Experimental study

How sequentially shown information facilitates performance of working memory: A comparison between single- and multimedia presentations

A group project for the course *Human Cognition in Human-Computer-Interaction*, held by Robert Ramberg, at the Stockholm University Department of Computer and System Sciences during the autumn semester 2019, by Group 1 - Sebastian Veuskens, Thomas Castro, Thibaud Marchand and Lena Spitz

Abstract

Former working memory studies identified two executive systems of our working memory, the phonological loop and the visual sketchpad [1]. To compare the capacity of these two systems of human memory, we conducted an experimental study where the participants had to remember as much items as they could, with the items being presented either as words or pictures. We further included a third type of presentation that combined both words and pictures in order to investigate if this use of multimedia facilitates remembering or not. The results of our experimental study suggest that showing the items in only one way supports meaningful processing of information and thus leads to a better memory performance. In contrast to that, the use of multimedia seemed to inhibit the “meaningful interpretation of information” [2] and as a consequence led to poorer performance.

Introduction

In this day and age, the digitized and multimedia world produces an enormous amount of visual information each second that has to be processed and memorized by humans. Therefore, many aspects of our daily life depend on the capacity of human memory, making it a broad research field with many subfields of interest. One of these subfields is short-term memory - a limited storage for information which will be forgotten as new information comes in, unless the information is rehearsed, repeated or processed in depth with high concentration. This leads to a higher probability of committing it to long-term memory. The capacity of the short-term memory used to be equalled with the memory span, meaning the amount of information one is able to repeat back immediately after hearing, and is about seven to eight items [1].

In 1986 Baddeley developed a theory about working memory, which refers to “a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning” [3]. It consists of a central executive that controls two slave systems: the visuospatial sketchpad and the phonological loop [4]. The visuospatial sketchpad refers to visually remembering

information, meaning having a picture of the information in the mind. The phonological loop refers to an inner voice, the articulatory loop, that repeats information to oneself in the mind which can be heard, processed and stored by an inner ear, the phonological store [1].

Experiments by Baddeley, Thomson and Buchanan in 1975 and Vallar and Baddeley in 1982 provided evidence for the phonological loop, as participants could remember more short words than long words at a time. Length of a word was measured by the number of syllables. The amount of words remembered strongly correlated with the amount of syllables participants were able to say in a certain amount of seconds. According to Baddeley, humans can thus keep information worth 1.5 to 2 seconds in the working memory [1].

The distinction between short-term and working memory is often imprecise since these memory types are based on different theoretical concepts and are supposed to overlap each other to a large extent [5]. The two are hard to differentiate, but ultimately working memory is meant to keep information readily available to perform a task, and not to process the information in order to commit it to long-term memory [1].

From the concept of the central executive and its slave system it follows that pictures are processed via the visuospatial sketchpad and written words via the phonological loop. The question arises if pictures or written text is easier to retain in the working memory and consequently if phonological loop or visuospatial sketchpad is more effective, and how far an anticipated overlap extends. Thus we developed an experimental study to compare the working memory's capacity for pictures and for written words. Since modern multimedia often uses a combination of both, we also compared the results to the capacity for annotated pictures.
The basic setup of our experiment is illustrated in *Figure 1*.

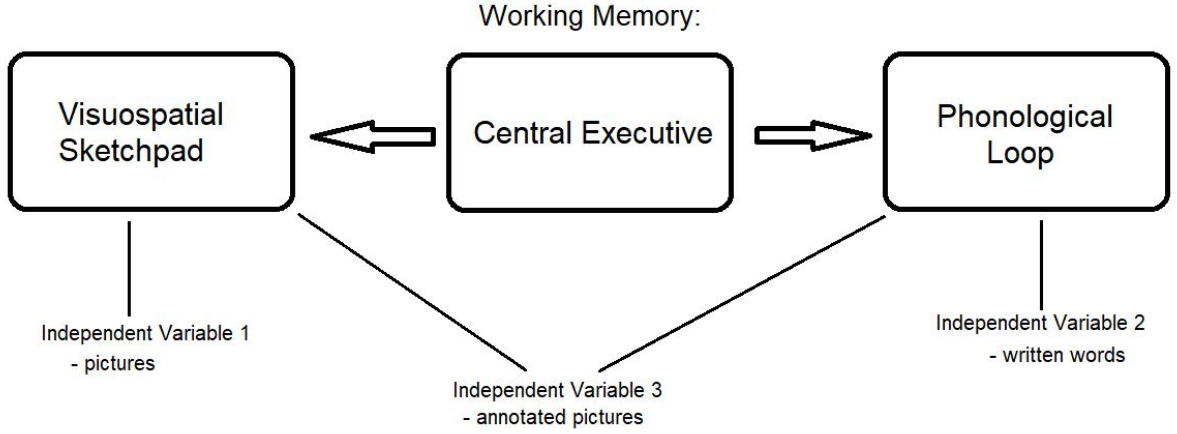

Figure 1: The basic setup of the presented experiment, where three values of the independent variable (pictures, words and annotated pictures) were tested for their effectiveness in regards to processing in the working memory.

There have been studies related to our research question.
The use of multimedia over single media intuitively leads to the assumption that it encourages more ways of engaging with the presented material, thus enabling deeper

processing and therefore heightening memory performance. However, studies have so far not found evidence that this is the case [6].
Another study by Shepard in 1967 indicates that the visual memory has a greater capacity than the verbal one. However, it did not focus on the working memory [1]. In contrast, our approach focused more on the ability to process information in a limited period of time. Furthermore we reached a higher level of comparableness since the same information was presented both in picture form and in written word form.

Research question

How does the type of presentation of items (words, pictures, or both) affect the human working memory, and is one type better to memorize as much items as possible?

Hypothesis

We hypothesized that annotated pictures would be easiest to process and recall for the working memory, as it encourages the use of both phonological loop and visuospatial sketchpad by providing an image that can be mentally visualized and the written name of what is shown in the picture, which can be repeated in the mind.

Methodology

The Experiment

We tested 30 individuals. Each individual was shown a set of items, presented either as pictures, words, or pictures and words combined.
Results were then compared to see if there is a difference in how many items an individual can remember on average depending on how they are presented to them. The amount of remembered words functioned as an indicator that reported how much the different types of presentation facilitate the capacity of human memory.

Design

Variables and criteria for selection

1. **Dependent variable:**
    - Number of correctly, incorrectly and not remembered items (working memory).
2. **Active independent variable:**
    - Specified number of items presented as pictures, written words or annotated pictures.
3. **Control variables:**
    - The English level of the participants had to exceed a minimum threshold equivalent to B1 (Meaning they “can understand the main points of clear

standard input on familiar matters regularly encountered in work, school, leisure, etc." [9]).
- The pictures and words represented the same items so each group had to remember the same items regardless of the independent variable.
- The items were shown to the participants for a set amount of time namely two seconds (for details consult the "Pilot" section).
- The items were always presented in exactly the same way since the instructions and presentation were shown on a computer and there was no interaction with the conductors of the experiment (except the welcome and in the case of questions referring to the experiment.)

Our experiment was structured into three different types of presentations, namely presentation of words, pictures or pictures and words combined. Thus the design of our experiment consisted of three experimental groups whose results were compared to each other, following the design concept of Fred N. Kerlinger [7].

We chose a **Between Subject Design**, meaning each participant was part of one group assigned to one value of the active independent variable. Thus, everyone had to memorize the same items (control variable) and there was no learning effect as would occur if each participant had been presented with all three different values of the independent variable in a within subject design [8].

Procedure

We assigned the independent variable of the experiment in repeating order to make sure each independent variable value had a comparable sample size consisting of ten participants. Half of the participants were friends or acquaintances of ours, who could book timeslots in a schedule we provided. The other half were strangers we asked directly at the place where we conducted the experimental study to fill in the holes in our schedule and to get more participants. The experiment was conducted in closed group rooms at the Frescati Library and the DSV in Stockholm.

There were 3 different groups :
- First group was shown a set of written words
- Second group was shown a set of pictures which represent the same words
- Third group was shown the same pictures which were captioned with the words they represented

The items were shown sequentially, each for a set amount of time before they were removed. The exact amount of time was determined within our pilot study.
All individuals were then asked to fill out a survey with the items they remembered from what has being shown to them (either pictures, words, or annotated pictures). They were allowed to write the items in their mother tongue if they didn't remember them in English. Furthermore, they were asked to describe how they remembered the words as an open question. They were subsequently asked if they imagined the items as a picture in their mind

(visuospatial sketchpad) or used an inner voice to repeat them to themselves (phonological loop) to memorize the items.

Material

The items were presented as an automatic slideshow in Microsoft Office PowerPoint on a **computer** in an **isolated room**. This ensured reproducibility and that all pictures and words were represented in the same way and size and for an exact amount of time.
A **slide** on the computer contained the items in the three different types of presentation (words, pictures and annotated pictures). The items were all presented one at a time in random order to ensure comparableness and to eliminate bias based on the order of the items.
A digital **survey** created in Google Forms was used to fill in the remembered items of the participants, to collect general information like age, gender, English level and whether they were currently enrolled at a university, as well as to inform participants about the ethics agreement and to confirm their consent to participation.

Pilot

We needed to calculate the exact arrangement and number of items in the used presentations. We did this in a pilot study as following:

1. For the period of time for which the items were presented, we originally proposed a period of two seconds, referring to what Baddeley indicated of the capacity of the working memory [1]. A pilot study was conducted to test if this amount was appropriate, comparing participants' reactions and memory when items were presented for one second or two seconds. Participants in the two second pilot remembered 12, 13, 17, 20 and 21 items, while participants in the one second pilot remembered 9, 17, 16, 13 and 16 items. One participant in the one second pilot remarked without being prompted that they found the time the pictures were presented too short and referred to how they were unable to repeat the name of the item to themselves in only one second. Since we wanted to test for use of the phonological loop, we decided to stay with the two second exposure time to encourage the inner voice and to go along with Baddeley's theory.
2. We needed to test the appropriate number of items to present to the participants. In the pilot study participants remembered between 9 and 21 items, though the highest number was treated as an outlier as the participant remarked on their exceptionally good memory. Since the surrounding circumstances of the pilot study were not as tightly controlled as the final study, we decided on 20 pictures for the experiment.
3. To minimize ambiguities, a preselection of 28 pictures was used to test whether they were associated to one or more words by presenting the pictures to eleven different persons, who then had to name the objects in the pictures. All pictures that were associated with two or more different words by different people were eliminated in the process. We decided to mainly use words with one or two syllables as these words tended to be less ambiguous, and to ascertain that the name could be repeated in

the two second expose period, thus enabling use of the phonological loop. The final 20 items can be seen in *Figure 2.*

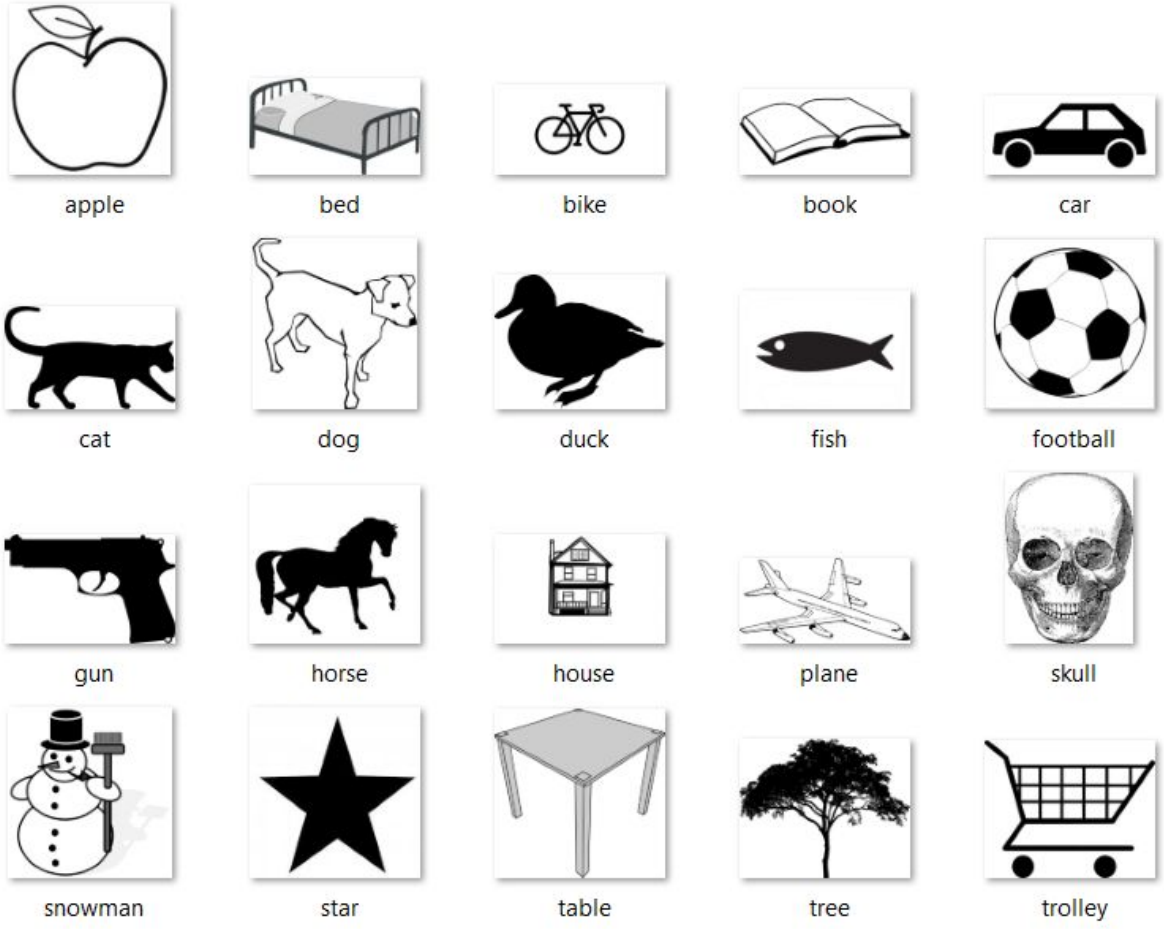

Figure 2: The 20 items used in the final test in picture and word form.

Research ethics

Participants were informed that they were participating in an experimental study about memory at the beginning of the survey and “that they can withdraw from a study at any time without penalty” [10].

Consent was collected with a yes-or-no question after the explanation of the ethics agreement, in which we explained that the data we collect would be anonymized and there would be no way to link a response to an individual, relating to “Fairness, Responsibility, and Informed Consent” [10] and the “Clarification [...] to ensure that the client provides informed consent before and during intervention“ [11] (see Appendix A). No individual data was included on the report.
Time stamps of when each response was recorded in the survey were deleted to comply with the “Adequate storage and handling of information and records, in any form, to ensure confidentiality, including taking reasonable safeguards to make data anonymous when appropriate”[11], and also, “every precaution must be taken to protect the privacy of the research subjects and the confidentiality of their personal information” [12]

After the test, participants were informed about the theory behind the experiment and what data we compared as referred to debriefing [10].

Results

In the final experimental study 30 participants took part, ten for each independent variable. All participants were currently enrolled at a university. Slightly more than half of our participants were in age group 18-22, while one third was in age group 23-26. Two participants were in age group 27-31 and one was in age group 31-35. With 17 female and 13 male participants the results should be fairly unbiased in regards to gender. All participants understood the instructions, consented to taking part in the experiment and voluntarily gave us their results of the experiment to evaluate.

Overall participants correctly remembered between 6 and 18 items with a total average of 11.73. More detailed results and statistics can be seen in *Table 1*.

	Mimi-mum # of remem-bered items	**Maxi-mum # of remem-bered items**	**Average**	**Average without min/max value**	**Median**	**Varianc e**	**Stand-ard Devia-tion**
total	6	18	11.73	11.68	11.5	9.72	3.12
Pictures	7	15	11.8	12	13	7.96	2.82
Annotat ed Pictures	6	17	11.4	11.38	10	13.6	3.69
Written words	9	18	12	11.63	11.5	9.56	3.09

Table 1: Statistics of the experiment, including minimum and maximum number of remembered items, average, average without counting the highest and lowest results, median, variance and standard deviation of the total items and items presented as pictures, annotated pictures and words.

Participants presented with annotated pictures overall remembered the least amount of items, with an average of 11.4 and a median of 10. Depending on which statistic measure is taken into account, pictures performed better than written words with a median of 13 compared to a median of 11.5, while the average number of correctly remembered items of words is 12 and the average of pictures is 11.8. According to the relatively high variance in comparison to the average, a reliable result would require a lot more participants, especially

since our subjects were chosen and assigned to the independent variable values randomly and not by matching [13].

The English level of participants seemed to have a big influence on the result. On average, participants with a higher level of English also remembered a higher number of items correctly, regardless of how the items were presented. 53% of participants had a English level of C1 and remembered 12.19 items correctly, while the 40% with an English level of B2 correctly remembered 11.25 items. The remaining 7% of participants with a B1 English level recalled 11 items correctly on average. The Histogram in *Figure 3* illustrates the difference, showing that people who had a higher English level remembered higher amounts of items more frequently than lower levels.

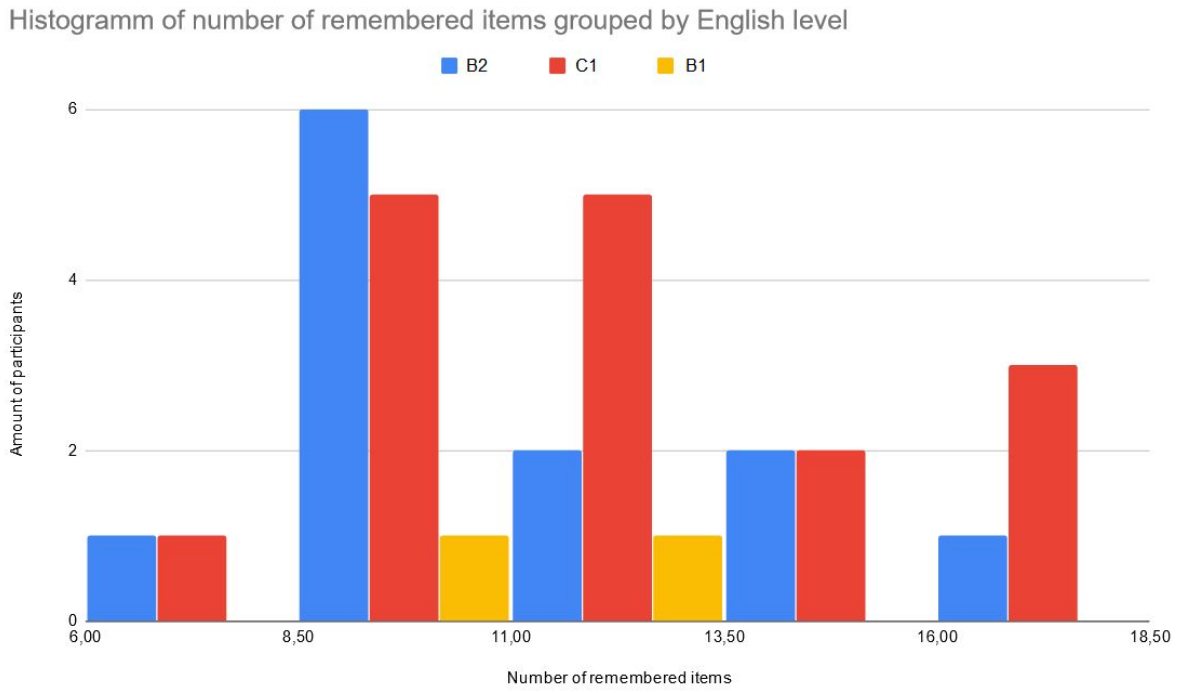

Figure 3: The Histogram of the number of remembered items by English Level. The X-Axis shows the number of remembered items. The Y-Axis shows the amount of participants who remembered the corresponding number of items. Yellow is for participants of English level B1, blue for English level B2 and red for English level C1.

When asked, 70% of all participants confirmed they used the phonological loop to remember the items, and 76,6% used the visuospatial sketchpad at least partially. Yet when looking at the different variables, only half of the participants presented with words indicated that they used the phonological loop, while 70% of participants presented with annotated pictures and 90% of participants presented only with pictures said they did repeat the names of the items in their mind to memorize them.
60% of participants presented with written words said they imagined the items visually in their mind. For participants who were presented with either just pictures or annotated pictures it was 90%.

In the open question we asked participants to explain how they tried to remember the items. Six participants noted that they were trying to come up with a story that included all the items. These participants remembered a much higher number of items on average, it being 14.67 items, while all other participants had an average of 11 items.
13 participants categorized the items or remembered them via associating them with each other or personal feelings about them. Three participants repeated the items in their mind in sequence, though some noted that it quickly became too many items to do this. A few single notes included remembering the first letters of the items, concentrating really hard with their eyes closed or writing the last items down first. One participant who was presented with annotated pictures said they remembered more image details than the text. Six participants specifically mentioned that they remembered the items as “images in [their] head[s]”, though half of these participants were presented with written words and not with pictures. Five participants explicitly noted that they repeated the words in their mind.

Discussion

Contrary to our hypothesis, both the average and median amount of remembered words was smallest for the group shown annotated pictures and highest for the group shown only words. However, removing the extreme cases (the minimum and maximum amount of remembered words) influences the order of the average of remembered words towards the presentation consisting of only pictures (see *Table 1*). As a consequence it is hard to categorize if either presentations consisting only words or only pictures result in a higher capacity.

Although the experimental study we conducted does not precisely quantify the capacity of the different types of memory systems (phonological and visual sketchpad) and there might be slight changes, our results suggest that multimedia contents are not necessarily superior to single media usage in facilitating the capacity of information. Therefore, our results support common reviews that point out that there is no evidence that a multimedia approach has advantages for learning or memory performance. Instead, any advantages that have been found in past studies could be explained by other surrounding factors unrelated to the multimedia presentation [14, 15].

An interesting aspect of our results is the distribution of methods used by the participants to remember the items. Contrary to our previous expectations, the most frequent use of the phonological loop seemed to occur in the group shown only pictures and the least use seemed to occur in the group only shown words, as described in our results. Since the phonological loop and the visual sketchpad are very complex and their mechanisms are still subject of modern investigations [16], we do not have the possibility to verify the answers of the participants regarding their used type of memory. However, this might support research results that indicate that written words are not necessarily bound to the phonological loop but can also be processed in the visual sketchpad of working memory [17].

People who wrote they created a story in their mind performed significantly better than the rest. This indicates that learning methodology influenced the amount of items participants could remember. Embedding a list of items into a meaningful context also helped people to remember. This is in accordance with Anderson [1] and a study by McAninch, Austin and Derks where participants remembered more nonsense pictures that had a humorous caption than pictures with a descriptive one [18]. This shows that processing information on a deeper level (like finding it funny, or making up a story) facilitates memory performance [1].

Furthermore the same study provides an explanation why the group with annotated pictures performed worse than the group shown either pictures or words only. Our participants tended to give back the exact same words which they saw during the presentation, even if they wouldn't have used those exact words to describe the presented pictures. "The condition with no captions benefitted from the subject's freedom to develop their own mnemonics" [18] (see Appendix B). Therefore, Austin's and Derks' findings are consistent with our results, because the presentations with only words or pictures left more space to create some meaningful context for each participant.

When evaluating the results, it turned out that the average English level of the participants that were shown the annotated pictures was lower than the English level of the two other groups. Therefore the relatively poor performance of the group that was shown the annotated pictures could to some extent be caused by coincidentally lower English level. Conducting the experimental study to participants with their mother tongue equal to the language of the experimental study would eliminate the influence of processing information in a foreign language.

Limitations and Future Work

Due to the relatively small number of participants, the findings and especially the quantitative differences manifested in the discussion are not completely determined. To verify our findings that the combined (multimedia) presentation does not facilitate memory performance compared to single media contents it would be interesting to conduct a similar experiment with a much larger number of participants.

As mentioned in the discussion, our results indicate that the phonological loop was not frequently used by the participants shown only written words. Therefore, a comparison between auditory and visual presentations would help to compare the phonological loop and the visual sketchpad and make our investigation more complete and comprehensive.

As we have seen during our study, participants who were making a story in their mind performed better. To refine out results it would be interesting to conduct a similar study where the participants are instructed to only use one method of remembering the items, thus turning it into a control variable and increasing comparableness of the presentation types, as was the original goal.

Conclusion

We conducted an experimental study about the working memory's performance in regards to remembering pictures, written words and annotated pictures. 30 participants took part in the study, ten for each type of presentation. While all groups performed similarly, the picture and word groups tied for the first place based on the statistical value chosen for evaluation. The group exposed to annotated pictures performed the worst, despite the intuitive assumption that multimedia presentation would make items easier to remember.
However, these results are congruent with those of other studies that indicate that multimedia presentations take away the freedom to form an individual way of processing what is presented. The importance of processing information rather than just registering it also showed itself in the significantly higher performance of participants who remembered items by combining them in a story in their minds, rather than just trying to remember the items on their own.
Another influence factor was the English level of the participants. Both that and the strategy for remembering items would be valuable additional control variables for further studies.

References

[1]	Anderson, John R. (2015) *Cognitive Psychology and Its Implications*, Eighth Edition, New York, Worth Publishers
[2]	Anderson, John R. (2015) *Cognitive Psychology and Its Implications*, Eighth Edition, New York, Worth Publishers, p.103
[3]	Baddeley, A. D. (1992) *Working Memory*, Science, Vol. 255, No. 5044, 1992, p.556-559.
[4]	Baddeley, A.D., & Hitch G.J. (1974). Working memory. In G.A. Bower (Ed.). *The Psychology of Learning and motivation: Advances in Research and theory*, Vol. 8, p.47-89. New York: Academic Press.
[5]	Aben, Stapert and Blokland (2012) *About the distinction between working memory and short-term memory,* In frontiers in Psychologie
[6]	Mayer, R. (2005) *Cognitive theory of multimedia learning*, In Mayer, R (Ed.) Cambridge Handbook of Multimedia Learning. Cambridge University Press
[7]	Kerlinger and Lee (2000) *Foundations of behavioral research*, Fourth Edition p.487, Wadsworth
[8]	Keren, Gideon & Lewis Charles (2014) *A Handbook for Data Analysis in Behaviorial Sciences: Volume 1: Methodological Issues Volume 1: Statistical Issues*, Psychology Press, p.257-272

[9]	TrackTest (2019) *English language levels (CEFR)*, English Proficiency Test Online & Express Publishing, https://tracktest.eu/english-levels-cefr/, accessed 01.10.2019
[10]	Kerlinger and Lee (2000) *Foundations of behavioral research*, Fourth Edition p.444-445, Wadsworth
[11]	EFPA (2005), *3.1 Respect for Person's Rights and Dignity*, Revised by General Assembly in Granada, http://www.efpa.eu/ethics/meta-code-of-ethics-, accessed 01.10.2019
[12]	Swedish Research Council (2017) *Good Research Practice*
[13]	Kerlinger and Lee (2000) *Foundations of behavioral research*, Fourth Edition p.511, Wadsworth
[14]	Clark, Richard E. & Feldon, David F. (2005) *Five Common but Questionable Principles of Multimedia Learning*, In Mayer, R (Ed.) Cambridge Handbook of Multimedia Learning. Cambridge University Press
[15]	Clark, Richard E. & Craig, Terrance G. (1992) *Research and Theory on Multi-Media Learning Effects*, Springer-Verlag, University of Southern California
[16]	Hepner, Christopher R. & Nozari, Nazbanou (2019) *Resource allocation in phonological working memory: Same or different principles from vision?* Journal of Memory & Language, Vol. 106, p.172-188. 17p.
[17]	Kruley, Peter & Sciama, Sonia C. & Glenberg, Arthur M. (1994) On-line processing of textual illustrations in the visuospatial sketchpad: Evidence from dual-task studies, Memory & Cognition, 22 (3), p.261-272
[18]	McAninch, Cecile & Austin, Joy & Derks, Peter. (1992). *Effect of caption meaning on memory for nonsense figures*. Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 11(4), 315-323

Appendix

Appendix A:

In the meta - code of ethics of the EFPA, the definition of a client is given :
“In this code the term ‘client’ refers to any person or persons with whom a psychologist interacts on a professional basis. For example, a client may be an individual (such as a patient, a student, or a research participant)” EFPA, Meta-Code of Ethics
Hence our use of the citation saying “client”, when in our case, means “research participant”.

Appendix B:

A mnemonic:
“A system such as a pattern of letters, ideas, or associations which assists in remembering something.” *(Lexico.com, powered by Oxford, accessed 01.10.2019)*

Appendix C:

KP-Results.xlsx