



university of
 groningen

faculty of mathematics
 and natural sciences

artificial intelligence

Can mind wandering improve retrospective reevaluation in sequential decision making?

Bachelor's Project Thesis

D.A. Schilling (s2199041), d.a.schilling@student.rug.nl

Supervisors: dr. M.K. van Vugt & S. Huijser, MSc

Humans spend a lot of time thinking about things that are not going on around them. This process is called mind wandering and previous work shows that these events can be regarded as a distraction from a task at hand, interfering with your concentration, while other research shows positive impact on future goals in less demanding cognitive tasks. In our project we have researched the influence of mind wandering on sequential decision making. It is shown that retrospective reevaluation takes place during unrelated tasks between sequential decision making tasks, and we intend to show that the reevaluation is actually credited to the cognitive process of mind wandering and has a beneficial effect on future goals. From the experiment we did not find enough evidence to conclude that there is a relation between mind wandering and retrospective reevaluation.

Introduction

Humans spend approximately 46.9% of their time thinking about what is not going on around them, contemplating events that happened in the past, might happen in the future, or will never happen at all, according to Killingsworth and Gilbert (2010). This description is an accurate definition of the cognitive process known as mind wandering. Some might regard this cognitive process of mind wandering purely as a distraction from a task at hand, as it can interfere with your concentration, or is even stated to come at an emotional cost (Killingsworth & Gilbert, 2010). On the other hand, researchers found that when performing cognitive low demanding tasks and wandering off of the current task are seen as a beneficence (Rummel & Boywitt, 2014). Your mind is, in other words, not using all its cognitive resources and exploits this opportunity to wander about personal goals or concerns, to be able to achieve a future benefit by processing these events during mind wandering episodes (Mooneyham & Schooler, 2013).

Episodes of mind wandering can have a positive effect on future-oriented tasks, while it also can have negative effects on present tasks. Whether mind wandering is positive or negative mostly depends on the kind of task, as for example attention-based tasks, like reading or planning tasks such as

autobiographical planning (Mooneyham & Schooler, 2013), on the workload of this task (Rummel & Boywitt, 2014) and the consequences of neglecting the task, as for example in driving (Galera et al., 2012), where dozing off could cause accidents. In tasks, where creative thinking (Baird et al., 2012) and problem solving skills (Ruby, Smallwood, Sackur, & Singer, 2013) are required, mind wandering seems to enhance the outcome of the tasks, potentially by increasing unconscious associative learning (Baird et al., 2012). Whereas in other daily-life events, requiring more direct focus on the present task, demanding attention-based tasks, that require lots of resources (Leszczynski et al., 2017), as for example reading (Smallwood, McSpadden, & Schooler, 2008), might result in negative outcomes. As for example being distracted from the cognitive high demanding task of driving it could result in severe casualties.

However, what may be a function of mind-wandering? One of the important things it could be involved in is simulating the consequences of different actions to predict which one is most profitable. To better understand how this could work, we turn to the theories of reinforcement learning (RL), the study of how an agent can learn to choose actions that maximize future rewards (Sutton & Barto, 1998), as a guidance. Reinforcement learning is used as it is the dominant

theoretical framework for operant learning in humans and animals (Niv, 2009). Operant learning is the process of learning through reinforcing and punishing behaviour, simulating rewards for choosing certain actions. We regard this kind of learning, since it provides us with the tools of being able to compare the positive and negative outcomes from mind-wandering in decision making.

In the study of Gershman, Markman and Otto (2014) a similar experiment design was used, where participants were suppose to perform a decision making experiment, where they had to learn hidden state-transitions and learn from the beneficial actions received from the environment (Gershman, Markman, & Otto, 2014). In this experiment, a subject had to learn the underlying state-action-pairs and from these pairs a reward-structure (deducted from performing actions in certain states). This form of learning from decision making can be seen as learning hidden associations (Si & Wang, 2001), where the underlying neural activity is linked to reinforcement learning in computational approaches (Shteingart & Loewenstein, 2014).

When a person is performing a repetitive task and has to learn from its experiences, old experiences and new gathered experiences need to be integrated. However when the earlier experiences and newly gathered experiences are interfering with each other, there has to be a reevaluation of the believes of the person, since there seems to be evidence that the experiences are reflecting different truths. This reevaluation can take place in online processing, where humans sort and comprehend the experiences, that they gathered and try to make senses of these differences by actively solving these issues.

Another approach however is offline processing, where the brain has to connect different pieces of information and integrate these pieces to make a model of the experiences (Momennejad, Otto, Daw, & Norman, 2017). Humans use offline processing, since considering all possible steps in decision making is too resource-expensive for computational devices (and therefore also the brain). In other words, we do not always have the time to argue over all the experiences, before being able to process the information and basing our decisions on these experiences (Gershman et al., 2014). This offline processing or reevaluation of earlier experiences from the same or similar decision making tasks is called retrospective reevaluation. During reevaluation, an integration of initial experience about a task with later experience about a change in its goals takes place (Momennejad et al., 2017). Retrospective reevaluation can therefore be seen as some form of planning.

Retrospective reevaluation can be regarded as, the change in the associative strength of an item in memory, while the

item (as a stimulus) is absent. Which also can be thought of as changes in the retrievability of items in memory (Le Pelley & McLaren, 2001). This phenomenon is, as in reinforcement learning, a replay of earlier experiences to make proper inferences, which can be attributed to events of mind wandering.

This replay or retrospective reevaluation is implemented in the DYNA architecture (Sutton, 1991), a model where online and offline processing learning mechanisms are integrated. In the DYNA architecture the environment provides an agent with experience in the form of transitions and rewards. These transitions and rewards are then learned and saved in some model of the world. With new experience the system can learn more about its environment, while the system is also able to simulate experience by reevaluation of state-action pairs from memory and then using these simulations to create an updated model of the environment, while learning effectively (Gershman et al., 2014). In the research of Gershman et al., this system was compared to retrospective reevaluation in humans and it was shown that there is a cooperative architecture between online and offline processing in humans, as it also is implemented in the DYNA architecture. However in this research there is no explanation where this cooperation comes from or how it works, where we think that this might be accounted to by the process of mind wandering.

It was shown that the combination of online and offline processing results in efficient learning. However there is still no explanation what could be the relation between offline processing and the process of retrospective reevaluation in the human mind. We therefore wanted to research whether this retrospective reevaluation can be attributed to the mind process of mind wandering in human learning. Mind wandering should give the mind a period to restructure gathered experiences, while performing an unrelated task, thus initiating some kind of offline processing of this information.

Since decision making tasks cannot always be determined immediately upon exposure and unrelated tasks give the mind a chance to solve these problems (Medea et al., 2016), we have constructed an experiment to try to elaborate on that retrospective reevaluation might actually be the process of mind wandering. This experiment therefore focuses on researching whether performing an unrelated task, giving the mind the ability to wander, triggers retrospective reevaluation. To determine if there are occurrences of mind wandering, we use thought probes to be able to express the focus of the mind and categorize task-unrelated thoughts as mind wandering (Weinstein, 2017). Through the results of from these thought probes, we want to research whether there is a

link between retrospective revaluation and mind wandering.

We suspect that the retrospective revaluation as shown in the results of the sequential decision making task (Gershman et al., 2014), also occurs through episodes of mind wandering and are therefore an explanation of the occurrence of retrospective revaluation. For our research we therefore wanted to investigate if mind wandering influences retrospective revaluation in sequential decision making. To this end, we have combined several components from experiments in other research. The first being the sequential decision making task experiment performed by Gershman, Markman and Otto (2014), from which we have used the reward structure design (where rewards are given for choosing actions), and combining it with a back-story based on the design of Momennejad, Otto, Daw and Norman (2017). In this experiment participants had to learn a reward structure from decisions. From learning this structure they could learn to receive the highest rewards for making certain decisions, that are determined by an underlying reward structure. The decision making task in this experiment was based on localization through the use of faces, scenes and object stimuli (Momennejad et al., 2017) and by choosing one of the shown stimuli participants received a reward.

Secondly we have implemented an n-back task, combined with thought probes, as performed in the experiment of Steindorf and Rummel (2017), in which participants were supposed to remember a shopping list, while performing other tasks. We used an n-back task, because it is stated in earlier research that in working memory tasks, people experience episodes of mind wandering (McVay & Kane, 2012). We chose the two-back task version of the n-back task, because it allowed for the occurrence of mind-wandering, while still being challenging enough as a working memory task (Smallwood, Nind, & O'Conner, 2012; Steindorf & Rummel, 2017). When performing the two-back task we give participants the opportunity for their minds to wander and reevaluate their decisions made earlier in the decision making task. Finally we use thought probes to measure mind wandering, as suggested by Weinstein (2017).

Our hypothesis is that mind wandering improves retrospective revaluation in sequential decision making. We wanted to test this by researching the relation between the amount of retrospective revaluation and the amount of mind wandering in participants, when performing a sequential decision task.

Methods

Participants

A total of 29 volunteers (male = 13, female = 16; mean age = 22.83, SD = 3.129) participated in the experiment. All participants were compensated for their participation, of which the amount depended on performance in the sequential decision making task (minimum = €7.50, maximum = €11.00; average = €10.13, SD = €0.33). The participants all received the same instructions and performed the experiment under the same conditions. All experiments were conducted in accordance with the Declaration of Helsinki, and all participants signed an informed consent prior to the start of the experiment. All participants were Bachelor, Master or PhD students. We excluded the data of one participant, since the data was incomplete.

Tasks

We built the experiment with the graphical experiment builder OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) (version=3.1.9) and for more advanced features the programming language Python (version=3.6.3).

To initiate the effects of mind wandering we used the two-back task. In the two-back task participants were presented a series of 12 letters, each letter presented for 500 ms and having an inter-stimulus interval of 2 seconds, in which the participant could respond by pressing 'y' for having a matching two-back or 'n' for a non-match. When a key was pressed, the participant received feedback that their response was submitted by showing a check mark.

After the two-back task we wanted to measure mind wandering. To measure mind wandering we used thought probes, which were presented in the form of a questionnaire. These thought probes contained five to six thought categories, depending on the phase in the experiment and were based on categories mentioned in earlier research (Weinstein, 2017). The questionnaire consisted out of the following options:

- A:** I was totally focused on the memory task
- B:** I was thinking about my performance or the duration of the memory task
- C:** I was externally distracted
- D:** I was daydreaming/thinking about unrelated things to this experiment
- E:** I was feeling blank/drowsy
- F:** I was thinking about aspects related to the decision making task (only after the first decision making block)

These thought probes will indicate whether a participant was mind wandering or was just focused on performing the

2-back task. For categories B-F we count the categories as mind wandering, while option A means that there was no mind wandering taking place.

Finally to measure retrospective revaluation we used a sequential decision making task, in which we used the underlying structure of the experiment performed by (Gershman et al., 2014), but informed by the story line of Momennejad et al. (2017). In this storyline the participant took the role of robber and had to explore different locations in different cities and steal money from people. For this part of the experiment we used different background colours indicating the state in which the participant was in. The initial background colour was black, while state 2 (after picking the left vault) and 3 (after picking the right vault) were respectively blue and red. Participants were told they were taking on the role of a burglar, being situated in a house containing two vaults. Since it takes some time to open a vault, they would receive a limited time, and could therefore only open one of the two vaults. To open a vault, the participant had to press either 'z' to choose the left vault, or 'm' for the right one.

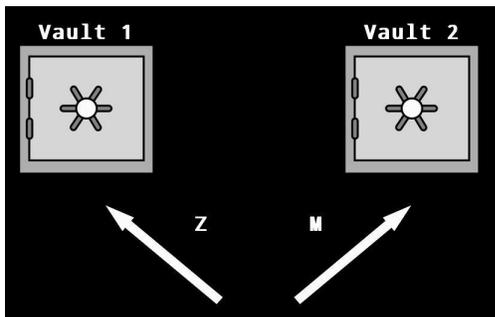


Figure 1. Screenshot taken from the experiment. The participant is in the initial state and by pressing either 'z' or 'm' opens the corresponding vault, receiving the reward from inside the chosen vault.

When one of the options was chosen, the participant moved from the initial state to either state 2 (left) or state 3 (right), indicated by the colour of the background. For these decisions the participant received 2 seconds to respond. For the transition of the states in the sequential decision making task (SDT) and the indication of the rewards, we used images of an open vault, a closed vault, four differently coloured money boxes and images of euro cent coins. The images of the coins were obtained from copyright free sources, while the vaults and money boxes were self made clip art images.

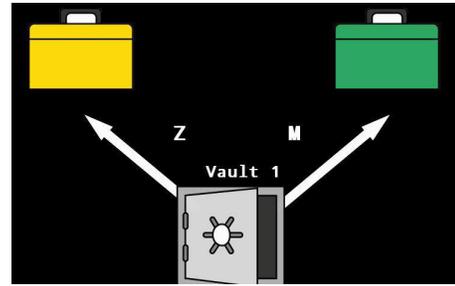


Figure 2. Screenshot taken from the experiment. The participant is situated in Vault 1 and by pressing either 'z' or 'm' opens the corresponding money box, receiving the reward from inside the chosen money box. This situation only takes place during the two-step horizon part of the experiment.

Procedure

The experiment was built into four stages, existing of three different components. These components were the 2-back task or Working Memory Task (WMT), the Thought Probe (TP) and the Sequential Decision making Task (SDT). Participants first had to sign a consent form and were shown the overall procedure of the experiment and the detailed instructions of the separate tasks.

Participants did a sequence of 12 trials of the WMT after which they were shown a thought probe. The first three letters of the sequence were always non-matching, while the remaining letters were a random mix of 5 non-matching and 4 matching two-backs. In the first WMT phase the thought probe consisted of the first five possible thought categories. In the second WMT phase the extra category was included, which involved the revaluation of the SDT.

The SDT contained two phases or three stages (see Figure 3). In the first stage participants would have to choose between two vaults and thereby learned the deterministic underlying structure, while receiving rewards for their choices. These rewards were represented by euro coins (5 cents in state 2 and 2 cents in state 3; see the image) printed on screen, after a choice was made. After performing 15 iterations the participant received feedback indicating how much money they earned.

After this stage participants directly go to the second stage of the SDT, which contains a two-step horizon (still in the first SDT phase), which chains another decision making point after the first decision. In this stage, where the two-step horizon was performed, the participant was already inside one of the vaults and now had to choose between two differently coloured money boxes. Alternately they were in the first and then the second vault. The best decision for choosing the left vault earlier switches to the best decision

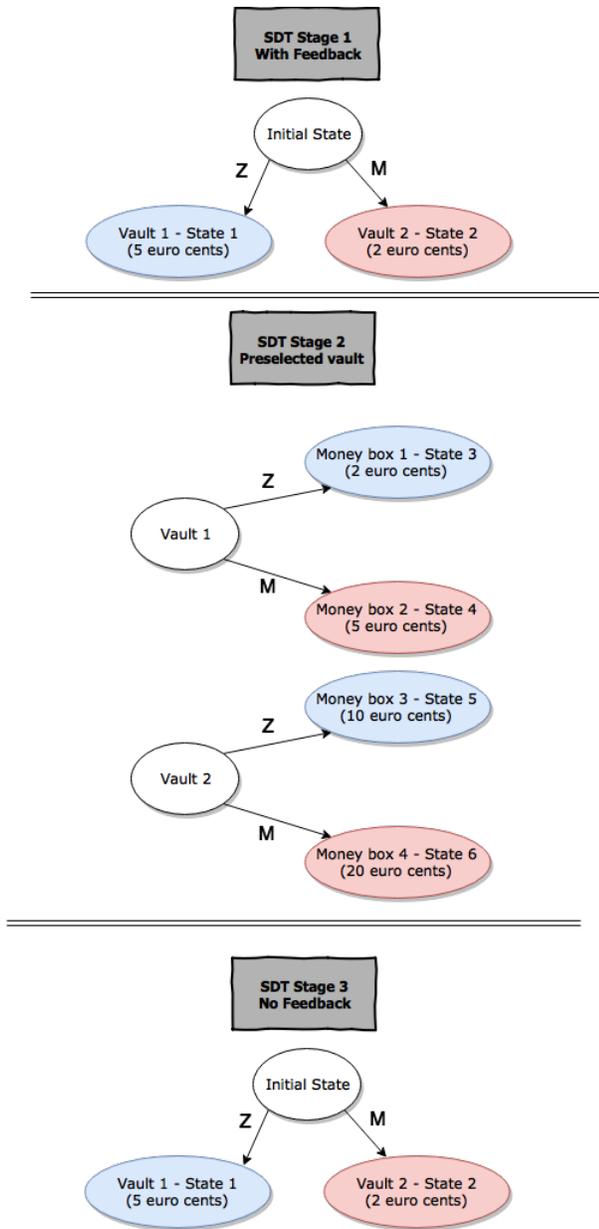


Figure 3. Overview of the reward structure in the sequential decision making task stages. The letters next to the arrows represent the keys pressed for the transition to the next stage. The colours of the balloons indicate the background colour shown in the corresponding state.

being the right money box, to gain the highest reward. This was implemented like this, to create an interference with the earlier associated learned structure from the first stage of the SDT. For these actions participants would receive a reward of 10 euro cents for choosing the left box and a reward of 20 euro cents for the right. All states have the

same corresponding background colours for choosing the left or right as before. This stage iterated 8 times, after which another WMT phase started. Finally the participant had to perform the final SDT phase, in which there were rewards for the decisions as given, however the participants would not receive any feedback, to avoid biasing. In this case the same deterministic structure as in the first phase was used. Before this phase started, they were presented with the following instructions:

Lastly, you are cracking open vaults again. In these vaults there are loot boxes, but these will be chosen at random. Keep in mind the rewards you have received in earlier stages.

You should make this decision on the basis of what you have learned over the course of the experiment.

You will receive money for making these choices, but you will not be able to immediately see how much each choice for a vault has earned you. We want you to use your previous knowledge to decide what vault you want to open.

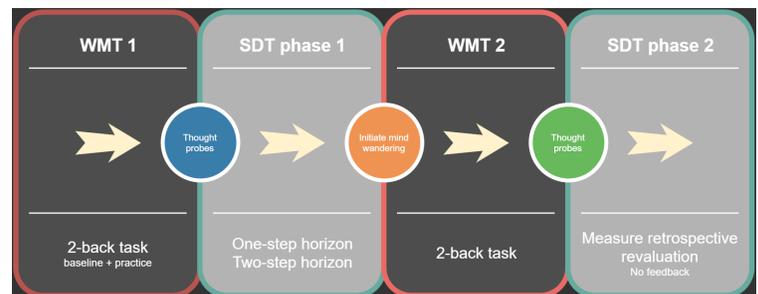


Figure 4. Overview of the experiment, separated into four phases. In the first phase of the sequential decision making task, the participant first had to perform 15 one-step horizon decisions and after that 8 iterations of the two-step horizon decisions.

The overall procedure of the experiment finally consists of four phases (see Figure 4), combined taking approximately 45 minutes. The first phase of the experiment is a practice phase of the WMT, combined with the thought probes, in which the participants received 240 trials, with after every 12 trials a thought probe. The second phase contains the first two stages of the SDT, followed by another WMT phase with 240 trials, including the extended thought probe. Finally participants had to perform the fourth phase (the third SDT stage), and at the end of the experiment they received feed-

back of the total amount of money they have earned. After the experiment participants received one final questionnaire with demographic questions on their sex, age and highest education and questions, in which the participants could elaborate on what they thought the experiment was about and feedback on the experiment.

Data analysis

We wanted to know, whether the amount of mind wandering affected the retrospective reevaluation on the choices made in the SDT. To analyse this we first compared the outcome of the choices made in the first SDT stage with the choices made in the third SDT stage. In the case that there was retrospective reevaluation, the data should show that there has been a change in decisions made between the two stages, while no change in decisions between the two stages means that participants did not reevaluate. The probability of choosing the most rewarding action from these stages are then compared to express the magnitude of retrospective reevaluation. We used a one sampled chi-square test to check whether the difference between stage 1 and stage 3 in probabilities for choosing the most rewarding action are significantly different.

To analyse the influence of mind wandering on the retrospective reevaluation we compare the thought categories chosen in the thought probes with the difference in the probability from the retrospective reevaluation, measured in the second WMT phase. In order to analyse the relation between these two variables we used an poisson regression, where the categories marked as mind wandering are compared with the outcome on the retrospective reevaluation. The poisson regression predicts whether a higher magnitude of retrospective reevaluation is linked to the thought categories, that are marked as mind wandering categories.

Results

Retrospective reevaluation

In the first and third stage in the SDT, participants greatly preferred the left vault over the right one. In the first stage participants pressed an overall percentage of 61.2% times left (highest reward), 20.1% times right (lowest reward) and 18.7% they got a time-out and therefore chose no option. In the third phase participants had an overall percentage of 64.0% left, 32.9% right and 3.1% no options submitted. In the second stage participants received the two step horizon choices and pressed 56.5% for the right (and now higher rewarding) option, 37.0% the left option and chose neither for 6.5% of the cases.

These results confirm that participants seemed to have learned the hidden rewards structure behind the decision task, since they were able to determine the highest rewarding

Stage	Most rewarding action	Least rewarding action	No action chosen
Stage 1	0.612	0.201	0.187
Stage 2	0.565	0.370	0.065
Stage 3	0.329	0.640	0.031
Stage 1-3	0.605	0.302	0.093

Table 1

Choice probabilities of actions chosen in each stage of the sequential decision task, reported as the mean over all participants. The highest rewarding option is in stage 1 the left option, while in stages 2 and 3 it is the right option. The last row indicates the probability over the whole experiment.

decisions. In Table 1 the choice probability for choosing the highest rewarded action is shown.

To test the whether participants significantly preferred the highest rewarding option over choosing a random option, we performed a one sampled chi-squared test using the proportions from Stages 1-3. We compared the amount of the highest rewarding option chosen with, including the factor of not choosing an option. We found that participants did not base their actions purely on chance and they thus seemed to have learned the hidden reward structure $\chi^2(1) = 64.946$, $p < 0.05$. When disregarding the option of not choosing an action in the proportions, we have also found enough evidence $\chi^2(1) = 143.47$, $p < 0.05$, that participants did not base their actions on chance, comparing the difference in proportions between choosing left and right, while leaving out the timed-out data.

Furthermore have we tested whether retrospective reevaluation has taken place during the experiment. To test the effect of retrospective reevaluation we compare the magnitude of reevaluation between the first and third stage in the SDT. This magnitude is given by subtraction of the probability of choosing the highest rewarding action in stage 3 and the probability of choosing the highest rewarding action in stage 1. There we saw that there was a magnitude of approximately 0.283, which is compared to the reevaluation magnitudes in the research of Gershman et al. (2014) with a 40% reevaluation, a low magnitude of retrospective reevaluation. We performed a one sampled chi-squared test to test, whether the difference in probability of choosing the highest rewarding option in stage 1 compared to stage 3 is significantly different or not. We found that there is not enough evidence to conclude that there is a significant difference in the actions made in stage 1 and stage 3 $\chi^2(1) = 0.08511$, $p = 0.7705$, and can therefore not conclude that participants experienced retrospective reevaluation.

Mind wandering

To measure mind wandering we used the thought probes in the second WMT phase to find if mind wandering is affecting retrospective reevaluation. From these thought probes we can see that a moderate amount of time, participants were experiencing mind wandering episodes. Approximately 43% of the options chosen in the questionnaire are considered to be mind wandering categories. The probabilities of the chosen options in the thought probes experiment can be seen in figure 5.

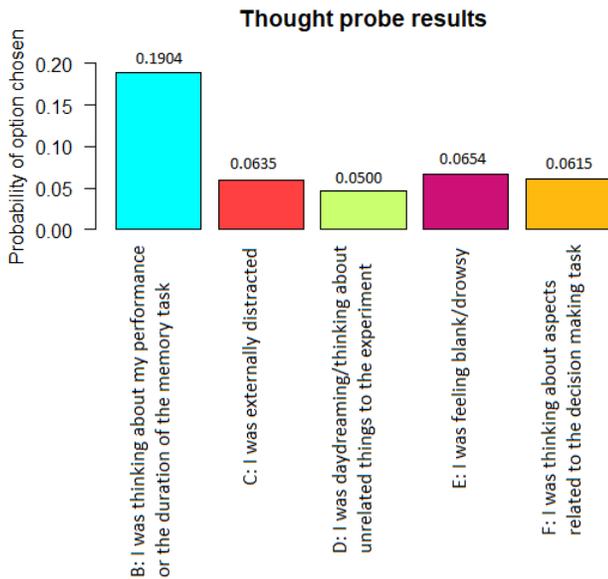


Figure 5. Thought probe results from the second working memory task phase. Results shown are the options chosen in the questionnaire and the categories considered to be mind wandering categories. The meaning of the letters corresponding with the description of the category can be found in the Methods section in the description list on page 3.

To answer the research question we have investigated whether there is an effect of mind wandering on the retrospective reevaluation. In order to test this we look at the relationship between the categories marked as mind wandering and the actions chosen in the first and third stage of the SDT. We have run a Poisson regression to predict, whether there is a larger amount of higher rewarding actions chosen, when there were more episodes of mind wandering submitted in the thought probes. From this regression we can conclude that there is not enough evidence that mind wandering has an effect on retrospective reevaluation with a regression coefficient of -0.8654 , $p = 0.619$ and 27 degrees of freedom. Since we cannot conclude that participants experienced retrospective reevaluation, we cannot conclude anything from the regression about the relation between mind wandering

and retrospective reevaluation and if it has an effect on it.

Working memory task

In the working memory task we found that in the first phase of the experiment, participants had an average score of 61.25% correct. A lot of participants seemed not to understand that after every letter presented, an answer had to be submitted on whether the letter they saw was a 2-back or not. After several trials they however seemed to understand the task and did give a response for every letter. In the second stage participants showed an average score of 81.13%. We performed a one-sided two sampled t-test, to test whether the score in the second stage was significantly higher than in the first stage and found that with $t(47.34) = 3.1305$, $p = 0.00149$, there the score in the second stage is indeed significantly higher.

Since the score in the second stage is higher, participants should be able to mind wander more than in the first stage, since they already should understand how the task works and require less focus on learning the 2-back task. As can be seen in the results from table 2, there is indeed an increase in mind wandering responses from the thought probes when comparing the first and second stage of the WMT. We performed a two sampled chi-squared test to find a significant difference in probabilities. We found that there is indeed a significant difference in the amount of thoughts on task versus mind wandering in the first and second phase of the WMT, with $\chi^2(1) = 13.845$, $p = 0.0001985$, and can therefore conclude that participants indeed experienced more mind wandering in the second WMT phase and therefore shows that the set-up of this task is suitable to promote the amount of mind wandering occurrences.

Stage	Focused on task	Mind wandering
Stage 1	60.0%	40.0%
Stage 2	52.4%	47.6%

Table 2

Percentage of responses given for mind wandering or focus on task in the thought probes from the two working memory task stages.

Discussion

In our experiment, we explored whether mind wandering has an effect on retrospective reevaluation in sequential decision making. We measured the retrospective reevaluation through comparing the change in choices made in stage 1 and 3 of the task. In stage 2 of the task the hidden reward structure of the task was swapped, to create an interference with the associated reward structure. We measured mind wandering by using thought probes between the first

two stages and the final stage of the sequential decision making task. From these mind wandering categories we measured whether mind wandering had an influence on the retrospective revaluation.

Since there was no evidence that participants experienced retrospective revaluation it was hard to draw conclusions from our experiment on the relation between mind wandering and retrospective revaluation. From the experiment of Gershman et al. (2014), where there was already proved that there was retrospective revaluation taking place in the participants in their experiment, we suspected that with the alteration of the design, we would find the same findings. However from our results, we see, that the proportion of choices made in stage 1 and 3 in the SDT are nearly equal, meaning that there is no change in decision making strategy.

From the question form (that we gave participants after they have performed the experiment) we found on the question on what basis the participants based their decisions in the final stage of the decision making task, that they thought that the situation in the final stage looked very similar to the first stage. In both of these stages, participants only had to choose a vault and there were no money boxes involved, like it was the case in the second stage. This caused participants to keep sticking to the strategy they used in stage 1, of what seemed to be the most rewarding choice.

Another possibility of the very low difference in probability between stage 1 and 3 is that it seems to be that participants did not associate the money boxes from the second stage with the rewards in the third stage, although the instruction stated that they had to keep these in mind. With the lack of an association with these money boxes, the interference we suspected to cause with the second stage seemed to have no effect on the decision making strategy. From the question form, we also found that this was indeed for 15 participants the reason they kept the strategy from the first stage.

For further research it would therefore be advised to keep the presentation of, for example vaults or money boxes, as similar as possible for all stages and therefore not use different images to indicate the two-step horizon. In our experiment 15 participants made decisions based on association of the vaults only and therefore never associated the money boxes from the two-step horizon with vaults of the final stage. This will make sure that participants then will experience an interference in their optimal decision strategy, and thus resulting in a larger difference in probability between stage 1 and 3, and thus in the revaluation magnitude.

We did find that participants experience more mind

wandering after the first sequential decisions making phase and it is therefore still worthwhile to research the effect of mind wandering on retrospective revaluation in sequential decision making. To explore this effect again, it is however necessary to redesign the experiment, having less factors that influence the learning of the decision making strategy. There also seemed to be no evidence that there is a relation between mind wandering and retrospective revaluation. However the results from the regression are not very reliable, since we could not find evidence that there was retrospective revaluation taking place. The experiment could be more equally designed to the experiment of Gershman et al. (2014), where there is a more abstract choice by choosing pictures of fractals. Another option is to completely redesign the experiment, but keeping the earlier mentioned problems of our experiment in mind, so using only one kind of image for all stages.

Conclusion

Mind wandering as a cognitive phenomenon could contribute to beneficial results in sequential decision making tasks. In our experiment we aimed to fill the gap, where retrospective revaluation can be linked to mind wandering in sequential decision making as a tool for the human mind for learning hidden associations. We found no reliable results however, but we believe that these results may however be the result of the experimental design or analyses, rather than the absence of the expected effects. The design of the experiment should have a stricter separation between the first and last phase in choosing the vaults. Now it seemed that subjects did not fully understand that there was a relation between the money boxes from the two-step horizon and the final phase. Furthermore, there were a lot of timed-out options in the beginning, probably due to the misunderstanding of the task. Some participants mentioned after the experiment, that they did not completely understand in the first case, that they had to press a button for every stimulus shown. They were waiting until something was happening on the screen, prompting them for a response. After a while they pressed one of the buttons to see what would happen and then found out how they could get rewards. It was however explicitly stated in the experiment manual that they had to press a button for every stimulus and should have therefore already been clear. If this experiment will be repeated, it should first include a practice round for both the SDT as the WMT to make sure that both the tasks are understood, even if the manual already states how the tasks work.

Gershman et al. (2014) showed already that retrospective revaluation in human decision making is consistent with the cooperative structure of the DYNA architecture (Sutton,

1991). From our experiment we are however not able to compare mind wandering to this architecture. Otherwise we could have been able to pinpoint more or less the function of mind wandering through comparison with the algorithms in the DYNA structure, being able to understand more of this complex mental process. The origin of retrospective reevaluation in humans however stays unknown and still requires more attention in research. From our regression we can not conclude whether there is a link between retrospective reevaluation and mind wandering and therefore requires more research to be able to research the origin of retrospective reevaluation.

Other research does show that mind wandering has a positive influence on decision making strategies. Instead of using deliberate strategies to come to a decision, mind wandering does not only give future beneficence, but even appears to be more mentally satisfying to humans compared to decision making by using a deliberate strategy (Giblin, Morewedge, & Norton, 2013). Secondly, functional MRI data shows that mind wandering is associated with specific regions in the brain and is a unique mental state, which allows otherwise opposing networks to work in cooperation (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009). This cooperation can also be useful for decision making strategies, where mind wandering allows for better decisions.

Although the results did not give enough evidence that there is a relation between mind wandering and retrospective reevaluation, we still suggest to keep researching this link. In our experiment design we were unable to conclude that retrospective reevaluation was taking place during the experiment, making the conclusion about the relation between mind wandering and retrospective reevaluation unreliable. Regarding the results from the two research projects of Giblin et al. and Christoff et al. and the absence of retrospective reevaluation in our experiment, we advice to continue researching the phenomenon of mind wandering and the link with retrospective reevaluation. Furthermore because earlier research shows that both retrospective reevaluation and mind wandering plays a big role in decision making and therefore a relation might still exist.

References

- Baird, B., Jonathan, S., Mrazek, M. D., Kam, J. W. Y., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction. *Psychological Science*, 23(10), 1117-1122.
- Christoff, K., Gordon, A., Smallwood, J., Smith, R., & Schooler, J. (2009). Experience sampling during fmri reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences*, 106, 8719-8724.
- Galera, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribereau-Gayon, R., ... Fort, A. (2012). Mind wandering and driving a responsibility case-control study. *BMJ British medical journal*, 345(7).
- Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). Retrospective reevaluation in sequential decision making: A tale of two systems. *Journal of Experimental Psychology: General*, 143(1), 182-194.
- Giblin, C., Morewedge, C., & Norton, M. (2013). Unexpected benefits of deciding by mind wandering. *Frontiers in psychology*, 4.
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science Magazine*, 330, 932.
- Le Pelley, M., & McLaren, I. (2001). Retrospective reevaluation in humans: learning or memory? *The Quarterly journal of experimental psychology. B Comparative and physiological psychology*, 54.
- Leszczynski, M., Chaieb, L., Reber, T. P., Derner, M., Axmacher, N., & Fell, J. (2017). Mind wandering simultaneously prolongs reactions and promotes creative incubation. *Scientific Reports*, 7.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). Opensesame: an open-source, graphical experiment builder for the social sciences. *Behavior research methods*, 44, 314-324.
- McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? on the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, 141, 302-320.
- Medea, B., Karapanagiotidis, T., Konishi, M., Ottaviani, C., Margulies, D., Bernasconi, A., ... Smallwood, J. (2016). How do we decide what to do? resting-state connectivity patterns and components of self-generated thought linked to the development of more concrete personal goals. *Experimental Brain Research*.
- Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2017). Offline replay supports planning: fmri evidence from reward reevaluation. *bioRxiv*.
- Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology*, 67(1), 11-18.
- Niv, Y. (2009). Reinforcement learning in the brain. *The Journal of Mathematical Psychology*, 53, 139-154.
- Ruby, F. J., Smallwood, J., Sackur, J., & Singer, T. (2013). Is self-generated thought a means of social problem solving? *Frontiers in Psychology British medical journal*, 962(4).
- Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review*, 21(5).
- Shteingart, H., & Loewenstein, Y. (2014). Reinforcement learning and human behavior. *Current Opinion in Neurobiology*, 25, 93-98.
- Si, J., & Wang, Y.-T. (2001). Online learning control by association and reinforcement. *IEEE Transactions on Neural Networks*, 12, 264-276.
- Smallwood, J., McSpadden, M., & Schooler, J. (2008). When attention matters: the curious incident of the wandering mind. *Memory & cognition*, 36(6).
- Smallwood, J., Nind, L., & O'Connor, R. (2012). When is your head at? an exploration of the factors associated with the tempo-

- ral focus of the wandering mind. *Consciousness and cognition*, 18, 118-125.
- Steindorf, L., & Rummel, J. (2017). I should not forget the apples! - mind-wandering episodes used as opportunities for rehearsal in an interrupted recall paradigm. *Applied Cognitive Psychology*, 31, 424-430.
- Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning and reacting. *ACM SIGART Bulletin*, 2(4), 160-163.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: an introduction*. Cambridge, Massachusetts: The MIT press.
- Weinstein, Y. (2017). Mind-wandering, how do i measure thee with probes? let me count the ways. *Behavior Research Methods*.