Bachelor Thesis

Remembering and forgetting: Testing the limits and structure of Visual Working Memory and whether memory decay can be observed

Marius Husmann

S212084

First supervisor: Dr Rob van der Lubbe

Second supervisor: Dr Simone Borsci

Faculty of Behavioural, Management and Social Sciences

University of Twente

June 2021

Testing the limits and structure of Visual Working Memory and whether memory decay can be observed

Abstract

Visual Working Memory (VWM) is a concept that is used to describe where and how the memorization and manipulation of visual information happens. Great amounts of research led to the development of different conflicting models that try to explain important mechanisms and aspects of this concept. Two of these aspects are the role that time plays in the precision of the memories as well as what the units of these memories are (whole objects or single features of these objects). There are two main theories about how information is stored, the first being the discrete capacity model which predicts that whole integrated objects fill a limited number of slots. The second model is the limited resource model, according to which there is a limited pool of resources that can be allocated to remember certain features of objects, here more precise memories are said to require more resources. For the role of time there are several theories. The most widely known is the decay theory. As the name suggests it predicts that memories gradually decay over time. To get further insight into these topics an experiment was conducted in which 19 participants, all of whom were students studying in Germany or the Netherlands, had to remember specific features of up to four objects which differed in their size, orientation and colour. First they only needed to remember each of these three feature dimensions separately, followed by a block of trials where they needed to attend to all dimensions. The participants had randomly chosen decay times of either one or three seconds between item memorization and recall, to see whether a decay effect could be observed. A repeated measures ANOVA showed that the number of presented stimuli as well as the number of attended dimensions had a significant effect on the recall precision of the participants, as was predicted by the limited resource model, leading to a rejection of the discrete capacity model. The decay time had no significant effect on the memory precision at any point or level of the experiment, but due to the limitations of the study no real claims about the role of time can be made.

Testing the limits and structure of Visual Working Memory and whether memory decay can be observed

Introduction

One very important skill of humans is the storing and manipulating of information. It allows us to make informed decisions, evaluate information and compare and combine new and already stored information. The commonly accepted concept that tries to explain where all of this happens is the concept of working memory (Diamond, 2013; Malenka, Nestler, Hyman, Sydor, & Brown, 2009; Miyake, & Shah, 1999). The current models differentiate between different components of working memory. One such component is the Visual Working Memory (VWM). As the name suggests, it is concerned with the storing and handling of visual information (Vogel, Woodman, & Luck, 2001). The classic and currently widely accepted model of VWM suggests that it has a limited capacity and can hold a certain amount of discrete items, therefore it is known as the discrete capacity model. The estimation of the number of items that can be stored changed over the course of time. Miller (1956) believed it to be 7 plus or minus two. More recent research suggests this number to be smaller, namely four (Cowan, 2001). According to these models, the VWM consists of this number of slots, items that are supposed to be remembered just fill these slots with all their features as one unified object each until there are no slots left. The complexity of the items does not play a role in these models and they suggest that all items can be recalled accurately with all of their features, and if all the slots are filled additional items cannot be remembered at all (Zhang, & Luck, 2008). Zhang and Luck (2008) compared participants recall precision of one, two, three and six unidimensional item trials and found that the performance of participants got way worse for the six item condition while the standard deviation remained similar, indicating that humans are not able to remember six or more items due to a capacity limit. The big gap between three and six items leaves a lot of room for speculation though as well as their use of unidimensional items. Including trials with four and five items might reveal more about the big decline between three and six items, while including further dimensions might lead to big dips in performance when less than four items are shown, indicating that the decrease in precision is not just due to a limited number of memory slots.

This would be in line with some more recent research findings which challenge the classic view on VWM(Bays, Catalao, & Husain, 2009; Huang, 2010; Zhang, & Luck, 2008). These Findings suggest that there is no preordained number of slots just waiting to be filled, but rather a limited amount of resources that can be distributed, giving it the name limited resource model. According to this model. Humans have a limited pool of resources that are used by the VWM to either remember more items with their features less precisely, or less items with more precise features. In this model, items are not considered as discrete objects but rather as a composition of different features. The more features are to be remembered the more resources have to be used (Bays, Catalao, & Husain, 2009; Huang, 2010). The limited resource model, therefore, suggests that already fewer items than four can produce errors in recall precision if they are complex enough. Generally, the decline in recall precision is predicted to be more gradual with an increasing number of items and item features according to the limited resource model, whereas the discrete capacity model predicts consistent performance for set numbers of up to four items with a steep decline in recall precision afterwards. In summary, while the discrete capacity model is concerned with the number of items the limited resource model is more about the quality of the item representations in VWM (Bays, Catalao, & Husain, 2009; Huang, 2010; Zhang, & Luck, 2008).

One further concept that is closely related to working memory and its functionality is attention. Attention is considered to be a selecting factor, meaning that items or information that one attends to will enter working memory for further processing. At the same time it is used to hold items in working memory active (Awh, Vogel, & Oh, 2006). For visual attention, this means that objects which are attended to go into the VWM and are held there. Regarding visual attention, there are again two main models that try to explain the functionality. The older and widely accepted model suggests that the units of attention are whole integrated objects with all their features(Zhang, & Luck, 2008; Luck & Vogel, 1997; O'Craven, Downing, & Kanwisher, 1999; Park et al., 2017). This model fits the discrete capacity model of VWM, as both suggest that objects are handled as single coherent units, no matter how complicated they are.

Like it is the case for VWM, more recent research findings seem to oppose this model. They suggest that certain features are remembered with higher precision if people are instructed to attend to these features specifically than they are without any specific attention instruction. Accordingly, the recall precision of some features suffers if more attention is paid to another specific feature of an object(Brummerloh et al., 2019; Park et al., 2017; Bocincova & Johnson, 2019; Ma et al., 2014; Markov, Tiurina, & Utochkin, 2019; Park et al., 2017; Shin & Ma, 2017).
These findings are more in line with the limited resource model as they stress recall quality over quantity and single item features over whole integrated item accounts.

The present study aims to compare the two opposing VWM models and their ability to predict the accuracy with which participants are able to recall item features of a single randomly selected item, consisting of three dimensions, out of a set of 1, 3 or 4 items. If the discrete capacity model holds true the performance is expected to be constant throughout all of the sets. The limited resource model on the other hand predicts a gradual decline in recall precision with an increased number of stimuli and item features that need to be attended to. As the more recent research findings seem to be in line with the limited resource model, participants recall precision in this study is expected to gradually decline with an increased number of features and items.

Yet another factor that is believed to influence working memory performance is time. There are different theories about the influence of time on working memory performance with conflicting views, but what they all agree on is that time plays a crucial role for the precision with which items can be recalled or if they can be recalled at all. Understanding the influence of time might help to gain deeper insight into the working mechanism of VWM and reveal new questions.

One widely known and accepted theory of the effect of time is that items in VWM decay over time, meaning that they get worse the longer a person has to hold on to them. This theory would suggest that there is an observable gradual decline in recall precision with the increased time between encoding and recalling (Barrouillet, Portrat, Vergauwe, Diependaele, & Camos, 2011). Like the limited resource model this theory is concerned with the quality of memory representation, creating a possible link between the two models.

Another opposing theory suggests that VWM representations “die a sudden death” (Zhang, & Luck, 2009). Thus, instead of decaying gradually, item representations are held in memory until a certain threshold is reached and then they vanish abruptly. This theory would suggest that up to a certain point there is no decrease in recall precision when the time between encoding and recall gets increased. However, once this point is reached there should be a big dip in performance as the item gets erased from working memory. This would fit the discrete capacity model as both theories predict either total recall precision or no recall precision at all.

The third theory that is concerned with the effect of time on VWM performance is the temporal distinctiveness model. According to this model, preceding memories produce noise which in term decreases working memory performance. According to this theory, bigger time intervals before the item presentation and encoding allow for better/less interrupted encoding, leading to higher recall performance(Souza, & Oberauer, 2015). Therefore this theory again thinks about memory representations as being qualitative.

There is an ongoing debate between researchers on which theory best explains the influence time has on working memory, with a lot of research findings being provided by all parties involved. As the decay theory is the most known and widely accepted theory, we aim to control if there is an effect of time between item presentation and item recall. This was done in order to control whether a decay effect on VWM items can be observed. The time differences were chosen to be rather small as bigger time differences might lead to the sudden death effect, again leading to ambiguity of the results. Also, if decay time has a gradual effect on memory performance this effect should also be consistently observable for smaller time intervals. It is important to note that the intervals that were chosen are longer than one second, as smaller intervals might lead to Iconic Memory (IM) having an effect on the recall precision. IM is believed to have a duration of less than one second (Sperling, 1960). Participants' performance is expected to be worse for the longer decay time interval, as predicted by the decay theory.

Methods

Participants

Nineteen students from different universities (one located in the Netherlands and five in Germany) and disciplines (mathematics, psychology, political science, etc.) participated in the study (Mage=22.36, range: 20-25, 4 female and 15 male, 15 right-handed). All participants had a normal colour vision, confirmed by an online colour blindness test (“ishihara 38 plates”) and all had normal or corrected normal visual acuity, which was confirmed by using Landolt C stimuli (Bach, 2007). All data could be included in the analysis. The experiment got approved by the University of Twente Faculty of Behavioural, Management and Social Sciences ethics committee. All participants gave written informed consent.

Materials

The experiment was conducted using a Microsoft Surface Pro 3 with a 12.00 inch touch display, a resolution of 2160x1440 pixels and a screen ratio of 3:2, running windows 10. Participants sat between 50 and 60 cm away from the screen. The code was programmed in python 3.8 using a Spyder 4.1.5 development environment.

The Stimuli consisted of rectangles with a side ratio of 16:9. The stimuli differed in colour, orientation and size. The values of these dimensions that were used in the main experiment were generated by the participants in a pretest to correct for individual differences in perception. A more detailed account of the stimulus generation is given in the procedure section.

Procedure

Before the experiment started, all the materials participants used were disinfected and the room was ventilated. The participants were then asked to wear a mask, informed about the risks affiliated with covid 19 infections and asked to fill in a questionnaire about common Covid 19 symptoms to make sure that the virus does not spread among researchers and participants. No participant reported any symptoms in the critical time frame.

After all Covid 19 related steps were taken, participants were provided with written instructions after which they could ask questions or voice concerns. In the following, they were handed a consent form, informing them about their rights and how their data will be handled. They then performed the colour blindness and the visual acuity test. Lastly, participants had to finish a pretest to generate stimuli that each participant could consistently distinguish visually before starting the main experiment.

The pretest consisted of three main blocks. To finish each block participants had to collect 5 respective values. In the first block, they had to distinguish between colours. They sat down in front of the screen and had to focus on the center of the screen. After a countdown there appeared a fixation cross in the middle of the screen. At the same time, two rectangles were presented. Participants had to indicate whether they looked the same or different to them. For the first block, they had the same size and orientation. The colours were generated by increasing the hue angle. The default colour had a hue angle of 0°, a saturation of 100% and an intensity of 100%. The first colour they had to distinguish had a hue angle that was increased by 5°. To add this colour to the stimulus set they had to successfully answer five sets. In three sets the colour differed and in two it was the same. If the participants answered one set wrong the hue angle got increased by another 5°. The stimuli were generated randomly in one of two positions each. If they answered all 5 sets correctly the colour got added to their pallet and became the new default colour for the next 5 sets. Additionally, the increment got decreased to 5° again.

They directly continued with the orientation sets. The same procedure was used with the rectangles having the default colour. The sets started with a default orientation of 0°. For orientation, the increments that were used to increase the difference when a wrong answer was given were 3°. So the very first set was made up of 0° and 3° orientations.

In the last block of the pretest, they had to distinguish sizes. Stimuli colour and orientation had the default values. The default size was 80x45 pixels. The second stimulus was generated by using a size multiplier. By default, this multiplier was 1.2 and it was increased by 0.2 for every wrong answer until a participant got 5 trials correct, then it was reset to 1.2 to get the next size.

After finishing the pretest participants were asked to take a break of 10 to 15 minutes before starting the main experiment.

For the main experiment, they first had to work through a tutorial, during this phase the researcher assisted them and answered questions about the procedure. Next, they had to work through 4 different blocks. Each block had short attention instructions, first only colour needed to be attended to, then only orientation, followed by only size and lastly they had to attend to all three dimensions.

Each block had 18 trials. Each trial consisted of a three second timer, the stimulus presentation, a decay time, a short mask and then a recall screen (Figure 1.). After the timer was shown the screen was filled with a fixation cross in the middle of the screen and either one, three or four stimuli (Figure 1B.). The stimuli were randomly generated out of the sets of values that the participants generated in the pretest in a way that each stimulus looked different in respect to each of the dimensions. The positions of the stimuli were selected randomly out of a set of four possible positions. This screen was shown to them for two seconds.

After the presentation followed a decay time period of either one or three seconds in which only the fixation cross was shown on the screen. The times were selected randomly in a way that each time was chosen nine times. After the decay time, one position where previously a stimulus was presented was randomly chosen and marked by a mask, which was a light grey rectangle of the size 180x180 pixels, indicating that the stimulus needs to be recalled (Figure 1C.).

The recall screen was different for the first three blocks. It consisted of instructions for the participant as well as a pallet of seven different answer choices, which were seven default rectangles differing in one dimension according to the attention instructions (Figure 1D).. Five rectangles were made up with the values collected in the respective pretest, the remaining two rectangles had extreme values. These extreme answer options were never correct. All answers were ordered from the lowest to the highest value. Participants were asked to select an answer by using the touchscreen to click on it. For colour, the extreme on the lower end had a hue angle of 345° (or -15°) and the one on the right was whatever the participants' last selected colour was plus another 15°. For the orientation block, the extremes were -9° rotation and again the highest rotation angle value selected plus another 9°. The size extremes were 64x36 pixels on the low end and whatever the biggest collected size was times 1.4.

The last block was the mixed condition. Here the participants were not told beforehand what they needed to attend to. Instead, they were just presented with one of the three recall screens. They then had to recall the respective stimulus dimension. Which screen got shown was selected randomly.

After finishing the mixed condition participants were able to take another break. The whole main experiment was repeated another two times so that participants, in the end, did each condition three times. After finishing the last repetition they were informed about the extreme answers never being correct. Additionally, they were asked if they experienced any particular difficulties or had any additional questions. Each experimental session took around one and a half hours to complete.

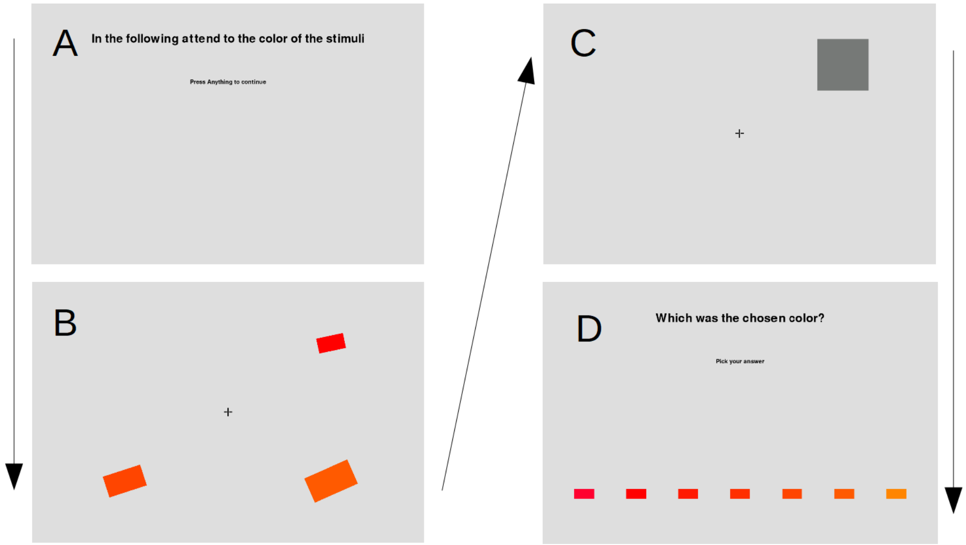

Figure 1. Example of a colour trial of the main experiment, starting with the attentional instructions (A), followed by the two-second stimulus presentation (B). After a delay time period the mask was shown (C). Lastly, the participants were presented with the recall screen where they needed to pick their answer (D).

Data Analysis

To make sure that the data can be analysed in an appropriate way they were tested for normality, homoscedasticity and sphericity. There were no critical violations found for any of the assumptions. The original dataset consisted of the position of the stimuli, the number of stimuli and the actual values of the critical dimensions, meaning colour was recorded as the hue angle, orientation as the angle, and size as the actual size in pixels. Therefore it needed recoding so that analysis of the aspects of interest was made possible. Each of the raw dimension values got recoded into ordinal values between 0 and 6. So if a participant had for example 345°, 0°, 5°, 10°, 15° and 20° as hue angles for his colour pallet, these values got recoded into 0, 1, 2, 3, 4, 5 and 6. The new variable error distance was calculated by taking the absolute value of the difference between the given and the required answer. Following the mean error distance for all participants was calculated for every possible combination of trial features (e.g., mixed attention, 1 second decay, set size 3, color).

It was then tested whether the accuracy performance of the participants was above chance. To do so the mean error distance for each combination was tested against the average distance value. A one-sample t-test was conducted for each of the possible combinations. To get the average distance value, all possible distances from the correct answer to the given one were determined and the average of these distances was calculated.

A 4 (Focus on colour, focus on orientation, focus on size and mixed focus) x 2 (decay time) x 3 (set-size) repeated measures ANOVA was conducted to see whether there is an overall effect of set size and attention condition and if there are differences between the different stimulus dimensions. Additionally, to test the effect of the different conditions and set sizes on the performance for the respective dimensions three repeated measures ANOVAs were conducted. Each repeated measures ANOVA analysed the error distances for one of the three stimulus dimensions (colour, orientation and size). The ANOVAs had a 2 (focused attention, mixed attention) x 2 (one second or three seconds decay time) x 3 (set size one, three or four) design.

Results

Above-chance Performance

The t-test showed that the mean error distances for all of the possible conditions were significantly lower than the average distance value (=2.666), indicating that the participants did not guess during the experiment. The largest divergence from the expected chance value occurred for the color condition with one stimulus and a decay time of 3 seconds, Mdiff = 1.807, t(18) = -50.8, p<0.001, while the smallest difference from guessing behaviour was found in the mixed attention condition with a decay time of 1 second and a set size of 4 when asked about orientation, Mdiff = 0.939, t(18) = -7.5, p<0.001.

Visual Working Memory task

The repeated measures ANOVA analysing the influence of the attention conditions (colour, orientation, size, mixed) in combination with decay time and set size showed that the attention condition had a significant effect on the observed error distance, F(3,54) = 11.430, p<0.001, $\eta_p^2 = 0.388$. Moreover, a significant effect was identified for set size, F(2,36) = 36.549, p<0.001, $\eta_p^2 = 0.670$. Decay time on the other hand had no significant effect on the answer accuracy, F(1,18) = 0.409, p = 0.531, $\eta_p^2 = 0.022$. A significant effect was also found for the interaction of attention condition with set size, F(6,108) = 5.133, p<0.001, $\eta_p^2 = 0.222$, while all other interaction effects were not significant. The mean error distances increased with increasing set size as was predicted. For set size one, the overall mean ED was 0.3264, 0.6077 for set size three and lastly 0.7000 for set size four. The difference between set size one and set size three is significant with t(18) = -9.880, p<0.001, as well as the difference between set size three and set size four, t(18) = -2.161, p = 0.044.

The mean error distances for the attention conditions were 0.3585 for colour, 0.4566 for orientation, 0.5571 for size and 0.6320 for the mixed condition. A significant difference from the mixed condition was measured for colour, t(18) = -6.796, p<0.001, and for orientation, t(18) = -4.328, p<0.001. The mean error distance of size on the other hand does not significantly differ from the mean error distance of the mixed condition, t(18) = -1.385, p = 0.183. Interestingly there there was a significant difference between the colour and the orientation condition, t(18) = -2,438, p = 0.025, colour and size, t(18) = -4.935, p<0.001, and orientation and size, t(18) = -2.479, p = 0.023.

The three repeated measures ANOVAS that were conducted to see whether the attentional instructions (only one dimension or mixed), set size and decay time had a significant effect on the ability to remember specific dimensions showed results which correspond with the first ANOVA for color (Figure 2.) and orientation (Figure 3.), with attentional instruction, set size and their interaction having a significant effect on the accuracy of the responses, while decay time at no point had any significant effect. The effect size of attentional instructions was larger for the orientation condition with $\eta_p^2 = 0.569$ compared to $\eta_p^2 = 0.396$ for colour. The same was true for the effect of set size with $\eta_p^2 = 0.633$ for orientation and $\eta_p^2 = 0.524$ for colour. For size on the other hand, the attentional instructions seemed to have no significant effect, F(1,18) = 2,840, p = 0.109, as well as the interaction between attentional instruction and set size, F(2,36) = 1.168, p = 0.322. Only the set size itself significantly affected the accuracy of the participants in recalling stimulus size with F(2,36) = 6.293, p = 0.005,$\eta_p^2 = 0.259$ (Figure 4.).

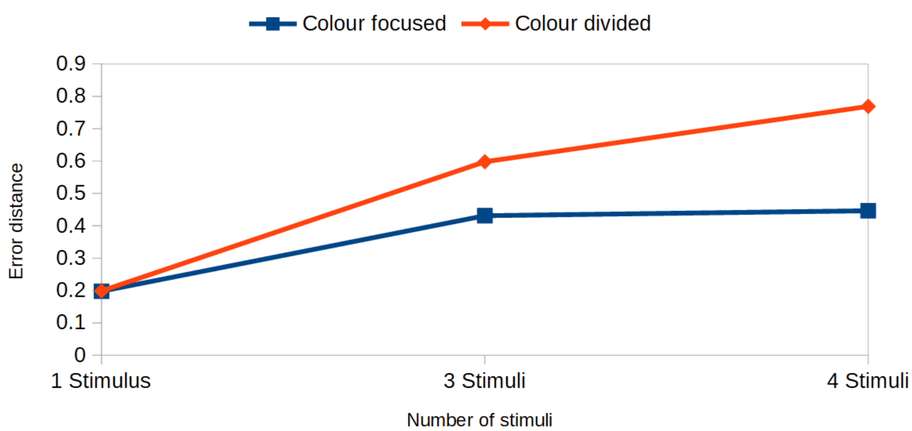

Figure 2. The mean error distance for the feature colour for one, three and four stimuli, with one line representing the focused and the other representing the divided attention condition

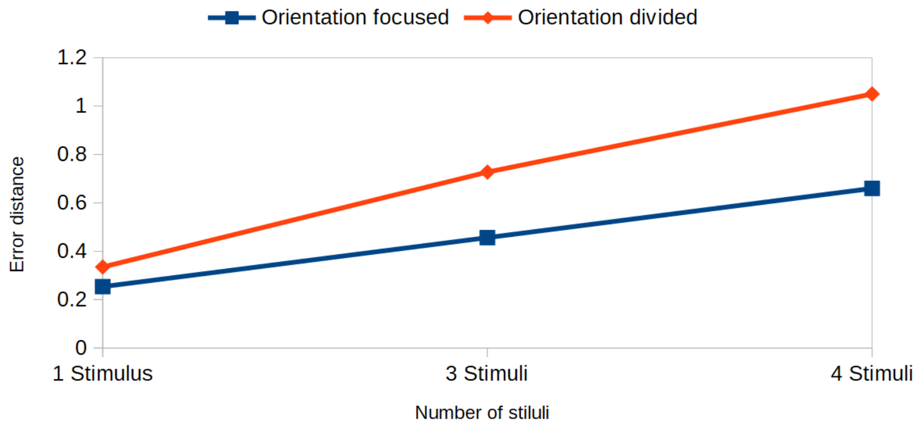

Figure 3. The mean error distance for the feature orientation for one, three and four stimuli, with one line representing the focused and the other representing the divided attention condition

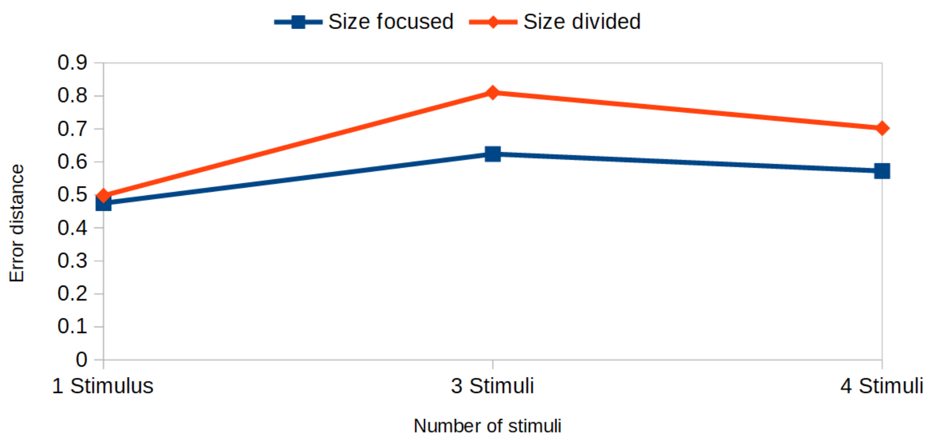

Figure 4. The mean error distance for the feature orientation for one, three and four stimuli, with one line representing the focused and the other representing the divided attention condition

Discussion

The goal of this study was to gain a better understanding of the Visual working memory. In order to do so two different questions were examined. First if there is an observable difference in recall accuracy when the time between stimuli encoding and stimuli recall gets increased from one to three seconds, and second whether recall performance decreased when more stimulus features needed to be remembered, to see which model of VWM predicts participants performance best.

Regarding the first question, the statistical analysis showed no significant results at any level of the experiment. The fact that there is no effect in these relatively small time frames of one and three seconds might support the sudden death theory of visual working memory which would be in line with the findings of Zhang and Luck (2009) rather than a gradual decay which might have been expected at least for the high memory load trials. This could support the discrete capacity model, as it as well takes an all or nothing approach to explain the functionality of VWM. But to really make any meaningful claims research that focuses on really comparing all feasible models has to be conducted. For this more time conditions should be added. These should increase gradually to see the nature of the effect time has on memory performance. All that can be said from the findings in this study is that a time difference of two seconds does not seem to lead to consistent differences in recall accuracy of items held in the visual working memory.

The second question was about confirming whether the discrete capacity model or the limited resource model of VWM predicts better people performance in recalling the features of the stimuli. According to the discrete capacity model whole integrated objects get encoded into the VWM until a certain number of slots, around four in total (Cowan, 2001), is filled (Cowan, 2001; Luck & Vogel, 1997). Inline with this model a consistent performance should be expected regardless of the complexity of the stimuli that need to be remembered. Accordingly, the recall accuracy should not change, no matter if the participants have to only remember one dimension or all three. However, this was not the case in the present study. In fact, attention condition and set size shown a clear effect on the performance, moreover, the interaction between the two, suggesting that an increase in memory complexity influences the number of stimuli that can be remembered and vice versa, so an increased number of stimuli leads to less precision in feature recall and more complex stimuli lead to a lower total number of stimuli that can be recalled with high precision.

These findings are in line with the limited resource model of VWM which claims that every human has a certain amount of resources to allocate, the more features have to be remembered the more resources are used. Precise memories require more resources than vague memories, therefore at a certain point there will be a tradeoff where a person has to decide to either remember a greater amount of features with less precision or only a limited amount of features, but therefore with greater precision (Bays, Catalao, & Husain, 2009).. In the mixed condition, where more features had to be remembered, the accuracy suffered notably, the same goes for an increased number of stimuli. As predicted by the limited resource model, the highest mean error distance was observed for the four stimuli mixed attention condition, but the results were still more accurate than if the participants would have guessed, suggesting that there was a decrease in memory precision rather than a total memory loss, just like the model predicts.

A result that neither of the two models can initially explain is the difference in recall accuracy for the three different dimensions. Colour was recalled more precisely than orientation, and orientation was more precisely recalled than size. The remarks of some of the participants might help to shed some light on this phenomenon. Several participants reported having difficulties recalling the size of the stimuli because they had to rotate the mental image they had of them, the reason being that all of the possible answers for size had the same orientation, which could differ from the orientation of the stimulus that had to be remembered. This also holds true for the orientation condition, where the answer stimuli possibly had a different size than the stimulus that had to be recalled, even though only two participants reported difficulties with recalling orientation because of size differences. This might suggest that the participants ability to mentally rotate, and stretch or shrink images affected, somehow, the performance of the participants in this study. This would also be in line with the findings of studies which suggest that VWM and mental imagery, manipulating mental images, happen in the same brain areas or at least tap into the same resource pool (Tong, 2013; Thomas B. Christophel, Radoslaw M. Cichy, Martin N. Hebart, & John-Dylan Haynes, 2015; Keogh, & Pearson, 2014). Tong (2013) even reported that VWM performance gets worse when additional mental rotation needs to be performed, fitting the results of this study. This could fit the limited resource model, as mentally rotating an image would require resources that cannot be used to hold and recall accurate representations of all the stimuli in the visual working memory. But at this point, these are only speculations which should be tested in future studies.

Another theory which could be used to explain the present findings is that different features use different resource pools (Wang et al., 2017). This theory would also reject the discrete capacity model as it suggests that different features are weighted differently. Accordingly, different features are remembered separately and not accumulated into one single unit and each uses its own resource pool. But this theory does not align with the present results which clearly showed that the accuracy of color and orientation recall get significantly worse when all dimensions have to be remembered. Wang's et al. (2017) model predicts that there should be no significant effect of the mixed attention condition as color and orientation both supposedly have their own resource pool.

Limitations and future research

As discussed above, some of the participants reported difficulties with recalling the size of the stimuli due to them having a different angle in the presentation and the recall section of a trial. This could have influenced their recall precision and therefore the results. One way that might be feasible to tackle this problem could be to use circles for the stimuli instead of rectangles. This way participants will not need to rotate the image mentally for it to fit the possible answers. On the downside, future researchers would need to come up with a different way of testing the participants for their ability to recall orientation. Maybe a fixed pattern on the stimulus which then can be rotated, like a single line dividing the circle, could work. This way the shape of the stimulus would not change but only its print.

Another limitation is that participants could have tried to use their semantic memory to remember for example colours instead of merely holding visual representations in their visual working memory. The use of the pretest to get stimuli with individual just noticeable differences helped to minimize the application of such strategies as the stimuli features were quite similar and made it hard to differentiate between stimuli by using semantic memory (for example when they were presented with different shades of red for the colour condition). Still, the possibility that participants used these kinds of strategies can not be excluded completely. Maybe similar studies with additional neuroimaging, for example with fMRI or EEG scans, could help observe the brain areas that get activated during the different attention conditions, giving more insights and ideas whether such strategies might be used to remember certain features or not, but this is outside of the scope of this study.

Another possible future direction for research is the interplay between mental imagery and visual working memory. Even though a lot of research was conducted on both areas respectively, only relatively few studies seem to focus on the interplay of the two. For such future studies researchers could provide participants with two tasks, comparable to the experiment that was conducted for this study. In one of the tasks, the orientation of the stimulus that has to be remembered differs from the orientation in the recall screen and in the second task the presented stimuli always have the same orientation as the stimuli in the recall screen. This way for the first task participants have to mentally rotate the stimulus to compare it to the recall stimuli while in the second task they can just recall what they initially encoded into their VWM when they attended to the stimuli. If the hypothesis holds true there should be significant differences between the results of the two tasks. The findings of such studies might shed more light on the functionality of human visual working memory and lead to even more sophisticated and detailed models.

To get more insights into the factor time and its role in visual working memory future researchers could include more time conditions, like it was already proposed in the discussion above. While bigger time differences should be more likely to produce a significant decay effect (if this theory holds true) the additional steps allow us to see the nature of the effect that time might have (e.g., if it is linear or if the memory just dies off after some time).

Conclusion

Other than expected the decay time between stimulus presentation and recall had at no point or level of the experiment a significant effect on the recall accuracy. This finding might be due to the limitations of the study. Further research should be conducted in this area.

The other results show clear evidence against the discrete capacity model as the recall accuracy significantly decreased with an increased number of features and the requirement to remember more different features, even though the stimulus amount was at no point higher than four. Instead, they support the limited resource model, which predicts that memory is a resource and not slot based and that an increased number of features per stimulus requires more resources and therefore has an influence on the recall accuracy (Lilburn, Smith, & Sewell, 2019). The results clearly showed that sets with the same amount of stimuli but more features to be remembered lead to worse recall accuracy. Still, the limited resource model is not sufficient to explain all the results, as different dimensions were recalled with different amounts of accuracy, indicating that either different types of features allow for different strategies to remember them, that they feed out of different resource pools (Wang et al., 2017) or that underlying processes like mental rotation take a toll on recall precision (Tong, 2015). Especially the latter theory gives room for a lot of exciting future research in the fields of Visual Working Memory and Mental Imagery and how the two interact and could result in models that have a way broader applicability and give more nuanced predictions for VWM performance than current models do.

References

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*, 201-208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Barrouillet, P., Portrat, S., Vergauwe, E., Diependaele, K., & Camos, V. (2011). Further evidence for temporal decay in working memory: reply to Lewandowsky and Oberauer (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 37, 1302–1317. doi: https://doi.org/10.1037/a0022933

Bach, M. (2007). The Freiburg visual acuity test-variability unchanged by post-hoc re-analysis. *Graefes Archive for Clinical and Experimental Ophthalmology, 245*(7), 965-971. doi:10.1007/s00417-006-0474-4

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, *9*(10), 7-7. https://doi.org/10.1167/9.10.7.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-114. doi:10.1017/s0140525x01003922

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135-168.

Tong, F. (2013). Imagery and visual working memory: one and the same?. *Trends in Cognitive Sciences, 17*, 489-490. https://doi.org/10.1016/j.tics.2013.08.005.

Huang, L. (2010). Visual working memory is better characterized as a distributed resource rather than discrete slots. *Journal of vision, 10*, 8-8. https://doi.org/10.1167/10.14.8.

Malenka, R. C., Nestler, E. J., Hyman, S. E., Sydor, A., & Brown, R. Y. (2009). *Molecular neuropharmacology: a foundation for clinical neuroscience.* NY: McGraw-Hill Medical.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97. doi:10.1037/h0043158

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Keogh, R., Pearson, J. (2014). The sensory strength of voluntary visual imagery redicts visual working memory capacity. *Journal of Vision*, *14*, 7. https://doi.org/10.1167/14.12.7.

Lilburn, S.D., Smith, P.L., Sewell, D.K. (2019). The separable effects of feature precision and item load in visual short-term memory. *Journal of Vision*, *19*(1), 2. https://doi.org/10.1167/19.1.2.

Souza, A. S., & Oberauer, K. (2015). Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. *Psychonomic Bulletin & Review, 22*(1), 156-162. https://doi.org/10.3758/s13423-014-0652-z

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied*, *74*(11), 1-29. https://doi.org/10.1037/h0093759

Christophel, T.B., Cichy, R.M., Hebart, M.N., Haynes, J.D. (2015). Parietal and early visual cortices encode working memory content across mental transformations. *NeuroImage*, *106*, 198-206, https://doi.org/10.1016/j.neuroimage.2014.11.018.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of experimental psychology: human perception and performance*, *27*(1), 92. https://doi.org/10.1037/0096-1523.27.1.92

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 226-236. doi:10.1037/xlm0000295

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233-235. https://doi.org/10.1038/nature06860

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological science*, *20*(4), 423-428. doi:10.1111/j.1467-9280.2009.02322.x