Martine Aune

Effect of response set-up and task design on response times

Master thesis in psychology – learning: brain, behaviour, environment

Main supervisor: Ute Gabriel

Co-supervisor: Jonathan Kim

Trondheim, October 2017

NTNU

“Cogito ergo sum”

- René Descartes

Acknowledgments

I would like to express my sincere appreciation to my Supervisor, Professor Ute Gabriel, for her constant guidance and encouragement, without which this work would not have been possible. For this support, I am truly grateful. I am also grateful for all written feedback from my co-supervisor Jonathan Kim, and to the SCaLa group at NTNU for their encouragement and moral support.

The outmost gratitude and thanks to the research project "What's so sexy about degenderizing language? Investigating gender representations of readers and listeners in Norwegian, Finnish and French" for letting me modify their experimental design and compare my results directly with their results. Especially to Anton Öttl, for always being available, positive minded and very, very helpful. You are great!

Thanks to friends and family for never giving up on me. *I did it*.

Preface

This thesis is submitted in partial fulfilment of the requirements for the degree Master of Science in Psychology, Learning: Brain, Behaviour, and Environment, for the course PSY3914 Master Thesis Learning – brain, behaviour, environment.

This thesis has utilised an experimental task developed by Öttl et al. (2016; presentation) in a research project named "What's so sexy about degenderizing language? Investigating gender representations of readers and listeners in Norwegian, Finnish and French" (funded by NFR, FriPro project nr. 24088, Principal Investigator: Ute Gabriel). The experimental response setup/layout has been developed and implemented by Martine Aune with support by Anton Öttl. All data has been collected by Martine Aune. Anton Öttl provided an R-Script to convert the data into an SPSS-compatible format. Data analysis and interpretation was performed by Martine Aune.

TABLE OF CONTENTS

Abstract

This thesis utilised an experiment originally developed and applied by a research project investigating gender representations of readers and listeners in Norwegian, Finnish and French" (funded by NFR, FriPro project nr. 24088), referred to as the original experiment. The aim of this thesis was to further explore the role of task design and response set-up on response time (RT) in a modified version of that experiment. The response set-up was modified from a combination of bimanual- and unimanual responses to a full unimanual response design. 34 participants were recruited. They had to complete two different tasks over five experimental blocks: a button task (BT), where they had to press the corresponding response button to the stimuli presented (MM, FF, FM/MF), and a face pair decision task (FPDT), where they had to categorise face pairs according to gender (male faces, female faces, or both male- and female faces) before pressing the response button corresponding the presented face pair. Common for both tasks were the use of one overlapping category (FM/MF in the BT, a male- and a female face pair in the FPDT). The main focus of the thesis was the first BT and the first FPDT. In both task the overlapping category had significantly longer RT than the non-overlapping categories ($p < .001$). Current response set-up provided faster RT than the original response set-up in the BT ($p = .010$, contrast of 53 ms), RT between the original response set-up and the current response set-up in the FPDT was non-significant ($p = .087$), contrast indicates that the original response set-up provided faster RT by 93 ms. In conclusion, RT is dependent on both task design and response set-up.

Key words: response times, response set-up, experimental psychology, task design, cognitive psychology, information processing

The research project “What’s so sexy about degenderizing language? Investigating gender representations of readers and listeners in Norwegian, Finnish and French” (funded by NFR, FriPro project nr. 24088) is examining cross-linguistic differences between Finnish, French and Norwegian in order to determine how language activates gender representations. The aim of this project is to further fill the gaps in the literature on degenderizing by neutralisation on mental representations of gender through a variety of experimental research on written and spoken language comprehension. A secondary aim of the project is to uncover language mechanisms at different levels of cognitive processing, ranging from verbal to perceptual levels of representations.

A substantial part of the experiments in the FriPro-project is focused around the development of a face pair decision task (FPDT, Öttl et al.; 2016). The experiment undertaken in the course of this thesis is intended to continue exploration into the best form of response execution for the FPDT.

Background: Face Pair Decision Task

The FPDT was originally developed by Öttl, Behne, Gygax & Gabriel (2016) to be used in a series of experiments examining sentence to image priming of gender information, and is a complex categorisation task with a three-alternative forced-choice response time (RT) test. A three-alternative forced-choice RT test is a choice RT task (CRT), and is used when participants are presented with one of two (or more) stimuli to which they must make a coordinated, compatible response to before progressing to the next trial (Burke et al., 2016). The primary goal for the FPDT was that it should be adequate for registering RT for three categories in the same response set-up: male, female or male- and female face pairs. The solution adopted by Öttl et al. (2016) was to combine unimanual responses with bimanual responses. Where there was two unimanual responses (responses had to be given with either the left hand or the right hand) and one bimanual response (responses had to be given with both hands). When conducting the FPDT, Öttl et al. (2016) had two overall aims. Firstly, to investigate the motoric aspect of the response registration and task implementation. Here the main issue was the use of two unimanual and one bimanual response. As one could expect that responses given only by one hand would be easier than giving responses with both hands simultaneously due to the motoric coordination of responses. The third goal was to evaluate the gender priming from text to picture. In addition to the RT obtained in the task, eye tracking data were also collected during the experiment (to obtain gaze contingency, to be able to analyse where on the screen participants were looking and for how long).

In the FPDT participants were instructed to classify face pairs according to the gender of presented face pairs, i.e., two male faces (male condition), two female faces (female condition) or one male and one female face (mixed condition) (See figure 1 for an illustration). Each face pair had a corresponding response button, and participants were instructed to press the button corresponding to the face pair presented on the screen in front of them. The male condition and the female condition required unimanual responses, that is two simultaneous button presses from either the left- or right hand. The mixed condition is an overlapping category which means that it is a combination of the female condition and the male condition, and it required a bimanual response. And as such, the bimanual response reflected both face genders in the response execution.

The FPDT was presented in three blocks that varied in prime type. In the first block no prime was presented prior to the face pair decision task. In the second block by role nouns (stereotypically male/female, such as “Sykepleierne” [Nurses], “Fotballspillerne” [Football players], “Mekanikerne” [Mechanics]) were used to prime gender, and in the third block person names (male/female, such as “Henrik” and “Ida”) were used to prime gender.

Figure 1. Illustration of the screen as shown to the participants during the FPDT, showing the mixed condition.

Since the FPDT required that each hand was associated with one response category, one potential weakness was that participants' handedness might interfere with the results. To avoid this problem an additional baseline task was utilised, referred to as the button task (BT). The BT utilised the same response set-up as the FPDT, and this will be further discussed later in this thesis. Participants had to complete the BT at the start of the experiment, before starting on the FPDT, as well as after finishing the last FPDT. Thus, it also provided the opportunity to evaluate how consistent participants' handedness was during the course of the experiment. The BT can be considered practice for the FPDT as the response set-up was the same. In addition, the letters applied as stimuli (M and F) are also generally used to mark gender, thus preparing the participants for the categorisation in the FPDT.

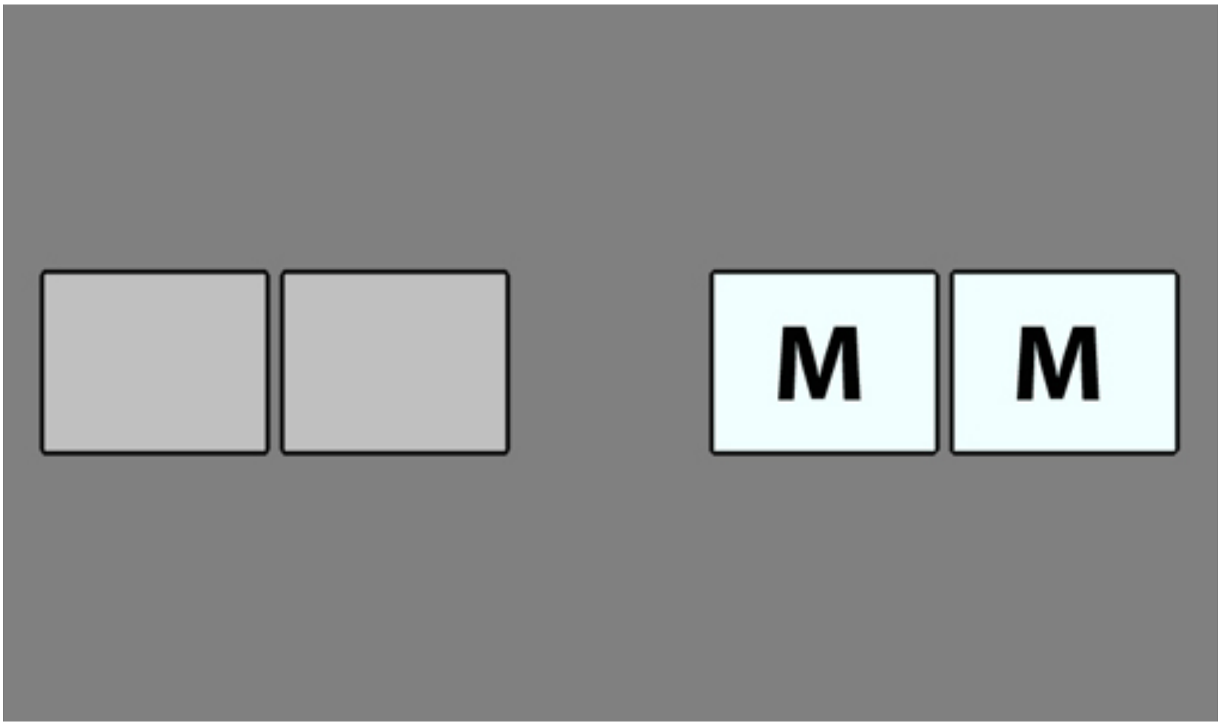


Figure 2. Illustration of the screen as shown to participants during the button task, showing response alternative "MM".

In the BT participants were presented with an image illustrating the response box. On the presented image of the response box one response alternative would be shown. The response alternatives were either MM, MF or FF (or FF-FM-MM, the order was counterbalanced across participants) (see figure 2 for an illustration). When presented with a response alternative they were told to press the button corresponding the that response alternative on the response box, thus link the visual stimuli with the correct motoric response.

In the BT the results showed a significant difference between response alternative “MM” and response alternative “MF” (443 ms vs 479 ms respectively, $p < 0.0001$). There was also a significant difference between response category “FF” and response category “MF” (448 ms vs 479 ms respectively, $p < 0.0001$). Meaning that when participants were presented with response category “MM” or “FF” they had shorter RT than when presented with response category “MF”. There was no significant difference between response alternative “MM” and response category “FF”. In the FPDT, the results showed similar results. Both the female condition (889 ms) and the male condition (890 ms) was significantly different from the mixed condition (1035 ms, $p < 0.0001$ for both contrasts), but no significant differences were found between the male condition and the female condition. As such, when participants were presented with a same gender face pair they had shorter RT than when presented with a mixed face pair. Öttl et al. (2016) concluded that the results from the BT (at the start- and at the end of the experiment) shows that the bimanual response was more difficult than the unimanual responses, and that this effect was not reduced by practice. In addition, when taking this effect into account, the bimanual response was more difficult in the FPDT compared to the unimanual responses. Therefore, categorisation of mixed faces was more difficult than categorisation of only male faces or only female faces. Importantly, taking these baselines tasks into consideration in the final analyses, the gender priming was deteceted, demonstrating that the developed task is adequate for the intended purposes. For the purposes of the experiment undertaken in this experiment, only results from the first BT and the FPDT without a prime will be explained and analysed.

Background: Response Time

RT is a measurement of the time elapsing between the onset of a stimulus and a response to it (Baayen & Milin, 2010; Jain, Bansal, Kumar, & Singh, 2015). RT was introduced by the Dutch ophthalmologist Franciscus Donders (Baayen & Milin, 2010; Donders, 1868). RT is utilised to measure reactions to stimuli with the goal of gaining external measurements of internal processes (Semmelmann & Weigelt, 2016). RT is assumed to reflect the time it takes to interpret a stimulus, get information from memory, and initiate a muscular response to the stimuli. For an individual to make a muscular response, such as a finger press, the nervous system must recognise a stimulus, after which the sensory neuron will relay a message to the brain[1]. When the message reaches the brain it travels from the brain via the spinal cord, which then reaches the individual's hand and fingers. The motor neurons tell the hand and fingers how to react. This time it takes from recognising the stimulus to making a response is used as a measure of information processing and/or processing speed (Godwin, Menneer, Cave, Thaibsyah, & Donnelly, 2015).In order to make conclusions about information processing and/or processing speed it is important to minimise and eliminate other factors and variables that can influence the RT. As RT is influenced by a range of factors it is close to impossible to attribute RT only to the factor of interest to a study (Kyllonen & Zu, 2016). This is a major concern with RT research.

RT is used within a broad range of experiments within experimental psychology and its design varies accordingly. Because the design of the experiment varies depending on the task it is important to be aware of some of those variables causing the variation. For the FPDT the design included a complex response set-up with a combination of bimanual and unimanual responses, and current experiment aims to simplify this response set-up.

Inter individual Variability: Age and Gender

Individual features of the participants participating in RT experiments can influence their RT results (Baayen & Milin, 2010). It is well-established that participants' response speed is influenced by age (Deary, Liewald, & Nissan, 2011), as RT to a stimuli becomes markedly slower with age beginning from young adulthood, accelerating after middle adulthood (Deary & Der, 2005; Der & Deary, 2006).

The influence of the participants' gender on RT is not as well-established. Some researchers argue that in general responses from male participants are faster than responses from female participants (Adam et al., 1999; Der & Deary, 2006; Jain et al., 2015). However, in an experiment similar to the FPDT, Godard and Fiori (2010) found that both male and female participants were equally capable of making accurate responses, but that male participants had longer RT than female participants.

To summarise, it is important to recruit participants from only one age group because the participants age can influence the RT obtained from that participants. Having participants from a wide age range can produce results that reflects the experimental effect as well as an

[1] or to the spinal cord, but reactions that only travel to and from the spinal cord are often called spinal reflexes or cord-mediated reflexes, and will not be discussed here.

increase/decrease in RT due to age, thus making it difficult to draw conclusions from the results. To avoid any ambiguity or bias that could result from having no control over the age of participants', current experiment will only test participants between the age of 18 and 30. In addition, as it is hard to argue for having only male or only female participants considering the inconsistencies in the literature, the current experiment aims to have an equal number of male and female participants to avoid a bias in results due to differences between the sexes.

Complexity of Task Design

There are a number of factors that need to be considered and accounted for when designing an RT experiment to avoid making incorrect conclusions and comparisons. Experiments with only one stimulus and one response, such as in simple RT (SRT) procedures and recognition RT procedures, provide faster responses than if there are more response options, such as in CRT procedures. Increasing the number of response options will respectively give slower responses. This is known as Hick's law. It is used to estimate the time it takes for a subject to make a decision as a result of the possible choices available. Under Hick's law, increasing the number of choices increases the decision time logarithmically until it gets to a point where RT remains constant even if more responses are added to the design. Hick's law is described through the equation '$RT = a + b \log^2 (n)$', where "a" is the total time that is not involved with decision making, "b" is an empirically derived constant based on the cognitive processing time for each option, and "n" is the number of equally probable alternatives. Hick's law assesses cognitive information capacity in choice reaction experiments, where the amount of time taken to process a certain amount of information capacity is known as the "rate of gain of information" (Usher, Olami, & Mcclelland, 2002).

Hick's law is dependent on a few factors. Participants can trade speed for accuracy, referred to as the speed-accuracy trade off. The speed-accuracy trade off states that if a participant is instructed to focus on giving correct answers, RT will be slower but responses will be more accurate than if the participant is instructed to respond as quickly as possible, indicating that an individual's willingness to respond slowly and make relatively few errors compared to their willingness to respond quickly and make relatively more errors is dependent on the instructions given before starting the experiment. This occurs regardless of the number of response alternatives. This trade-off is also an important consideration for the methodological design and in the interpretation of the findings. If the goal of the experiment is to determine the minimum amount of time that is required to produce a correct response on a given task, and the instructions were given with an emphasis on giving correct responses, it can be assumed that the results would look different than those obtained if the participants had been asked to respond as quickly as possible. In experiments where the goal is to measure processing speed and information processing, such as the FPDT, one wants the participants to maximize performance on both speed and accuracy. As such, the importance of speed over accuracy (or vice versa) needs to be explicitly stated by the experimenter or in the experimental instructions provided to the participants to avoid the participants increasing their response time at the cost of reducing the accuracy of their responses, or to decrease RT in order to increase their overall accuracy level (Proctor & Vu, 2003). The length of a participants' RT can tell us something about the participants' information processing. However, if the participants received no instructions as to speed or accuracy one cannot clearly draw conclusions about the information processing, as it can also reflect carefulness (Kyllonen & Zu, 2016).

It has been observed in RT research with different types of stimuli, similar to the FriPro FPDT, that participants have faster RT when there is a "run" of several identical stimuli compared to when the different types of stimuli appear in a mixed order (Sanders, 1998). This effect of a given sequence of experimental trials is referred to as trial-by-trial sequential effects (Baayen & Milin, 2010), and in terms of the changes in the course of the experiment the trial-by-trial sequential effects is the most important one in terms of the design.

To summarise, the role of Hick's law is important when considering the FPDT, as it is dependent on the speed-accuracy trade-off. The speed-accuracy trade-off is not straightforward in the FPDT as participants are asked to respond as quickly as possible while still trying to give correct responses. In addition, the FPDT is designed to avoid a trial-by-trial sequential effects by having randomised stimuli, where the program software (E-prime) is coded to present the face pair stimuli in a random order. This thesis aims to further explore the response set-up in the FPDT, which requires the FPDT to be tested on new participants to avoid the role of practice.

Task Type

RT is used within a broad range of experiments resulting in different categories of RT procedures. The most common research paradigms involving RT are SRT and CRT. For experiments examining SRT responses participants are presented with a single stimulus and have to make a single response. The FPDT utilises CRT. The response to CRT stimulus requires more time for stimulus identification and response selection than the single response in SRT (Miller & Low, 2001). A CRT procedure can be further subdivided into type of task. The type of task is related to what the participants actually have to do, for example in the FPDT participants have to categorise face pairs. When classifying by type of task we are interested in what the participant has to do to decide on a response and how it affects the RT results. More specifically, we are interested in the decision the participants have to make regarding the presented stimulus and the response execution. Common for all CRT experiment is that participants cannot skip one trial; they have to give a response to progress to the next trial. However, the decision participants have to make regarding the presented stimulus and the response execution varies between CRT experiments. This is interesting because a cognitive information processing event takes place during the time interval between being presented with a stimulus and given a response (Kyllonen & Zu, 2016).

In a study conducted by Antonova et al. (2016) participants were presented with four white outlines of squares on a black background positioned in a horizontal line approximately in the middle of the screen, four keys on four-key response board corresponded to the different squares (Figure 3).

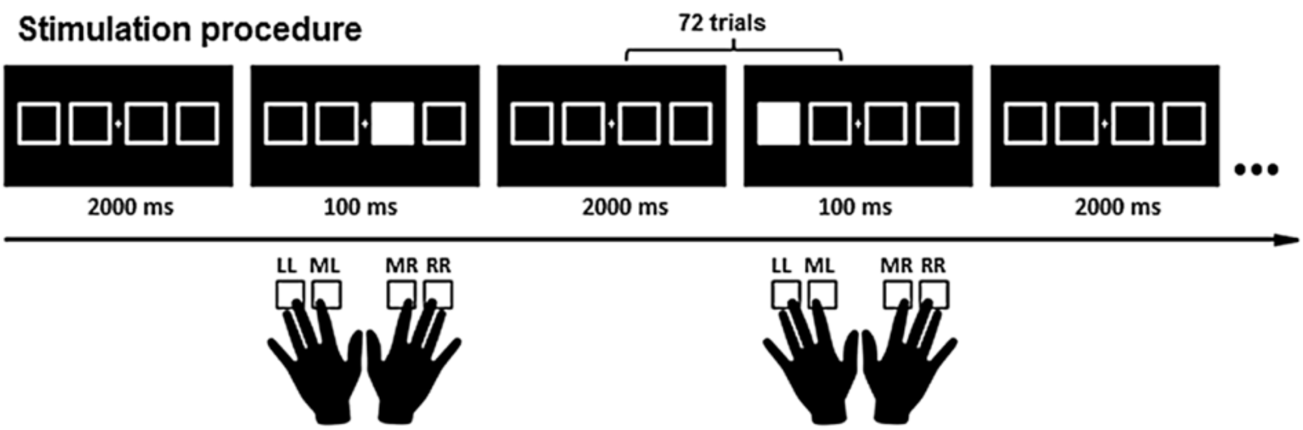


Figure 3. Example of stimuli and stimulation procedure from Antonova et al. (2016).

In each trial one square was filled with white to indicate the target, and participants were instructed to press to corresponding key on the keyboard. Participants in this experiment only had to make two decisions; firstly, which square was filled with white, and secondly, what button to give the response on. Participants were instructed to press the button corresponding to the white square in front of them. Similar task was applied in a study conducted by Deary, Liewald & Nisssan (2011). In their experiment four white squares were positioned in a horizontal line approximately in the middle of the screen and in each trial a cross appeared in one of the four squares. Participants then had to press the button corresponding to the square where the cross was presented. Thus, in both of the above mentioned experiments participants had to link a visual signal with a specific motoric response.

Figure 4. Examples of the schematic happy, neutral, and sad faces used in Leppänen and Hietanen (2004).

But, not all experiment is as straightforward as those ones. Leppänen & Hietanen (2004) utilised a three-choice RT paradigm, and instructed participants to categorise faces according to schematic drawings of facial expression (happy, neutral, and disgust; Figure 4). The responses were obtained through participants' pressing correspondingly marked response buttons. This task involved participants making a conscious decision as to which facial expression was presented, then pressing the corresponding button to the facial expression presented.

The experiment conducted by Leppänen and Hietanen (2004) can be argued to be more difficult than the experiment conducted by Antonova et al. (2016). In the latter participants only had to decide where the stimulus was presented, whereas in the former participants had to decide which facial expression of the three possible facial expression was presented. This can be discussed in terms of current experiment, both the FPDT and the BT.

The BT, outlined in section 1, requires participants to press a button corresponding to the stimuli presented in front of them. There are strong similarities between the BT and the experiment conducted by Antonova et al. (2016). Both experiments require the participants to press a button corresponding to a stimulus presented visually. However, in the BT the presented stimuli, instead of being filled with white as in the study by Antonova et al. (2016), has the additional marking with letters (MM/MF/FF). Meaning that all stimuli in the experiment conducted by Antonova et al. (2016) are visually the same, and that the stimuli in the BT are visually different requiring participants to categorise the stimuli before responding. Thus, adding an additional factor to consider in terms of the information processing used in that task. One can therefore argue that the BT needs more information processing than the experiment by Antonova et al. (2016). The FPDT, outlined in section 1, on the other hand is a much more complex task than both the BT and the task by Antonova et al. (2016). In the FPDT participants are presented with face pairs who are either both male, both female, or one male- and one female. It can be compared to the experimental task conducted by Leppänen & Hietanen (2004). They have in common that they have some sort of face as stimuli. But where participants in the experimental task by Leppänen & Hietanen (2004) have to decide on which facial expression is presented before responding, participants in the FPDT have to decide which gender is presented and whether both faces shown are the same or not before responding. Thereby adding another factor for the participants to evaluate before responding, which can arguably need longer information processing and processing speed.

Overlapping Categories as a Challenge in Response Execution in Choice Response Time Paradigms

The fair majority of RT research has been conducted with unimanual responses. A common feature of this type of experiments is that they involve a simple decision about some feature of a stimulus that is expressed as a choice between two alternative responses (Ratcliff & Smith, 2004). However, an experiment's complexity is not limited to easy and traditional simple SRT or CRT. By increasing the number of response alternatives more complex experiments can be designed. Where they in simpler paradigms purely apply unimanual responses, an increase in response alternatives can open up for a combination of both unimanual- and bimanual responses, or simply more options for either a bimanual- or a unimanual response.

Three and four Response Alternatives

The process of participants' responding during CRT experiments can be subdivided into different parts. Firstly, a participant sees the stimulus, then they categorise it and decide on what response to give, then finally give their selected response. To give a response requires participants to press a button corresponding to the stimulus presented, but the response set-up applied varies. For RT studies with four response alternatives given with two hands, the trend is to use both hands and two responses on each hand to indicate the responses. Thus, equally distributed between hands. This way of designing the response set-up was outlined above in the experiment conducted by Deary, Liewald & Nissan (2011) and Antonova et al. (2016). However, the colour scheme and how the target is indicated to the participant varied in these two experiment. Both Deary, Liewald & Nissan (2011) and Antonova et al. (2016) instructed participants to press a key (on a keyboard or on a response board) when something changed on the screen (either cross or light change).

But as we saw earlier in the experiment by Leppänen & Hietanen (2004) experiments are not limited to tasks where there are an even number of response alternatives and where the response alternatives are equally divided between both hands. In the study conducted by Leppänen and Hietan (2004) participants were instructed to consciously make a decision about what they saw before responding to presented stimuli. This experiment examined the recognition speed advantage for happy faces. They utilised a three-choice design to explore a number of elements related to the judgment of facial expressions. Of interest to this thesis is the response set-up design. Their primary aim was to investigate differences in the processing speed of happy and disgusted facial expressions. While a two-choice paradigm would have been sufficient for this aim, the authors opted to include a third option (emotional neutrality) due to methodological reasons. Specifically, they argued that a two-choice RT paradigm might prompt participants to use a strategy where they classify the faces as either “happy” or “not happy”, and subsequently do not process the “disgust” emotion. Therefore, they used a three-choice RT paradigm, stating that this would make it less probable for participants to utilise such strategies. Participants were requested to decide what facial expression was presented on a screen, and make a response reflecting their answer. Responses were to be given with either the right- or left hand, where the index finger of the left hand was used to press the left-hand key, the index finger on the right hand the middle key and the middle finger on the right hand the right-hand key. A design where the response options are not equally divided between the hands, rather predominantly applying right handed responses.

To summarise, this section looked at the response set-up in CRT research with four- and three response alternatives. It found that when there are four response alternatives the trend is to use four response keys/buttons equally divided between the right- and left hand. With three response alternatives one the other hand, response alternatives were not equally distributed across the right and left hand, but rather split into two responses with one hand and one with the other hand.

Two + One Response: The Case of an Overlapping Response Category

Up until now we have seen how different tasks affects the RT and how the response set-up has been applied in these tasks. But, as opposed to the other experiments introduced both the BT and the FPDT has the additional component of an overlapping response alternative. The responses in the BT and the FPDT (See figure 5 for an illustration) consisted of two unimanual responses, and one bimanual response. The two unimanual responses were used to respond to either the two male faces or two female faces in the FPDT, and to response categories “MM” and “FF” in the BT. The unimanual responses required participants to press two buttons located next to each other on a response box with the second and third finger of one hand simultaneously.

Figure 5. The response set-up for the button task and the FPDT. From the left, third- and second finger presses for response category “MM” in the button task, or for two male faces in the FPDT. To the right, second- and third finger presses for response category “FF” in the button task, or for two female faces in the FPDT. For the bimanual response, responses were given by pressing one finger of each hand (second finger of left hand, second finger of right hand).

The bimanual response was for one male face and one female face in the FPDT, and to response category “MF” in the BT. Bimanual responses were given by pressing a button with one finger of each hand simultaneously (second finger of left hand, second finger of right hand). The use of two button presses for the unimanual responses was applied in effort to make the response set-up and thus result more similar to the set-up and results from the bimanual responses. In addition, the consistent use of two response buttons is due to the use of the overlapping categories, where the bimanual response incorporated one element each from both unimanual responses.

The goal of the current research is to further explore the methodology used by Ottl et al. (2016). Altering an experimental methodology can provide valuable insights into the topic(s) investigated or explored by revealing different aspects and new ways to approach the same experimental topics. It can serve as a mean to discredit or support research, and can reveal more useful and reliable ways of conducting research on specific topics or with specific techniques. For new innovative methods, such as the use of overlapping categories and overlapping responses in the FPDT and in the BT, it provides support and a more nuanced understanding of the methodology. In the experiment undertaken in the course of this thesis the FPDT and the BT will remain the same, but response set-up will be altered in order to explore a different way of conducting an experiment with an overlapping category, both in stimuli and in response design. The current experiment aims to explore the use of an alternative, namely unimanual response set-up in the FPDT and in the BT.

A response set-up will be tested in which all responses are given with the participant's dominant (here: right) hand. The response box layout will be modified as follows: Firstly, current experiment implemented one button press, as opposed to two in original experiment. For the unimanual responses the corresponding buttons were either button 1 or button 3. The corresponding button for the original bimanual response were button 2. Thus making the full modified response set-up unimanual. The corresponding keys were assigned to fingers accordingly; MM – second finger right hand, MF – third finger right hand, and FF – fourth finger right hand. Secondly, as a consequence of changing the response set-up to only unimanual responses the stimulus material for the BT were changed to fit with the modified response box set-up. Thirdly, because of the full unimanual response set-up only right handed participants were tested to make sure that participants responded with their dominant hand. Fourthly, all instructions to both the BT and the FPDT were modified to fit new response box set-up.

In the original experiment participants had to press two buttons simultaneously, either unimanually with one hand or bimanually with two hands (as outlined above). The current experiment changed the response set-up from a combination of unimanual- and bimanual responses to only unimanual responses. The argument for having overlapping buttons for the bimanual response was that it should reflect and match the overlapping use of visual stimuli, and such make it easier for the participants. I argue that by having the middle button be the bimanual response the same logic can be applied without adding to the motoric noise and level of difficulty for the participants. Thus, arguing that the original experiment can be conducted with a simpler set-up. In addition, it has been found that bimanual responses are more difficult than unimanual responses (cf. Shen 2005). Therefore, a full unimanual response set-up can possibly reduce coordination difficulties associated with using a combination of unimanual-and bimanual responses.

As stated, the current experiment will not change the experiment itself, only the response set-up. Changing the response set-up can possibly reduce the information processing associated with response registration, as the participants will potentially focus more on actually categorising face pairs correctly and less on remembering the correct buttons to press. As outlined earlier, in RT experiments there is a time interval between being presented with a stimulus and making a response. Where the less the participants have to do, the less information processing is needed. So for the BT participants have to identify which response category is presented, then press the corresponding response button. However, in theory participants do not actually have to identify the response category before responding. Because the response categories shown are always in MM-MF-FF order and the corresponding response button would always mirror the stimuli presented, so that response category "MM" would be the left response button, response category "MF" would always be in the middle response button, and response category "FF" would always be the right response button (or opposite, FF-FM-MM, the order is counterbalanced across participants). Overall the BT should require less information processing than the FPDT.

In the FPDT participants have to decide which gender is presented and whether both faces shown are the same or not before pressing the response button corresponding to the presented face pair. It requires more information processing than the BT, as participants have to decide on which face pairs are presented in front of them, requiring participants to identify face pair in terms of gender before pressing the corresponding button to the identified face pairs.

I expect that the modified response set-up provides an easier response set-up for participants than the original experiment because participants no longer need to coordinate motoric responses from two hands but rather coordinate three fingers on one hand. Thus, cause less motoric difficulties, hand- and finger coordination and moving of hands, resulting in faster overall RT compared to the set-up in the original experiment. Therefore, my first hypothesis is that if the significant difference between the unimanual and bimanual responses in both the BT and the FPDT can be explained by the motoric difficulties associated with the bimanual response, using current set-up instead, should lead to results showing no significant differences at alpha level above .20 ($p > .20$.). Further, my second hypothesis is that if the modifications made to the original experiment removes the motoric bias associated with a combination of both unimanual and bimanual responses, using current set-up instead, the results should provide faster RT in each response alternative in both the BT and the FPDT compared to the original BT and FPDT.

Method

Participants

34 adults (15 women, 17 men; mean age 23.9, SD = 3.6) participated in this experiment. Participants were a convenience sample recruited at the Norwegian University of Science. All participants were self-reported as being right handed based on the results of a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal vision. Before the experiment all participants gave their informed written consent. Each participant was compensated with a 26 NOK coffee voucher from the university coffee shop for participating in the experiment.

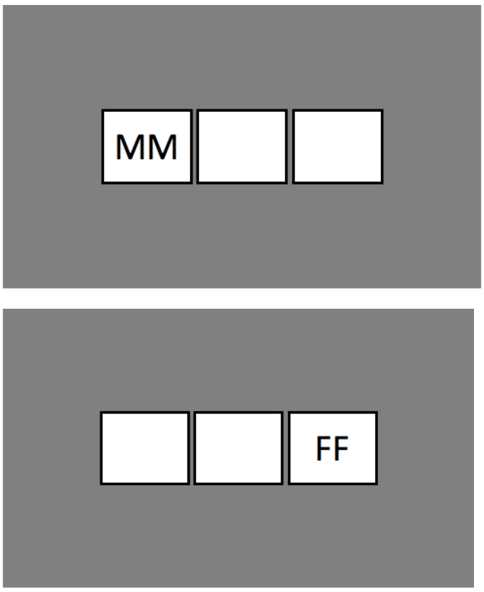
Materials

Materials were adapted from the research project "What's so sexy about degenderizing language? Investigating gender representations of readers and listeners in Norwegian, Finnish and French (NFR FriPro 24088; PI: Ute GABRIEL)". As part of their research they conducted an experiment that used a FPDT and BT, as well as a recognition task (Öttl, Behne, Gygax, & Gabriel, 2016).

In the face pair decision task participants were instructed to classify face pairs according to the gender of presented face pairs, i.e., two male faces, two female face or a male and a female face. The face pair decision task was presented in three blocks. In the first block no prime was presented prior to the face pair decision task. In the second block gender information was primed by role nouns (stereotypically male/female, such as "Sykepleierne" (Nurses), "Fotballspillerne" (Football players), "Mekanikerne" (Mechanics)). In the third block gender information was primed by person names (male/female, such as "Henrik" and "Ida"). Person names were pretested for masculinity/femininity (Öttl ... 2016: presentation). Role nouns were based on an earlier study by Misersky et al. (2013). Faces had been obtained from the Chicago Face Database, cropped to oval shapes to remove global cues and further pretested to account for how rapidly the images were perceived as male or female. See figure 6 for example pictures of faces as presented to the participants during the face pair decision task.

Figure 6. Example of the pictures as presented to the participants in the face pair decision task. From the left; two male faces (MM), a male- and a female faces (MF), two female faces (FF). Images adapted from Chicago Face Database (http://chicagofaces. org/).

In the BT participants were asked to press the response alternative displayed on the screen in front of them on a response box. See Figure 7 for an example of how the response alternative were displayed to the participants.



A)

B)

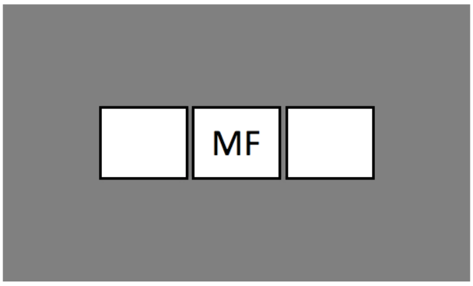
C) 

Figure 7. Example of how the response alternatives were displayed to the participants in the button task. A) response category "MM", B) response category "FF", C) response category "MF".

The recognition task consisted of control questions and was presented at the end of the experiment. Contrary to the rest of the tasks in the experiment participants were asked to take the time they needed to respond, and not to respond as quickly as possible. They tested participants' recognition of words used as primes. Participants were shown person names and role nouns that had or had not been presented earlier in the experiment as the primes for the face pair decision task. The participants were asked to press "yes" if the role name or person name had been presented earlier in the experiment, or "no" if they had not seen the role name or person name previously in the experiment.

Apparatus

This experiment used a modified response lay-out as compared to the original response lay-out. In contrast to a combination of bimanual- and unimanual response options (Öttl et al, 2016: presentation) the current experiment applied right hand unimanual responses only. Each response alternative had a corresponding finger press from either the right index finger, right middle finger or right ring finger (see figure 8). Responses were recorded using Chronos response box, a multifunctional response and stimulus device which allows the accurate collection and verification of tactile, visual, and analog responses.

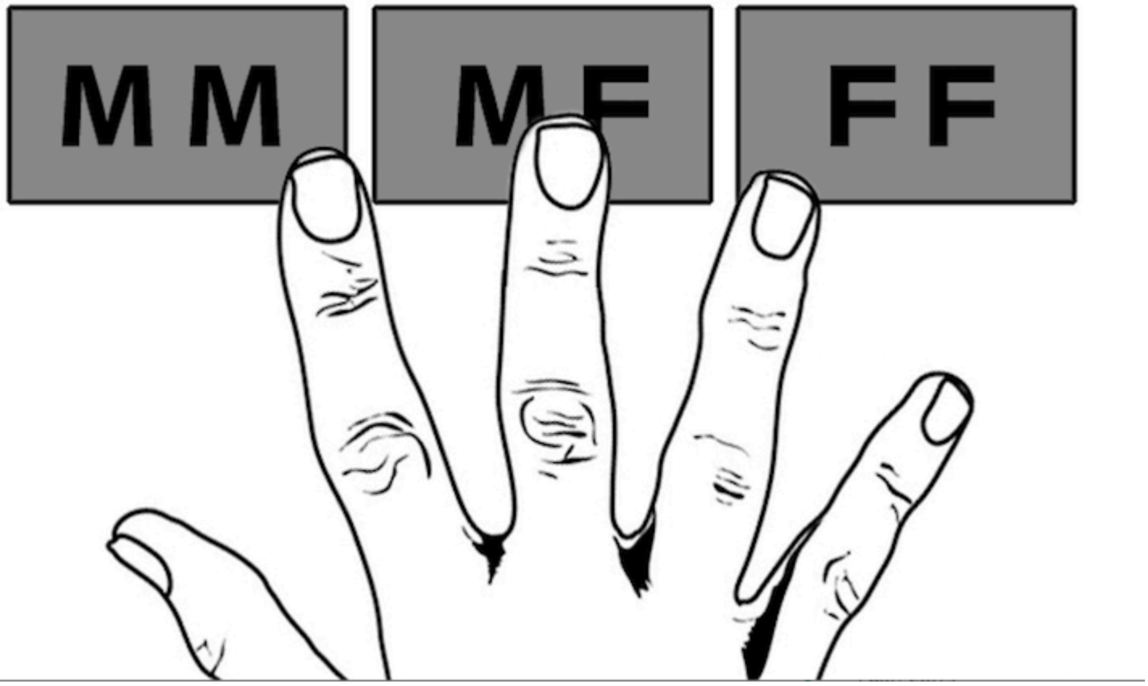


Figure 8. An illustration of the correct set-up of the fingers of the response box. With the right second finger on response alternative “MM”, right third finger on response alternative “MF”, and right fourth finger on response alternative “FF”.

Stimulus presentation and response recording were controlled by the E-prime Version 2.0 software. Participants’ eye movements were recorded with an iView X 2.8 SMI RED500 eye-tracking system. Stimuli were presented in the centre of 23-inch display (Name, model). The display was viewed from a distance of between 50 and 70 cm and the head was stabilised using a chin- and forehead rest. Participants completed a seven-point eye-tracking calibration. To accept the calibration, the participant should be able to follow the computer cursor around the screen with their eyes. If the participant was not able to do this at a level that was considered good enough by the experimenter, i.e. being able to successfully follow the cursor around the screen to the top left, top right, middle, button left and button right while keeping their gaze very close to the cursor, the calibration was rejected and had to be repeated. On average the participants had to complete the calibration procedure between two and four times, depending on factors including glasses, make-up, fatigue and lenses. The calibration procedure was repeated until the researcher accepted the calibration procedure.

Procedure

Participants were tested individually in a small testing room (x m by x m). Participants were informed that the experiment was about “change of attention”, and were instructed to read and sign the consent form prior to starting the experiment and to ask if they had any questions. Next, they completed the self-report Edinburgh Handedness Inventory. Participants were then asked to rest their hands on the table in front of them, and instructed to rest their fingers on the buttons corresponding to the response alternatives in the experiment. Respectively, MM- right second finger, MF – right third finger, and FF - right fourth finger (for condition A, for condition B the opposite instructions were given). See figure 3 for an illustration. The participants' arms were bent at the elbows to minimise motoric movement and to ensure that the participants used only their fingers to respond. Speed of responding was emphasised as well as giving the right response. Throughout the experiment participants were asked not to correct errors. Response button set-up was counterbalanced across participants.

Participants completed six experimental blocks (see table. 1). The first block consisted of the BT with 45 trials (The first 15 trials were practice trials). The second, third and fourth block consisted of the face pair decision tasks with 45 trials each, preceded by no prime (block 2), role noun prime (block 3) or person name prime (block 4). The fifth block was a repetition of the first block with the BT. The experiment ended with a recognition task (block 6).

Table 1

Experimental outline

Block	Trial	Prime	Task
1	15+30	No prime.	Button task
2	45	No prime.	Face pair decision task
3	45	Role nouns (male/female) as primes.	Face pair decision task
4	45	Person names (male/female) as primes.	Face pair decision task
5	30	No prime.	Button task
6	40	No prime.	Control questions in a recognition task

For the BT (block 1, block 5) a fixation point appeared in the centre of the screen and had to be gazed at for 500 ms to continue to the BT. The participants' task was to respond to the target as fast as possible whilst still trying to provide the right answer. A BT trial is illustrated in figure 9. The answer (MM/FF/FM) was given as a unimanual right index finger press, unimanual right middle finger press or a unimanual right ring finger press depending on which alternative was shown on the screen. The target remained on the screen until an answer was given.

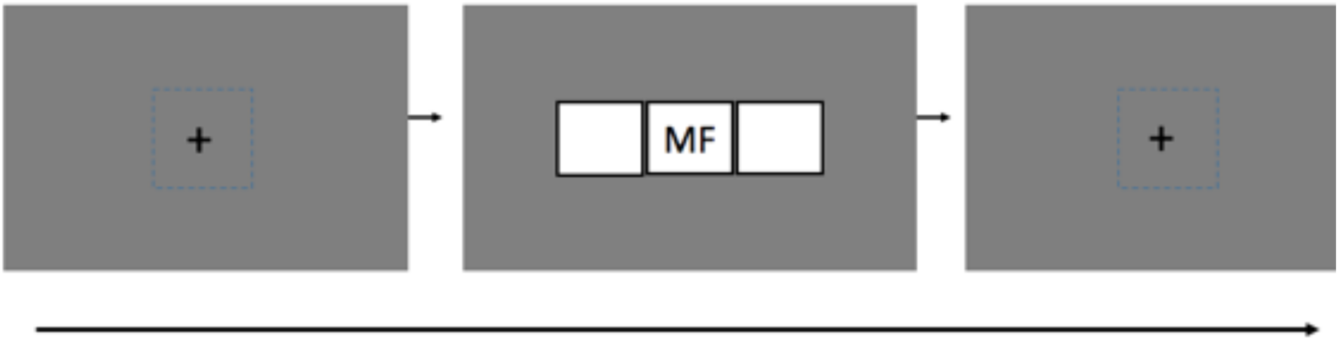


Figure 9. Illustration of a button task sequence of events showing one trial. From left: Screen with fixation point, had to be gazed at for 500 ms to continue to the button task. Screen as shown to participants after fixation point, showing the button to press on the response box. New fixation point after having given a response to the button task.

The face pair decision task in block 2 followed the same structure as the BT outlined above. The fixation point appeared and had to be gazed at for 500 mm before continuing to the face pair decision task. In the face pair decision task two faces were presented next to each other and the participant had to make a decision as to whether there were two male faces, two female faces or one male face and one female face, indicating which ones by pressing the corresponding button on the response box. The target remained on the screen until an answer was given (MM/FF/FM). A face pair decision task trial without a prime is illustrated in figure 10.

Figure 10. Illustration of a face pair decision task without a prime trial. From the left, screen with fixation point, had to be gazed at for 500 ms to continue to the face pair decision task. Screen as shown to participants after fixation point, showing the face pair decision task. New fixation point after having given a response to face pair decision task.

Before start of block 3 and block 4 participants were instructed to not respond to the prime, but to give a response in the same way as mentioned above to the target pictures. The target remained on the screen until an answer was given (MM/FF/FM). After finishing the five blocks and finishing the recognition task, participants were debriefed about the real purpose of the study. The experiment lasted for approximately 20 minutes per participant.

Results

Reaction times (RT, in ms) and accuracy (correct responses, in %) were recorded. Reaction time refers to the time between the target's appearance on the screen and pressing the button. Only RTs for correct responses were analysed. The analysis was conducted using repeated measures one-way ANOVA and 2 x 3 independent measures ANOVA. This differs from the approach used by Ottl (2016), who utilised mixed models. The data analysis involved three parts; firstly, an analysis of the data gathered in the course of this experiment. Secondly, a re-analysis of the data collected by Öttl et al. (2016) to allow for a comparison between the two datasets. Thirdly, a comparison between the data gathered in current experiment and the original experiment.

Analysis of My Data

Before analysing the data, incorrect responses and responses with RT falling outside the window of individual mean response time ± 2.5 SD were removed. On average 9.81 % of responses were removed. All tested participants performed above 50 % correct, thus according to our preregistered criteria, there was no need to exclude. Average accuracy was 77.5 (SD = 6.8). Although female participants performed better than male participants (79.3% versus 75.9% correct), the difference was not significant (*p* = 0.16).

Button task.

A repeated measures one-way ANOVA (SPSS) demonstrated a significant difference between the condition means ($F(2,62) = 25{,}121$, $p < .001$ [sphericity assumed]). This represents an effect size (η eta^2) of .45, showing that 45% of the variation in RT was accounted for by response category (descriptive statistics are given in table 2).

Table 2

RT for each response alternative in the button task

Response Category	M(SD)	95% CI
FF	397.01 (13.24)	[370, 424]
Mixed	420.78 (13.33)	[393, 447]
MM	397.49 (11.87)	[373, 421]

Note. CI = Confidence interval

Post hoc comparisons applying Bonferroni correction found that response category "female" resulted in significantly faster RT than response category "mixed" ($p < .001$), and that response category "male" resulted in significantly faster RT than response category "mixed" ($p < .001$). These differences demonstrated large effect sizes (d = 1.79 and 1.85, respectively). The difference in RT between response category "female" and response category "male" was not statistically significant ($p = 1.00$), and this comparison demonstrated a small effect size ($d = -0.038$). The results are illustrated in figure 11. These results do not support the hypothesis stating that there should not be any significant differences between response categories ($p = .20$).

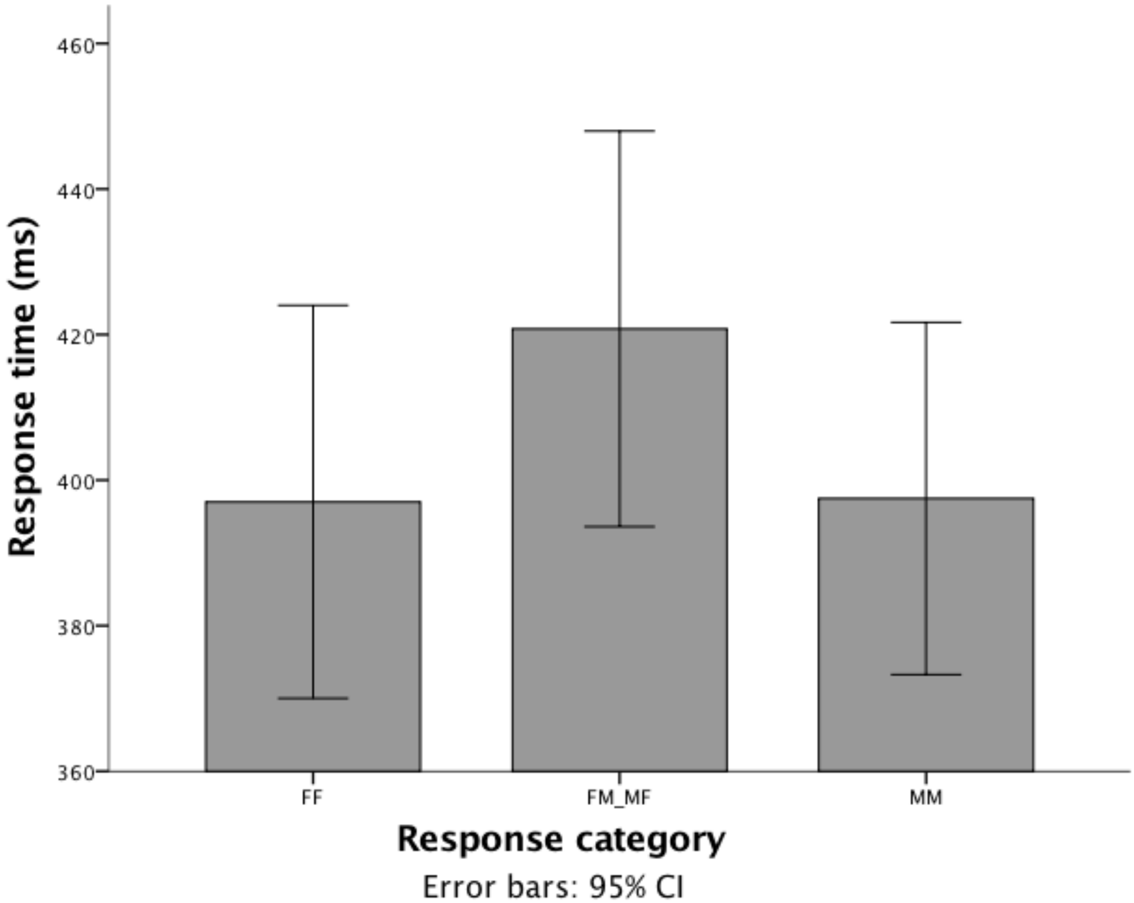


Figure 11. Bar plot showing mean RT for each response category in the button task. Error bars represents 95% confidence intervals. Note: Low numbers indicate slower RT.

Face pair decision task

Descriptive statistics are presented in table 3. A repeated measures one-way ANOVA (SPSS) demonstrated a significant difference between the condition means, F (2,62) = 35.93, *p*<. 001 (spericity assumed). This represents an effect size (η eta^2) of .54, showing that 54 % of the variation in RT was accounted for by response category.

Table 3

RT for each response alternative in face pair decision task

Response Category	M(SD)	95% CI
FF	962.16 (220.46)	[883, 1042]
Mixed	1175.88 (267.14)	[1080, 1272]
MM	954.82 (277.57)	[855, 1055]

Note. CI = Confidence interval

Post hoc comparisons applying Bonferroni correction confirmed that response category "female" resulted in significantly faster RT than response category "mixed" (p <.001), and that response category "male" resulted in significantly faster RT than response category "mixed" (p < .001). These differences demonstrated large effect sizes (d = -0.87 and -0.81, respectively). The difference in RT between response category "female" and response category "male" was not statistically significant (p = 1.000), and this comparison demonstrated a small effect size (d = 0.03). An illustrating of these result can be seen in figure 12. These results do not support the hypothesis stating that there should not be any significant differences between response categories (p = .20).

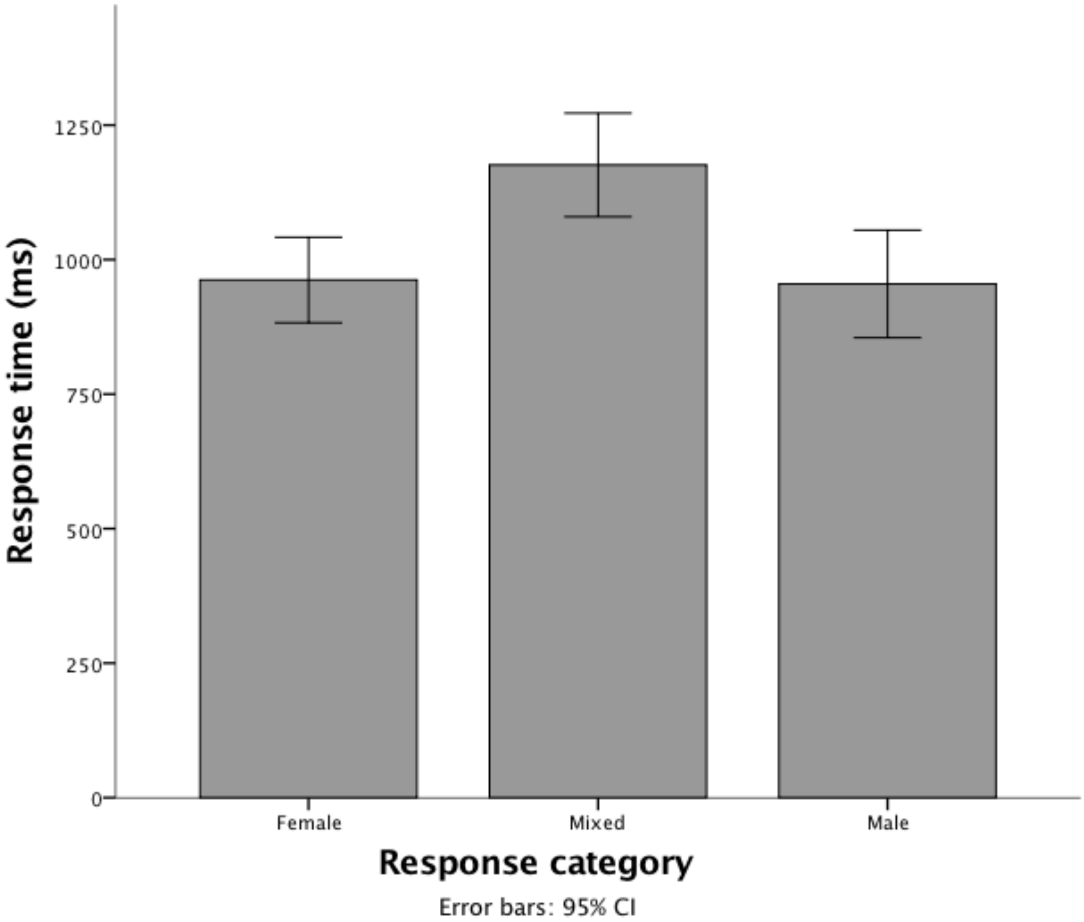


Figure 12. Bar plot showing mean RT for each response category in the FPDT. Error bars represents 95% confidence intervals. Note: Low numbers indicate slower RT.

Reanalysis of Data from Öttl et al. (2016)

Before the reanalysis of the original data incorrect responses and responses with RT falling outside the window of individual mean response time ± 2.5 SD were removed. In addition, as the original response set-up had a total of four response buttons, responses were participants pressed more than the required two buttons were also removed. The reanalysis was conducted using repeated measure one-way ANOVA (SPSS).

Button task.

A repeated measures one-way ANOVA (SPSS) demonstrated a significant difference between the condition means, $F(2,58) = 17.889$, $p<.001$ (Greenhouse-Geisser correction). This represents an effect size (η eta^2) of .38, showing that 38% of the variation in RT was accounted for by response category (descriptive statistics are given in table 4).

Table 4

RT for the reanalysis of the original data, showing each response alternative in the button task

Response Category	M(SD)	95% CI
FF	449.45 (88.44)	[416, 483]
Mixed	480.99 (84.36)	[450, 513]
MM	443.88 (91.98)	[410, 478]

Note. CI = Confidence interval

Post hoc comparisons applying Bonferroni correction confirmed that response category “female” resulted in significantly faster RT than response category “mixed” (p <.001), and that response category “male” resulted in significantly faster RT than response category “mixed” (p < .001). These differences demonstrated medium effect sizes (d = -0.36 and -0.42, respectively). The difference in RT between response category “female” and response category “male” was not statistically significant (p = 1.00), and this comparison demonstrated a small effect size (d = -0.06). An illustration of these result can be seen in figure 13.

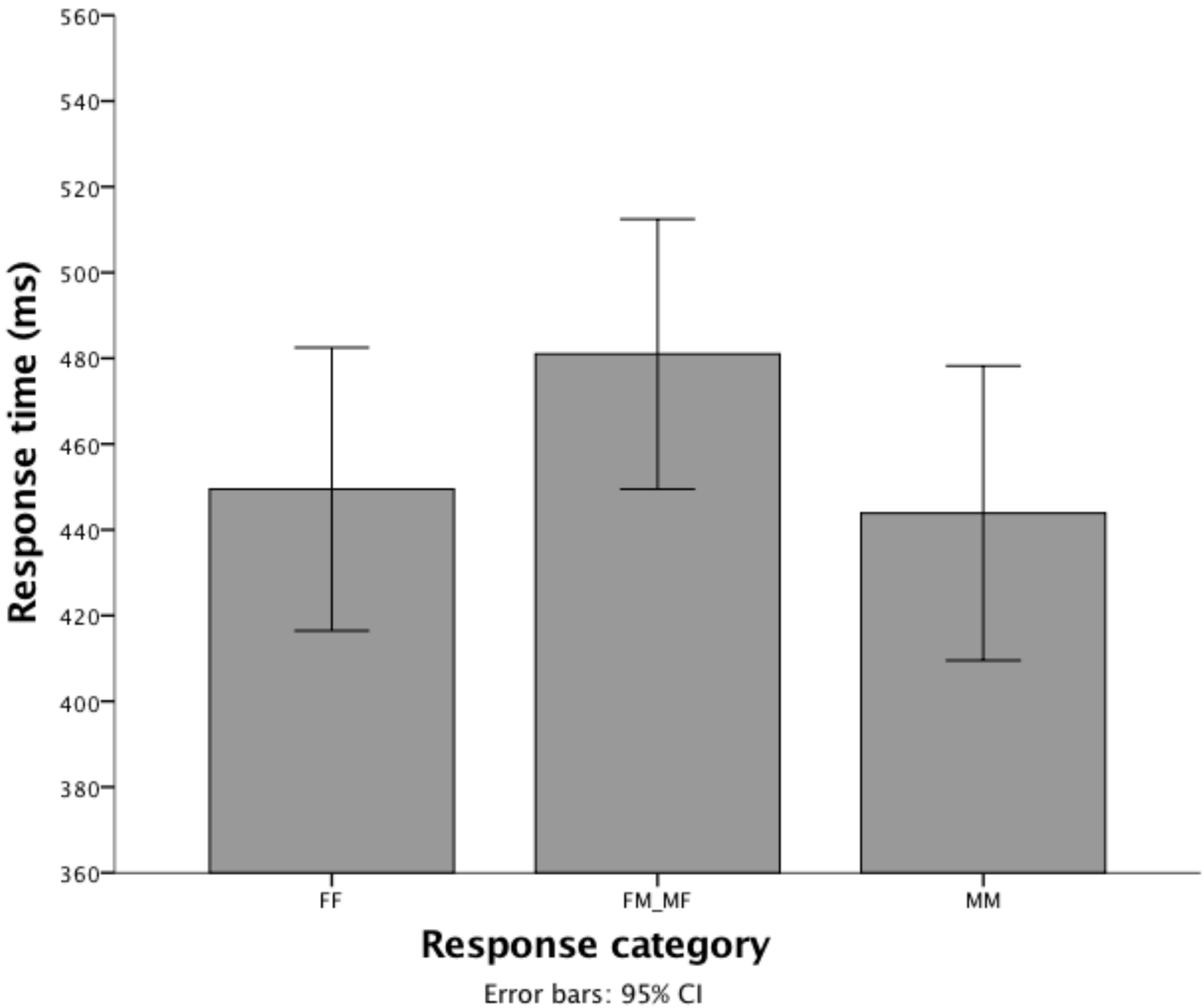


Figure 13. Bar plot showing mean RT for each response category in the button task in the original experiment. Error bars represents 95% confidence intervals. Note: Low numbers indicate slower RT.

Face pair decision task.

A repeated measures one-way ANOVA (SPSS) demonstrated a significant difference between the condition means, $F(2,58) = 28.099$, $p <. 001$ (Greenhouse-Geisser correction). This represents an effect size (η eta^2) of .49, showing that 49% of the variation in RT was accounted for by response category (descriptive statistics are given in table 5).

Table 5

RT for the reanalysis of the original data, showing each response alternative in the FPDT

Response Category	M(SD)	95% CI
Female	889.14 (152.36)	[832, 946]
Mixed	1035.50 (208.43)	[958, 1113]
Male	890.56 (201.52)	[815, 966]

Note. CI = Confidence interval

Post hoc comparisons applying Bonferroni correction confirmed that response category “female” resulted in significantly faster RT than response category “mixed” (p <.001), and that response category “male” resulted in significantly faster RT than response category “mixed” (p < .001). These differences demonstrated large effect sizes (d = -0.8 and -0.7, respectively). The difference in RT between response category “female” and response category “male” was not statistically significant (p = 1.00), and this comparison demonstrated a small effect size (d = -0.007). An illustration of these result can be seen in figure 14.

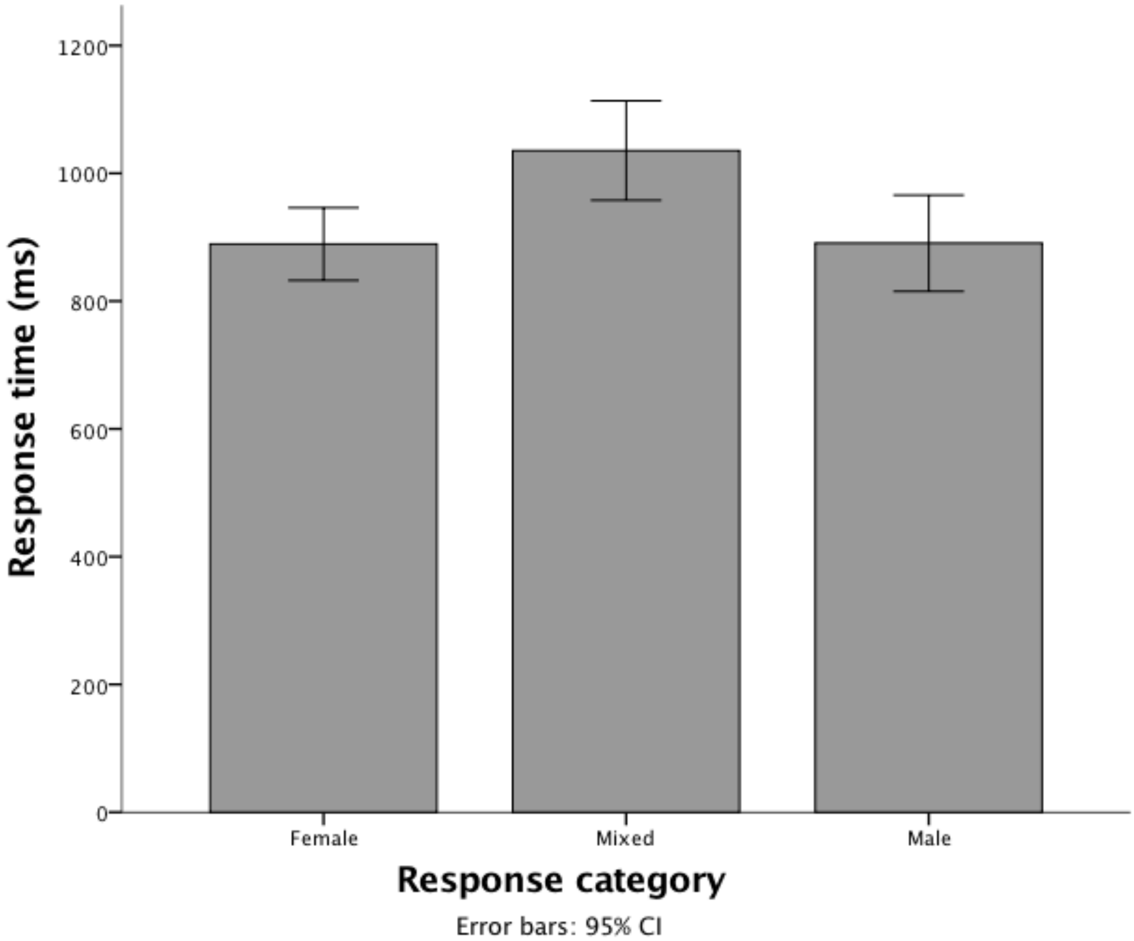


Figure 14. Bar plot showing mean RT for each response category in the FPDT in the original experiment. Error bars represents 95% confidence intervals. Note: Low numbers indicate slower RT.

Comparison Between Mean Response Time in Button Task and the Face Pair Decision Task by Response Set-Up

In addition to the individual analysis of the BT and FPDT from the current dataset and the reanalysed dataset, further analysis was conducted to establish whether there were any differences in mean RT between the two response set-ups in current experiment and the original experiment. This was done by conducting a mixed ANOVA individually for the BT and the FPDT. Where response set-up (original vs current) and response categories (MM, MF/FM, FF [BT] or male, female, mixed [FPDT]) was investigated.

Descriptive statistics for the BT is shown in table 6. The mixed ANOVA for the BT revealed a significant main effect of response category, $F(2, 120) = 38.792$, $p < .001$. This effect tells us that if we ignore the type of response set-up applied, some response categories had longer RT than other response categories. The pairwise comparison for the main effect of response category corrected using Bonferroni adjustments revealed that the significant main effect of response category reflects a significant difference (*p* < .001*)* between response category “FF” and response category “FM/MF”, and response category “MM” and response category “FM/MF”, but not between response category “MM” and response category “FF”. There was also a main effect of type of response set-up, $F(1, 60) = 7.039$, $p = .010$. This effect tells us that if we ignore the different response categories, the RT obtained with current response set-up was significantly different than the RT obtained with the original response set-up. The pairwise comparison for the main effect of response set-up corrected using Bonferroni adjustments revealed that the current response set-up obtained shorter RT than the original response set-up, $p = .010$, with a contrast of 53 ms. There was not a significant response set-up x response category interaction, $F(2, 120) = 1.698$, $p = .197$. The lack of a significant interaction effect tells us that the RT in the three response categories did not significantly differ in the original response set-up and in the current response set-up.

Table 6

Descriptive statistics for response categories with the current response set-up and the original response set-up in the button task

Response category	Response set-up	M (SD)	95% CI
MM	Current	396.94 (75.90)	[370, 424]
	Original	449.47 (88.46)	[416, 483]
MF/FM	Current	420.81 (75.40)	[394, 448]
	Original	481.03 (84.32)	[450, 513]
MM	Current	397.53 (67.14)	[373, 422]
	Original	443.97 (92.01)	[410, 478]

Note. CI = Confidence interval

Descriptive statistics for the FPDT is shown in table 6. For the FPDT, the mixed ANOVA revealed a significant main effect of response category, $F(2, 120) = 62.314$, $p < .001$ (Greenhouse-Geisser correction). This effect tells us that if we ignore the type of response set-up applied, some response categories had longer RT than other response categories. The pairwise comparison for the main effect of response category corrected using Bonferroni adjustments revealed that the significant main effect of response category reflects a significant difference (*p* < .001*)* between response category female and response category mixed and response category male and response category mixed, but not between response category male and response category female. There was a non-significant main effect of type of response set-up, $F(1, 60) = 3.018$, $p = .087$ (Greenhouse-Geisser correction). The lack of a significant effect tells us that we cannot be sure that the RT obtained in current experiment was significantly different than RT obtained in the original experiment. However, despite not a significant different, the pairwise analysis (with Bonferroni adjustments) revealed that the RT obtained from the original experiment was 93 ms faster than the RT obtained in current experiment. There was not a significant response set-up x response category interaction, $F(2, 120) = 2.472$, $p = .094$ (Greenhouse-Geisser correction). The lack of a significant interaction effect tells us that the RT in the three response categories did not significantly differ in the original response set-up and in the current response set-up. The results from the analysis of the BT and the FPDT offers partly support for the hypothesis stating that RT to current experiment should be faster than RT to the original experiment.

Table 7

Descriptive statistics for response categories with the current response set-up and the original response set-up in the FPDT

Response category	Response set-up	M(SD)	95% CI
Female	Current	962.16 (220.53)	[883, 1042]
	Original	889.23 (152.27)	[832, 946]
Mixed	Current	1175.97 (267.19)	[1080, 1272]
	Original	1035.47 (208.52)	[958, 1113]
Male	Current	954.84 (277.54)	[855, 1055]
	Original	890.60 (201.57)	[815, 966]

Note. CI = Confidence interval

Discussion

The aim of this study was to investigate the effect of two tree alternative response set-up and two types of task design (the BT and the FPDT) on RT. Hypothesising that the results from the current experiment should not in either the BT or the FPDT reveal any significant differences between response categories ($p = .20$), and that the results from the current experiment should have significantly shorter RT compared to the results from the original experiment. This was done through an exploration of the RT in the BT and the FPDT with the current response set-up, a re-analysis of the original data, and through am investigation into the difference in RT between the original response set-up and the current response set-up in both the BT and the FPDT.

The Current Set-Up: Differences Between Response Categories Were Found

The results of the current response set-up showed that in the BT participants responded faster to response category “MM” and response category “FF” than to response category “MF/FM”. Further, the results showed that in the FPDT participants responded faster to face pairs with only female- or male faces than to face pairs with one male face and one female face. Thus, there were differences in RT between the response categories in both the BT and the FPDT. Therefore, the results failed to support the first hypothesis stating that there should be no significant differences between the response categories in the BT and in the FPDT. These results imply that the motoric difficulties, such as the hand coordination and finger coordination in the original experiment, is not the reason why one response category takes more time than the two other response alternatives.

A possible reason why participants take longer to reply to the mixed category (response category MF/FM) in the BT could be that the stimuli presented in response category “MF” has two different letters as opposed to two identical letters in response category “FF” and response category “MM”. Suggesting that when participants are presented with response category “MF” the identification of stimuli takes longer because of the additional identification of two letters before making a response. Thus, participants have to do some sort of categorisation of stimuli before responding. This categorisation is not the main goal of the BT, but rather a consequence of the task design. In the FPDT on the other hand, the categorisation of stimuli is the main goal. A possible reason why participants take longer to reply to the mixed category (female- and male face) in the FPDT can be the level of difficulty associated with this stimulus categorisation. Categorising a face pair with a male- and female face appears to require longer information processing than the categorisation of only male- or female faces. Arguably, this can be because a face pair with a male- and female face requires the participants to identify each gender separately, thus drawing up on different linguistic terms and definitions before making a response. Whereas in the only male- or only female face pair condition participants only had to identify one gender, make sure that the stimuli only present one gender, and then make a response.

In both the BT and the FPDT participants could have spent more time on the mixed condition because of carefulness: they could have spent more time looking back and forth between the faces to be sure that they had identified correctly. This could mean that participants chose accuracy over speed of responding in the mixed condition, and as such providing longer RT in the mixed condition compared to the other two conditions.

Comparison Between Results from Current Experiment and Original Experiment: Response Time is Dependent on Both Task Design and Type of Response Set-Up

The reanalysing of the original data shows that, at least for the first BT and the first FPDT, the current experiment replicated the results from the original experiment. Indicating that a full unimanual response set-up can be used in those tasks. However, whether or not it is the best response set-up alternative is not as clear as the comparison of results from the current experiment and the original experiment revealed some interesting aspects of the RT obtained with the different response set-ups.

The results from the comparison between the original response set-up and the current response set-up in the BT revealed a main effect for both response categories and response set-up, but it did not reveal an interaction effect between them. The difference between response categories were found between response category “MM” and response category “MF/FM”, and between response category “FF” and response category “MF/FM”, but did not reveal a difference between response category “MM” and response category “FF”. Interestingly, the main effect of response set-up revealed that the current response set-up provided faster overall RT than the original response set-up. These results prove that parts of the RT from the BT was due to the motoric aspect of the response set-up in the original experiment, as participants were faster in the current full unimanual response set-up. However, the motoric aspect associated with the combination of unimanual- and bimanual responses were not the reason for response category “MF/FM” to require longer RT than response category “MM” and response category “FF”. The results from the BT also tells us that current response set-up possibly removes parts of the information processing associated with the response execution in the original experiments (i.e. the information processing related to remembering which response button the response category presented correlated to). And as such the RT obtained is more probable a result of the information processing needed to categorise and decide on what response to give. Thus, the BT results provide a more accurate measure of the task at hand when the response set-up is full unimanual.

Comparing the results from the current experiment and the original experiment in the FPDT revealed a main effect for response categories, but not for response set-up. Even though the difference between type of response set-up in the FPDT was not significant, it does reveal that the original response set-up provided faster RT than the current response set-up (by 93 ms). This is fascinating because one would expect participants in current experiment to be faster in the FPDT than participants in the original response set-up because they are faster in the BT, but this is not the case. For the response set-up in current experiment the results indicate that participants find the FPDT more difficult than the BT, whereas the results from the original experiment indicate that participants found the FPDT easier. The results from the original experiment seems to indicate that the motoric aspect of the RT disappears in the FPDT. One reason for this could be that the response set-up in the original experiment is helpful for participants when categorising face pairs. Where each hand represents one gender, and thereby making the distinction between face pairs more prominent in the response execution. In addition, the response execution for the mixed response is not only a combination of a male- and a female face, but also a combination of a male- and a female response. Therefore, one can assume that the reason participants had longer RT in the BT because the combination of unimanual- and bimanual responses took longer to learn, thus the results resulted reflects a period of practice of the response set-up in the original experiment.

The results show that task design and response set-up together effects the RT. In tasks requiring less information processing, such as the BT, a full unimanual response set-up provides shorter RT without changing the difference between response categories. Therefore, the use of a full unimanual response set-up can be argued to be the best response set-up design for such tasks. However, for tasks requiring a higher level of information processing, such as the FPDT, there seems to be a benefit of having a more complex response set-up with a combination of unimanual- and bimanual responses. Based on this one can argue that when choosing response set-up, the task at hand should be considered as we have shown that different types of tasks give different results depending on response set-up design. Importantly, when an experiment consists of different types of tasks the decision on which response set-up to apply needs to be carefully considered and pre-tested to avoid a bias in the results due to difference in RT trend between the tasks.

Limitations of the Present Study

The theoretical framework on overlapping categories and the combination of bimanual- and unimanual responses in RT research is very limited. As a result of this lack of knowledge about overlapping categories the introduction part of this thesis does not have a straightforward reasoning behind the choice of response set-up and task design. Therefore, both the original experiment and current experiment is part of developing a better understanding about how to design experiment with overlapping categories.

The data analysis conducted could have used more complex models, such as general linear models as conducted by Öttl et al. (2016), in order to provide an even deeper understanding of the results by taking into account trial number and block number on RT results. And even though eye-tracking data was collected it was not analysed in this study. Had it been analysed it could have provided a deeper understanding about why the "MF" condition in the BT had longer RT than response categories "MM" and "FF", as well as an indication of different looking patterns in the FPDT. Further research should include an analysis of eye tracking data when looking at overlapping categories as it can provide more information on what participants are actually doing when participating in such experiment, and thereby add valuable information to use when designing new experiments with overlapping categories in RT research.

Future Prospects and Conclusion

Future experiment working with overlapping categories and RT should include an eye tracking analysis of gaze pattern to establish whether participants spend more time looking at and identifying the "MF/FM" response category than the "MM" and "FF" response categories. If such an analysis provided results supporting the notion that participants spend more time looking at the "MF/FM" response category than the "FF"- and "MM" response categories, an easy solution would be to remove the letters from the stimuli and apply an appearance of a cross or light change as used in previous research (Antonova et al., 2016; I. J. Deary et al., 2011). Such method could possible reduce the time spent identifying stimuluses. However, keeping in mind that the use of letters is a mean to help the participant prepare for the categorisation of gender in the following task (FPDT), removing them would make part of the reason for having the BT disappear. Regardless, it would be interesting to see whether the significant difference in RT between response categories would disappear if the stimulus where changed to a stimulus that require less information processing, such as light change or appearance of a cross, as it could possibly provide results indicating if the significant RT difference found in the BT is due to either the response set-up itself, the stimuluses used or both. Therefore, in order to account for more variation in the results an additional RT task with appearance of light or cross should be added to the experiment. Thus, if the experiment consists of different tasks, such the experiment conducted in this thesis, the additional RT task with appearance of light (or cross etc) could be used as a practice task for participants to get used to the response set-up before starting the main body of the experiment.

In conclusion, RT is both task dependent and response set-up dependent; tasks requiring minor information processing having a benefit of a full unimanual response set-up, and tasks requiring higher level information processing having a benefit of a combination of bimanual- and unimanual response set-up. Subsequently, experiments implementing different types of tasks should be aware of the differences between them and the amount of practice needed for the response set-up by participants before deciding on which response set-up to apply.

References

Adam, J. J., Paas, F. G., Buekers, M. J., Wuyts, I. J., Spijkers, W. a, & Wallmeyer, P. (1999). Gender differences in choice reaction time: evidence for differential strategies. *Ergonomics*, *42*(2), 327–335. https://doi.org/10.1080/001401399185685

Antonova, I., van Swam, C., Hubl, D., Dierks, T., Griskova-Bulanova, I., & Koenig, T. (2016). Reaction Time in a Visual 4-Choice Reaction Time Task: ERP Effects of Motor Preparation and Hemispheric Involvement. *Brain Topography*, *29*(4), 491–505. https://doi.org/10.1007/s10548-016-0473-7

Burke, D., Linder, S., Hirsch, J., Dey, T., Kana, D., Ringenbach, S., … Alberts, J. (2016). Characterizing Information Processing With a Mobile Device : Measurement of Simple and Choice Reaction Time. https://doi.org/10.1177/1073191116633752

Baayen, R. H., & Milin, P. (2010). Analyzing Reaction Times. *International Journal of Psychological Research*, *3*(2), 12–28. https://doi.org/10.1287/mksc.12.4.395

Deary, I., & Der, G. (2005). Reaction time, age and cognitive ability: Logitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology, and Cognition*, *12*(2), 187–215. https://doi.org/10.1080/13825580590969235

Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. *Behavior Research Methods*, *43*(1), 258–68. https://doi.org/10.3758/s13428-010-0024-1

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, *21*(1), 62–73. https://doi.org/10.1037/0882-7974.21.1.62

Donders, F. C. (1868). Die Schnelligkeit psychischer Processe. *Archiv Für Anatomie, Physiologie Und Wissenschaftliche Medicin*, 657–681.

Godard, O., & Fiori, N. (2010). Sex differences in face processing: Are women less lateralized and faster than men? *Brain and Cognition*, *73*(3), 167–175. https://doi.org/10.1016/j.bandc.2010.04.008

Godwin, H. J., Menneer, T., Cave, K. R., Thaibsyah, M., & Donnelly, N. (2015). The effects of increasing target prevalence on information processing during visual search, 469–475. https://doi.org/10.3758/s13423-014-0686-2

Jain, A., Bansal, R., Kumar, A., & Singh, K. D. (2015). A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied & Basic Medical Research*, *5*(2), 124–7. https://doi.org/10.4103/2229-516X.157168

Kyllonen, P. C., & Zu, J. (2016). Use of Response Time for Measuring Cognitive Ability. https://doi.org/10.3390/jintelligence4040014

Miller, J. O., & Low, K. (2001). Motor Processes in Simple , Go / No-Go , and Choice Reaction Time Tasks : A Psychophysiological Analysis, *27*(2), 266–289.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Ratcliff, R., & Smith, Philip, L. (2004). A Comparison of Sequential Sampling Models for Two-Choice Reaction Time. *Psychological Review*, *111*(2), 333–367. https://doi.org/10.1037/0033-295X.111.2.333

Sanders, A. F. (1998). *Elements of Human Performance: Reaction Processes and Attention in Human Skill*. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.

Semmelmann, K., & Weigelt, S. (2016). Online psychophysics: reaction time effects in cognitive experiments. *Behavior Research Methods*. https://doi.org/10.3758/s13428-016-0783-4

Usher, M., Olami, Z., & Mcclelland, J. L. (2002). Hick's Law in a Stochastic Race Model with Speed–Accuracy Tradeoff. *Journal of Mathematical Psychology*, *46*, 704–715. https://doi.org/10.1006/jmps.2002.1420

Öttl, A., Behne, D. M., Gygax, P., & Gabriel, U. (2016). No Title. In *Measuring response times using three response alternatives in a decision task.*