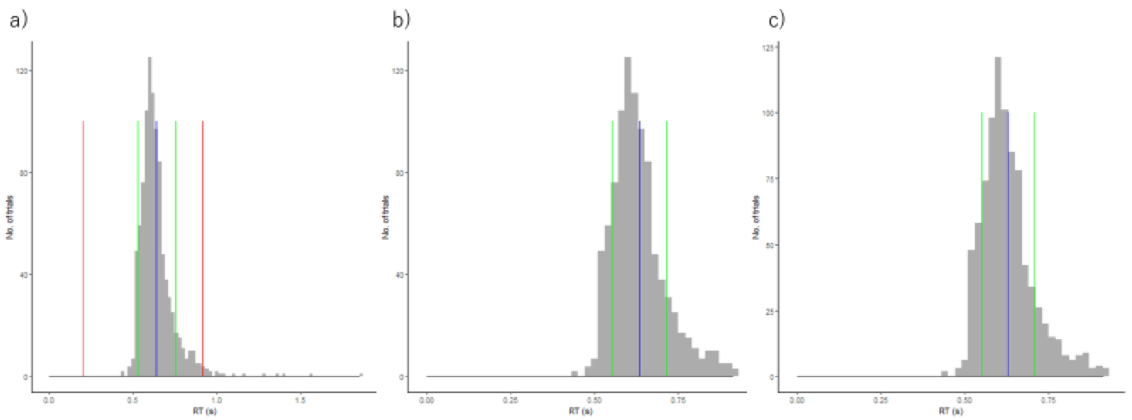

FIGURE A.8. Reaction time distribution for Participant I (male; took part in Experiment

2). a) Distribution over all trials. N = 960; mean = 641 ms; standard deviation = 102 ms. b) Distribution of non-outlier trials. N = 943; mean = 632 ms; standard deviation = 81 ms. c) Distribution of non-outlier and correctly answered trials. N = 866; mean = 628 ms; standard deviation = 79 ms. Shapiro-Wilk test of normality p-value: 6.92e-19. Hit rate for voluntary trials: 85.5%. Hit rate for involuntary trials: 98.1%.

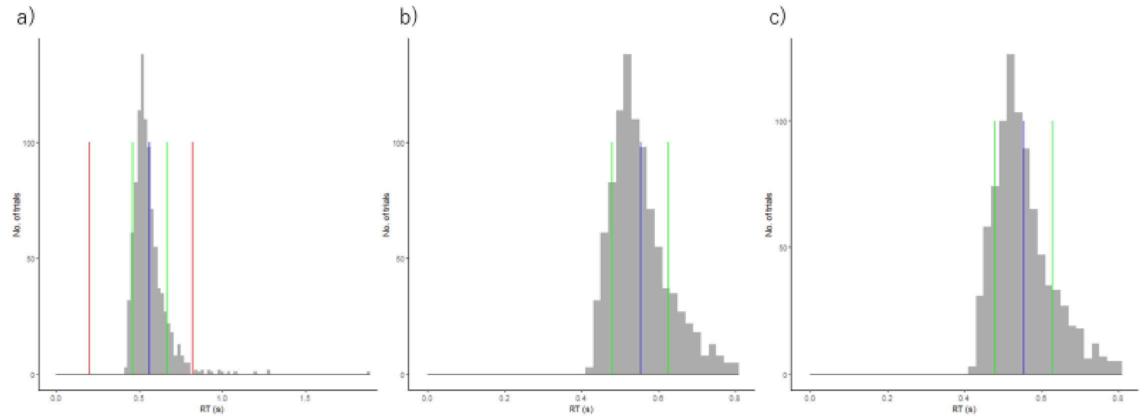

FIGURE A.9. Reaction time distribution for Participant K (male; took part in Experiment 2). a) Distribution over all trials. N = 960; mean = 561 ms; standard deviation = 104 ms. b) Distribution of non-outlier trials. N = 943; mean = 552 ms; standard deviation = 74 ms. c) Distribution of non-outlier and correctly answered trials. N = 863; mean = 552 ms; standard deviation = 75 ms. Shapiro-Wilk test of normality p-value: 5.51e-18. Hit rate for voluntary trials: 85.0%. Hit rate for involuntary trials: 98.1%.

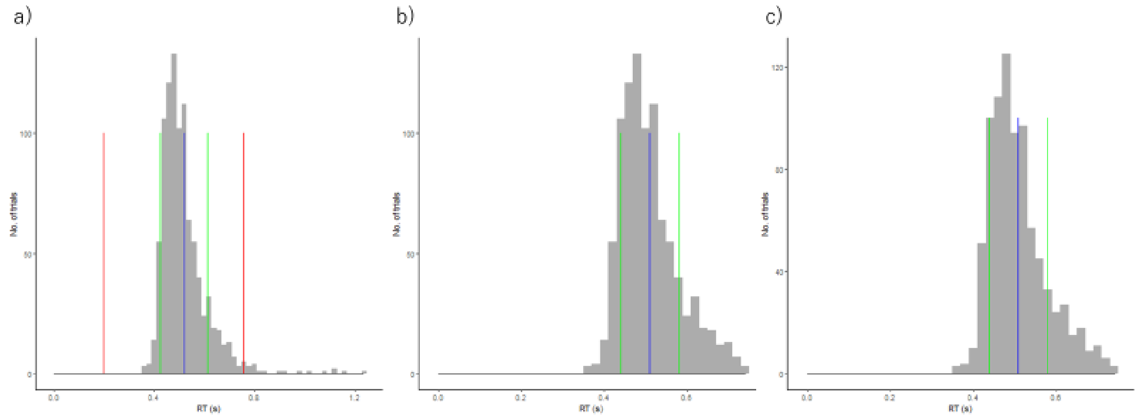

FIGURE A.10. Reaction time distribution for Participant N (male; took part in Experiment 2). a) Distribution over all trials. N = 960; mean = 519 ms; standard deviation = 94 ms. b) Distribution of non-outlier trials. N = 937; mean = 510 ms; standard deviation = 71 ms. c) Distribution of non-outlier and correctly answered trials. N = 840; mean = 509 ms; standard deviation = 70 ms. Shapiro-Wilk test of normality p-value: 3.74e-18. Hit rate for voluntary trials: 87.3%. Hit rate for involuntary trials: 92.1%.

A.3. Exploratory run results – deciding a participant training criterion

The exploratory run was conducted in the exact same way as the confirmatory run, including the manner in which its results were analyzed. As such, refer to Chapters 2 and 3 for the procedure and method of analysis.

The purpose of running an exploratory session before the main session was to discover which participant removal criteria applied during training would result in a good selection of participants whose data were suitable to the present analysis. The individual data of each participant who took part in this exploratory run is available in Figures A.1 to A.7. No participant data was removed for the analysis, and all seven of the participants belonged to the same research team as the author. In this session, first the overall results will be shown, including the reaction time by different conditions plot and a summarized ANOVA table for each analysis, and then these results will be compared to Suizu's results, with new restrictions. Finally, the participant removal criteria will be considered.

Exploratory results. Figures A.11 to A.13 show respectively the outlier application process, the reaction time distribution by each factor of the 3-way ANOVA, and the reaction time distribution by attention in each type of trial considered in the 2-way ANOVAs. Table A.8 shows the results of the 3-way ANOVA and Table A.9 the results of the 2-way ANOVAs.

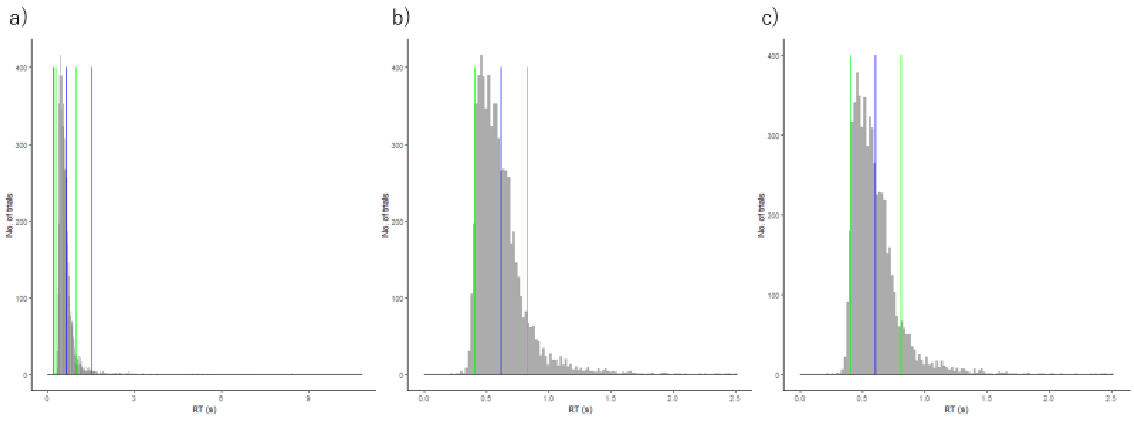

FIGURE A.11. Reaction time distribution before and after application of outlier criteria on the exploratory run data. In all histograms, the blue line represents the mean and the green lines are the limits of the one standard deviation interval from the mean. All histograms share a bin width of 20 ms, and the line along the horizontal axis stretches up to the maximum value of each distribution. a) Raw reaction time data distribution. The red lines represent where the criteria would be applied if this were a single participant data. Number of trials = 6720; mean = 641 ms; standard deviation = 353 ms. b) Non-outlier reaction time distribution. Number of trials = 6601; mean = 615 ms; standard deviation = 211 ms. c) Non-outlier and correctly answered reaction time distribution. Number of trials = 5693; mean = 606 ms; standard deviation = 202 ms.

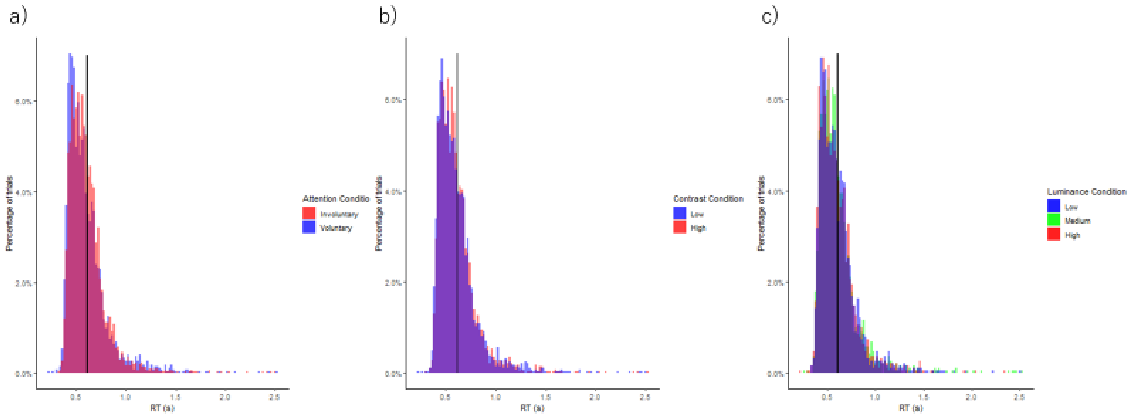

FIGURE A.12. The normalized histograms of reaction time in the exploratory run. The vertical black lines represent the mean of each distribution and the bin width is 20 ms. a) Distributions by attention condition. The red is of involuntary trials (N = 3065; mean = 609 ms; standard error = 3 ms) and the blue is of voluntary trials (N = 2628; mean = 603 ms; standard error = 4 ms). b) Distributions by contrast condition. The blue is of low contrast trials (N = 2853; mean = 606 ms; standard error = 4 ms) and the red is of high contrast trials (N = 2840; mean = 606 ms; standard error = 4 ms). c) Distributions by luminance condition. The blue is of low luminance trials (N = 1903; mean = 700 ms; standard error = 5 ms), the green is of medium luminance trials (N = 1893; mean = 600 ms; standard error = 4 ms), and the red one is of high luminance trials (N = 1897; mean = 609 ms; standard error = 5 ms).

TABLE A.8. 3-way ANOVA of all trials in the exploratory run (participant N = 7; trial N = 5693).

Factor	F	p	η^2
Attention	0.441	0.531	0.0005
Luminance	0.254	0.663	0.0004
Contrast	0.126	0.734	0.0001
Attention:Luminance	0.040	0.867	0.00002
Attention:Contrast	0.258	0.629	0.0001
Luminance:Contrast	0.453	0.636	0.0003
Attention:Luminance:Contrast	1.270	0.315	0.0007

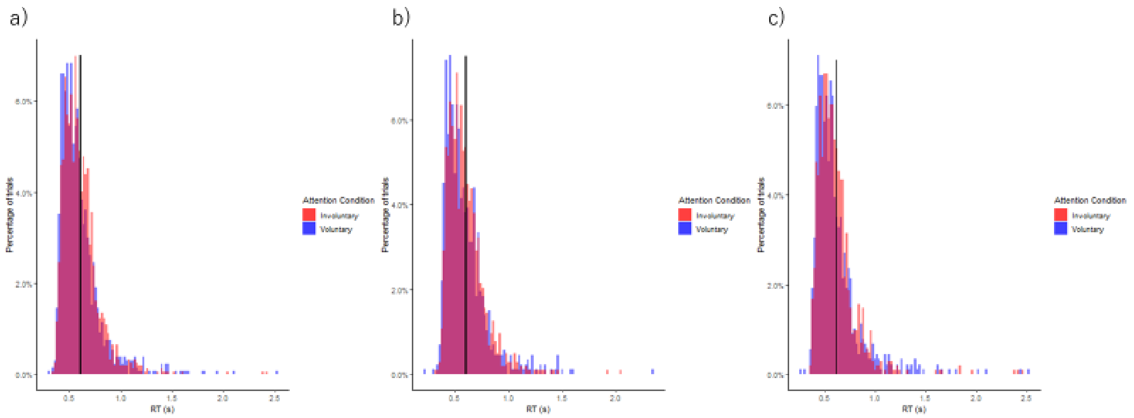

FIGURE A.13. Normalized histograms of reaction time categorized by attention condition on difficult visual recognition trials in the exploratory run. In all histograms, the bin width is 20ms, involuntary and voluntary attention trials are in red and blue, respectively, and the vertical black lines represent the corresponding mean value of the distributions. a) low contrast condition trials (involuntary distribution: N = 1548, mean = 609 ms; standard error = 5 ms; voluntary distribution: N = 1305, mean = 602 ms, standard error = 6 ms). b) medium luminance condition trials (involuntary distribution: N = 1023, mean = 603 ms, standard error = 5 ms; voluntary distribution: N = 865, mean = 596 ms, standard error = 7 ms). c) low luminance condition trials (involuntary distribution: N = 1016, mean = 610 ms, standard error = 6 ms; voluntary distribution: N = 887, mean = 609 ms, standard error = 8 ms).

TABLE A.9. 2-way ANOVA of difficult visual recognition trials in the exploratory run.

Factor	F	p	η^2
Low contrast trials only			
Attention	1.201	0.315	0.001
Luminance	0.933	0.397	0.003
Attention:Luminance	0.512	0.564	0.0007
Medium luminance trials only			
Attention	0.296	0.605	0.0004
Contrast	2.684	0.152	0.001
Attention:Contrast	0.626	0.458	0.0002
Low luminance trials only			
Attention	0.108	0.753	0.0001
Contrast	0.161	0.701	0.0004
Attention:Contrast	0.043	0.841	0.00001

As can be seen from the Figures and Tables above, not only were any attention effects not expected from the plots, they also were in any way statistically significant. This contradicts Suizu's finding discussed in the end of Chapter 1: that at least in darker conditions attentional effects are to be found. Few were the alterations done from Suizu's experiment, so to what could be such a discrepancy be attributed?

A first hypothesis was that this disarray was due to either a difference in participants or a difference in methodology, more specifically, a difference in the criteria used to judge a reaction time as an outlier or not. To test this, three additional analyses were conducted, altering the data that was used as input to the ANOVAs: 1) Rerun the analysis with only data from overlapping participants of Suizu and the present study (N = 4); 2) Rerun the analysis using Suizu's outlier criteria (described in Chapter 2) (N = 7); 3) Rerun the analysis with both only overlapping participants and Suizu's criteria (N = 4). These resulted in: 1) a marginally significant attention main effect in the 3-way ANOVA ($F = 9.201$; $p = 0.056$; $\eta^2 = 0.021$), a significant attention main effect in the low contrast trials 2-way ANOVA ($F = 19.700$; $p = 0.021$; $\eta^2 = 0.022$) and a marginally significant attention main effect in the medium luminance 2-way ANOVA ($F = 9.292$; $p = 0.055$; $\eta^2 = 0.025$); 2) no significant effects; 3) a significant interaction between all factors in the 3-way ANOVA ($F = 5.993$; $p = 0.039$; $\eta^2 = 0.031$).

In this manner, it was possible to confirm that just by changing the participants to the overlapping ones some statistical effects started to appear, and that if the experiment was to be done with participants similar to these, a clearer investigation could be expected.

The overlapping participants were participants A, B, E and F, and thorough comparison between their reaction time distributions suggested that a good describer of a suitable participant would be: 1) hit rate of 80% or more in both voluntary and involuntary attention trials; 2) upper bound outlier removal criterion ($\mu + 2.5\sigma$) equal to or smaller than 1.3s. These two conditions where then used as the participant removal criteria in the confirmatory run. Since the application of these criteria to the present participants would decrease the total N from seven to three, seven more participants were called to take part in the experiment, from which three were eligible to go through the training session to the testing sessions. Their reaction time distributions are presented in Figures A.8 to A.10.

A.4. ANOVA R source code

The R source code for running the ANOVAs in the present study is as follows:

List A.1. 3-way repeated measures ANOVA

```
ezANOVA(data=trials, dv=RT, within=c(Attention,L,Contrast),
wid=Participant, type=3, detailed=TRUE)
```

List A.2. 2-way repeated measures ANOVA (low contrast trials as an example)

```
ezANOVA(data=filter(trials, Contrast=="Low"), dv=RT,
within=Attention, wid=Participant, type=3, detailed=TRUE)
```

The source code above requires the “ez” package, the “tidyr” package and the “dplyr” package. Each ANOVA was run 101 times with random sampling of the trials for balancing purposes. The results presented in the main text of this thesis are the ones with the median attention F-value.