# Auditory Stimuli and Their Influence on Mind Wandering in a SART

**Merlijn Mulder (m.q.mulder@student.rug.nl)**
Faculty of science and Engineering, University of Groningen
Broerstraat 5, 9712 CP, Groningen

## Abstract

Mind Wandering (MW) decreases performance in various attention tasks. We investigate how a task-relevant (cue) and a task-irrelevant (distraction) auditory stimulus affect MW in a sustained attention response task (SART). Using the ACT-R architecture, we build a simple MW model in which we simulate three conditions, consisting of a condition without an auditory stimulus and the different stimuli previously stated. Although MW remained near 50% across conditions, our results reveal that the auditory stimuli affects the timing of MW. Cues significantly reduced MW at moments when a button press was required in response to a visual stimulus. The task-irrelevant condition displayed a similar but less pronounced effect. A one-way ANOVA confirmed significant differences between the conditions in MW frequency upon the visual stimulus onset in the SART, indicating auditory stimuli increase attention when needed in a task. Therefore, we cannot conclude that the auditory stimuli influences the total time spent MW, but it does influence when you mind wander.

**Keywords:** Mind Wandering, ACT-R, SART, Auditory Stimuli, Distractions, Cues

## Introduction

You may have experienced this: while listening to a podcast, lecture, or video, your mind drifts away. You might not notice it at first, but when you are addressed, you are startled. You realize you ended up in a mind-wandering (MW) loop and are not focused on your primary task anymore. Research shows that MW occurs about 50% of the time (Mooneyham, Benjamin, & Jonathan, 2013). You could view this process as successfully getting distracted.

Such a cognitive model in which both task-irrelevant distractions and MW are defined has been proposed by Taatgen et al. (2021) through three experiments. According to the paper, a distraction is a task that has secondary priority, external stimuli and/or unrelated thought processes. The external stimuli can be both auditory or visual, and the unrelated thought process can be considered MW. However, they argue that there is no distinction between them as all compete for what is next mentally. The extent to which a task is disrupted depends on the strength of the goal, its ability to trigger the right response and the prominence of the distraction. They will succeed when, at first, they make use of an unused cognitive resource, allowing a subject to notice the task-irrelevant stimuli. Goal-related tasks do not inherently dominate distractions as the interaction of task goals, sensory inputs, and memory activation shapes cognitive processing.

However, the threaded cognition theory states that separate cognitive resources can operate in parallel (Niels A. Taatgen, 2008). While competition can cause errors, distraction and MW may also coexist. Both can be triggered by internal and external resources, where the most active operator determines the outcome. When a focus (or attend) goal/operator weakens, the wander task starts to retrieve memories. When elaborating on these memories, you enter a sustained MW process leading to a decrease in accuracy and response times (Taatgen et al., 2021).

In relation to this, Robinson and Unsworth (2015) examined the impact of MW and distraction in silent and noisy environments using working memory capacity tasks. Their findings support a distinction between MW and external distraction. They measured this using thought probes. The self-reported number of MW was equal in both conditions, whereas the distraction rate of the participant in the noisy condition was higher. This contradicts Vannucci, Pelagatti, and Marchetti (2017), concluding an increase in self-reported MW when task-irrelevant verbal stimuli where presented on the screen in a vigilance task compared to a control condition. Notably, the type and prominence of task-irrelevant stimuli differ between studies. Furthermore, the participants became less distracted by external stimuli overtime during the task. Furthermore, a sudden loud unexpected beep can be considered more prominent than continuous background noise from a noisy environment. As described, the prominence of a task-irrelevant stimulus can define its effectiveness (Taatgen et al., 2021), providing a possible explanation for why it did not lead to an increase in MW.

Another prominent factor is cognitive exhaustion. Whilst this can be one of the explanation for why participants became less distracted over time it supports research by Helton and Russell (2011) investigating the impact of working memory load on vigilance decrement. The study finds that both verbal and spatial working memory tasks worsen performance over time. These results support the resource depletion model, as multitasking depletes cognitive resources and thus reduces attention. The results provide evidence against the boredom hypothesis, as multitasking would make the task more difficult. This suggests that vigilance failure results from cognitive exhaustion and not disengagement. As Taatgen et al. (2021) defined a distraction as a secondary task, one would effectively try to multi-task when it occurs as you try to stay focused on the primary task, not letting the distraction succeed. Therefore, you could reason that a prominent distraction is more effective over time, including MW. A study conducted by Thomson, Besner, and Smilek (2015) examined this relation. They introduce a resource-control theory by integrating resource depletion and MW theories to explain vigilance failures. They argue that neither theory explains experimental results well enough and argue that attention lapses arise from both cognitive exhaustion and executive control failures. Based on their theories, they suggest that attentional control decreases as resources deplete, which leads to MW and disrupted performance on the primary task. Therefore, we have now established a theory that an external task-irrelevant stimulus could potentially lead to an increase in MW, leading to the question if an external task-relevant stimulus would result in the opposite. A task-relevant stimulus, or cue, can be defined as Intentional sounds to capture attention and facilitate task performance (Brock, Halverson, Hornof, & McClimens, 2006). Within classification and tracking tasks for military purposes, a sound played when in the task a "blip" changed and needed classification. This led to reduced gaze shifts, faster response times on tasks and a reduced attention shift. The reduced gaze shifts could be defined as a reduced frequency of MW, although the article does not explicitly report on this. Furthermore, you could theorize that such a cue could also help you learn or recognize a regular interval faster in tasks like the sustained attention response task (SART). Identifying such a regular interval can reduce the frequency of MW when a stimulus in the experiment needs attending. To support this theory Seli et al. (2018) implemented an experiment in which participants had to respond to stimuli that appeared at regular predictable intervals (20 s). Through intermittent probes, they tested the MW kind (intentional or unintentional) and frequency of MW. The results showed that the MW frequency peaked just between the intervals and decreased before the next interval was due. From this, we can deduct that participants subconsciously planned their attention effort necessary for the task and adjusted it based on the intervals.

We have not identified a cognitive model explaining both task-relevant and task-irrelevant auditory stimuli using a SART at the time of writing. This leads to the question if *the presence of a prominent auditory stimulus in both task-relevant and task-irrelevant conditions in a sustained attention response task decreases the frequency of mind wandering in comparison to no auditory stimulus.* Based on the presented literature we hypothesize that a task-relevant auditory stimulus reduces MW (Brock et al., 2006), while a task-irrelevant stimulus increases it (Vannucci et al., 2017).

We aim to test this using a SART and a prominent tone serving as a distraction or cue. Building from the MW model proposed by van Vugt, Taatgen, Sackur, and Bastian (2015) in ACT-R. A production rule that can fire when the auditory stimuli is perceived will lead the model to either go to the attend part of the model or the MW part depending on whether a visual stimulus is presented. To test both hypothesis we will test this using a condition with a task-relevant, task-irrelevant or no auditory stimulus.

## Model

For the model implementation, we will be using ACT-R, a cognitive architecture that models human cognition. This computational framework, which consists of procedural and declarative memory interacting through a production system, is grounded in psychological theories. Therefore it can accurately mimic human cognitive processes on an abstract and task specific-level depending on the implemented model. Due to its validity and predictive power we follow previous studies by utilizing this architecture as shown. As seen in Figure 1, we build on the model proposed by van Vugt et al. (2015). They proposed two cognitive roads, which you can access via a production where you check your current goal, wander or attend. One such road identifies the presented stimulus and then retrieves and executes the correct response. The other defines MW as a loop of retrieving random memories. When the memory's content reminds you to attend, you will go back to checking your current goal. Additionally, this loop incorporates a standard response, which can fire when a stimulus is presented, independent of its content.

We added an auditory-stimulus production setting the state in the retrieval buffer to *ATTEND* and empties the goal state and aural-location buffer. Theoretically, this results in the model firing the check-current-goal production if there is no scene change and otherwise firing the identify-stimulus production if a visual stimulus is registered as seen in Figure 2. This marks the difference between the possible productions following auditory-stimulus production. The state of the visual location buffer is dependent on the visual stimulus onset.

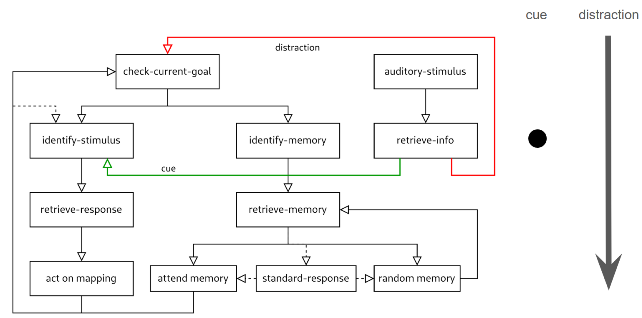


Figure 1: An overview of the extended model. The red and green lines portray which production has the highest chance of being redirected to when the state is set to *ATTEND*.

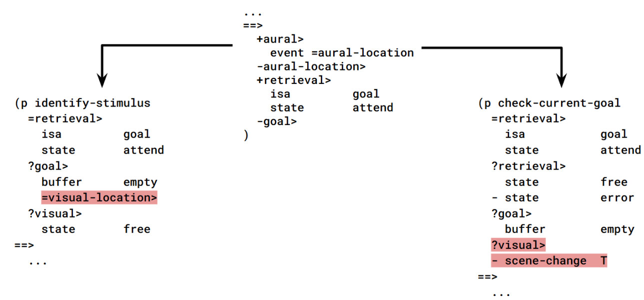


Figure 2: The possible productions firing after the auditory-stimulus production depending on the visual stimulus onset. The essential difference in pattern matching between the productions is marked in red.

The base-level chunk activation of *random-memory* is set `(12000 -10000)`. This is chosen so that the base model will spend 50% of its time MW during the task (Mooneyham et al., 2013). This percentage is confirmed through an analysis of the trace file. All other base-level activations are set to `(10000 -10000)`, keeping the default activation.

Within our study, we define a cue as a prominent auditory stimuli of `0.5 seconds` that only appears when a visual stimulus is presented. A distraction will appear randomly between `0.5 seconds` and `2 seconds`. This will ensure the distraction is presented within the trail that has a default duration of `2.5` seconds. A cue and a distraction, therefore, differ in their timing of presentation. This results in the different productions previously described firing after the auditory stimulus has been acted on, as outlined in Figure 1. All auditory stimuli are a tune at `1000 Hz`, arbitrarily chosen within the range of human perceivable frequencies.

According to the presented theory, resource depletion can play a major role in how a stimulus affects the frequency of MW. However, since resource depletion was difficult to implement in our model, we assume that the task is brief enough for this effect to be negligible. Furthermore, as we are only interested in MW and attending, we assume a world where MW and *ATTEND* are the only two possible main states. Lastly, within our model, we assume that an auditory stimulus can either be a cue or a distraction, but it cannot become background noise, as this complements the previously defined conditions.

## Methods

With the auditory stimulus as our independent variable, we define three conditions: task-relevant, task-irrelevant, and no auditory stimulus. The model remains the same in each condition, but the task is altered by introducing an extra auditory stimulus. The task that participants must complete in each condition is the SART. Participants are presented with two visual stimuli. If the target *'O'* appears, they must press a button." This is also defined as the standard response for this task. When the non-target stimulus *'Q'* appears, participants must refrain from pressing the button

Since responses to task-relevant visual stimuli define task performance, we measure the frequency of MW during these responses. We measure this frequency based on the firing of the standard-response production during MW. This determines how often a subject responds while wandering compared to when attending. This is especially relevant in the task-relevant condition, as a cue should theoretically help direct a participant's attention to the task.

Furthermore, to analyze emergent behavior across conditions, we measure MW frequency in each condition, defined as the time spent in either the *ATTEND* state or the *WANDER* state. We will not count the start of each state as this will only differ by one due to the two-condition environment rendering the measure useless. To compare conditions and test for significant differences, we will use a one-way ANOVA.

In each condition, we will run 200 trials per 25 participants. These numbers are arbitrarily chosen but are large enough to reduce noise and provide a better representation of the population.

## Results

Figure 3 shows the comparison of MW frequency across conditions at the times a subject responds. The model appears less frequent in the MW state when responding to visual stimuli if an auditory task-relevant stimulus ($M = 3.92$, $SE = 0.26$) or a task-irrelevant auditory stimulus ($M = 6.60$, $SE = 0.37$) is present compared to when no auditory stimulus is given ($M = 7.74$, $SE = 0.39$). A one-way ANOVA reveals a significant difference between the three conditions ($F(2) = 31.09$, $p < .001$).

To analyze MW frequency throughout the entire SART, we examined the model trace by tracking the time that a participant spent in the *WANDER* state and in the *ATTEND* state. The average result for all participants is presented in Figure 4. This shows the mean frequency of MW is 49.2% with no auditory stimulus present, 49.4% with a task-relevant stimulus present and 48.5% with a task-irrelevant stimulus present. A one-way ANOVA on these results indicates a significant difference between conditions ($F(2) = 6.80$, $p = .002$).

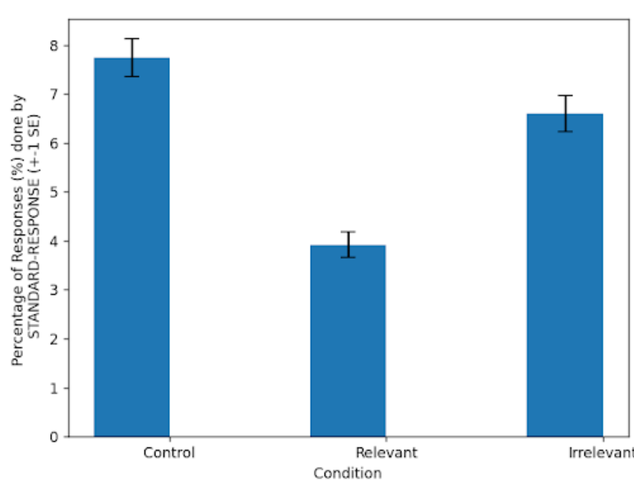


Figure 3: The frequency of responding to a visual stimuli whilst MW according to the standard response production

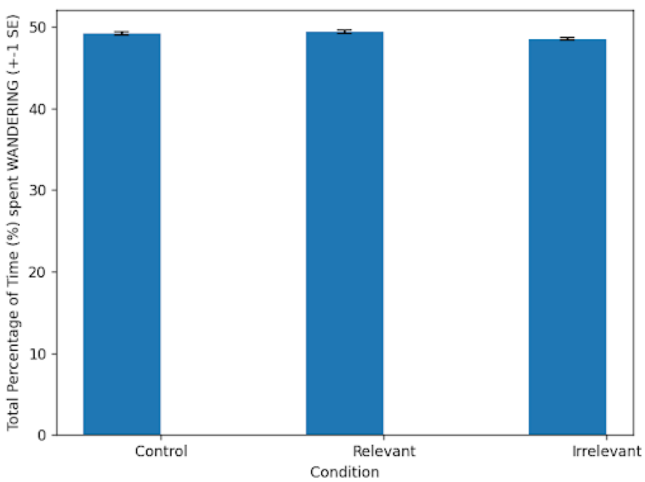


Figure 4: Relative frequency of MW throughout the trail

## Discussion

In this study, we investigated whether the presence of an auditory stimulus in both task-relevant (cue) and task-irrelevant (distraction) conditions in a sustained attention response task decreases the frequency of MW compared to having no auditory stimulus.

MW frequency decreases for both conditions when a visual stimulus is presented and the model has to respond, compared to when no auditory stimulus is present. This effect is most pronounced in the task-relevant condition, confirming our hypothesis and aligning with the findings of Brock et al. (2006). On the other hand, we expected the opposite for the task-irrelevant stimuli. Potentially, some task-irrelevant stimuli might have served as a cue, ripping you out of MW and allowing more opportunity to attend. This could lead from the fact that the MW is interrupted once the auditory stimuli is noticed. Thus, the likelihood of retrieving random memories decreases if the task-irrelevant auditory stimulus is played close to the visual stimulus onset, making it more likely that the model checks its current goal instead of staying in the MW loop. If the decrease can be attributed to this fact, this supports the hypothesis from Robinson and Unsworth (2015), stating there is no change in MW frequency upon the presentation of task-irrelevant stimuli. However, this contradicts the hypothesis from Vannucci et al. (2017), expecting an increase compared to no auditory stimuli.

As shown in Figure 4 the relative MW frequency is approximately 50% across all conditions, supporting Mooneyham et al. (2013) hypothesis that MW occurs about half of the time. That this is the case for the condition without an auditory stimulus increases the model's plausibility, assuming their hypothesis holds true universally. Although the test reveals a significant difference between conditions, our observations suggest that the effect size is very small, as the results show minimal variation. Thus, we cannot conclude that an additional emergent behavior arises throughout the task by changing the way the model responds to the visual stimulus when it perceives an auditory stimulus. This is possibly due to the model switching between MW and attending multiple times a second, whereas the state is only reset once after the auditory stimulus is perceived, meaning it has a small impact on the overall frequency.

Taking all these results into account, we argue that we cannot conclude that the auditory stimulus influences the total time spent MW, but it does influence the frequency of MW at the onset of the task-relevant visual stimuli.

However, our model does not account for resource depletion, which may influence MW throughout a task, as noted in previous studies (Thomson et al., 2015) (Helton & Russell, 2011).

Additionally, our model does not account for the possibility that the auditory stimulus becomes background noise over time. Therefore, these potential improvements can be a promising way forward for future research to more completely represent the influential cognitive processes. Nevertheless, a key strength of our model is its simplicity. It only resets the state in which a subject finds themselves, proving a minimal adjustment that allows all other cognitive processes to be the next mental step as it defaults back to checking its goals, being able to enter both cognitive roads. This also contributes to the plausibility of our model as it does not alter the base model of MW as proposed by van Vugt et al. (2015). This minimal implementation relies on the previously stated assumptions presented in the Model section. Lastly, as we have shown, our model generalized to both task-relevant and task-irrelevant stimuli. This model can also be applied to variations of the task, as the implementation is not specific to the SART. Currently, the model only applies to detectable tones. If words were presented as a task-irrelevant stimulus, they could trigger specific memories, allowing for further elaboration based on memory content leading to more sustained MW (Taatgen et al., 2021). This is not part of the current model. Future research could explore whether our mechanism generalizes to such additions.

In conclusion, we introduced a simple mechanism that builds upon the model of MW proposed by van Vugt et al. (2015). Our findings indicate that task-relevant stimuli (cues) reduce MW when a task-related visual stimulus is presented, while total MW duration throughout the task remains unchanged. The same, but less pronounced, decrease was observed to task-irrelevant stimuli (distractions). Since the model can switch states multiple times per second, additional MW outside the stimulus onset is minimal, if any, after the auditory-stimulus production resets the state. Future research could validate the assumptions of this model and incorporate missing cognitive processes, such as resource depletion.

## References


Brock, D., Halverson, T., Hornof, A. J., & McClimens, B. (2006). Cognitive models of the effect of audio cueing on attentional shifts in a complex multimodal, dual-display dual task. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *28*. Retrieved from `https://escholarship.org/uc/item/70d1x38m`

Helton, W. S., & Russell, P. N. (2011, July). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*(3), 429–437. doi: 10.1007/s00221-011-2749-1

Mooneyham, Benjamin, W. S., & Jonathan, W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology*, *61*(1), 11–18. doi: 10.1037/a0031569

Niels A. Taatgen, D. D. S. (2008). *Threaded Cognition: An Integrated Theory of Concurrent Multitasking.* doi: 10.1037/0033-295X.115.1.101

Robinson, M. K., & Unsworth, N. (2015). *Working memory capacity offers resistance to mind-wandering and external distraction in a context-specific manner.* doi: 10.1002/acp.3150

Seli, P., Carriere, J. S. A., Wammes, J. D., Risko, E. F., Schacter, D. L., & Smilek, D. (2018). On the clock: Evidence for the rapid and strategic modulation of mind wandering. *Psychological Science*, *29*(8), 1247-1256. doi: 10.1177/0956797618761039

Taatgen, N. A., van Vugt, M. K., Daamen, J., Katidioti, I., Huijser, S., & Borst, J. P. (2021). The resource-availability model of distraction and mind-wandering. *Cognitive Systems Research*, *68*, 84-104. doi: 10.1016/j.cogsys.2021.03.001

Thomson, D. R., Besner, D., & Smilek, D. (2015, January). A Resource-Control Account of Sustained Attention: Evidence From Mind-Wandering and Vigilance Paradigms. *Perspectives on Psychological Science*, *10*(1), 82–96. doi: 10.1177/1745691614556681

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). University of Groningen.

Vannucci, M., Pelagatti, C., & Marchetti, I. (2017). Manipulating cues in mind wandering: Verbal cues affect the frequency and the temporal focus of mind wandering. *Consciousness and Cognition*, *53*, 61–69. doi: 10.1016/j.concog.2017.06.004