Influence of task length and task complexity on mind-wandering

Ivana D.M. Akrum (i.d.m.tjong.a.hung@student.rug.nl)
Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Abstract

Everyone mind-wanders. It is the process where attention strays from a current task in favour of thoughts unrelated to external events. This research looks into the joined effect of task complexity and task length on the mind-wandering process during a Sustained Attention to Response Task (SART). A model built in the ACT-R cognitive architecture is used to test whether a task that is both long *and* complex cancels out the individual effects of task length and task complexity. We define task length as the number of trials in a task and task complexity as the task's strain on working memory. Generally, we see task length leads to an increase in mind-wandering, and task complexity leads to a decrease in mind-wandering. A combination of both should then result in the same amount of mind-wandering as a standard SART would. The experiment reveals that there is no evidence there is a significant difference in how much mind-wandering occurs in the long, complex task versus the standard SART. We conclude that a combination of length and complexity do balance each other out. From this we can say that spending a long time on a complex task will have the same result in terms of how well you can pay attention to said task (on average) as spending a normal amount of time on a task of normal difficulty would. If a task in daily life is more difficult, it is thus beneficial to spend more time on it.

Keywords: mind-wandering; task complexity; task length; SART task; ACT-R

Introduction

In daily life, people spend up to 50% of their waking hours on mind-wandering (Smallwood, McSpadden, & Schooler, 2007). Mind-wandering is a process where attentions strays from a current task in favour of thoughts unrelated to external events. It has been modelled by various scientists in cognitive architects (van Vugt, Taatgen, Sackur, & Bastian, 2015). In this paper, we will use such a cognitive model to take a closer look at the circumstances that influence mind-wandering.

More specifically, we will look at how the length and complexity of a task play a role. Previous evidence suggests more complex tasks will lead to less mind-wandering and more on-task thoughts (Smallwood & Andrews-Hanna, 2013). Here complexity is defined as tasks that require continuous attention and more cognitive resources (compared to simple tasks). For our experiment, a complex task is considered one that requires more use of our working memory, which has been identified as necessary for mind-wandering (Levinsion, Smallwood, & Davidson, 2012). With the working memory occupied with task-related activities, it will be less free for mind-wandering.

Furthermore, evidence suggests task-unrelated thoughts increase over the duration of an experiment (McVay & Kane, 2009). This leads us to believe that a longer task will lead to more mind-wandering as opposed to a normal length task (what constitutes as *normal* and *long* will be discussed in *Methods*).

With ideas of how the length and complexity of a task influence mind-wandering, we are interested in the effect of combining these two features. In this paper, we try to answer the question: **How does a combination of complexity and length of task influence the mind-wandering process?**

For the sake of thoroughness, we will check the effect of complexity and length separately to see if this aligns with the previous research mentioned. This leads to two sub-hypotheses:

a. *An increase in task complexity will lead to a decrease in mind-wandering (negative relation).*

b. *An increase in task length will lead to an increase in mind-wandering (positive relation).*

If we then combine complexity with length, as one leads to a decrease in mind-wandering and the other to an increase, we form the following hypothesis to the research question:

1. *In a task with increased complexity and length, the two effects will balance each other, leading to no significant difference in average mind-wandering* [1].

The hypothesis will be tested with a cognitive model that performs a Sustained Attention to Response Task (**SART**) under four different conditions. These will be further explained in the *Methods* section, but are meant to test two different task lengths (*normal* and *long*) and two different levels of task complexity (*simple* and *complex*).

Model

The cognitive model that will run the experiment is build in a cognitive architecture known as ACT-R (Adaptive Control of Thought-Rational), developed primarily by John R. Anderson (Anderson, 2007).

ACT-R works on the basis of buffers, chunks, and production rules. The buffers represent different brain regions/processes. For our experiment, we make use of four different buffers: visual, retrieval, goal, and manual. The visual buffer deals with the processing of visual information on the screen. The retrieval buffer is in charge of memory retrieval from the model's declarative memory. The goal buffer is used to keep track of the model's current goal. The final buffer known as the manual buffer processes any motor movements.

In the model's declarative memory, memories are represented as chunks of information. Each of these chunks has an activation level that needs to be above a certain threshold in order for the retrieval buffer to recall it. This is similar to how neurons work in our brain. If a memory is not recalled for a while, its activation level will decay over time. Generally, each time the memory is recalled, its activation is strengthened.

The buffers and chunks together help the model process information. The production rules determine the model's behaviour. These rules tell the model what actions to take, given some current internal state. A production rule might say: Given we see some number on the screen, press a key. The model will then check its visual buffer for information, is there a number on the screen? If its visual buffer confirms that is the case, the associated production rule will fire, telling the model (its manual buffer) to press a key. Production rules thus

[1] As compared to average mind-wandering of the standard task.

start with a condition, and result in the model taking some action (usually a manipulation of one of its buffers).

Our model is set-up to model mind-wandering during a task. The task is performed very straightforwardly. The model is presented with an experiment in the same way a human participant would be, and it gives a certain response to what it sees according to its visual buffer. What response to give is encoded in one of the model's memory chunks known as stimulus-response mappings. In the case of our experiment, the response is either to press a key using its manual buffer, or to withhold on pressing. In summary, the model sees something, processes what it has seen, and then dives into its memory to find the appropriate response. Finally, the model gives that response.

As was said in the *Introduction*, complexity has been modelled as a greater strain on the declarative memory. Instead of immediately getting a response based on a visual input, the model takes an intermediary step where it retrieves some fact about this input. The model subsequently uses this fact to retrieve the right stimulus-response mapping. The model's retrieval buffer is thus busier than it would have been in a simple task.

During the task process, there is a chance the model will start to mind-wander. The mind-wandering process had been designed as a series of memory retrievals (van Vugt et al., 2015). This means that, instead of completing the task, the model will start to continuously retrieve memories from its declarative memory. The associated memory chunks are meant to represent some of the benefits of mind-wandering (Mooneyham & Schooler, 2013). The model will stop mind-wandering when it retrieves a memory chunk that tells it to refocus onto the task at hand.

The model has two goals: `attend` and `wander`. We said before that generally, when a chunk is called, its activation strengthens. In the case of the goal chunks, we actively prevent this from happening such that each of these goals has the same chance of being chosen throughout the experiment trials (and any difference is attributed fully to decay and noise). The task length is simply a parameter in the model that can be adjusted as desired, but what we hope to see in the model is that over time the `attend` chunk has a lower activation than the `wander` chunk, thus causing the model to mind-wander more often (have `wander` as its goal) as a result of the decay.

The ACT-R architecture does operate under certain assumptions. One of the most debated assumptions is the fact that in ACT-R only one production rule can fire at a time, resulting in a serial bottleneck. Other cognitive architectures, such as SOAR (Laird, 2012), allow multiple production rules to fire in parallel (although still only one action can be performed at a time). Similarly to the production rules, each individual buffer can also only contain one chunk at a time, and while a buffer is busy (e.g. the retrieval buffer is in the process of retrieving some memory chunk), it cannot be used for another action. Despite of these assumption, the ACT-R architecture has proven itself successful in modelling/predicting human data across various cases (Anderson, 2007) and is thus widely used to create cognitive models.

In designing our model, we made the added assumption that a participant that is in the process of mind-wandering (and thus not having task-related thoughts) will not respond to an on-screen visual input. The alternative is that the participant gives some response because of habituation, but for the sake of simplicity, our model does not take that into account.

Methods

In order to investigate the influence of complexity and length of task on mind-wandering, we set up a Sustained Attention to Response Task (SART) (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997) under four different task environments. In this task, participants are shown a number ranging from 1 to 9 on a screen. All numbers are considered targets (require a response), except number 3. Number 3 is a non-target (requires no response). For our experiment, we have the model perform this same SART, with one addition: What to give as a response depends on the task environment. Across all task environments, each trial has a 90% chance of being a target, and a 10% chance of a non-target. The following task environments have been modeled:

- **standard SART**: 180 trials. Press *f* on target[2]. Do not respond on non-target.
- **long SART**: same procedure as standard SART, but with 540 trials (3 times longer).
- **complex SART**: 180 trials. Press *f* on all even targets and *j* on all odd targets. Do not respond on non-target.
- **complex, long SART**: same procedure as complex SART, but with 540 trials (3 times longer).

In each task environment, the independent variables manipulated are the duration of the task (number of trials) and its complexity (response difficulty i.e. the amount of steps necessary before a response can be given). The dependent variable we are interested in is how much the model mind-wanders in each task environment. How much the model mind-wanders is represented by the cross participant average activation of the `wander` chunk binned to seconds in one trial (across all trials). Concretely, the dependent variable is thus the activation of the `wander` chunk.

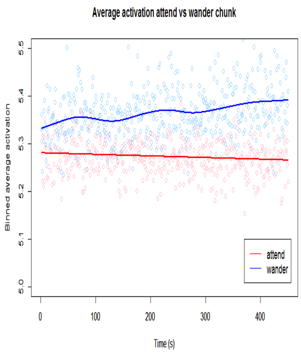
As stated previously in *Introduction*, the long SART environment and the complex SART environment are only modelled to confirm the sub-hypotheses (if one, or both, of these cannot be replicated, then our hypothesis is likely to be incorrect as well). The most important comparison is that between the average activation of `wander` chunk of the standard SART (s) and the complex, long SART (cl). If there is no significant difference in average mind-wandering, we should see $H_0 : \mu_s - \mu_{cl} = 0$. Else, we can claim $H_A : \mu_s - \mu_{cl} \neq 0$.

[2] Recall from *Model* that this is a key-press, thus performed by the manual buffer.

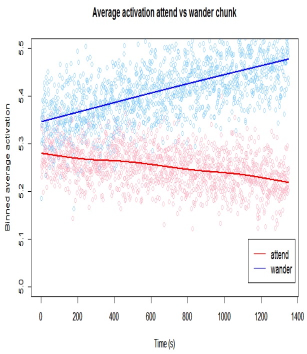
The significance of any difference will be checked using a two-tailed t-test with a significance level of 0.05[3].

Results

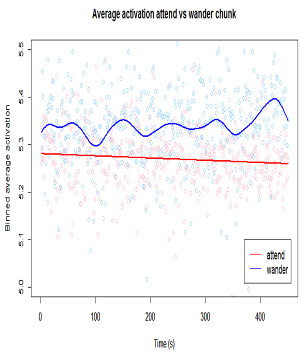
We have created four mean averages, one for each task environment. These are cross-participant and over all trials. They consist of an array of activation values mapped to the seconds of one trial length (binned average activation). For the 180 trials length, there are 500 seconds ($\approx$ 500 data points). For the 540 trials length, there are 1400 seconds ($\approx$ 1400 data points). Some NAN-values were removed from these arrays, which were seconds for which no mean exist because there is no data for the activation of the `wander` chunk at that second.



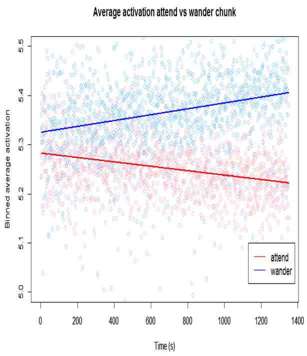
(a) Standard SART



(b) Long SART



(c) Complex SART



(d) Complex, long SART

Figure 1: Plots of the binned average activation (y-axis, binned over time) of the `wander` chunk versus the `attend` chunk over time (s) (x-axis).

In the plots in figure 1, the general trends of the data can be seen. Here we see for the task environments of 180 trials (1a and 1c), the activation of `wander` fluctuates. In all task environments there is a decrease of `attend`, and in all but plot 1c there is a clear linear increase of `wander`. We also see in all plots (1a to 1d) that the `attend` chunk starts with a lower activation than the `wander` chunk.

Since the plots do not allow for an easy comparison of the means, we determined the mean average for each task environment (and the associated standard error). This is the mean (sum of all data points / number of data points) of the original arrays of the binned average activation.

Table 1: The mean activation of wander chunk cross participant over all trials (rounded to 5 significant digits).

task environment	mean average	standard error
standard SART	5.3649	0.0023962
long SART	5.4125	0.0017008
complex SART	5.3399	0.0048074
complex, long SART	5.3651	0.0027503

Finally, in order to state something about the significance of the difference in means, we performed an unpaired t-test comparing the standard SART (M=5.3649, SD=0.0023962) with the complex, long SART (M=5.3651, SD=0.027503); $(t(25) = -0.065, p = 0.9478)$[4]. With $\alpha = 0.05$, we find there is no evidence that the difference in means is significant, as $p \geq 0.05$. Note that the arrays of binned average activation were given as input for the t-test, not the means in table 1.

Discussion

In graph 1b it is quite clear that there is more mind-wandering, the longer the experiment has been running. This fits with the sub-hypothesis:

b. *An increase in task length will lead to an increase in mind-wandering (positive relation).*

Thus we can state **there is a positive relation between task length and mind-wandering**.

In graph 1c the linear increase from the standard and long SART is no longer visible. Nonetheless, there does not seem to be a decrease in mind-wandering compared to the standard SART either. Checking table 1, it shows the mean average of the complex task is lower. We carefully confirm the hypothesis:

a. *An increase in task complexity will lead to a decrease in mind-wandering (negative relation).*

There is a negative relation between mind-wandering and task complexity. This relations is, however, less strong than the positive relation associated with hypothesis b.

Finally, we can conclude something about our main hypothesis. The acceptance of our sub-hypotheses would suggest our main hypothesis should also be accepted. In the t-test performed, the null hypothesis was not rejected, meaning there is no evidence that there is a significant difference in means between the standard SART and the complex, long SART. We thus accept the hypothesis:

1. *In a task with increased complexity and length, the two effects will balance each other, leading to no significant difference in average mind-wandering.*

[3] All statistical analysis is done using R (R Core Team, 2013).

[4] Since the results of the long and complex SART are not the main focus of the experiment, no significance test was performed to compare these to the standard SART.

Task complexity and task length do balance each other out. Thus, there is not a significant difference in average mind-wandering of a standard SART compared to a complex, long SART.

What does accepting our hypothesis mean on a larger scale? Since mind-wandering takes up 50% of our waking hours, it is important to understand the conditions under which it occurs. We modelled task complexity as an increase in how much memory retrieval is necessary for the task. Our results suggest this assumption is at least partly correct as it does support our data. Alternatively, it could be modelling complexity in this way works only because we modelled the mind-wandering process as a series of memory retrievals as well, which then logically causes these two processes to interfere with one another. However, since both these assumptions fit previous research results (van Vugt et al., 2015), (Smallwood & Andrews-Hanna, 2013), (Levinsion et al., 2012), we suggest the former interpretation is more likely than the latter.

Furthermore, the conclusion that task complexity and task length balance each other out indicates that spending a long time on a complex task will have the same result in terms of how well you can pay attention to said task as spending a normal amount of time on a task of normal difficulty would. Thus, if a task in every day life is more difficult, it is beneficial to spend more time on it.

There is a very important aspect of human cognition our model does not account for though, which is mental exhaustion. It is quite possible for a person to spend a long time on a boring task that involves pressing a button every now and again. Doing the same for a task that requires a lot of mental power is likely to be much more taxing. Studying the effect of exhaustion on the mind-wandering process is an interesting question to pursue in future research. We expect that exhaustion would lead to a decrease in task performance over time and perhaps longer episodes of mind-wandering.

In regards to task performance, our model only looks at the effects of the task environments on the amount of mind-wandering engaged in. We do not spend any time looking at the task performance, as it was irrelevant to our hypothesis. It would, of course, also be interesting to see how each of the task environments influence task performance in terms of response time and accuracy. We expect task length to lead to an increased response time and task complexity to lead to a decreased accuracy.

The fact we do not check these also has as drawback that we cannot use them as a measure of how accurate our model as a whole is. The effects of the task environments on mind-wandering may fit our hypothesis, but this means little if the model's overall performance does not fit with human data. However, comparison with human data is already difficult, since our model was primarily built on previous researches (as opposed to human trials) and we thus do not have human data that fits our model (and the way we model task complexity) specifically. We do see our model has a much higher accuracy compared to human data (97% as opposed to human 60% $\sim$ 80% for the standard SART)[5]. A final idea for future research is to perform our experiment with human participant, so that the model can be adjusted to fit human performance better.

Despite of this, the model was built in what is widely considered a reliable cognitive architecture (ACT-R) (Anderson, 2007), and most aspects of the model design were based on research as well. Thus, although our model does not keep every aspect of human cognition in mind, it is an accurate enough portrayal of the SART task and the mind-wandering process it intends to model. Considering our research question, these are the most important factors the model needs to portray accurately, and it does.

References

Anderson, J. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Laird, J. E. (2012). *The soar cognitive architecture*. MIT Press.

Levinsion, D., Smallwood, J., & Davidson, R. (2012). The persistence of thought: Evidence for a role of working memory in the maintenance of task-unrelated thinking. *Psychological Science*, *23*, 375–380.

McVay, J., & Kane, M. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, *35*, 196-204.

Mooneyham, B., & Schooler, J. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology*, *67*, 11–18.

R Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`

Robertson, I., Manly, T., Andrade, J., Baddeley, B., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*, 747–758.

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: the importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology*, *4*, 441.

Smallwood, J., McSpadden, M., & Schooler, J. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, *14*, 527–533.

van Vugt, M., Taatgen, N., Sackur, J., & Bastian, M. (2015). Modeling mind-wandering: a tool to better understand distraction. In *Proceedings of the 13th International Conference on Cognitive Modeling* (p. 252). Groningen: University of Groningen.

[5]This was omitted from *Results* at a lack of relevancy for the hypothesis.