The processing of semantic contents of verbal working memory information during rehearsal

Myrthe Braam

Master Thesis - Applied Cognitive Neuroscience

3786722
June 2022
Department of Psychology
University of Groningen
Examiner/Daily supervisor:
Dr. Muhammet Ikbal Sahan

Abstract

The direct link between long-term memory (LTM) and verbal working memory (WM) remains controversial, even after years of research. Verbal working memory was thought to be primarily built on the processing of phonological aspects. Still, more recent research has demonstrated a possible interaction between LTM and verbal WM, wherein semantic processing of memoranda is involved. Mathôt et al (2017-2019) have made use of the eyes to get a closer look at the attention-based processes of WM. They discovered that pupils constrict while being presented with brightness-related words, and dilate while being presented with darkness-related words. These involuntary reflexes suggest the involvement of semantic processing in WM. In the current study we examined whether the pupils can reveal the semantic contents of verbal working memory information during rehearsal. We conducted a memory task wherein participants were verbally presented with brightness- and darkness-related words. The participants were instructed to subvocally rehearse the stimuli while their pupil size was being measured. The results revealed no significant difference in pupil size during the rehearsal of brightness- and darkness-related stimuli. This indicates that we did not find evidence for the contribution of semantic processing in verbal WM during rehearsal.

The processing of semantic contents of verbal working memory information during rehearsal

Working memory (WM) is an essential cognitive function that is in charge of maintaining information for short periods of time. It serves as an attentional control system that is responsible for the retention of limited capacity. According to one of Baddeley and Hitch' (1974) noteworthy models, WM is divided into verbal and visual subsystems serving the short-term storage of modality-specific information. While the storage of visual information is dependent on the visuospatial sketchpad governed by spatial attention, verbal WM is conceived as a storage system for verbal information that is dependent on articulatory processes in the phonological loop. The retention of verbally presented word sequences has been broadly examined within the phonological loop framework. The phonological loop is, among other components, the storage system subserving verbal information via subvocal rehearsal processes, also known as inner speech. This modality-specific storage system was understood to be a distinct component of long-term memory (LTM), since verbal information was thought to be stored as a sound-based code which purely relied on phonological aspects. Nevertheless, striking research has demonstrated a possible connection and co-operation between LTM and verbal WM (Hulme et al., 1991; Camos et al., 2018). For instance, it has been shown that lists of existing words are memorized better than lists of non-words (Hulme et al., 1991), which suggests the involvement of semantic processing during the maintenance of subvocal information. Whereas, more recently, Kowaliewski and Majerus (2018-2020) found that the effect of semantics in verbal WM is rather limited. They indicated that phonological knowledge and the associative nature of item sequences are primarily involved in the maintenance of verbal memoranda, with no dependence on item-level semantic processing.

In the current study we wanted to investigate whether the semantic contents of -item specific- verbal memoranda are activated during the rehearsal. To this end, we made use of eye movements (pupillary responses) to track the semantic contents during rehearsal.

Literature review

WM models

To keep information active in WM and to facilitate item encoding, attention is necessary. Attention is a recourse that helps with selecting and retaining essential information in our memory. In Cowan’s (1999) embedded-processes model of WM, the link between memory and attention is emphasized. The model represents WM as a two-phase process wherein WM information is hierarchically arranged. Instead of defining WM and LTM as two distinct components, Cowan described WM as an accumulation of embedded-processes from attention and LTM. Input could either be driven to an activated part of LTM, also known as short-term memory, or to the focus of attention. Whenever a stimulus is voluntarily attended to or novel, it will proceed to the focus of attention, which results in greater processing of the item. Relatedly, Oberauer‘s (2002) concentric model distinguished 3 components in WM; an activated portion of LTM, a focus of attention, and a capacity-limited area of direct access. The focus of attention is thought to process exclusively one representation at a time. This single representation in the focus of attention could either be a newly encountered stimulus selected from the region of direct access, or an item from the activated LTM to prevent its decay. The refocusing of attention on memoranda to keep mental representations active is referred to as ‘refreshing’. Johnson (1992) was one of the first researchers that coined the term refreshing, which has been elaborated on in recent research. Within the Multiple Modular memory framework (MEM), Johnson (1992) describes the mechanisms of recollection wherein refreshing operates to prolong the activation of representations. It is considered that a quickly rotating focus of attention could (re)activate multiple items rapidly one after another, resulting in the contemporary retention of various items (Vergauwe & Cowan, 2014). These reactivated WM representations benefit from being shifted to a heightened state of accessibility, as a function of attentional refreshing.

The time-based-resource-sharing model (TBRS-model) also postulates that item processing and storage relies on executive attention (Barrouillet et al., 2004). According to the TBRS-model, the maintenance of verbal stimuli involves an executive loop and a phonological loop. The executive loop includes a procedural system and an episodic buffer wherein WM representations are held. Whenever items need to be retained, their retrospection is induced via attentional refreshing. In addition, the maintenance of verbal memoranda involves articulatory rehearsal, which is also described in Baddeley's Phonological loop model (1986). This non-attentional mechanism uses inner speech to reactivate phonological memory traces. The domain-specific influences of the phonological loop, and the attention-based mechanisms of refreshing in the executive loop depict a clear distinction in verbal WM mechanisms. Whether one of these separable mechanisms plays a greater role in verbal WM remains a frequently asked question. Still, the domain-general mechanisms of attentional refreshing suggest some form of LTM contribution.

The interaction between LTM and verbal WM

Camos et al. (2018) investigated the possible association between refreshing and semantic processing in verbal WM. They examined the influence of semantic versus phonological aspects of memoranda on WM recall. In one of their experiments, participants were instructed to complete a memory span task wherein they continuously had to remember and recall five words. Whenever the participant had forgotten a word, they could ask for help by using a cue. The cues for the to-be-remembered stimuli were either a phonological cue or a semantic cue. The results showed that WM recall was greater when semantic retrieval cues were used compared to phonological cues. This demonstrates a possible contribution of LTM in WM processes, where refreshing seems to effectuate the retrieval of LTM knowledge.

Moreover, there is supplementary evidence suggesting that semantic LTM effects are observed in verbal WM processes. Multiple studies have put focus on a variety of factors that were known to impact the retrieval of semantic information from LTM. Examples are the frequency effect (better recall for high-frequency words than low-frequency words) (Popov & Reder, 2020), the lexicality effect (better recall for words than non-words) (Hulme et al., 1991; Loaiza et al., 2015) and the concreteness effect (better recall for concrete words than abstract words) (Bourassa & Besner, 1994). Respectively, Hulme et al. (1991) studied the lexicality effect by administering a memory span task to participants. Lists of word and non-word pairs were verbally presented to participants who were instructed to recall them in the correct order. It was evident that WM performance was consistently higher for words compared to non-words. Specifically, during rehearsal and recall, the existing words presumably benefited from LTM storage in addition to being held in the phonological loop. This LTM contribution demonstrates that rehearsal is not purely build on phonological aspects but possibly also relies on semantics.

Nevertheless, others claim that the interaction between LTM and verbal WM is limited to the associative nature within the item sequence rather than the item-level LTM associations. Kowaliewski and Majerus (2020) showed that word sequences wherein the individual words share semantic features, tend be better recalled than sequences that do not share semantic features. The same was true for sequences containing words that were phonologically similar to each other. Specifically, semantically and phonologically related words stimulate one another via activation between the semantic and lexical levels of representations, which could lead to greater inter-item associations. This suggests a rather limited involvement of item-level semantics in verbal WM. Furthermore, Kowaliewski and Majerus (2018) did a similar experiment to Hulme et al. (1991) where the maintenance of non-words was compared to the maintenance of existing words. They observed greater WM performance for words compared to non-words as well. Still, they interpreted this phenomenon rather differently. Instead of relating this outcome to item-level semantic knowledge, they postulated an extensive involvement of morphological and phonological knowledge. Words exist out of well-known phonemes that make them recognizable and easier to recall, compared to non-words, which is unrelated to semantic processing. Besides this, highly imaginable words did not lead to detectable benefits on recall performance, compared to less concrete, abstract words in the study by Kowaliewski and Majerus (2020). This suggests that item-level semantic knowledge in itself does not bring about any differences for the recall of memoranda.

The present study in relation to existing literature

In the current study we wanted to investigate whether semantic processing of verbal WM contents is involved during rehearsal. According to the previously stated WM models the maintenance and reactivation of WM representations is thought to rely on attentional refreshing, governed by executive functioning, where LTM contribution is suggested (Cowan, 1991; Oberauer, 2002; Barrouillet et al., 2004). Still, the role of rehearsal and refreshing in verbal WM has been controversial in the literature. The potential influence of semantic processing in verbal WM could offer valuable information regarding the direct link between verbal WM and LTM. To this end, we wanted to study the involvement of attention-based rehearsal mechanisms via the eyes. There is evidence showing a convincing involvement of the oculomotor system in attentional shifts (Postle, 2006; Van der Stigchel & Hollingworth, 2018). Therefore, eye-movements provide great opportunities to study attention-based mechanisms since they are part of a tightly integrated system involving attention and pupillary responses.

Previous work by Sahan et al. (2021) has revealed the involvement of eye-movements in retrieving spatial information of verbal WM. They tested whether systematic eye-movements were made while participants were instructed to recall item sequences that were verbally presented. Their study showed that the recollection of a sequence of items resulted in leftward eye-movements when recalling begin items and in rightward eye-movements when recalling end items. This finding demonstrates the assumption that sequential inputs are bound to so-called position markers, and thus reveals the involvement of attention-based processes in verbal WM. While this study reveals the spatial context, it is not clear yet whether visiting the spatial context, to which memoranda are bound, also activates the semantic contents.

The investigation of pupillary response fluctuations is an interesting approach to reveal the semantic contents during verbal WM rehearsal. A series of eye-tracking studies conducted by Mathôt et al. (2017-2019) has shown that words related to brightness evoke pupillary changes. It is known that pupils constrict as a consequence of light and dilates as a consequence of darkness, which is called the pupil light response (PLR) (Mathôt, 2018). Just like PLR is evoked by natural light, Mathôt et al observed that the semantic brightness of words induces similar responses. Participants were visually and verbally presented with words associated with brightness and darkness (e.g. sun and night) while their pupil sizes were being measured. The eyes' pupils were smaller when people read or listened to words related to brightness than when people read or listened to words related to darkness. Here, the involvement of semantics in WM is clearly depicted. They showed that word meaning can cause involuntary (pupillary) responses. This illustrates that eye-tracking can be used as a resource to investigate WM. Using this methodology, we wanted to address whether similar semantic effects can be evoked in the retention of verbal WM contents in particular.

Specifically, the present study investigated whether attention-based processes can be tracked during the active maintenance of verbal memoranda that is achieved via subvocal rehearsal. We observed and compared pupil sizes during the rehearsal of darkness and brightness-related words. Based on the earlier work of Mathôt et al (2017-2019), the presentation of words related to brightness evokes pupils to constrict and the presentation of words related to darkness evokes pupils to dilate. We expected to find similar results as to when these words are rehearsed in verbal WM. Since WM is thought to be an activated part of LTM, we predicted activation of semantic LTM knowledge during the rehearsal of verbally presented stimuli. In our experiment, the dependent variable was the pupil size during the rehearsal of verbally presented words in a memory task. The independent variables were the semantic categories to which the words belonged to (i.e., dark vs bright). We hypothesized that (1) during the rehearsal of words conveying sense of brightness, the pupil size will be smaller as compared to the pupil size during the rehearsal of words conveying sense of darkness, and (2) during the rehearsal of words conveying sense of darkness, the pupil size will be larger as compared to the pupil size during the rehearsal of words conveying sense of brightness.

Method

Participants

29 students from the University of Groningen (13 males, 16 females) with a mean age of 19.5 years (SD = 1.9, range = 18-25) participated in the experiment in return for SONA-credits. Our sample size (N = 29) was based on the analogous study of Mathôt et al. (2019) that executed a similar investigation based on the effect of cognition on the PLR. The data of one participant was deleted and replaced as the system crashed midway the task. Vision was normal or corrected-to-normal for every participant. All but 3 participants were right-handed. Every participant was a native speaker of Dutch, and all those participating gave written

informed consent. The research adhered to the guidelines of the local ethics committee of the University of Groningen.

Apparatus

The stimuli were presented on a 22-inch LCD monitor (1,920 × 1,080 resolution) at a 0.60 m distance from the participants. The presentation of the stimuli was constructed by using MATLAB (The MathWorks, Natick, MA) with Psychtoolbox-3 extensions (Brainard, 1997). The auditory stimuli were presented via speakers (output hardwire devices) that were connected to the computer. We made use of an EyeLink 1000 eye-tracker (SR Research, Canada) for the recording of the pupillary responses. The pupil size was continuously measured during the experiment, and was recorded with a sampling frequency of 1000 Hz. To reduce the head movements, we made use of a chin rest. A strap was used to support the forehead. The participants were instructed not to wear mascara or any related eye-make-up. The experiment took place in a dimly-lit laboratory setting.

In advance of every practice trial and every full session of 160 trials, the eye tracker was calibrated to the screen. The right eye of every participant was selected and calibrated by using an in-built 5-point calibration protocol. If needed, recalibration was done during the session. Drift-correction was applied before the start of each trial.

Stimulus selection

We selected 20 darkness-related words, 20 brightness-related words, 20 control words (i.e. non- related to brightness or darkness), and 20 animal names that were used in Mathôt et al.’s (2019) study (see appendix 1). Every word that was used in the experiment was Dutch. The darkness- and brightness-related words were matched on their amount of letters (bright M = 6.10; dark M = 6.00; range 3 – 11) and lexical frequency. We made use of the Mac OS textto-speech synthesizer (say). The animal names and control words were not matched on number of letters or lexical frequency. Therefore, only the darkness- and brightness-related words were considered in the statistical analyses.

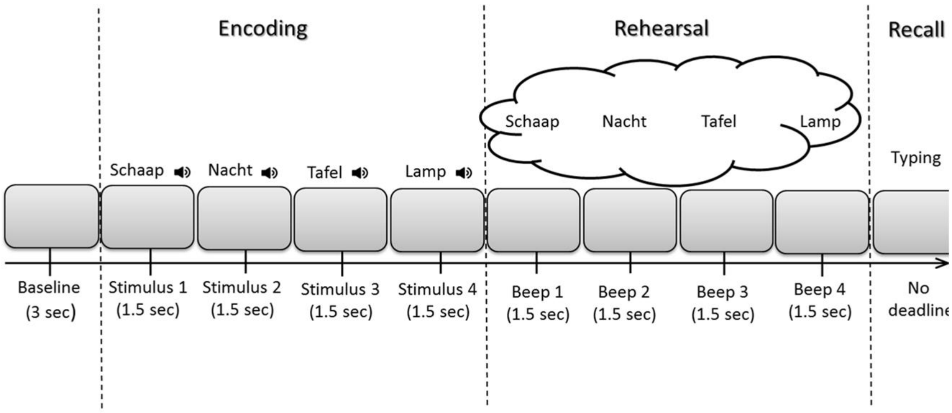
Procedure and design

The participants were instructed to start each trial by pressing the spacebar while being presented with a central dark gray fixation-dot on a gray background. It was only possible to start each trial when the pupil's focus was on the fixation dot. Each trial consisted out of three phases (see figure 1): In the encoding phase (phase 1) the participants were verbally presented with four words with an interval of 1.5 seconds between each word. Each word was sampled from a different semantic category (i.e., brightness-related, darkness-related, animal name and control word). Participants were instructed to memorize these words in the right order. During the rehearsal phase (phase 2), four auditory beep tones were presented at the same pace as the spoken words were presented (with an interval of 1.5 seconds between each tone). Participants were instructed to mentally rehearse the words at the same pace of the tones, with each tone serving as a rehearsal cue for every word in the memory sequence. More specifically, the first word had to be rehearsed on the first tone, the second word on the second tone and so on. During the recall phase (phase 3), the participants were presented with a central question mark. This was an indicator for them to type the four words, one by one, in the correct order. There was no deadline in time during the recall phase. Participants were instructed to type an 'x' whenever they were unable to recall a word.

The experiment began with two practice trials. Subsequently the participants completed 160 trials (8 blocks of 20 trials). The order of the semantic categories was pseudo-randomly presented across the trials such that each category was presented at each WM position equally often (40 times). This, together with the number of participants (N = 29) ensured a reasonable number of observations per participant.

Figure 1.

Trial interval lay-out



Note. Each trial consisted out of three phases: an encoding phase, a rehearsal phase and a recall phase. The words that are used in the experiment are in Dutch. Ordered translation: sheep, night, table, and lamp.

Preprocessing and analyses

In order to correctly organize the analysis we performed the following pre-processing. Since the data were recorded at a high sampling frequency, we downsampled the frequency to 100 Hz. Blinks were reconstructed by eliminating the data 100 MS before and after the blink. Subsequently, the missing data were linearly interpolated. We excluded the trials where the range of missing data exceeded 500 ms (based on Salvaggio et al., 2019). For each trial the pupil sizes were z-transformed to acquire comparable measures across participants and trials. Improbable pupil values were removed from the data by specifying a lower (-5 SD) and upper (5 SD) pupil boundary.

The time course of pupil data during the rehearsal and encoding period was extracted for each subject and trial. Subsequently the pupil time courses were grouped in line with the semantic categories at each WM position separately and averaged over the time course. The result of these steps is that we ended up with a mean pupil size at each WM position for each semantic category. The semantic category of the presented word (e.g. darkness- or brightness-related) was the independent variable, and the mean pupil size as a consequence of that word was the dependent variable. Throughout the analysis a significance level of 0.05 was applied. We analyzed the mean pupil size during the rehearsal and encoding phase. These mean pupil sizes were subjected to a 4 x 2 repeated measures analysis which was conducted with 4 WM positions (1, 2, 3, and 4) and 2 semantic categories (dark and bright). We analyzed the response accuracy by comparing each given stimulus in the encoding phase with the participants’ responses in the recall phase. We computed the response accuracy for the complete trials, and the response accuracy for each WM position.

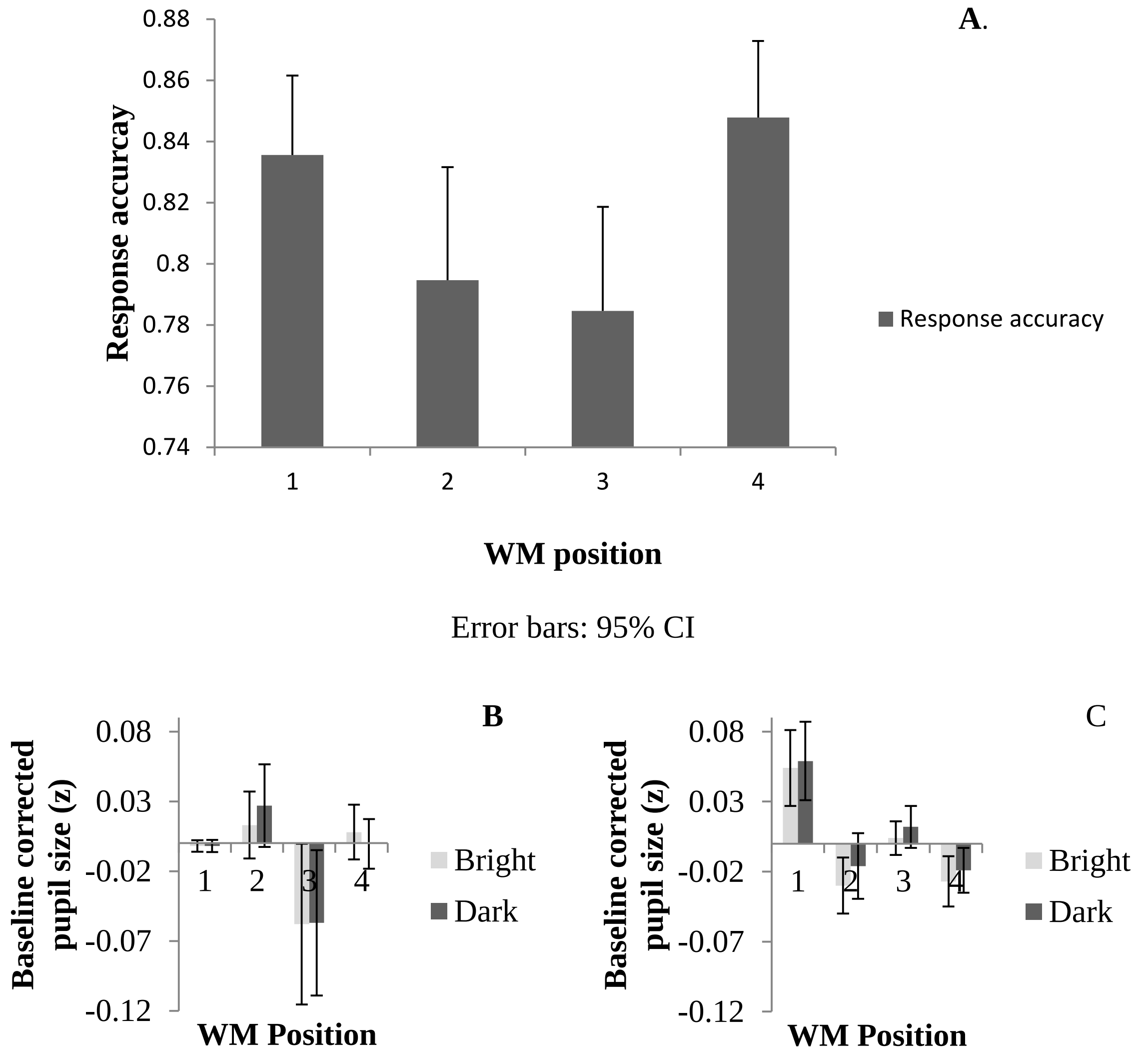
Results

Behavioral results

The response accuracy for the overall trials was 52%. The response accuracy for the stimuli on each WM position is given in figure 2a; 84% for position 1, 80% for position 2, 79% for position 3, and 85% for position 4.

Figure 2.

Analyses results



Note. (a) Response accuracy was the highest for WM position 1 and 4. During (b) the rehearsal phase, and (c) the encoding phase, the pupil sizes did not significantly differ for the different semantic categories (bright and dark).

Pupil-size results

A 4 x 2 repeated measures ANOVA was conducted to compare pupil sizes between the semantic and the WM position condition in the rehearsal phase. As determined by the repeated measures ANOVA, semantics had no statistically significant effect on pupil size (F(1, 28) = 0.03; p = .87, η2 = .0007). This indicates that there was no significant difference in pupil size during the rehearsal of brightness- and darkness-related words (see figure 2b). This was confirmed by the 95% confidence interval of the mean difference including zero (*M*: -0.001, *SE*: 0.01, 95% CI [-0.02, 0.02]). We did find a significant difference in pupil size for WM positions (F(3, 84) = 5.59; p = .002, η2 = .13). Regarding the post hoc Bonferroni test, there was a statistically significant difference in pupil size on position 3 compared to position 1 (p = .044), position 2 (p = .001), and position 4 (p = .019). Furthermore, there was no significant interaction between semantics and WM position (F(3, 84) = 0.38; p = .77, η2 = .002), which suggests that presenting the semantic categories on different positions had no effect on pupil size.

Secondly we examined the pupil size data from the encoding phase. As determined by the 4 x 2 repeated measures ANOVA, semantics had no statistically significant effect on pupil size during the encoding period (F(1, 28) = 1.79, p = .13, η2 = .005). This indicates that there was no significant difference in pupil size during the rehearsal of brightness- and darkness-related words (see figure 2c). This was confirmed by the 95% confidence interval of the mean difference including zero (*M*: -0.01, *SE*: 0.01, 95% CI [-0.02, 0.01]). We did find a WM position effect on pupil size (F(3, 84) = 18.00; p = < .001, η2 = .30). The post hoc Bonferroni test indicates that there was a statistically significant difference in pupil size on position 1 compared to position 3 (p = < .001), position 2 (p = < .001), and position 4 (p = < .001). Furthermore, there was no significant interaction between semantics and position ($F(3, 84)$ = .13; p = .94, $\eta 2$ = .0006), which suggests that presenting the semantic categories on different positions had no effect on pupil size.

Discussion

In the current study we wanted to reveal whether the semantic contents of -item specific- verbal memoranda are activated during rehearsal. We made use of pupillary responses to track the semantic contents. Our first hypothesis was that pupil size will be smaller during the rehearsal of brightness-related words, compared to darkness-related words. Secondly, we hypothesized that pupil size will be larger during the rehearsal of darkness-related words compared to brightness-related words.

The present results did not confirm our hypotheses. No significant difference in pupil size was detected during the rehearsal of brightness- and darkness-related stimuli. This outcome could have been caused by various reasons. A possible explanation could be a lack of attentional engagement during the maintenance of the words. It was predicted that attentional engagement contributes to the semantic processing of the presented stimuli. Here, the eyes were used as a resource to reflect attention-based processes from verbal WM. Nevertheless, the average pupil sizes did not reveal significant fluctuations due to the differing semantic categories. Hence, the study seems to imply that no attentional engagement was present during the maintenance of the verbal memoranda. More specifically, the retention of the words might have been purely dependent on articulatory rehearsal. It has been proposed that the rehearsal and refreshing of verbal information are distinct cognitive operations and robustly differ in functional connectivity (Raye et al., 2007). Relatedly, the TBRS-model suggests that the maintenance of verbal stimuli involves two separable mechanisms; an executive loop and a phonological loop (Barrouillet et al., 2004). We expected interplay between the two operations, where both attentional refreshing and phonological rehearsal are engaged in the maintenance of verbal memoranda. While the maintenance and processing of memoranda within the executive loop is profoundly dependent on attention, the domain-specific phonological loop requires no attention for the maintenance of verbal information (Barrouillet et al., 2004). Since attention is thought to be essential for the retrieval of semantic aspects of WM contents from LTM (Chen & Cowan, 2009), our results suggest that solely the phonological loop was activated.

The study by Vergauwe et al. (2014) implies that both operations (attentional refreshing and articulatory rehearsal) can be used jointly, but also independently from each other for the maintenance of verbal information. When people are asked about their maintenance strategy in a memory task, articulatory rehearsal has been the most frequently used (Oberauer, 2019). Refreshing, on the other hand, has been reported on a way smaller extent as a rehearsal strategy. This indicates that participants could favor one of the systems over the other. The superiority of one of these systems could be dependent on the instructions. It may be the case that the inequality in instructions resulted in differences between our study results and those of Mathôt et al. (2019). Mathôt et al. found significantly smaller pupils while reading, and importantly, hearing brightness-related words, compared to darkness-related words. The participants were informed to press spacebar whenever they saw or heard an animal name. These instructions ensured attentional engagement during the task, since semantic processing of the stimuli was inevitable. In our study we only instructed the participants to memorize the presented words by mentally repeating them in the rehearsal phase. This may explain the lack of attentional engagement and the dominance of phonological rehearsal during the task. In future investigations, it might be necessary to ensure attentional engagement during the maintenance of verbal memoranda. This could be achieved by replacing one of the four tones in the rehearsal phase with a higher pitched beep, together with giving the instructions to press spacebar whenever this deviating beep is an animal name.

Since we made use of item sequences in our task, it is assumed that there is some involvement of serial order processing. According to Abrahamse et al. (2014) the idea of spatial organization of serial input is explained by means of the mental whiteboard hypothesis. Position markers are seen as coordinates for memoranda within a mental spatial system, where spatial attention is essential for the search and selection of these memoranda. These to-be-remembered items are represented and maintained on a supposed 'mental whiteboard'. The mental whiteboard hypothesis seems to be connected to attentional refreshing, since arranging verbal stimuli in a mental space will endorse spatial attention to scan across the items for the refreshing of their representations (Vergauwe & Cowan, 2015). This assumption indicates an involvement of attentional refreshing in our experiment, which seems unlikely according to our results. Moreover, our results revealed a certain primacy effect (better recall for the first item of a memory list) and recency effect (better recall for the last item of a memory list) considering the higher response accuracy for stimuli on the first and last WM position (Ebbinghaus et al., 1913). While the recency effect is thought to be stronger when articulatory rehearsal is predominant, the primacy effect is thought to be related to influences of refreshing (Lemaire et al., 2021). Attentional refreshing always starts with the reactivation of the first item in a sequence, also when distractions take place. This describes that the first item of a sequence is likely reactivated to a greater extent, hence the primacy effect. This, once more, suggests the contribution of attentional refreshing in our task, which seems to be contradicting our results, since attentional refreshing is expected to bring about some form of LTM contribution. A possible explanation brings us back to

Kowaliewski and Majerus' (2018/2020) point of view on the relation between verbal WM and LTM. They postulated that the retention of stimuli were dependent on the semantic and phonological associations between the items in a list. The inter-item connections between words that are sharing semantic and phonological features could lead to better WM performance. Furthermore, Kowaliewski and Majerus (2020) discussed that the retention of stimuli could be highly dependent on the phonological knowledge of the memoranda, instead of the semantic processing. This frame of reference indicates that there might be LTM contributions in verbal WM processes, yet these contributions seem to be limited to inter-item associations and phonological knowledge instead of item-level semantics. In that matter, the involvement of attentional refreshing in our experiment seems conceivable.

While the previously stated justifications seem to provide reasonable explanations for our inability to reveal semantic processing during the rehearsal phase, it remains questionable that we did not find a semantic effect in the encoding phase. Mathôt et al.'s (2017/2019) series of eye-tracking studies put focus on pupil sizes during the encoding of words. Therefore, we decided to analyze the data from the encoding phase in our task as well. Nonetheless, we still did not find a difference in pupil size during the presentation of the words conveying sense of brightness and darkness. Consequently, we started to question our methodology. A possible explanation could be the overlapping of pupillary responses by virtue of the consecutive stimuli. We used four different semantic categories in every word sequence. The immediate switch from word meaning (1.5 seconds) might provide inadequate time for the pupil to re-adjust and thereafter adapt according to the new stimulus. For example, whenever a brightness-related stimulus is directly followed by a darkness-related stimulus, the probable initial response of the pupil would be constriction, to be followed by dilation. The immediate succession of these stimuli presumably imbricates their pupillary responses. Wierdaa et al. (2012) examined the pupil dilation time course, and described that the pupillary responses to two closely succeeding stimuli are destined to overlap. This notion has also been discussed by Hoeks & Levelt (1993). They determined and investigated the linear time course of the pupil diameter. An evident overlap in pupillary responses had been detected while reacting to stimuli in close succession. In future investigations it would be important to take into account the possibility of overlapping pupillary responses. Therefore, the rapid succession of different semantic categories should be avoided. The usage of one semantic category per trial would be highly recommended. This ensures constancy in pupillary responses for each category, which enables more appropriate comparisons.

In summary, we did not find a significant difference in pupil size during the rehearsal of brightness- and darkness-related words. More specifically, pupil size was not smaller during the maintenance of brightness-related words compared to darkness-related words, and pupil size was not larger during the maintenance of darkness-related words compared to brightness-related words. These results determine that we did not find evidence for the contribution of semantic processing in verbal WM during the rehearsal of brightness- and darkness-related words.

References

Abrahamse, E., van Dijck, J.-P., Majerus, S., & Fias, W. (2014). Finding the answer in space: the mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience*, *8*(November), 1–12. https://doi.org/10.3389/fnhum.2014.00932

Baddeley, A. D. (1986). Working memory. Clarendon Press/Oxford University Press.

Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-90). Academic Press.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi-org.proxy-ub.rug.nl/10.1037/0096-3445.133.1.83

Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review*, *1*(1), 122–125. https://doi-org.proxy-ub.rug.nl/10.3758/BF03200768

Brainard, D., H. (1997). The Psychophysics Toolbox. *Spatial Vision*,*10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, 1424(1), 19–32. https://doi.org/10.1111/nyas.13616

Chen, Z., & Cowan, N. (2009). Core verbal working-memory capacity: The limit in words retained without covert articulation. *The Quarterly Journal of Experimental Psychology*, *62*(7), 1420–1429. https://doi-org.proxy-ub.rug.nl/10.1080/17470210802453977

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 62–101). Cambridge University Press. https://doi-org.proxy-ub.rug.nl/10.1017/CBO9781139174909.006

Ebbinghaus, H., Ruger, H. A., & Bussenius, C. E. (1913). *Memory: A contribution to experimental psychology*. Teachers College Press. https://doi-org.proxy-ub.rug.nl/10.1037/10011-000

Hoeks, B., & Levelt, W. J. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments & Computers*, *25*(1), 16–26. https://doi-org.proxy-ub.rug.nl/10.3758/BF03204445

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685–701. https://doi-org.proxy-ub.rug.nl/10.1016/0749-596X(91)90032-F

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, *4*(3), 268–280. https://doi-org.proxy-ub.rug.nl/10.1162/jocn.1992.4.3.268

Kowialiewski, B., & Majerus, S. (2018). The non-strategic nature of linguistic long-term memory effects in verbal short-term memory. *Journal of Memory and Language*, *101*, 64–83. https://doi-org.proxy-ub.rug.nl/10.1016/j.jml.2018.03.005

Kowialiewski, B., & Majerus, S. (2020). The varying nature of semantic effects in working memory. *Cognition*, *202*. https://doi-org.proxy-ub.rug.nl/10.1016/j.cognition.2020.104278

Lemaire, B., Heuer, C., & Portrat, S. (2021). Modeling articulatory rehearsal in an attention-based model of working memory. *Cognitive Computation*, 13(1), 49–68. https://doi-org.proxy-ub.rug.nl/10.1007/s12559-020-09791-9

Loaiza, V. M., Rhodes, M. G., & Anglin, J. (2015). The influence of age-related differences in prior knowledge and attentional refreshing opportunities on episodic memory. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, *70*(5), 729–736. https://doi-org.proxy-ub.rug.nl/10.1093/geronb/gbt11

Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary Responses to Words That Convey a Sense of Brightness or Darkness. *Psychological Science*, 28(8), 1116–1124. https://doi.org/10.1177/0956797617702699

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, 1(1), 1–23. https://doi.org/10.5334/joc.18

Mathôt, S., Sundermann, L., & Van Rijn, H. (2019). The effect of semantic brightness on pupil size: A replication with Dutch words. *bioRxiv*. https://doi:10.1101/689265

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411–421. https://doi-org.proxy-ub.rug.nl/10.1037/0278-7393.28.3.411

Oberauer, K. (2019). Is Rehearsal an Effective Maintenance Strategy for Working Memory? *Trends in Cognitive Sciences*, 23(9), 798–809. https://doi.org/10.1016/j.tics.2019.06.002

Popov, V., & Reder, L. M. (2020). Frequency effects on memory: A resource-limited theory. *Psychological Review*, *127*(1), 1–46. https://doi-org.proxy-ub.rug.nl/10.1037/rev0000161

Postle, B R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *43*(1), 135–145. https://doi-org.proxy-ub.rug.nl/10.1016/S0010-9452(08)70451-9

Sahan, M. I., van Dijck, J. P., & Fias, W. (2021). Eye-movements reveal the serial position of the attended item in verbal working memory. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/s13423-021-02005-9

Salvaggio, S., Masson, N., & Andres, M. (2019). Eye position reflects the spatial coding of numbers during magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1910–1921. https://doi-org.proxy-ub.rug.nl/10.1037/xlm0000681.supp

Van der Stigchel, S., & Hollingworth, A. (2018). Visuospatial working memory as a fundamental component of the eye movement system. *Current Directions in Psychological Science*, *27*(2), 136–143. https://doi.org/10.1177/0963721417741710

Vergauwe, E., Camos, V., & Barrouillet, P. (2014). The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1072–1095. https://doi-org.proxy-ub.rug.nl/10.1037/a0035779

Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval rate may describe many cognitive procedures. *Frontiers in Human Neuroscience, 8,* Article 126. https://doi.org/10.3389/fnhum.2014.00126

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: Evidence that refreshing and memory search are closely related. *Psychonomic Bulletin & Review*, *22*(4), 1001–1006. https://doi-org.proxy-ub.rug.nl/10.3758/s13423-014-0755-6

Wierdaa, S. M., van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *109*(22), 8456–8460. https://doi-org.proxy-ub.rug.nl/10.1073/pnas.1201858109

Appendix

The stimuli that were used in the experiment (based on Mathôt et al., 2019)

Bright	*Dark*	*Animal*	*Filler*
Dag	Avond	Aap	Beheer
Daglicht	Blinddoek	Beer	Gedroomd
Flits	Bruin	Capybara	Gestrooid
Glanzend	Dof	Ekster	Graan
Glans	Doffe	Geit	Jas
Helder	Donker	Hagedis	Juist
Lamp	Duister	Hond	Kloppen
Lamplicht	Gordijn	Inktvis	Leerling
Licht	Hol	Kat	Lof
Schijnen	Kelder	Kever	Nota
Sneeuw	Kuil	Konijn	Oom
Stralen	Modder	Kwal	Punten
Stralend	Nacht	Mug	Schouwburg
Straling	Nachten	Muis	Tafel
Wit	Somber	Olifant	Toetsen
Witte	Schaduw	Paard	Transport
Zomers	Schoorsteen	Poes	Trui
Zon	Winter	Schaap	Voordoen
Zonlicht	Winters	Tor	Vuisten
Zonning	Zwart	Varken	Wacht