university of groningen
faculty of science and engineering

Increasing Accessibility for Learners with ADHD in an Adaptive Learning System through the User Interface

Indy van Boven

university of groningen

faculty of science and engineering

University of Groningen

Increasing Accessibility for Learners with ADHD in an Adaptive Learning System through the User Interface

Master's Thesis

To fulfill the requirements for the degree of
Master of Science in Computational Cognitive Science
at University of Groningen under the supervision of
Prof. dr. J.P. Borst (Artificial Intelligence, University of Groningen),
T.M. Haile, PhD (Artificial Intelligence, University of Groningen),
and
Dr. M. van der Velde (MemoryLab)

Indy van Boven (s4340280)

July 16, 2025

Contents

Acknowledgments

Many thanks to my supervisors for your expertise, help and daily supervision along this Master Research Project! I would also like to thank my family and friends for their support and happy distractions along the way!

This thesis made use of the template made by author Manvi Agarwal with Creative Commons CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/). The original template can be found here: https://www.overleaf.com/latex/templates/master-thesis-template-university-of-groningen/txxpkkjmfksx. Besides the content (text and images), I modified the reference style through adding and omitting packages, added packages such as to include URL's and added, modified or omitted headers, which all also changed the table of content.

Abstract

Attention-deficit/hyperactivity disorder (ADHD) can be linked to academic underachievement, such as not obtaining expected degrees (Biederman et al., 2008). In the Netherlands, adulthood ADHD has an annual prevalence of 3.2% (ten Have et al., 2023). Learning has been optimized for many students, but it is not yet clear if they improve learning or are accessible for learners with ADHD.
This project aims to investigate if frequent breaks and self-monitoring prompts within MemoryLab's adaptive learning system, improves learning outcomes. Taking breaks has had varying results in literature, and self-monitoring, which entailed cues for assessing and recording on-task attention, was related to increased attention and performance.
Twenty-seven participants learned the flags of nations and were assigned to counterbalanced blocks of uninterrupted learning and learning with breaks or self-monitoring. Participants completed the Dutch version of the Self-Report Questionnaire on Attention Problems and Hyperactivity for Adulthood and Childhood and gave their perspective on the sessions in an interview.
Overall, both breaks and self-monitoring prompt conditions did not lead to better or faster responses. Even though more interview responses indicated preference for uninterrupted learning, variation occurred between participants. Future work is needed to further investigate these approaches with adaptive learning, where sustained attention is required.

1 Introduction

People with Attention-Deficit/Hyperactivity Disorder face more difficulties compared to their peers, in regular activities like completing assignments, organizing these and their items, keeping focus and attention during a task (American Psychiatric Association, 2013). In the Netherlands, ADHD has an annual prevalence of 3.2% in adults aged 18 to 75, according to a population screening program conducted from 2019 to 2022 (ten Have et al., 2023). There is a connection between diminished achievement, attainment and ADHD in work and academics, according to the American Psychiatric Association (2013). In Biederman et al. (2008), fewer individuals had obtained a graduate degree or graduated from college than estimated. Advokat, Lane, and Luo (2011) found lower GPA's of students with ADHD in comparison to control students, despite those taking medication. Individuals with ADHD underachieve in educational and occupational environments (Kooij, 2022). While new learning technologies constantly innovate to improve learning conditions, their effectiveness for students with ADHD needs further investigation. In this project, I aim to test how previously studied approaches improve fact learning for individuals with ADHD symptoms with-in an adaptive learning system.

Learning via an online adaptive learning system is advantageous for better fact recall for learners in general (van Rijn, van Maanen, & van Woudenberg, 2009). However, not everyone is a typical learner. 6% of higher education graduates in the Netherlands, from the academic year 2019-2020 reported having ADHD, ADD or concentration problems and 4.6% reported being diagnosed in this category. Of these graduates, 48.6% stated hindrance in their education (Ramaekers & van der Mooren, 2022). Kooij (2022) provides a table of symptom presentation examples related to inattention, hyperactivity and impulsivity that an individual with ADHD might have. Among these are being easily disinterested, losing attention fast, being impulsive, being disorganized, difficulty recalling, hard time completing tasks or assignments and staying seated and still (Kooij, 2022). However, it is not that individuals with ADHD are never able to focus, but that they are not able to focus in necessary moments (Kooij, 2022). Nonetheless, being able to keep attention on the task, completing assignments and being able to sit still and be organized are necessary skills for academic and professional success, and coupled with the prevalence of ADHD and its hindrance in education, indicate the importance of improved educational tools for students with ADHD.

1.1 Improving on-task attention and performance

Various approaches have been investigated to aid children with ADHD in an academic setting (e.g. Mathes and Bender (1997); Reid, Trout, and Schartz (2005); Harris, Friedlander, Saddler, Frizzelle, and Graham (2005); Wills and Mason (2014); Pritchard et al. (2016); Harrison et al. (2022); Blood, Johnson, Ridenour, Simmons, and Crouch (2011)), in order to improve disruptive behavior, attention and/or academic performance. One of these approaches is self-monitoring. Self-monitoring is a self-regulation strategy, where the latter has the purpose of building and establishing behavior that is self-regulated by the individual themself (Reid et al., 2005). Self-monitoring and self-management interventions for children with ADHD have shown positive results, especially in on-task attention or behavior (e.g. (Mathes & Bender, 1997; Blood et al., 2011; Harris et al., 2005; Wills & Mason, 2014)) and performance (e.g.(Harris et al., 2005)). This intervention in turn also improved performance for adults, through an increased GPA (Scheithauer & Kelley, 2017). More specifically, self-monitoring of attention, where the students evaluates and then records whether their attention was on- or off-task at set intervals has led to increased on-task attention and/or performance (e.g. Harris et al. (2005); Mathes and Bender (1997)). Self-monitoring of attention could be utilized to stimulate the perception or consciousness of whether she or he is paying attention on the work (Reid et al., 2005). Coming back to the purpose of self-regulation strategies in Reid et al. (2005), the purpose of self-monitoring of attention could then mean that the individual self-evaluates and records whether their attention was on the task and self-regulates this attention. In addition to self-monitoring (of attention), rewards can be applied as well as a form of reinforcement from something or someone else due to positive/changed behavior (Reid et al., 2005). The self-monitoring sheet utilized in Blood et al. (2011) show emoticons, next to the yes/no answer of being on-task. We have based our self-monitoring page on this sheet, where we apply positive emoticons as a form of positive reinforcement, when the learner responses affirmatively to the on-task attention question.

Another approach that has been investigated for learners with ADHD is an accommodation, namely breaks (e.g. Hall-Ruiz (2016); Harrison et al. (2022); Pritchard et al. (2016)). The utilization of breaks to aid learners with ADHD has resulted in varying outcomes in literature. A break accommodation, as investigated in the form of one research(assistant)-initiated large break, did not result in the best outcome for children (Harrison et al., 2022). However, employing breaks with the addition of choice of duration and later different initiators in a continuous performance test (CPT see: Conners, Epstein, Angold, and Klaric (2003)) with college students, did show positive results (Hall-Ruiz, 2016). In a experiment with participant initiated breaks, they found that the number of false alarms between participants in the control groups versus the ADHD group were comparable when the latter took longer and three or more breaks (Hall-Ruiz, 2016). They also found that most of the participants in the control group favored breaks they could initiate themselves, whereas half of the participants in the ADHD group favored the breaks initiated by the experimenter through their performance (Hall-Ruiz, 2016). Breaks might allow learners with ADHD to counterbalance effects of a wandering mind (Hall-Ruiz, 2016). As mentioned above, individuals with ADHD might struggle with keeping focus (American Psychiatric Association, 2013; Kooij, 2022). Applying breaks, splits this duration into shorter pieces, which means that attention is required, but for shorter durations.

1.2 Adaptive learning and ADHD

Those approaches that have been investigated academically for learners with ADHD in previous work might aid in an online adaptive learning system as well. On the other hand, an individual with ADHD might present minimized or no symptoms of ADHD at all in certain situations (American Psychiatric Association, 2013). Examples are situations in which the individual is doing something particularity engaging, is in a situation which is new to them or receives regular inputs or stimuli from outside sources (American Psychiatric Association, 2013). An online adaptive learning system provides these regular inputs via a digital display. Therefore, an adaptive learning system could be beneficial for learners with ADHD as well. Adaptive learning systems are described in various forms in literature. The adaptive fact-learning system utilized in this project is adaptive in the form of modifying the spacing and order of facts, based on the response time and accuracy for each seen fact and for each individual learner (van Rijn et al., 2009).

However, response times might not be an accurate representation of memory retention for every seen fact. To illustrate, a learner who finds sustained attention challenging and has their attention drawn away more easily, might take longer to respond than is expected. Their response time for that individual fact might then not be in line with their actual estimated retention of that fact. It is therefore important to investigate whether this system optimizes learning and is accessible for neurodivergent learners, specifically those with ADHD characteristics. Yet with the inclusion of possibly attention-improving approaches, response times might be more in line with their actual retention and the online adaptive system might be beneficial for non-neurotypical learners as well. More specifically, adults

with ADHD do not test themselves as much as those without ADHD during studying for a learning task in Knouse, Anastopoulos, and Dunlosky (2012). Therefore, employing an online adaptive system that employs testing in learning is already an improvement for learners with ADHD. Including interventions that encourages individuals with ADHD to monitor their attention or performance could be even more advantageous.
In this project the adaptive learning algorithm by MemoryLab (`https://www.memorylab.nl/`) is utilized, with interface modifications designed to improve learning for students experiencing symptoms of ADHD. The MemoryLab algorithm uses reaction times and accuracy, as the learner practices facts, in flashcard style, to establish the degree of retention (speed of forgetting, SOF) of that specific fact for that specific learner to tailor learning. This algorithm takes the spacing effect (e.g. see (van Rijn et al., 2009)) and the testing effect (e.g. see (Roediger III & Karpicke, 2006)) into account. Once the algorithm estimates a speed of forgetting, it presents facts to the learner right before it is likely to be forgotten, which has been shown to improve retention, especially compared to simple flashcards without adaptive learning implemented (van Rijn et al., 2009; MemoryLab, n.d.). The use of online adaptive learning systems presents promising opportunities for students with ADHD who are already struggling academically.

1.3 Implementation

We aim to investigate self-monitoring and breaks in an adaptive learning system. The user interface of the adaptive learning system will be modified to allow investigation into the approaches self-monitoring of attention and breaks with optional duration on performance through reaction time and accuracy. Participants self-monitor their attention through self-assessing and recording whether they were paying attention to the exercises or not. Positive reinforcement is employed when the participants responds to the self-monitoring question with ”yes”. The online self-monitoring form in this experiment will then show a positive smiley, based on the sheet in Blood et al. (2011). These self-monitoring prompts are implemented at intervals of at least 90 seconds. Breaks are implemented at approximately the same intervals. In order to limit confounds of forced breaks for learners that are not in need of one, we opted for breaks with an optional duration, which can be clicked away at any moment.
We expect learners with high ADHD characteristics scores to be more accurate while receiving self-monitoring of attention prompts than those learners that have lower ADHD characteristics scores, as the attention of a learner with ADHD might wafer easily and therefore self-monitoring of their attention might have more impact than for a learner that does not suffer from losing focus more easily. We expect that learners will have a higher accuracy and be more on-task from added breaks as opposed to uninterrupted learning. We further expect that those on with higher ADHD characteristic scores benefit more from breaks than those learners with lower ADHD characteristics scores. As the latter individuals might be less in need of a break due to fewer self-reported inattention and/or hyperactivity/impulsivity characteristics.

1.4 Research Questions

Q1. What is the effect of self-monitoring of attention on performance and on-task behavior within the MemoryLab system, while taking ADHD characteristics score into account?

- What is the effect of self-monitoring of attention on experience within the MemoryLab system, while taking ADHD characteristics score into account?

Q2. What is the effect of breaks on performance and on on-task behavior within the MemoryLab system, while taking ADHD characteristics score into account?

- What is the effect of breaks on experience within the MemoryLab system, while taking ADHD characteristics score into account?

2 Background Literature

2.1 ADHD

Attention-deficit/hyperactivity disorder presents itself in the form of mainly inattentiveness, mainly impulsiveness and hyperactiveness or a combination, with mild, moderate or severe symptoms (American Psychiatric Association, 2013). An individual with a mainly inattentive form of ADHD might frequently struggle with keeping focus, having their attention strayed due to external stimuli, having a disfavor or disinclination for a prolonged mental effort task, seemingly being absent-minded and being organized, among other symptoms (American Psychiatric Association, 2013). An individual with mainly an impulsive and hyperactive form might in turn struggle to remain seated or be quiet, might display interruptive behavior and might have a preference or inclination for instant gratification (American Psychiatric Association, 2013). Part of the requisites of an ADHD diagnosis for adults is the exhibition of at least five symptoms for a minimum of six months, with adverse effects on social endeavors and education/work, where these symptoms are present in at least two environments (American Psychiatric Association, 2013). Wilens et al. (2009) found that the majority of adults in their sample had combined ADHD, followed by 31% with the inattention type and lastly a minority (7%) with the hyperactivity/impulsivity type. They found more often reported symptoms of inattentiveness than of hyperactivity/impulsivity.

Besides medication, an ubiquitous form of intervention for ADHD, other approaches should be utilized as well to advance the learning capabilities of the in-class student more (Mathes & Bender, 1997). Mathes and Bender (1997) investigated the usage of self-monitoring, besides medication, which increased behavior on the task. In the Netherlands, one in twenty-five adults have utilized ADHD medication at least once at some point during their lives, according to an estimated measurement in 2022 Nationale Drug Monitor (2025). In a study by Advokat et al. (2011), the majority (78.3%) of participating college students with ADHD were on medication. However, in order to maintain or increase academic performance for college students with ADHD, medication alone is not enough (Advokat et al., 2011). Even though the majority of their participants with ADHD on medication stated that medication helped them (92,4%, see Table 2 in Advokat et al. (2011)), their academic performance in terms of overall GPA was significantly less than the GPA of those in the control group without ADHD (Advokat et al., 2011).

Mathes and Bender (1997) found that that the on-task behavior of their young participants increased when utilizing self-monitoring, in addition to having medication. Not only self-monitoring has been tested on students with ADHD, other approaches have also been investigated, such as breaks and prompting (see: Hall-Ruiz (2016); Harrison et al. (2022)) and have let to interesting results. Knouse et al. (2012) studied whether adults with ADHD show shortfalls in monitoring recall correctness and strategy use during their task and analyzed what techniques lead to improved performance in retention for those with ADHD and those without. After having studied noun pairs and while seeing one noun, participants had to judge whether they would recall the other noun, on a scale of zero to hundred (Knouse et al., 2012). They found that those participants that they detected to be utilizing self-testing recalled a larger number of words than those who did not make use of self-testing.

More of self-testing and ADHD, Knouse et al. (2012) found that adults with ADHD were using self-testing less often than adults without ADHD, based on participant answers on what amount they utilized self-testing. Additionally, fewer adults with ADHD applied self-testing, even once, during the studying part of the task, than adults without ADHD. Considering the importance of self-testing, the adaptive algorithm applied in this project automatically makes use of the testing effect, meaning that individuals do not have to apply testing on themselves. This argues the case that an adaptive learning system might already be beneficial for individuals with ADHD. This project proposes to implement two approaches: breaks and self-monitoring, in an adaptive learning system to improve academic performance and motivation for learners with ADHD symptoms. These approaches are discussed below.

2.2 Adaptive learning

Adaptive learning systems have the upper hand over their non-adaptive counterparts in the sense that they are able to modify their behavior to each individual as the individual continues learning (van der Velde, Sense, Borst, & van Rijn, 2021). The adaptive system by van Rijn et al. (2009) has shown better performance through scores than utilizing a non-adaptive flashcard method. The adaptive system employed in this project is build upon (Pavlik Jr. & Anderson, 2005). The adaptiveness in the MemoryLab system is in its behavior of which facts are presented and in what order, based on the estimated memory retention of each seen fact for each individual to optimize their learning (van Rijn et al., 2009; Sense, Behrens, Meijer, & van Rijn, 2016; van der Velde et al., 2021). The system thus adapts itself to the individual learner and to each individual fact for that learner.

2.2.1 MemoryLab's algorithm

In this project, the MemoryLab algorithm is employed, which takes retrieval time of an item and accuracy into account for each item and each individual. This behavioral feedback is then taken into account to calculate an individuals activation. In order to improve learning, the spacing and testing effect was taken into account (van Rijn et al., 2009). The testing effect, constitutes greater later retention when being tested on the subject and accurately recollecting or identifying it, in contrast with studying but not participating in testing (Roediger III & Karpicke, 2006). The spacing effect, as explained in van Rijn et al. (2009) entails repeated facts learned with more spacing around itself, with for instance other facts, results in an increased likelihood of recollection, in contrast with repeatedly learning or encountering one fact before continuing to the next fact. The algorithm utilized employs several equations, based on those from ACT-R's declarative memory Anderson (2007).
The equation calculated for each item for each individual in the MemoryLab algorithm, as further explained in van Rijn et al. (2009); Sense et al. (2016) calculates the activation per item i based on the ACT-R (spacing) model from Pavlik Jr. and Anderson (2005); Anderson (2007).

$$A_i(t,n) = \sum_{j=1}^{n} (t - t_j)^{-d_j} \tag{1}$$

The activation of the item is influenced by each time the item i has been seen before, where an occurrence j of the previously seen item in time t is represented by t_j (van Rijn et al., 2009). The parameter $-d_j$ indicates the decay of an item in memory, which is a computation of the item's activation during the j^{th} occurrence when the item was seen (van Rijn et al., 2009). This decay is calculated by taking into account the relative contribution of the activation of all previous occurrence where item i was seen, (except for the current occurrence of item i) set by parameter c, called the decay scale (van Rijn et al., 2009; Sense et al., 2016). The first occurrence of an item i makes use of intercept value α as the decay, alpha also holds the minimum decay for an item i (van Rijn et al., 2009; Sense et al., 2016).. This parameter can be calculated with the following formula:

$$d_{ji} = ce^{A_i(t_j)} + \alpha \tag{2}$$

Alpha (α) is calculated differently in the MemoryLab algorithm (see van Rijn et al. (2009); Sense et al. (2016). The MemoryLab algorithm takes response time into account by looking into the difference between the observed and estimated response times (Sense et al., 2016). In order to adapt the model to situations when the individual has a faster retrieval for an occurrence of a specific item, the alpha is modified based on a difference between the observed retrieval time and the estimated retrieval time (van Rijn et al., 2009). The alpha parameter is altered by use of binary search (Sense et al., 2016).
The expected retrieval time can be calculated with the use of activation at time t, as explained in van Rijn et al. (2009); Sense et al. (2016). The equation to calculate the expected or estimated retrieval time can be seen below, as seen in Sense et al. (2016).

$$RT_i(t) = Fe^{-A_i(t)} + \text{fixed time} \tag{3}$$

Whether an individual will remember and recall a specific item is calculated by the activation (van Rijn et al., 2009). In the case that the activation is larger than the retrieval threshold, then the individual will be able to recollect the item (van Rijn et al., 2009). Otherwise, if below the threshold, then the item is most likely forgotten. Then the item is repeated just before the activation would fall below threshold and the individual would forget.
Thus, the activation for each item for the individual user is calculated for a number of seconds in the future, if this activation is below the threshold, then the item should be seen again (Sense et al., 2016). If no activation of the previously seen items falls below threshold in the near future, then a new item is displayed or if no new items are left, then the item with the lowest activation will be displayed again (Sense et al., 2016). After each occurrence of an item, the activation value is recalculated and updated (Sense et al., 2016). The visual description of the algorithm can be seen in Figure 2 in Sense et al. (2016). Thus, this adaptive algorithm takes individuals response times and accuracy into account to calculated an activation for each item occurrence for each individual, making the system adaptive to the individual (van Rijn et al., 2009; Sense et al., 2016).

2.3 Approaches: accommodations and interventions

Wrong answers on an assessment does not necessarily mean insufficient mastery of the material (Ketterlin-Geller, 2008). The student might have the ”target skills”, which are the skills or learning related to that what is aimed and planned to be assessed (Elliott, Braden, & White, 2001; Ketterlin-Geller, 2008; Ketterlin-Geller, Jamgochian, Nelson-Walker, & Geller, 2012). According to Elliott et al. (2001) ”access skills” are needed for the ingress of the tested material and therefore for the display of one’s possible ”target skills”. Lovett and Fienup (2023) provides a very straightforward example: a (U.S.A.) history test measures ”target skills”, which is the knowledge about history material in the test, but in turn might also measure ”access skills”, such as being able to remain focused for a prolonged period to be able to display one’s comprehension of the history material. Some students might have a deficiency in relation to ”access skills”, this could prevent or block valid assessment of their ”target skills” in standard tests (e.g. Elliott et al. (2001); Ketterlin-Geller (2008); Ketterlin-Geller et al. (2012); Lovett and Fienup (2023)).
Both interventions and accommodations could aid individuals with ADHD to demonstrate their knowledge or skills. An accommodation could diminish deficiency (e.g. of not keeping focus) to prevent requirement for this ”access skill” (Lovett & Fienup, 2023), by for instance applying breaks. Accommodations are defined as alterations made to educational methods (such as taking a test) that gives the student an advantage to moderate the effect of the disability on the curriculum (Harrison, Bunford, Evans, & Owens, 2013). However, all students conform to the same test criteria, regardless of disability status (Harrison et al., 2013). An example of an accommodation in use for students with

ADHD in the Netherlands is taking their final high school exams with extra time, depending on their school (Ministerie van Onderwijs, Cultuur en Wetenschap, n.d.). Interventions, on the other hand, aid the student in constructing and enhancing this ”access skill” that allows for later utilization in exams or assessments instead of using accommodations during these assessments (Lovett & Fienup, 2023). An example could be be letting participants monitor their own attention to the task. Therefore both accommodations and interventions can be employed, but they aid in demonstrating one’s knowledge or skills, from different angles.

However, when an individual falls short in performing a skill, but does have the skill, accommodations reduce the need for that skill which in turn might impede the building of this skill (Lovett & Fienup, 2023). Lovett and Fienup (2023) reason about ”skill deficits” and ”performance deficits”, where the latter is different from the former in the sense that it is not a deficit of skill but of not performing that skill. Lovett and Fienup (2023) made suggestions than can be employed to limit the need for accommodations. One of the suggestions is the application of self-management instead of (solely) accommodating breaks, to improve sustained focus (Lovett & Fienup, 2023). Another could be to disregard when learners are off-task and to reinforce when learners are on-task (Lovett & Fienup, 2023). Thus, in the current learning system, reaction times of individuals with ADHD might not be an accurate reflection of their retention, as their ”access skills” might not accurately display their ”target skills”. Both accommodations and interventions could be employed to aid learners with ADHD, but they approach these skills in different ways. In order to compare both approaches, we will discuss breaks and self-management next, as investigated in Harrison et al. (2022).

2.3.1 Comparing approaches: prompts, breaks, sensory items and self-management

Harrison et al. (2022) investigated three accommodations and one intervention for middle schoolers with ADHD. These four approaches include breaks, sensory elements (e.g. fidget objects), prompts (to turn back to the task) and self-management. For the break condition, a five minute break occurred at the half-way point of the task duration and induced by the research assistant (Harrison et al., 2022). The break could either be the full duration, part of the duration or no duration. Even though research suggests accommodations such as breaks do not significantly increase performance in math or reading (Pritchard et al., 2016), it is recommended to educators (Children and Adults with Attention-Deficit/Hyperactivity Disorder (CHADD), n.d.).

In a different condition, participants could opt for an item of their own choosing to utilize, as part of sensory proprioception (Harrison et al., 2022). Prompting entailed that the experimenter prompted the student to turn their attention back on their homework, after the student was not on task for five seconds in this condition. Lastly, self-management was conducted by the student, where she/he recorded their (in)attention to the task on each five minute period. In this condition, Harrison et al. (2022) measured whether participants recorded at every signal, some of the signals or none. They measured homework engagement as behavior of participants on the task, interruptive behavior and time between instruction and start of the task for all conditions. Harrison et al. (2022) further took participants’ opinions on the accommodations and intervention into account, their behavioral responses in terms of frequency or duration of engagement in each condition and research assistants’ fidelity.

They found significant differences between prompting and sensory proprioception, where prompting led to a higher average percentage of engagement with the task and a diminished occurrence of disruptive behavior compared to sensory proprioception. The same holds for prompting versus breaks, but not for the percentage of engagement with the task, only for interruptive behavior. However, participants favored the use of sensory elements before self-management, breaks and lastly prompting. Perceived usefulness resulted in a different order compared to preferences in this study.

Prompting is first in perceived usefulness, then breaks, then the use of sensory elements and lastly self-management. In contrast, Harrison et al. (2022) found that prompting to turn back to the homework and self-management resulted in the best outcomes, though the first was found to be irritating by participants. Interestingly, Harrison et al. (2022) states that there does not seem to be a correlation between the preferences of participants and the performance of participants for these approaches. Harrison et al. (2022) state how imperative it is to observe their behavior when utilizing an approach for social validity and suggest using interventions such as self-management over accommodations. Harrison et al. (2022) discuss that self-management might no longer be needed for a student in the future, whereas prompting would always be needed (as an accommodation). This reflects back to the different skills needed for the history test in, as accommodations omit the need for the ”access skill” whereas intervention built this skill (Lovett & Fienup, 2023).

2.4 Accommodation: breaks

As we have seen above, research seems mostly pointing towards first applying interventions over accommodations. Specifically breaks results in varying outcomes in literature. Pritchard et al. (2016) tested the impact of various accommodations such as, but not limited to, extra time, allowed use of calculator and more breaks on math and reading. Math or reading for students did not significantly improve when accommodations were available to them (Pritchard et al., 2016). As we have seen above, the utilization of breaks for children with ADHD did not lead to the highest average percentage of engagement with the task, nor to less frequent disruptive behavior or less time to start homework in comparison to the other strategies tested in Harrison et al. (2022).

However, it might not be so straightforward. In their limitation, Pritchard et al. (2016) state that in the case that children with an ADHD accommodation did not receive an accommodation, they could have possibly scored worse on the tests than other children with ADHD without an accommodation, due to worse abilities with verbal or linguistic reasoning skills. Moreover, Pritchard et al. (2016) also state that there is no knowledge of whether the accommodations were in fact given and used. Unfortunately, nor are the specifics of the accommodation breaks mentioned in (Pritchard et al., 2016). As mentioned before, Harrison et al. (2022) applied a break, but only one large break in the middle of the task. Reasoning for the design choices, such as length, frequency and experimenter initiation, surrounding this accommodation is not mentioned. Although not a homework or fact-learning task, a study that did investigate the use of breaks in adults, while also looking into ADHD scores and use of break initiator, did show positive results (Hall-Ruiz, 2016). This study is discussed below.

2.4.1 Choice in breaks

In the third chapter of Hall-Ruiz (2016), they investigated the effect of breaks during a modified version of Connor’s Continuous Performance Test (CPT) (Conners et al., 2003). They investigated the application of breaks where participants could opt themselves whether, when and for how long they would take a break. In a second experiment, they compared breaks initiated by the participant and breaks initiated by the experimenter. The breaks initiated by the experimenter consisted of a maximum of four and were based on performance, due to participant’s strategies from the first experiment in this chapter (Hall-Ruiz, 2016). Two breaks were initiated based on the first correct trial in their corresponding blocks and two breaks were initiated based on the first incorrect or false alarm trial (Hall-Ruiz, 2016). In the first experiment of this chapter, they found that taking more than three breaks meant a reduced difference and a comparable performance between college students with (or previously with) ADHD and control students. Students with ADHD especially gained from taking

breaks (Hall-Ruiz, 2016). More specifically, those students with ADHD that took fewer than three pauses performed worse by significantly making errors faster after a pause (Hall-Ruiz, 2016). They also found an extended durational effect on performance when breaks taken were longer in duration for students with ADHD.
Interestingly, they did not find a significant difference between the student initiated and the researcher initiated breaks. On the other hand, Hall-Ruiz (2016) state that they have a smaller sample size, which results in not being able to infer very strongly whether the initiated breaks by the experimenter are as advantageous as self-initiated breaks. On the other hand, there appeared to have a remainder of choice in breaks that were not initiated by the participant. Even though the first ten seconds of the breaks were mandatory, hereafter the participant could choose whether they wanted to take a longer break or continue with the task (Hall-Ruiz, 2016). We could be argue that the small quantity of ten seconds for a mandatory break and the choice to either take a longer break or continue with the task, might be less frustrating than a mandatory long break. Moreover, Hall-Ruiz (2016) state that fifty percent of their participants with ADHD preferred pauses initiated by the researcher. However, including those with ADHD and those without, most preferred self-initiated breaks. A reason for this was the control. It thus appears that experimenter imposed breaks can be utilized to keep breaks structured. However, some might prefer self-initiated breaks. We could therefore opt to omit a mandatory duration which might give the learner a little more control and might be less frustrating for participants who would prefer to continue with the task.

2.4.2 Motivation

A literature review about the consequences of choice for motivation and the elements which causes choice to be favorable (in class) (Katz & Assor, 2007), argue that it is not merely choice in itself that is motivating, but that it leads to increased motivation when it meets the learners needs based on self-determination theory (self-determination theory see e.g. Deci and Ryan (2000)). Reeve, Nix, and Hamm (2003) found that volition and internal locus, which are other parts of self-determination theory, increased intrinsic motivation more than choice. Giving participants power over time allocation within the task, such as stopping the break early or continuing with the break (e.g. choice over break duration in Hall-Ruiz (2016)), could perhaps induce volition in the sense that participants are free to opt to stop or continue. Alternatively, volition might perhaps not be as limited with a break taken in part, fully or clicked away, instead of a break that has a mandatory duration. In the third study in Reeve et al. (2003), they compared two forms of choices. In one form, the participant could choose the puzzle, but not when and for how long. In their other choice condition, the participant could choose the puzzle, make an estimate how long they wanted to puzzle for and after this estimate had expired, choose to continue with the current puzzle or with a different puzzle. Reeve et al. (2003) found that intrinsic motivation was highest for those that had choice over when to do which puzzle than only which puzzle to work on or working on a puzzle chosen by the researcher.

2.4.3 Probing thoughts

In a different chapter in Hall-Ruiz (2016), they were probing participants about their thoughts at random trials. The participants could choose from various answers on thoughts about: the task itself, their performance on the task, latest or future tasks or happenings in their life, their (physical) state, angst or worriment, fantasizing or imagining and lastly an option answer of thinking of other matters (Hall-Ruiz, 2016). In their second experiment with probing, they reduced the answer alternatives to three and gave participants a questionnaire. This questionnaire revealed that the majority of participants perceived probing positively. Their answers were categorized into thoughts related to the task and thoughts not related to the task in both experiments. Hall-Ruiz (2016) found that thoughts not related to the inhibition task occurred more often before a mistake than thoughts that were related to the task.

2.5 Self-regulation interventions

A meta-analysis by Reid et al. (2005) investigated four types of self-regulation on behavior (attention and interruptive or (in)appropriate) and performance (correctness or completeness) for minors with ADHD. Self-regulation includes various techniques in order to control, track and document or evaluate, or both, their educational performance or their way of behaving (Reid et al., 2005). These four types of self-regulation discussed in Reid et al. (2005) are self-monitoring, self-monitoring and reinforcement, self-reinforcement and self-management. They further mention two types of self-monitoring: self-monitoring of attention and of performance. In this project, with a paced learning system, the focus will lie on monitoring ones attention. Self-monitoring of attention is utilized to increase consciousness of the student's attention and usually entails the student evaluating their own on-task attention, after a designated signal (Reid et al., 2005). Self-monitoring with the addition of reinforcement could also be opted for. When the student alters their behavior, they will be reinforced (rewarded) by someone or something else, besides the application of self-monitoring (Reid et al., 2005).

A third type of self-regulation is self-reinforcement. Self-reinforcement is applied by students themselves with their chosen reinforcement, after meeting or surpassing the performance requirements (Graham, Harris, & Reid, 1992). The fourth self-regulation type is self-management. The distinction between self-monitoring and self-management is that the latter employs comparisons of the assessment of the student with another surveyor, instead of only self-evaluating. In addition, reinforcements are given when students' self-assessment lies close with the surveyors' assessment (Reid et al., 2005). As a result, self-monitoring with reinforcement was found to have a positive impact on both on-task and disruptive/improper behavior and on academic performance, whereas self-reinforcement also had a positive impact on the latter but seemingly not on behavior (Reid et al., 2005). There was a positive impact on these behaviors when self-management was utilized, but this was not investigated in any studies on academic performance (Reid et al., 2005). Additionally, almost half of the participants were utilizing medication in the time of the intervention experiments (Reid et al., 2005). The results indicate that behavior and academic performance can be enhanced for young (medicated) students by utilizing self-regulation strategies (Reid et al., 2005).

2.5.1 Self-monitoring

Mathes and Bender (1997) investigated the use of self-monitoring in order to increase on-task behavior, of three boys with ADHD and comorbid disorders, whom were also utilizing medication. The cues for the students to self-assess and self-record where tones from a piano recorded by the teacher with varying, average 45 second, intervals. The participants self-assessed through evaluating and recording their on/off-task behavior. This behavior was observed, besides their students' self-assessment frequency of their on-task behavior. Students were observed during their 10-minute sessions. These sessions were conducted over multitude of days where different phases were employed to measure and improve behavior. Fading was employed by eliminating signals to self-monitor and then eliminating the paper on which to self-record. Instead the participants were instructed to reinforce themselves through shoulder pats and positive phrasing, when their answer was ”yes” to the thought-about self-monitoring question of on-task attention. The participants were prompted to self-monitor by the teacher in the second phase of fading. The self-monitoring intervention employing in Mathes and Bender (1997) enhanced on-task behavior for each medicated participant and stayed elevated throughout the phases with fading. Mathes and Bender (1997) state that both participants and their teachers found betterment in on-task behavior.

Wills and Mason (2014) found an increase in on-task behavior for two high school students self-monitoring via the tablet application I-Connect. Every 5 minutes, the participants were cued by flashing screens to respond within six seconds to the binary question whether they were on-task. The sessions were in 55 minute classes and included a repetition of both baseline and intervention phases in a multitude of sessions (Wills & Mason, 2014). The duration the participants were on the task was also observed, besides disruptive behavior frequency (Wills & Mason, 2014). For both students, on-task duration increased with the use of the intervention, which also seems the case for reducing disruption, but the authors also state this frequency of disruption is more varying.

Vogelgesang, Bruhn, Coghill-Behrends, Kern, and Troughton (2016) employed the application SCORE IT and examined the educational engagement on three fifth-grade participants. These participants with ADHD diagnoses or scores indicated being at risk for ADHD, often showed off-task behavior. Beeps at each 10-minute interval, signalled to the participants to go to the tablet and self-record their behavior on three class appropriate behavior expectancies with five scaled response options (Vogelgesang et al., 2016). Their teacher then did the same for each participant. Vogelgesang et al. (2016) found positive impacts on the behavior of participants, such as paying attention, displaying appropriate behavior, and adhering to rules, even after different phases where SCORE IT was not or less utilized (Vogelgesang et al., 2016). Blood et al. (2011) investigated the use of both self-monitoring and skill modeling via video on a boy with ADHD and comorbidities, on disruptive and on-task behavior. Two-minute intervals of self-monitoring in addition to watching the modeling video before the session resulted in a positive impact on on-task behavior and reduced disruptive behavior (Blood et al., 2011).

A study that investigated self-monitoring on college students with ADHD, found that those that additionally received the self-monitoring intervention, besides other strategies, showed significant increased positive results on reading, taking tests and inattention (Scheithauer & Kelley, 2017). Each individual participant in their study self-monitored on specific behaviors, tailored to their objectives and agenda. Except for monitoring medication and showing up to class, which all participants in this condition self-monitored (Scheithauer & Kelley, 2017). This indicates that self-monitoring not only seems to work on children, but also on (young) adults.

Self-monitoring attention versus performance

Harris et al. (2005) compared self-monitoring of attention versus of performance on on-task behavior and the number of correct words spelled, in children with ADHD receiving medication. In the self-monitoring of attention condition, the child would hear a tone at random intervals between ten and ninety seconds where they had to think about and record whether their attention was on the practice spelling task. Afterwards, they had to graph to amount of ”yes” answers. Each session had a duration of 15 minutes, where the last 10 minutes were used for observations of the participants. Self-monitoring of performance entailed that the participants would graph the number of accurate spelled words after each session (Harris et al., 2005). Both types of self-monitoring increased participants' on-task behavior and accurately practiced words, but self-monitoring of attention led to a higher number of accurately practiced words (Harris et al., 2005). On the other hand, monitoring of performance was favored by the majority of the six participants instead of monitoring their attention (Harris et al., 2005). Even though their sample size was small, this indicates that self-monitoring of attention might be the most optimal form of self-monitoring to increase performance of learners with ADHD.

Breaks and self-monitoring in an adaptive learning system

Our aim is to investigate both the break accommodation and the self-monitoring intervention in a adaptive learning system for adults with ADHD (symptoms). Break are recommended for children and adults with ADHD (e.g. (Children and Adults with Attention-Deficit/Hyperactivity Disorder (CHADD), n.d.; McKnight, 2010; Adamou et al., 2021)) and show positive results for adults with ADHD in Hall-Ruiz (2016). Thus, we will employ experimenter-imposed breaks as this has shown positive results in Hall-Ruiz (2016). Choice over time to continue working on a puzzle or go to a next one could lead to feeling intrinsic motivation, whereas solely having the choice of which puzzle to solve without a choice of how long does not make a much impact in self-determination, according to Reeve et al. (2003). The experiment-imposed breaks in Hall-Ruiz (2016) were based on performance. We impose breaks at set interval to prevent punishing of those that perform well by not presenting a break and/or to prevent punishing of those that do not perform well by not presenting a break.
Research has shown self-regulation to be a beneficial strategy to employ for children with ADHD for on-task behavior, but also academic performance (e.g. see (Harrison et al., 2022; Harris et al., 2005; Mathes & Bender, 1997; Reid et al., 2005; Gaastra, Groen, Tucha, & Tucha, 2016)). Additionally, Reid et al. (2005) has shown that self-monitoring with the addition of rewards had a positive impact on all measures (attention to task, disruptive behavior and performance). The monitoring sheet in (Blood et al., 2011) made use of smileys, which we applied as a form of reinforcement. We only made use of the positive smileys, to keep positive reinforcement. Even though these were tested on young students, not on older students (Reid et al., 2005; Blood et al., 2011), self-monitoring also has a positive impact on adults with ADHD (Scheithauer & Kelley, 2017). We could not find as much research for adults with ADHD (symptoms) as for children. We will employ self-monitoring of attention and breaks to investigate their impact on a learning task for adult learners with and without ADHD characteristics.

3 Methods

We assigned participants to learn facts in a computerized and paced flash-card setup. More specially, we asked participants to memorize national flags in Dutch. The ordering of facts was determined by MemoryLab software. We investigated the effect of breaks and self-monitoring compared to uninterrupted learning, while also taking ADHD characteristics into account. In order to test this effect, we analyzed accuracy, reaction time and interview responses as dependent variables.

3.1 The learning environment

The experiment utilized MemoryLab's software. Their algorithm takes reaction time and accuracy into account and calculates the retention time or *speed of forgetting* of that item, for the individual learner (van Rijn et al., 2009; MemoryLab, n.d.; Sense et al., 2016). An item's activation for the individual learner is calculated a number of seconds seconds in the future (Sense et al., 2016). In this project the items where flags of nations. The algorithm then checks whether the activation of any of the previously seen flags falls below the threshold at the specified time in the future (Sense et al., 2016). If that is the case, then the "almost forgotten" flag will be repeated again immediately and their speed of forgetting and activation recalculated and updated (Sense et al., 2016). If no previous seen flags have an activation below the threshold, then a new flag is introduced (Sense et al., 2016). If no new flags are left, the flag with the lowest activation is repeated (Sense et al., 2016).

3.2 Participants

The participants were 27 native or fluent Dutch speakers over the age of 18, recruited through word-of-mouth, personal social networks, social media and from the University of Groningen through posters and slides. Participants' mean age was 24.4 years (SD = 7.9). Fifteen participants were women (55.56%) and twelve were men (44.44%). Three participants (11.11%) self-reported to have an ADHD diagnosis. Two participants with an ADHD diagnosis received self-monitoring, and the third received breaks. The compensation for participants was eight euros. A letter of no objection was obtained from the ethics committee at the Faculty of Science and Engineering.

	Self-monitoring (n=14)	Breaks (n=13)
Adulthood inattention	3.4 (SD = 3.3)	1.8 (SD = 2.7)
Adulthood hyperactivity/impulsivity	2.8 (SD = 2.3)	2.4 (SD = 1.8)
Childhood inattention	3.4 (SD = 3.0)	2.2 (SD = 3.3)
Childhood hyperactivity/impulsivity	2.5 (SD = 2.6)	1.8 (SD = 2.0)

Table 1: Rounded means and standard deviations of participants characteristics in each group. C stands for control, which means a learning session with uninterrupted learning. S stands for self-monitoring, which means a learning session with self-monitoring questions. B stands for breaks, which means a learning session with breaks of optional duration. The order of the letters (i.e. C-S is the order of the learning sessions. The scores could range from zero to nine.

The visual distribution of scores in adulthood can be found in Figure 1.

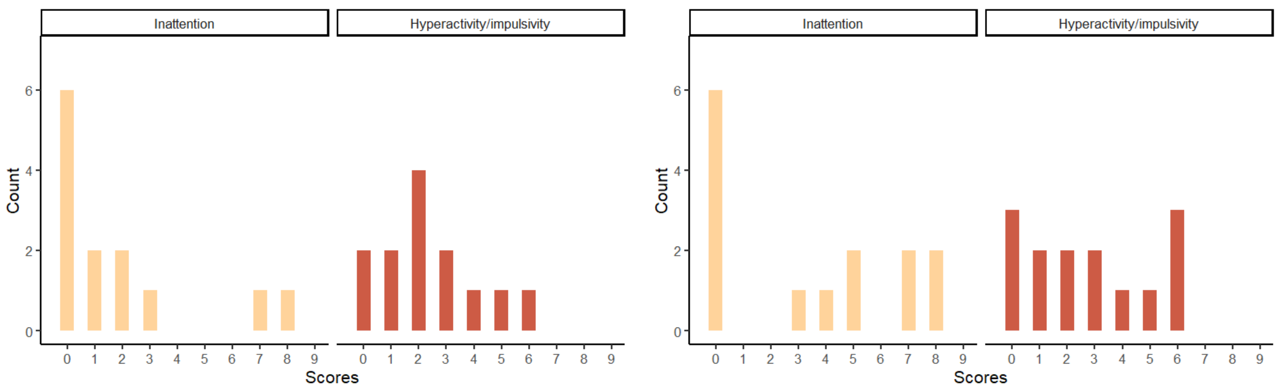
3.3 Tools

The experiment was delivered via a web interface, where front-end and back-end were modified through the basis of the API guide of MemoryLab. JavaScript, Typescript, HTML and CCS were utilized to create and adapt the front-end and adapt the back-end. The order and timing of the facts was determined by the MemoryLab algorithm through their API. Figure 2 shows a practice trial before and after a response has been made.

3.3.1 Questionnaire

Besides asking participants if they have an ADHD diagnosis, a screening questionnaire measuring ADHD symptoms was included. We utilized the Dutch version of the Self-Report Questionnaire on Attention Problems and Hyperactivity for Adulthood and Childhood (Kooij and Buitelaar, 1997 in Kooij (2009)). The questionnaire consists of 23 statements about symptoms related to inattention and hyperactivity/impulsivity during adulthood and the same 23 statements about the childhood of the participant, based on the 18 statements of DuPaul et al. (1998) as cited in Kooij et al. (2005). Within these two parts, 11 questions are about inattentiveness characteristics and 12 about hyperactivity/impulsivity characteristics (Kooij, 2009, 2022; Kooij et al., 2005). These questions are reduced to 9 criteria for inattention and 9 for hyperactivity/impulsivity (Kooij, 2009, 2022). For inattention, there are two criteria that each correspond to two questions and for hyperactivity, three criteria each correspond to two questions, which results in nine criteria for inattention and nine criteria for hyperactivity/impulsivity.

The participant could opt for one of four answers ranging from never or rarely to very often (Kooij et al., 2005; Kooij, 2009, 2022). These answers are indicated with 0 (never) to 3 (very often). In order to get a score, each of the 18 criteria is taken into account, if the participant responded with ”often” (2) or ”very often” (3). This means that scores range from zero to nine characteristics for inattention and for hyperactivity in both adulthood and childhood. We will only use these characteristic screening scores, without interpretation. The questionnaire was created and administered through Qualtrics (`www.qualtrics.com`), wherein data was collected as well. The distributions of participants and their scores in adulthood are illustrated in Figure 1a for participants in the breaks condition and in Figure 1b for participants prompted to self-monitor.



(a) Score distribution breaks condition.

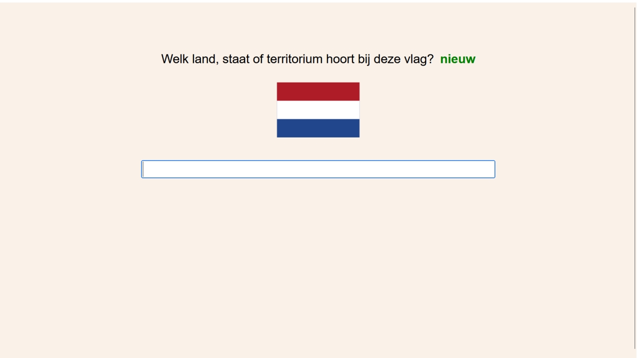
(b) Score distribution self-monitoring condition.

Figure 1: Score distribution for the participants, where each figure represents either the condition with breaks or self-monitoring prompts. The left subfigure denotes the scores for inattention and the right subfigure denotes the scores for hyperactivity/impulsivity.

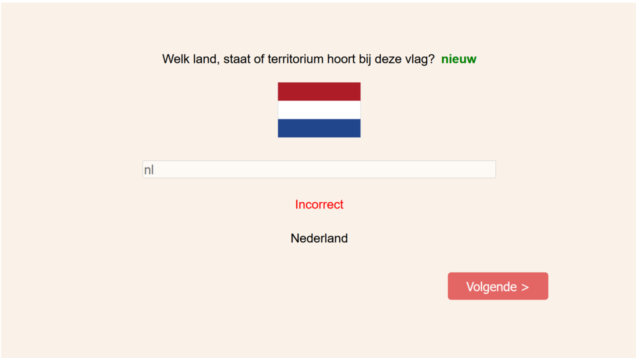
Instead of solely a binary metric on ADHD, the questionnaire provides us with a more continuous scale of ADHD characteristics. We can thus measure a more continuous effect of ADHD with or without an intervention or accommodation applied in learning. Another reason for utilizing the questionnaire is to enlarge the subject sample, as an ADHD diagnosis would no longer be a requirement to participate. Nor do we interpret these scores.

3.3.2 Stimuli

The to-be-learned facts were flags and the name of the country, state or territory. The flags and their names were originally in English and sourced through Wikipedia and their list of flags (see: https://en.wikipedia.org/wiki/List_of_national_flags_of_sovereign_states). These names were then translated to Dutch with the aid of the Microsoft Excel Translator (Microsoft Excel version 2502 and/or 2503), and then checked with various Wikipedia sites. Stimuli comprised of flags from all over the world. The flag of the Netherlands, the United States of America, France and Canada were visible in the practice trials of the experiments and were thus omitted from the learning sessions. Figure 2 shows a practice trial before responding in Figure 2a and after responding incorrectly in Figure 2b. These four flags were chosen for their difficulty, we assume that most Dutch speaking participants would know or recognize these flags. We also removed flags that were similar to a different flag or those that had their name present on the flag. Names with accents that might be difficult to type in a learning environment with only a keyboard were also removed (for example Åland). One flag was removed based on its size, which overlapped with the inputfield on the experiment page.



(a) Before responding.



(b) After responding.

Figure 2: Practice trial screenshots.

Furthermore, we omitted names that were longer than twenty characters, including whitespace, which might be harder to type correctly and thus more difficult than shorter names. For instance, *South Georgia and the South Sandwich Islands* versus *Spain*. To account for this difference in difficulty, we omitted these flags from the stimuli list. The parameter of twenty characters was chosen as this would take out the very long names, but at the same time limit the impact on the stimuli list, as these very long names appeared very infrequent. Spelling was taken in to account, including hyphenation, except for words with accents and capital letters. In the case that accents or capital letters were omitted from an answer then the answer was still counted as correct. However, when accents were placed on wrong letters then the answer would be marked as incorrect.

3.4 Experimental setup

Almost all, (92.6%) of participants completed the task in a lab, the remaining 7.4%, completed the task in a seperate location. Experimenter was present for both in-lab and at location. Participants in the lab made use of a mouse and if in-lab a separate keyboard. Participants were in a separate cubicle away from the experimenter in the lab. The monitor lab is a 24 inch BenQ XL2420T with a 1920x1080 resolution. The experiment was run on the Windows operating system, version 10 Pro Education. The experiments conducted in other locations used a laptop and mouse. A laptop with Windows 11 Home was utilized with Mozilla Firefox 138.0.4 and a 15.6 inch screen size.

3.5 Procedure

Participants were assigned to one of the groups based on the order of coming in to keep the number of participants in each condition as balanced as possible. If known beforehand, participants with an ADHD diagnosis were spread over different groups.

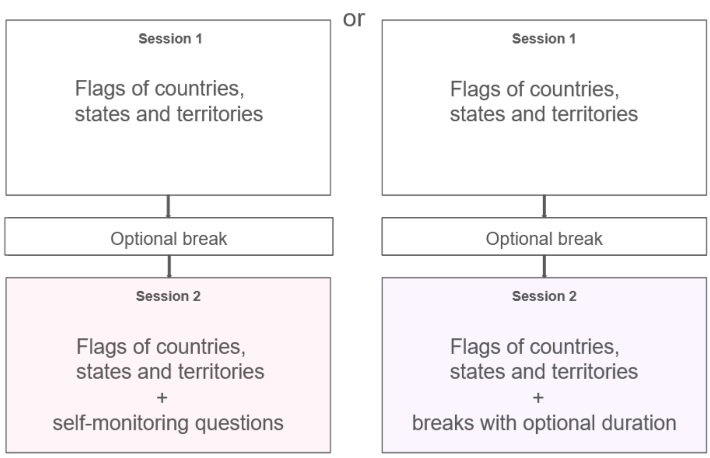


Figure 3: The design, where both approaches are visible in the second session. We counterbalanced this, which meant that either approach could also occur in the first session and the uninterrupted learning in the second session. This then entails a total of four groups.

Both between- and within-subject design was utilized. We investigated the effect of breaks or self-monitoring versus uninterrupted learning, which was a within-subject design and a between-subject design. A visualization of the conditions and its order is shown in Figure 3. The order of the learning sessions was counterbalanced, meaning that a participant could be assigned in one of four groups. The participants characteristics per group can be seen in Table 1. The experiment was set-up on the computer in the lab at the university building, linked to a second monitor. The participant was introduced to what they were to learn during the task (i.e. flags), that there was to be a questionnaire and interview. Hereafter when the information letter was read and the consent form was signed, the participant received an explanation of the task and the procedure. Then the participant could start with the learning sessions. Before each session, more detailed explanation was visible on the screen which the participant had to click through in order to start the practice trials and hereafter start the actual session.
Two modifications were made: the use of breaks with an optional duration and the use of self-monitoring. These modifications are visible in Figure 4. The self-monitoring page as seen in Figure

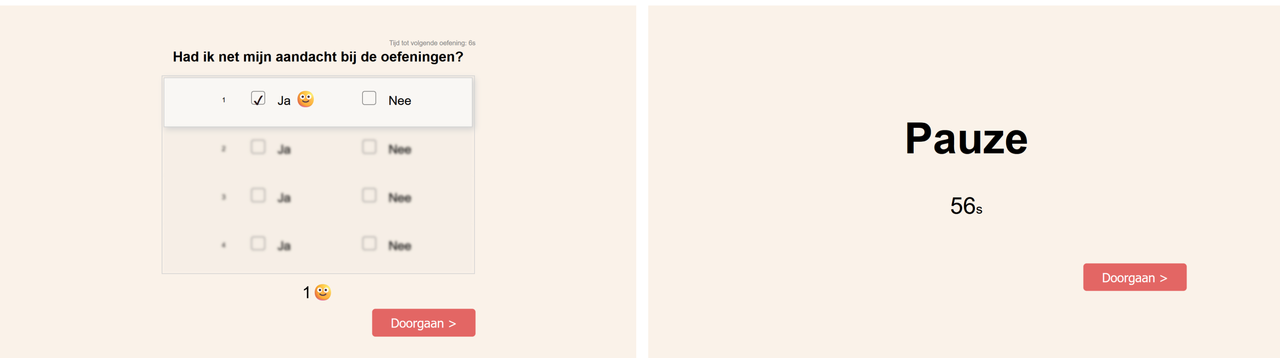
4a was based upon the sheet in Blood et al. (2011), except that only an emoticon was utilized when participants clicked ”yes”. A learning session could consists of uninterrupted learning, of learning with self-monitoring after each 1.5 minute or of learning with an optional duration of break after each 1.5 minute. The maximum amount of time for a break is 60 second and the maximum amount of time to answer the self-monitoring of attention question is 15 seconds. Each participants performed two learning sessions with an optional small break in between.

Each learning sessions took at least 8 minutes, but depended on the choices of the participants. For example, a participant could opt to immediately continue when being directed towards a break or wait in part of the full 60 seconds for the next trial. The same is the case for self-monitoring, a participant could answer the question whether he or she was on the task immediately and then click continue to go to the next trial or wait the maximum of fifteen seconds to automatically continue to the next trial.

In both interventions, a countdown timer was present on the page. These timers are visible in Figure 4. In the self-monitoring (of attention) intervention, the participants also received an explanation on the screen what *not* paying attention entailed for this task. In order to pace the experiment, a maximum of 30 seconds was available to enter the answer of the flag and a maximum of 30 seconds to look at the correct answer (after they had given an incorrect answer). Participants were not told of these limits nor was there a timer present on the page, to prevent pressure.

In the case the answer was correct, the word correct would appear for one second and then the participants were automatically directed to the next trial. In the case that the answer was incorrect, the word incorrect and the correct answer would appear for a maximum of 30 seconds. Participants could then either press ’Enter’ or click the button to continue to the next trial.

At the end of the first learning session, participants had time to take a small break. At the end of the second learning sessions the screens indicated the end of the sessions and a button that redirected the participant to the questionnaire. When participants were finished with the questionnaire, a verbal semi-structured interview was held. This interview consisted of ten questions where participants were asked their opinions, preferences and experiences of the interface and the sessions with its interventions. No audio recordings were taken, interview responses were transcribed in Microsoft Excel Version 2504 and 2505 while participants answered. The interview questions for both groups can be seen in Appendix A.



(a) Self-monitoring. (b) Break.

Figure 4: The intervention and accommodation during practice.

3.5.1 Self-monitoring of attention

In the self-monitoring condition participants were given instructions what not paying attention entailed. These were based on the unrelated to task options of the question what participants were thinking (on p. 23 of Hall-Ruiz (2016)). We defined not paying attention in this task as being occupied (i.e. doing something else) or thinking about anything other than the task. Where we mentioned examples such as daydreaming, looking out the window, thinking about what you will do tomorrow or what you will eat tonight or how you feel. Which we summarize as thoughts unrelated to the task. As mentioned above, those in the self-monitoring condition had fifteen seconds to answer the question whether they were paying attention. The English translated question was: ”Were you just paying attention to the exercises?”, based on the previous 90 seconds of learning flags. When the participant answered, a button appeared which when clicked would take them to the next trial. If the participant answered ”yes” a happy smiley would appear alongside the answer, based on the self-monitoring sheet of (Blood et al., 2011). Except, we did not make use of a sad or negative smiley when answered ”no”, to only keep positive reinforcement. This learning session consisted of four self-monitoring questions, which meant that regardless of speed of answering this question, each participant completed the same number of self-monitoring questions.

3.5.2 Breaks with choice of duration

Each (optional) break consisted of a maximum of 60 seconds, where a participant could continue to the next trial at any moment. Five seconds before the end of the break, the break text would change to: ”Break almost over” in Dutch. After the maximum amount of time had passed and participant had not yet pressed the button to go to the next trial, the page would automatically go to the next trial. This learning session consisted of four breaks, in order to compare this intervention with the self-monitoring intervention. Each participant in this condition was presented with four (optional) breaks, regardless of time taken in previous breaks.

3.6 Data analysis

Behavioral data and questionnaire responses were preprocessed and the interview responses were analyzed in Rstudio version (Version 2024.12.1+563) (Posit team, 2025) and R version 4.5.0 (R Core Team, 2025). Interview responses and annotations of these responses were combined with the csv module and the os module in Python version 3.7.6 in Microsoft Visual Studio Code (VS code) Version 1.87.1 (Microsoft Corporation, 2024)) to edit and run the code. Then the behavioral and questionnaire data was further analyzed with Rstudio (Posit team, 2025) (Version 2025.05.0+496) and R version 4.4.2 (R Core Team, 2024). Data was further processed using the dplyr R package version 1.1.4 (Wickham, François, Henry, Müller, & Vaughan, 2023), the stringr R package version 1.5.1 (Wickham, 2023) and the tidyr R package version 1.3.1 (Wickham, Vaughan, & Girlich, 2024). Plots were created with the ggplot2 R package version 3.5.1 (Wickham, 2016). For statistical analysis the lme4 R package version 1.1.36 (Bates, Mächler, Bolker, & Walker, 2015) was utilized.

3.6.1 Cleaning

Two participants received a bug during a break in the experiment. We opted to remove one break of one participant as this break duration was not in line with their other break durations. Therefore, this break was omitted for analysis, as it is not representative for this participant. The break of the other participant was not omitted, as this seemed in line with their other break durations. Furthermore, three outliers were omitted from the data that had reaction times below 50 milliseconds. Two other external events happened. One participant's session was truncated due to a press on a wrong button, as is further explained in the limitation. This participant was not omitted from the analysis. Another participant mention in the interview they felt distracted by a mosquito coming in the room and stinging, which is a confound. However, their reaction time boxplot and their accuracy averages do not appear to be at odds with other participants nor did the participant report to be otherwise distracted. Therefore this participant was in turn also not omitted from the analysis. The descriptive plots in the results show the data with outliers removed.

3.6.2 Analysis of complete sessions

We utilized a linear mixed effects model to investigate the effect on log-transformed reaction time, as participants performed both a control condition and a condition with either a break or self-monitoring prompts. Reaction time was log-transformed, as this distribution was right-skewed and only contained positive values. We opted for random intercepts for each level of session over random intercept participant and a random intercept for facts. We opted for random intercepts of participant, as there is variance between participants that we do not aim to predict. Furthermore, we opted for a random intercept of fact, because one fact might be slightly more difficult or easier than another. For example, *Zwitserland* (Switzerland) had a seemingly lower difficulty than *Chagosarchipel* (Chagos Archipelago). Session as a random effect was chosen as this explained more correlation than additionally including an adjustment for session as well. Other possible variations, such as random intercepts for each level of condition over random intercept participant explained less correlation/variation. A random intercept of different sessions or conditions over fact ids led to singularity or convergence errors.

We took condition, adulthood inattention score or adulthood hyperactivity/impulsivity score into account as main effects. Only adulthood scores were taken into account as we are most interested in a learners current characteristics and childhood scores increased complexity of the model. Spearman's correlation was utilized to test the correlation between the adulthood characteristics scores. Inattention and hyperactivity/impulsivity characteristics score in adulthood turned out to be highly correlated ($\rho = 0.79$, n = 27, $p < .001$), but due to ties exact p-values could not be calculated. We opted not to combine these scores as both scores related to different ADHD criteria. We instead utilized a model for each ADHD characteristic score. We further took the interactions of both adulthood scores with condition into account. Additionally, we employed a generalized linear mixed effect model to investigate the effect on accuracy as this has a binomial distribution.

3.6.3 Analysis of effects between breaks or self-monitoring prompts

Both models were further utilized to investigate the effect of the condition in the seconds after the either the accommodation or intervention, through binning the trials. The number of bins (four) represents the minimum number of trials performed in a block before approach occurrences. The first bin consists of trials until 22.5 seconds after the occurrence. These trials were averaged in mean for accuracy and median for reaction time. Then the next bin contains the next 22.5 seconds, until the next approach is recorded. This entails four bins of 22.5 seconds each, except for the last bin. The last bin consists of the trials after 67.5 seconds and before the next break or self-monitoring prompt. The same bins are calculated for the uninterrupted learning session, where four bins of 22.5 second sum up to 90 seconds per block. The 90 second mark would then resemble an intervention occurrence. Even though everyone at least performed four trials per block (time between each breaks or self-monitoring prompt), there are cases of participants not performing a trial within each 22.5 seconds. The maximum number of time a participant could take on a trial is approximately 30 seconds, therefore some participants are not represented in every bin in every condition. This is visualized by the size of the datapoints.

We employed a linear mixed effects model with bin, condition and either adulthood inattention or hyperactivity/impulsivity scores and the interactions between bin with score and condition and condition with score on log-transformed linear mixed effects. We also aimed to investigate these predictors on accuracy in a log regression model, but this led to convergence errors. In order to investigate the effect of bin and condition on log-transformed reaction time on its own, we also analyzed a model without ADHD characteristics scores. A general linear mixed effects model with family as binomial and with condition, bin and their interaction was employed to investigate its effect on accuracy.

3.6.4 Analysis of effects immediately after an interruption

In order to compare the effect of uninterrupted learning, breaks and self-monitoring immediately after a break or self-monitoring prompt, we analyzed the effect in of the conditions with ADHD characteristics scores in the 22.5 seconds after a break, self-monitoring prompt or in the case of uninterrupted learning after each 90 seconds. Only the trials in the first 22.5 seconds were taken into account in addition to the trials for each 22.5 seconds after a break, self-monitoring prompt or the first 22.5 seconds after 90 seconds has passed in the control condition. A linear mixed effects regression was utilized for log-transformed reaction time with this dataset. Condition, inattention or hyperactivity/impulsivity characteristic score and the interaction with condition and score were utilized as fixed effects. A random intercept for fact id and a random intercept for participants was employed. No other random intercepts or slopes were utilized as these led to converge or singular errors in both the model with inattention characteristic scores and the model with hyperactivity characteristics scores. For accuracy, we utilized logistic regression again. A general linear mixed effects model was employed with the same fixed effects and random effects, as additive random slopes or intercepts resulted in errors.

3.6.5 Analysis of uninterrupted versus interrupted sessions

Both break and self-monitoring prompts give participants a small break or interruption in their task. It is therefore also interesting to analyze the effect of an interruption alone versus uninterrupted learning. We therefore combined the break and self-monitoring condition into one condition called treatment. Again a linear mixed effects model was utilized to analyze the effect of condition (uninterrupted or interrupted) and inattention or hyperactivity/impulsivity characteristics score and the interaction of score and condition on log-transformed reaction time. The random effects structure entailed a random intercept of session over the random intercept participant and a separate random intercept of fact id. The linear mixed effects formula for inattention and hyperactivity/impulsivity characteristics score is the same, apart for their respective ADHD characteristics scores.

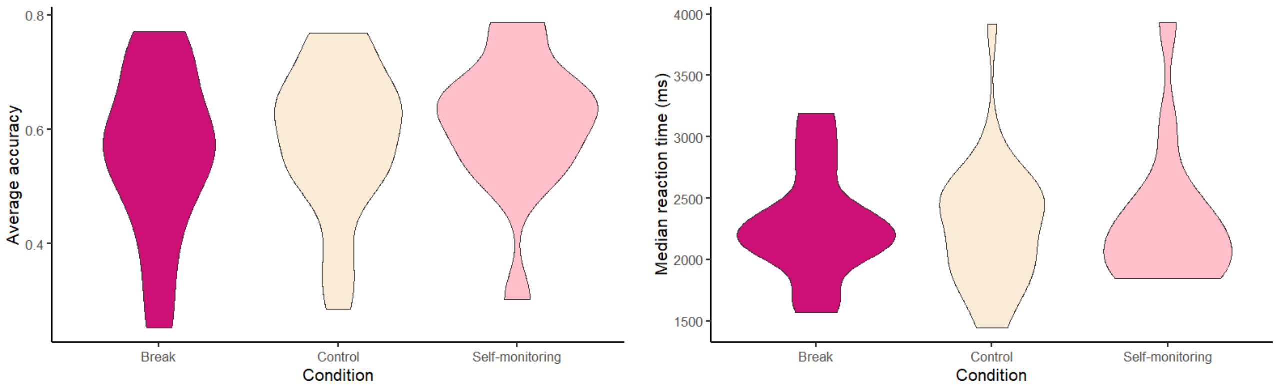
We further made use of a general linear mixed effects model. The effect of condition, inattention and this interaction on accuracy were investigated. The random effects structure consisted of a random intercept of participants and a random intercept for fact id. Other structures returned errors. The effect of condition, hyperactivity and this interaction on accuracy with this random effects strucuture was in turn also analyzed.

3.6.6 Interview analysis

Hereafter, the experiences of participants through a semi-structured interview were also investigated. The responses were annotated were relevant and possible and other responses were investigated through various themes of questions and responses and discussed where relevant. Responses of two annotated questions where compared between choices and their inattention and hyperactivity/impulsivity characteristic scores. The Kruskal-Wallis H test (Kruskal & Wallis, 1952) was employed for those three different groups of responses. When only two different annotated labels resulted from the responses, a Wilcoxon Rank-Sum test (Wilcoxon, 1945) was utilized.

4 Results

This project aimed to investigate whether there was an effect of applying breaks or self-monitoring on performance and experience for learners with ADHD characteristics. Performance was measured in three conditions, the distribution in median reaction times and accuracies over condition are visible in Figure 5.

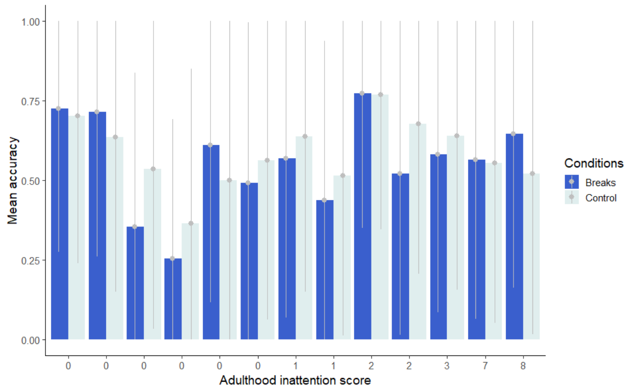


(a) Accuracy. (b) Reaction times

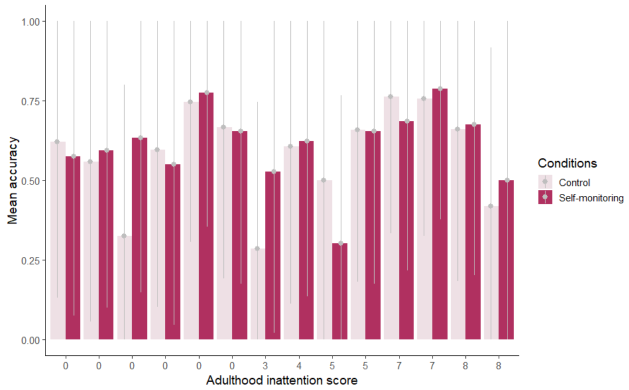
Figure 5: The variance in median reaction times and accuracies of participants in each condition. The control condition consists of both participants from the group that received breaks and the participants that received self-monitoring prompts.

Figure 5a shows that the participants in the self-monitoring condition seem to have more clustering around higher accuracy than the other two conditions. Especially the breaks condition seems to extend to lower accuracies, otherwise the three distributions appear quite similar. The average accuracy of the mean accuracy of each participant in the breaks condition is 0.56 (SD = 0.15). This average accuracy in the self-monitoring condition is 0.61 (SD = 0.12) and the average accuracy of the mean accuracy of each participant from the control condition is 0.58 (SD = 0.13). The distributions appear to be a little more different for median reaction time as seen in Figure 5b. There seems to be a large part of the participants in the breaks condition that have median reaction times of around 2200 milliseconds, whereas the control condition is more dispersed from faster reaction times to slower reaction times. The self-monitoring session also indicates more clustering of participants around faster reaction times, but these seem to start at a slower reaction time than the breaks session and is a bit more dispersed to slower reaction times. The average reaction time of each median reaction time of participants in the break condition is 2298.19 milliseconds (SD = 447.50). This average is 2322.14 milliseconds (SD = 577.03) in the self-monitoring condition and 2277.65 milliseconds (SD = 512.25) in the control condition.

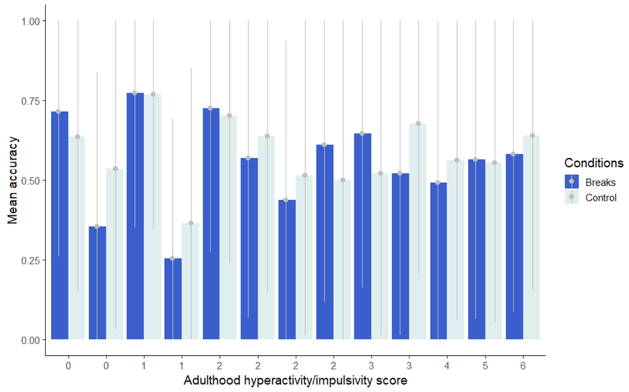
First, we investigated the overall effect of both conditions and adulthood ADHD characteristics score on reaction time and accuracy. Figure 6 shows the average accuracies of the control condition and either the breaks or self-monitoring condition for each participant and their ADHD characteristic scores. These plots shot the average accuracies per participant per condition for ADHD characteristic score for adulthood. The childhood scores can be fount in Appendix [to be added]. Figure 6a shows these average accuracies and inattention scores for the break condition. As mentioned in the methods, the scores range from zero to nine. What is noticeable is that there are a lot of zero's in both Figure 6a and Figure 6b, where the latter Figure includes more scores on the high-end of the range. The standard deviations in all four Figures are large. Both Figures mentioned above include higher scores than the scores in Figure 6c and Figure 6d.



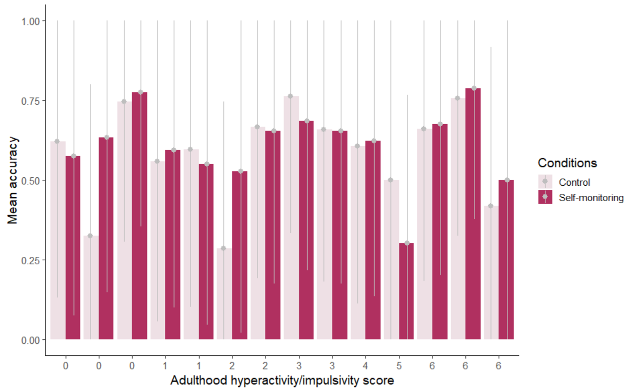
(a) Adulthood inattention scores.



(b) Adulthood inattention scores.



(c) Adulthood impulsivity/hyperactivity scores.

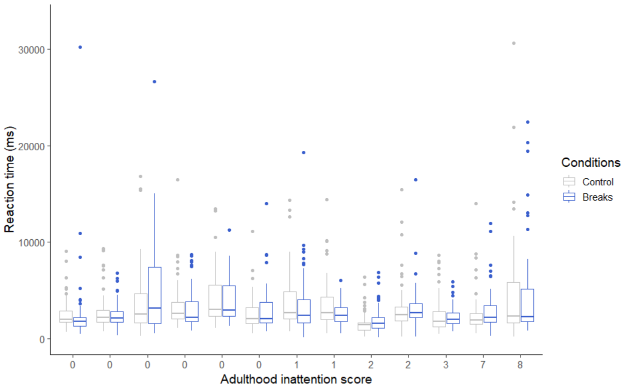


(d) Adulthood impulsivity/hyperactivity scores.

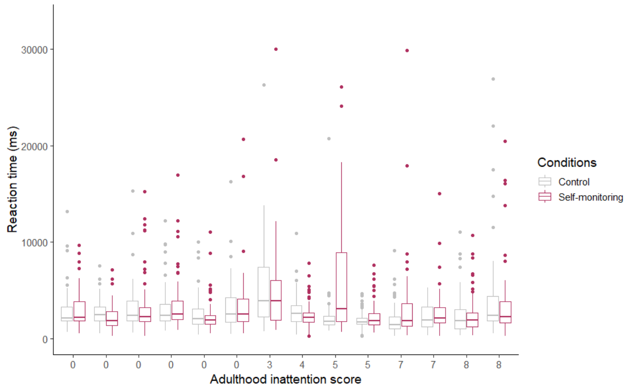
Figure 6: The mean accuracies and their ADHD characteristics score for each participant. Blue denotes the break condition and red denotes the self-monitoring condition. The lighter colors represent the control condition for the participant. Standard deviation has a maximum of 1 in the plot for readability, but is often larger.

Furthermore, those two scores on the high-end of the x-axis in Figure 6a, do indicate that the average accuracy of breaks is slightly higher than that in the control condition, but the standard deviations are very large. This difference is not in favor of the breaks condition in 6c for hyperactivity/impulsivity on the high-end scores, but in turn the hyperactivity/impulsivity scores do not reach the range of the inattention scores. The scores on the high-end of the x-axis in Figure 6b and 6d show variability in highest average accuracy between the self-monitoring condition and the control condition. There does not appear to be a visible trend in Figure 6. As inattention or hyperactivity/impulsivity scores go up, average accuracies appear to stay variable between conditions and between participants. This includes variations in average accuracy between the conditions of participants.

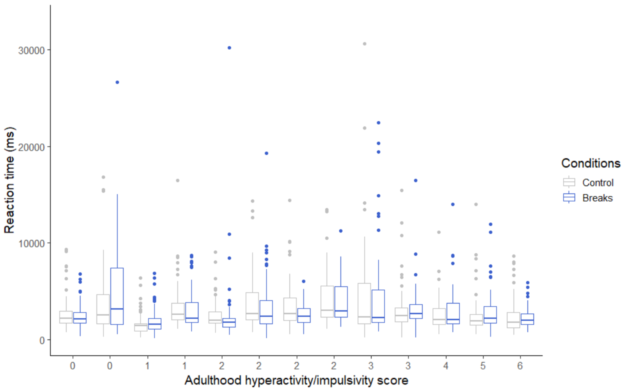
The median reaction times can be seen in Figure 7 for both conditions, both adulthood scores and for each participant. Participants are ordered along the x-axis on the basis of their inattention or hyperactivity/impulsivity score. Quite a few of those participants that had a score of zero for adulthood inattention in Figure 7a and Figure 7b appear to be slightly slower in the control condition, as can be seen by their median reaction times. The scores on the high-end of the four subfigures, which range from zero to nine, do not appear to show improvement in reaction time for the breaks or self-monitoring condition in comparison to the control condition. Some participants show more variability than others. There does not appear to be a clear difference between the break and self-monitoring approaches, nor between the scores related to ADHD characteristics.



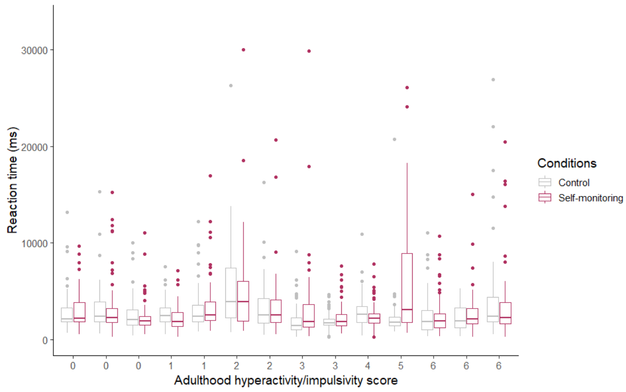
(a) Adulthood inattention scores.



(b) Adulthood inattention scores.



(c) Adulthood impulsivity/hyperactivity scores.



(d) Adulthood impulsivity/hyperactivity scores.

Figure 7: The reaction time for both conditions and their symptom score for each participant. Blue denotes the break condition and red denotes the self-monitoring condition.

No significant effect was found of either condition or either ADHD characteristics score with the linear mixed effects model on reaction time. Table 2a shows The rounded estimates, standard errors and t-values of the main effects and the interactions for predicting log-transformed reaction time, when taking the inattention characteristics scores into account. The results of the model with hyperactivity/impulsivity scores are visible in Table 2b Neither did the logistic mixed effects models show significant results on accuracy. These fixed effects while taking inattention characteristics score into account are in Table 3a and the fixed effects while taking hyperactivity/impulsivity characteristics score into account are in Table 3b.

	Estimates	SE	t-value
Intercept	7.77	0.06	140.09
Break	−0.05	0.05	−0.97
Self-monitoring	−0.03	0.06	−0.58
Inattention	−0.01	0.01	−1.06
Break x inattention	0.03	0.02	1.76
Self-monitoring x inattention	0.01	0.01	1.08

(a) The rounded estimates from the linear mixed effects model with inattention characteristics scores.

	Estimates	SE	t-value
Intercept	7.79	0.07	110.55
Break	−0.04	0.08	−0.57
Self-monitoring	−0.03	0.07	−0.45
Hyperactivity/impulsivity	−0.02	0.02	−1.06
Break x hyperactivity/impulsivity	0.02	0.02	0.87
Self-monitoring x hyperactivity/impulsivity	0.02	0.02	0.81

(b) The rounded estimates from the linear mixed effects model with hyperactivity/impulsivity characteristics scores.

Table 2: The rounded estimates from the linear mixed effects model on log-transformed reaction time with inattention or hyperactivity/impulsivity characteristics scores, the standard errors and the t-values. The reference was the control condition.

	Estimates	SE	z-value	p-value
Intercept	0.42	0.14	3.06	.002
Break	−0.15	0.13	−1.09	.274
Self-monitoring	0.13	0.15	0.88	.380
Inattention	0.02	0.03	0.64	.523
Break x inattention	0.02	0.04	0.48	.630
Self-monitoring x inattention	−0.02	0.03	−0.68	.498

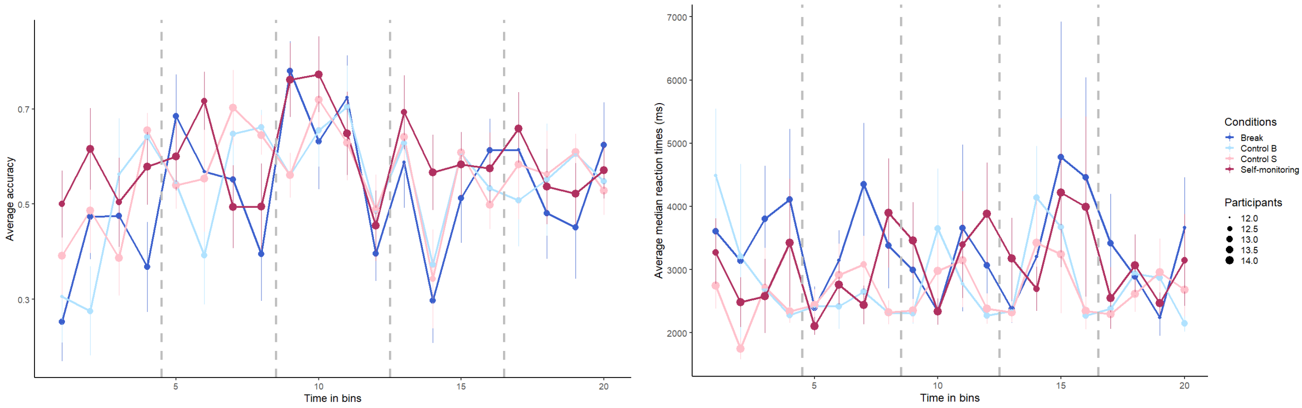
(a) The rounded estimates from the logistic mixed effects model on accuracy with inattention characteristics scores.

	Estimates	SE	z-value	p-value
Intercept	0.43	0.17	2.57	.010
Break	−0.08	0.18	−0.42	.678
Self-monitoring	0.17	0.16	1.02	.309
Hyperactivity/impulsivity	0.02	0.05	0.32	.752
Break x hyperactivity/impulsivity	−0.01	0.06	−0.26	.797
Self-monitoring x hyperactivity/impulsivity	−0.04	0.05	−0.84	.404

(b) The rounded estimates from the logistic mixed effects model on accuracy with hyperactivity/impulsivity characteristics scores.

Table 3: The rounded estimates from the logistic mixed effects model on accuracy with inattention or hyperactivity/impulsivity characteristics scores, the standard errors, z-values and p-values. The reference was the control condition.

4.1 Bins: analysis between interventions



(a) The average accuracy over bins.

(b) The average median reaction time over bins.

Figure 8: Behavioral data in accuracy and reaction time blocked between approaches with standard errors. Four lines are present: one for the self-montioring condition, one for the break condition, and two lines for the control condition, subdivided depending on which approach that participant encountered in the other session. The gray dashed lines represent each occurrence of an approach. Not all participants performed a trial in each bin, as trials could take up to 30 seconds, therefore the size of the points denote the number of participants that was averaged over.

We were further interested in the effect of conditions in the seconds after participants received a break or a self-monitoring prompt. The average accuracies and average median reaction times over binned time are illustrated in Figure 8. Both figures illustrate the general effect of each condition over all participants, regardless of ADHD characteristic score. Both figures consists of four lines, two lines for the participants that received breaks and their control condition and two lines for participants that were prompted with self-monitoring and their counter control condition. The effects on average accuracy in Figure 8a shows that both approaches increase in average accuracy in the seconds after a break or self-monitoring encounter (the gray dashed line). Accuracy then decreases again before the next approach encounter. Even though the control condition did not have any interventions, variance can be seen in increasing and decreasing average accuracies.

However it is apparent in the time between the first and second interruption from Figure 8a, that the control conditions of the participants illustrate an opposite effect from their approach condition. Instead of an increase in average accuracy, it decreases after the ninety second mark and later increases again. On the other hand, after at least 4 minutes and 30 seconds in bin 3, the self-monitoring and breaks conditions appear to fall in the same pattern as the control condition until around 67.5 seconds after the self-monitoring prompt or break. We were therefore interested to further investigate the effect of breaks and self-monitoring prompt in the time after its occurrence as opposed to the time before the next occurrence. We thus analyzed the effect of conditions over binned time.

Figure 8b shows a different effect on reaction time in comparison to accuracy. Both approaches, breaks and self-monitoring, illustrate higher increases in averaged median reaction times and more variance than their counter control conditions. No significant effect could be found in the linear mixed effects model of condition, ADHD characteristics scores and these interactions on log-transformed reaction time. These results of the statistical models of inattention, hyperactivity/impulsivity and without inattention or hyperactivity/impulsivity can be found in Appendix C. We further investigated bins, condition and their interaction (without ADHD characteristic scores) with a logistic mixed ef-

	Estimates	SE	z-value	p-value
Intercept	0.82	0.18	4.58	$<.001$
Break	−0.69	0.22	−3.17	.002
Self-monitoring	0.21	0.06	−1.56	.118
Bin 1	−0.45	0.16	−2.78	.005
Bin 2	0.21	0.29	0.74	.458
Bin 3	0.09	0.29	0.30	.764
Break x bin 1	0.99	0.26	3.74	$<.001$
Self-monitoring x bin 1	0.25	0.03	1.77	.077
Break x bin 2	0.04	0.36	0.10	.918
Self-monitoring x bin 2	0.01	0.35	0.01	.988
Break x bin 3	0.05	0.36	0.13	.893
Self-monitoring x bin 3	−0.12	0.35	−0.35	.725

Table 4: The rounded estimates from the logistic mixed effects model on accuracy, the standard errors, z-values and p-values. The reference was the control condition and bin 4.

fects model on accuracy.

According to this latter model, the control condition in bin 4 has a significantly higher accuracy than the breaks condition in bin 4 (β = -0.69, SE = 0.22, z-value = -3.17, $p < .001$), as can be seen in Table 4. Moreover, the accuracy in bin 4 of the control condition is also significantly higher than in bin 1 of this condition (β = -0.45, SE = 0.16, z-value = -2.78, p = .005), according to this model. Lastly, there is an interaction effect of the breaks condition with bin 1 (β = 0.99, SE = 0.26, z-value = 3.74, $p < .001$). According to the model, the breaks condition has a higher accuracy in bin 1 than the control condition and the breaks condition also has a higher accuracy in bin 1 than in bin 4.

4.1.1 Accuracy and reaction time immediately after an intervention

We further analyzed the effects of breaks and self-monitoring while taking inattention or hyperactivity/impulsivity characteristics scores into account on log-transformed reaction time and accuracy. We therefore investigates these effects in the first 22.5 seconds or right after an intervention until 22.5 seconds. We can see a significant main effect of the breaks condition on accuracy in these first 22.5 seconds (β = 0.48, SE = 0.21, z-value = 2.25, $p = .025$) when taking inattention characteristics score into account and when taking hyperactivity/impulsivity characteristic score into account (β = 0.75, SE = 0.29, z-value = 2.58, $p = .010$). This condition has a higher accuracy than the uninterrupted learning condition in bin 1 as can be seen from Table 5a when taking inattention score into account and from Table 5b when taking hyperactivity/impulsivity scores into account.

No significant main effects or interaction effects where found in both models on log-transformed reaction time. Those in the breaks condition were not significantly faster in these 22.5 seconds than in the uninterrupted learning condition when taking inattention into account (β = -0.06, SE = 0.06, t-value = -1.07) or when taking hyperactivity/impulsivity into account (β = -0.07, SE = 0.08, t-value = -0.79). The same holds for the main effect of self-monitoring when taking inattention into account (β = -0.04, SE = 0.07, t-value = -0.58) or when taking hyperactivity/impulsivity into account (β = -0.05, SE = 0.08, t-value = -0.72). The rounded estimates of these models can be found in Table 11a and Table 11b of Appendix C.

	Estimates	SE	z-value	p-value
Intercept	0.34	0.14	2.41	.016
Break	0.48	0.21	2.25	.025
Self-monitoring	0.14	0.23	0.61	.543
Inattention	0.01	0.03	0.24	.813
Break x inattention	−0.10	0.07	−1.44	.150
Self-monitoring x inattention	−0.01	0.05	−0.13	.900

(a) The rounded estimates with inattention characteristics scores.

	Estimates	SE	z-value	p-value
Intercept	0.38	0.17	2.16	.031
Break	0.75	0.29	2.58	.010
Self-monitoring	0.20	0.26	0.80	.426
Hyperactivity/impulsivity	−0.01	0.05	−0.12	.903
Break x hyperactivity/impulsivity	−0.18	0.09	−1.96	.050
Self-monitoring x Hyperactivity/impulsivity	−0.03	0.07	−0.39	.697

(b) The rounded estimates with hyperactivity/impulsivity characteristics scores.

Table 5: The rounded estimates from the logistic mixed effects models on accuracy, the standard errors, z-values and p-values with inattention or hyperactivity/impulsivity characteristics scores. The reference was the control condition the dataset is only comprised of bin 1.

4.2 Uninterrupted versus interrupted

We further investigated the effect of an interruption versus an uninterrupted session, regardless whether these are breaks or self-monitoring prompts. The results while taking inattention characteristics scores into account can be seen in Table 6a. Both models have a random intercept of session over participants and a random intercept of fact ids. The results of the model on log-transformed reaction time with hyperactivity/impulsivity scores can be found in Table 6b. Neither the main effects of treatment (interrupted or uninterrupted) and inattention nor the interaction of these fixed effects on log-transformed reaction time were significant. The main effects of treatment (interrupted or uninterrupted) and hyperactivity/impulsivity also did no have a significant effect on log-transformed reaction time.
We also analyzed the effect of treatment with inattention characteristics scores on accuracy with a random intercept of participants and a random intercept of fact id. as combinations with session and condition resulted in converge or singularity errors. There are no significant main effects or interactions, as can be seen in Table 7a. We also analyzed the effect of treatment on accuracy with adulthood hyperactivity/impulsivity characteristics scores and a random intercept of participant and a random intercept of fact id. There are no significant main effects or interactions, according to this model. The results can be seen in Table 7b.

4.3 Break and self-monitoring utilization

We recorded participants' engagement with the self-monitoring prompts through self-recorded answers and time to continue after answering and break utilization through break durations. For the group that received breaks, we recorded their break duration for each break. Participants received

	Estimates	SE	t-value
Intercept	7.77	0.06	137.71
Treatment	−0.04	0.04	−0.88
Inattention	−0.02	0.01	−1.13
Treatment x inattention	0.01	0.02	1.77

(a) The rounded estimates with inattention characteristics scores.

	Estimates	SE	t-value
Intercept	7.79	0.07	110.39
Treatment	−0.03	0.05	−0.65
Hyperactivity/impulsivity	−0.02	0.02	−1.06
Treatment x hyperactivity/impulsivity	0.02	0.02	1.12

(b) The rounded estimates with hyperactivity/impulsivity characteristics scores.

Table 6: The rounded estimates from the linear mixed effects models on log-transformed reaction time, the standard errors and t-values with inattention or hyperactivity/impulsivity characteristics scores. The reference was the control condition the dataset is only comprised of the control condition and the treatment or uninterrupted condition.

four breaks in this session with a maximum of approximately 60 seconds for each break.

	Estimates	SE	z-value	p-value
Intercept	0.42	0.14	3.06	.002
Treatment	−0.02	0.10	−0.23	.819
Inattention	0.02	0.03	0.63	.527
Treatment x inattention	0.00	0.02	0.03	.973

(a) The rounded estimates with inattention characteristics scores.

	Estimates	SE	z-value	p-value
Intercept	0.43	0.17	2.56	.011
Treatment	0.05	0.13	0.39	.694
Hyperactivity/impulsivity	0.02	0.05	0.31	.757
Treatment x hyperactivity/impulsivity	−0.03	0.04	−0.71	.478

(b) The rounded estimates with hyperactivity/impulsivity characteristics scores.

Table 7: The rounded estimates from the logistic mixed effects models on accuracy, the standard errors, z-values and p-values inattention or with hyperactivity/impulsivity characteristics scores. The reference was the control condition the dataset is only comprised of the control (uninterrupted) and treatment condition (interrupted).

For the group that received self-monitoring prompts, we recorded their responses and the time they took to respond. Participants received four self-monitoring of attention prompts during this session. The responses of the self-monitoring prompt: ”Did you have your attention on-task just now?” are illustrated in Figure 9. Every participant in this group responded to all self-monitoring prompts. We can further see that in the first two occurrences, all participants responded ”yes” and thus self-reported to have had their attention on the task of fact learning. In the third occurrence, one participant responded with "no" and this doubled to two participants in the last occurrence of the self-monitoring prompt.

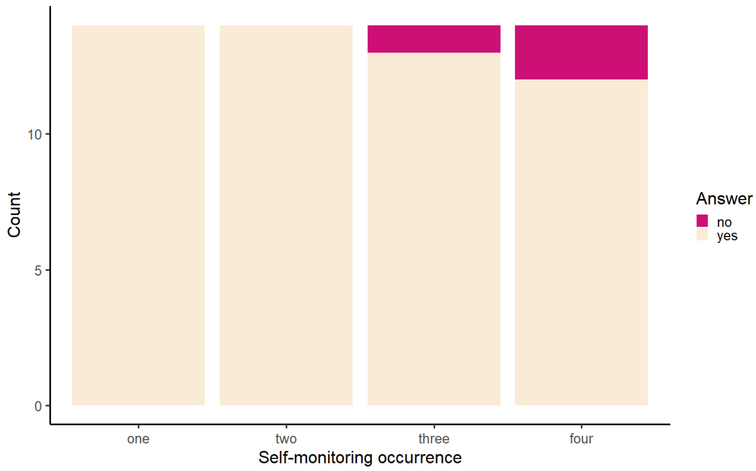


Figure 9: The responses of participants in the self-monitoring groups for each occurrence of the self-monitoring prompt. Responses are counted.

We further looked into the duration it took participants to respond to the self-monitoring question. This duration is counted until the moment they click continue to continue to the next trial. Participants could still change their answer before continuing. All participants responded within the maximum timeframe of fifteen seconds, as can be seen in Figure 10. The median self-monitoring time in the first occurrence was 3.44 seconds (IQR = 1.00). The second median self-monitoring time was 3.72 seconds (IQR = 1.21). The third self-monitoring time was 3.42 seconds (IQR = 0.90) and the last self-monitoring time was 3.32 seconds (IQR = 1.86), which shows a little more variation. Both the left figure, where inattention characteristic scores are shown and the right figure where hyperactivity/impulsivity scores are shown, indicate that at least half of participants responded within three seconds for a majority of the breaks. Two participants with high inattention and hyperactivity/impulsivity characteristic scores show a little slower responses in the second break than the other participants. However, this does not last to further breaks. Their self-monitoring times then appear to be faster in the last two self-monitoring prompts. Instead, other participants that appear to be a little lower on the inattention and hyperactivity/impulsivity characteristics score range take longer to self-monitor and continue to the next trial.

Figure 11 shows the length of each break that a participant took. Almost half of participants seem to have skipped the break or to have only taken around 10 seconds or less, which is also noticeable from the median break durations. In the first break the median break duration is 7.86 seconds (IQR = 25.71). The second break had a median break duration of 9.94 seconds (IQR = 30.15). The third break had a median break duration of 11.44 seconds (IQR = 16.49) and the last break had a median break duration of 10.78 seconds (IQR = 18.70) . There appear to be only seven participants that took rests of 10 seconds or more for some of the breaks. Furthermore, the four participants that have taken the longest breaks during at least one break also are quite low on the ADHD characteristic score range. Both for inattention as well for hyperactivity/impulsivity.

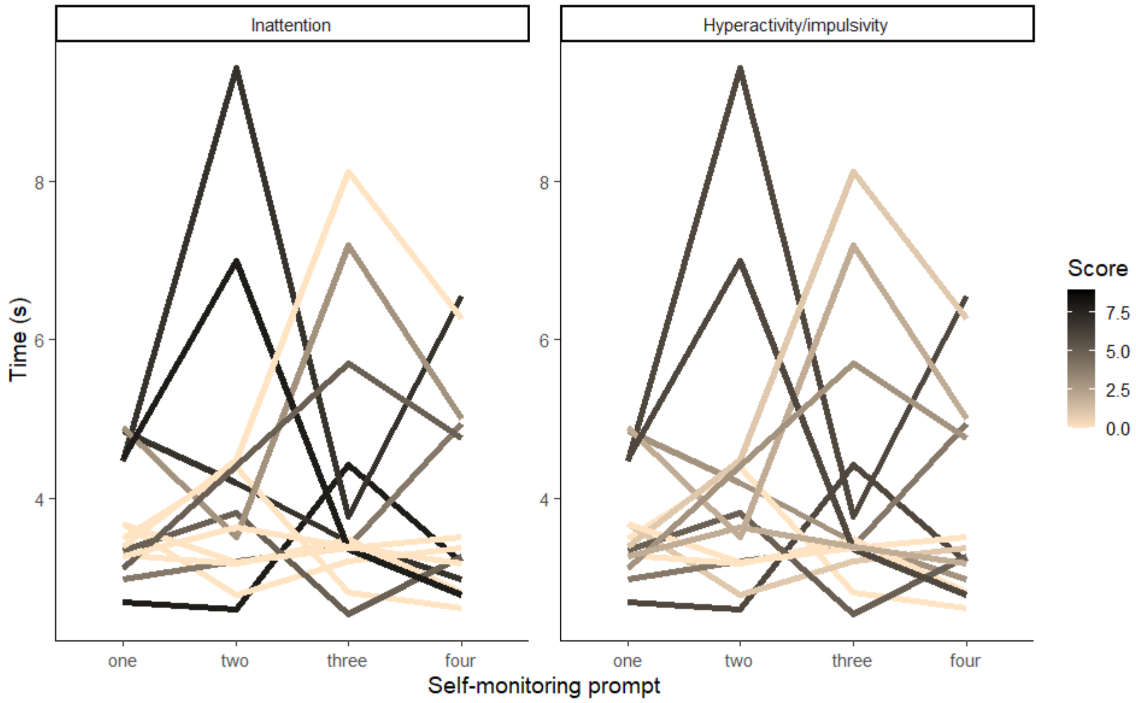


Figure 10: The self-monitoring durations of participants and their inattention and hyperactivity/impulsivity characteristic scores.

4.4 Experiences and preferences of participants

In order to analyze the experience of participants with these approaches while learning facts through an adaptive learning algorithm, a semi-structured interview was conducted and analyzed.

4.4.1 Self-monitoring preference

Four out of fourteen participants (28.6 %) that received self-monitoring prompts, preferred the session with these self-monitoring prompts. Half of these participants that preferred receiving self-monitoring prompts over uninterrupted learning, reported to have an ADHD diagnosis and/or has an adulthood inattention characteristics score above five. These latter two participants in turn mentioned that their attention wandered from the task at times, whereas this was not the case for many other participants. We did receive positive feedback on self-monitoring, from one of these participants. They stated that the self-monitoring prompts provided a pause to rest, feedback and organization/structure. They mention being in competition with oneself to do well and receive *"positive reinforcement*" afterwards, which is not provided in uninterrupted learning according to the participant. They report rewarding themselves which leads to an increase in *"motivation"* to continue, because of the small durations of learning (between self-monitoring prompts).

Three out of four participants that preferred the session with self-monitoring responded "yes" to the question whether they felt the self-monitoring helped in keeping their attention on-task during learning. The fourth participant mentioned already having on-task attention.

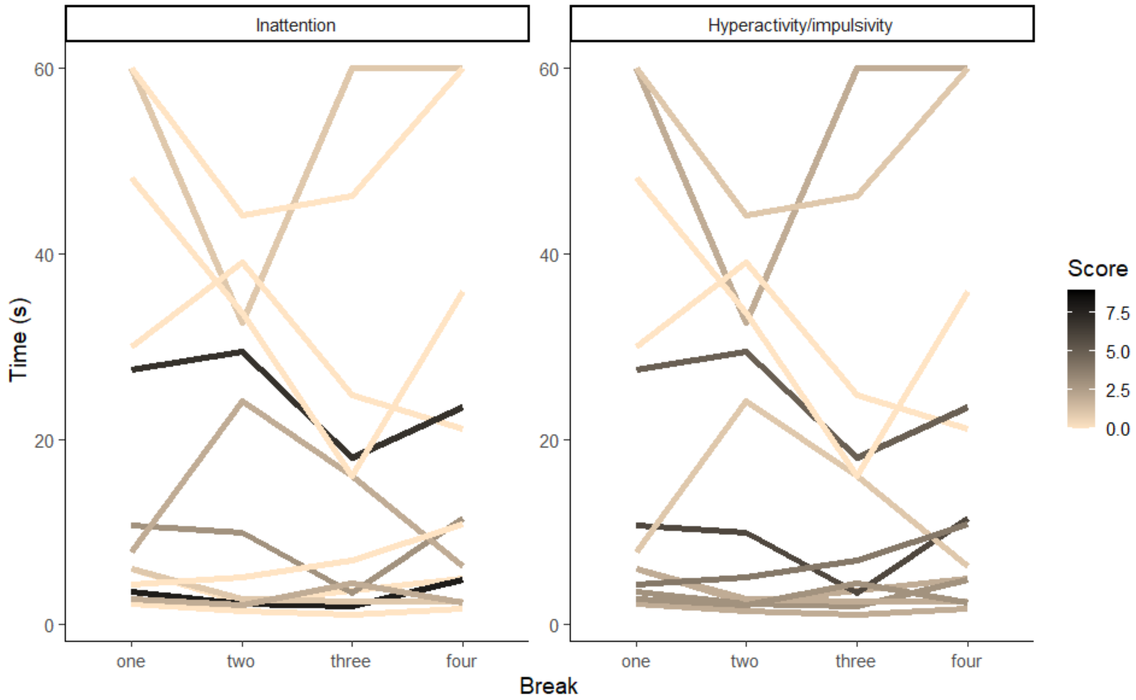


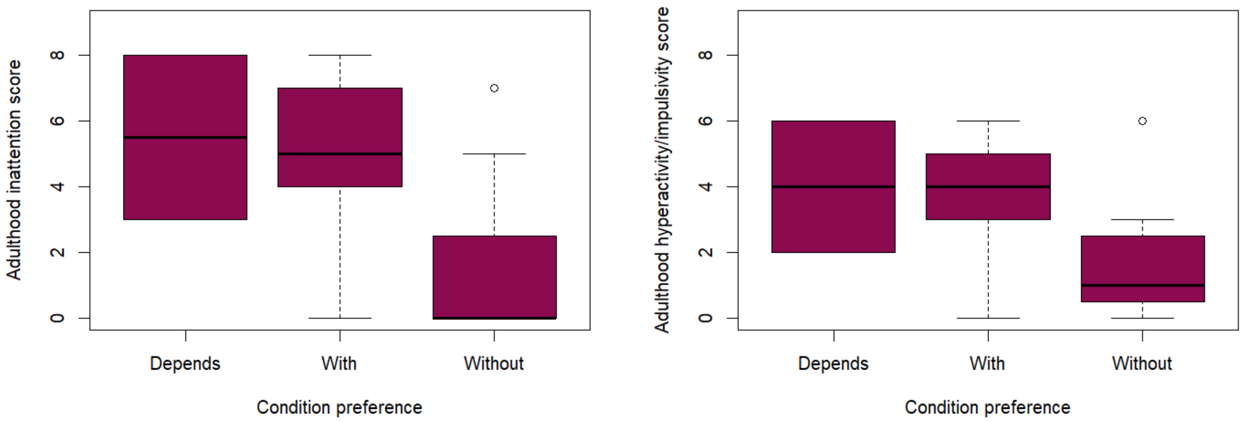
Figure 11: The break durations of participants and their inattention and hyperactivity/impulsivity characteristic scores.

4.4.2 No preference of self-monitoring

Furthermore, 50% of participants that received self-monitoring prompts preferred the control condition and the other 21.4% did not notice much of a difference between sessions or had no preference. These latter participants had varying responses (no, no difference and yes) to the question whether the self-monitoring questions helped keep their attention on-task during learning. We annotated these participants as a preference for self-monitoring, for uninterrupted learning or other.
We conducted a Kruskal-Wallis H test (Kruskal & Wallis, 1952) to investigate if the inattention characteristic scores are significantly different between preference groups of the fourteen participants that received self-monitoring prompts. This test shows no significant difference $H(2) = 0.661$, as the p-value (.719) is above the alpha of 0.05. The hyperactivity/impulsivity scores in turn do also not significantly differ between group $H(2) = 0.007$, p-value = .997.
Most of those that preferred the control condition either mentioned that they were not distracted during the sessions or that the self-monitoring prompt distracted them or removed them from the flow of the task. One participant that preferred the control condition mentioned distractions through spelling errors and one other responded negatingly sometimes having had wavering thoughs for very short moments. The frequency of self-monitoring prompts was also said to be quite fast. They were still on-task at this interval.

4.4.3 Self-monitoring frequency and screen

Participants were asked whether they noticed their previous self-monitoring answers on the self-motoring screen, and if yes, what they thought about that. Not all participants had noticed, most were fine with it but indifferent, two participants mentioned it was motivating. One participant stated that they noticed that there were more questions to come and one participant mentioned that it disturbed them. Additionally, participants were asked how they felt about the number of self-monitoring questions. Seven participants reported positively by use of ”good” or ”fine”. Five participants stated there were (too) many, one said it was fine but there could have been less and one participant responded with too little. Furthermore, there was also a time limit for answering the self-monitoring question, but participants said they either did not notice or were not bothered by the time limit. Overall, participants thought that the self-monitoring question was alright. Some mentioned that they liked the definition and explanation given before the start of the session on what paying attention entailed. A few participants mentioned preference for an in-between option or options on a scale, instead of the binary yes/no.



(a) Inattention characteristics scores.

(b) Hyperactivity/impulsivity characteristics scores.

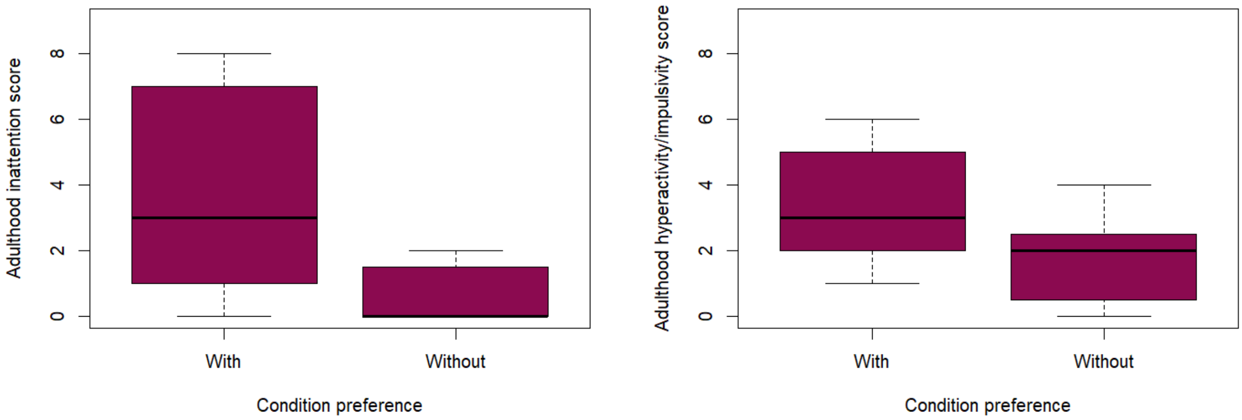
Figure 12: Responses and participants’ corresponding score to the question: *If you would do the task again, would you want to do it with or without self-monitoring questions?*

Lastly, participants received the question: ”if you would do the task again, would you want to do it with or without the self-monitoring questions?”. Four participants responded ”with”, one participant said to have no preference, but would choose with a self-monitoring pause if they had to choose. This participants was in turn annotated as ”with”. Seven participants responded ”without” and two participants responded that it depends on the duration and the time of day or the number of learning questions and the number of self-monitoring questions. These participants and their corresponding inattention and hyperacitivity/impulsivity scores are visible in Figure 12. We conducted a Kruskal-Wallis H test (Kruskal & Wallis, 1952) to investigate if the inattention characteristic scores are significantly different between the groups of participants that received self-monitoring prompts. The test shows no significant difference H(2)=3.590, p-value = .166. There is also no significant difference between the groups for hyperactivity/impulsivity characteristic scores H(2) = 2.273, p-value = .321.

4.4.4 Break preferences

For the thirteen participants that performed the task with optional breaks, 30.8% preferred this condition as opposed to uninterrupted learning in the control condition. Half of the participants that preferred breaks over uninterrupted learning, showed scores for at least adulthood inattention to be above of 5 or in addition have an ADHD diagnosis. Which is two third of all participants in the break condition with scores of five or more. Seven participants (53.8%) preferred uninterrupted learning.

The other participants were either unsure (they preferred a break after a number of learned facts), or said not to have a preference. We annotated these participants as a preference for breaks, for uninterrupted learning or other. We also analyzed whether there was a significant difference between the inattention score of the preference groups of those thirteen participants that performed the break condition. No significant difference was found $H(2) = 1.941$, p-value = .379. In turn, no significant difference between the hyperactivity/impulsivity scores were found between the preference groups of participants that received breaks $H(2)= 0.374$ p-value = .829.



(a) Inattention characteristics scores. (b) Hyperactivity/impulsivity characteristics scores.

Figure 13: Responses and participants' corresponding score to the question: *If you would do the task again, would you want to do it with or without breaks?*

However, when asked if they would opt for with or without breaks when performing the task again, 38.5% of participants opted for with breaks. One of these participants said ”in this form with, because you can thus click them away”. The other 61.5% opted for no breaks if they performed this task again. These participants and their inattention and hyperactivity/impulsivity scores are visualized in Figure 13. There were only two preference groups for participants that received breaks, namely ”with” and ”without” breaks. Therefore a Kruskal-Wallis H test was not utilized. Instead the Wilcoxon Rank-Sum test (Wilcoxon, 1945) was used. There was no significant difference according to this test ($W = 8$, $p = .076$), but there were ties in this test as well. Therefore, the p-value might not be precise according to the warning. The same warning occurred for testing a difference of hyperactivity/impulsivity characteristic scores between groups. In turn, no significant difference was found ($W = 10.5$, $p= .180$).

4.4.5 Break and distractibility

Participants were further asked whether they were distracted in one of the exercises and most responded with no. In turn, the majority of participants negated that breaks helped them with keeping their attention on-task. Mentioned reasoning was that they did not need to break because they had remained focused, that the break would stop their flow of learning, that breaks would distract them with other thoughts or that the task was too short so a break was unnecessary. One participant mentioned that they would take a break when they feel they need to, instead of a ”forced” break after a set duration. On the other had, two participants responded with ”yes” it helped them to regroup or to rest. Others were more unsure, did not notice a difference or had things to say for both sides (with or without) breaks.

4.4.6 Break duration and length

Moreover, participants were asked: ”what do you think about the length of the breaks?”. Nine participants mentioned they either felt the break was too long or that they skipped the breaks. The other participants were fine with the duration, one participant mentioned they were fine because they could skip/leave earlier.
In addition, I asked participants: ”what do you think about the number of breaks?”. Six participants thought these were good or fine, whereas three participants felt there were too many breaks. One participant responded that the breaks occurred often, but did not think these were too often. Two participants responded that they skipped the breaks. Lastly, one participant stated that these were fine, the breaks can be clicked away, but were not necessarily needed.

4.4.7 During the break

Participant also received a question about the look of the break screen, as is visible in Figure 4b. Nine participants provided positive feedback, such as neutral, fine, good, clear, calm without being distracted. One participant mentioned that they thought the last five seconds countdown where the text changed to: ”break almost over” was smart. Another participant felt this as pressure, with another participant mentioning that the timer on its own was ”stressful”. Two other participants either responded with ”boring” or ”nothing special”.
Participants were also asked what they were doing during the break. There answers were annotated based on two categories *task-related* and *task-unrelated* in Hall-Ruiz (2016). However as these categories are about thoughts and the question asked about what the participant was doing, anything other than thinking about the task, performance, as in Hall-Ruiz (2016), and additionally for how long they would like to break, is categorized as unrelated. Firstly, four participants responded that they skipped the breaks and another participant said to have thought of nothing. Four participants were having task-related thoughts, and the other four participants were categorized as having both task-related thought and having thought or doing task-unrelated activities. The duration of the task (at least 8 minutes per session) was also mentioned as a reason for not opting for taking (longer) breaks. The duration is too short, to be needing breaks according to one participant. If the task had more flags they would be in need for a break.

4.4.8 General

Spelling was mentioned as a distracting factor. One of the reasons that some participants mentioned for not wanting the self-monitoring questions or breaks is because they felt this would distract them from the task or because they were not losing attention. Some participants mentioned that the flags were quiet difficult, or that spelling was very strict. Such as knowing the flag, but not being able to precisely spell it correctly. One participant mentioned that the task was very fun and they would have like work on it for longer.

5 Discussion

5.1 Breaks

Contrarily to results in Hall-Ruiz (2016) and our hypothesis, applying breaks did not significantly impact overall performance. Reaction time nor accuracy were significantly or very visibly improved, nor was there a significant interaction between breaks and inattention or hyperactivity/impulsivity characteristics score. These results are however, not in contrast with other research that investigated breaks as an accommodation for learning (Harrison et al., 2022; Pritchard et al., 2016). In the comparison of breaks, self-management, sensory items and prompts, Harrison et al. (2022) found that breaks were not the most helpful in improving attention on-task, interruptive behavior or time to start the task. Furthermore, Pritchard et al. (2016) did not find a significant improvement in their meta-analysis of breaks on performance. Contrastingly, I only measured the accommodation of performance, not on attention. Moreover, the strategies employed in Harrison et al. (2022) were conducted over multiple sessions, which was not utilized in this project. I can thus not provide any conclusions on whether attention was improved or not. Moreover, participants were not required to have an ADHD diagnosis, nor did many participants have inattention and/or hyperactivity/impulsivity characteristic scores on the high-end. When looking at adulthood inattention and adulthood hyperactivity/impulsivity scores in a continuous manner, a lot of participants fall within the range of zero to five for these scores. Twenty-seven participants took part in the experiment, of which thirteen received breaks. Therefore, very strong conclusion and significance might be difficult to obtain.
Another possible explanation for not improved performance, could be that quite a portion of the participants skipped the breaks, or only took a few seconds (see Figure 11). As we have seen above, reasoning for this could be due to the short length of the task where participants did not loose focus in order to actually require a break. Finding an effect of breaks on learning might be quite difficult, in the case that these breaks are skipped. On the other hand, it might be argued that choice in break duration should not have been omitted, as otherwise these participants would have been ”forced” to ”wait”. This might have led to other negative effects that we wanted to prevent. Furthermore, individual differences, even though accounted for as a random effect in the model, could be part of the explanation as well. Individuals could have other disorders, difficulties or comorbidities that were not asked during the experiment. For example, the task of the flags might have already been hard for someone that has difficulties with spelling. Interestingly, even though literature does not always find positive effects of breaks on on-task behavior and performance, this accommodation is recommended or suggested for children (Children and Adults with Attention-Deficit/Hyperactivity Disorder (CHADD), n.d.; McKnight, 2010).
On the other hand, when zooming in on the seconds after a break occurrence, our model did indicate a significant effect. There appears to be a difference between the immediate seconds after and before an approach in the break condition and the control condition for accuracy while taking inattention or hyperactivity/impulsivity scores into account. This might suggest a short effect of breaks on learners. More research is needed to investigate breaks for adults in such an online adaptive learning system.

5.2 Self-monitoring

An unexpected and surprising outcome of this project is the result of self-monitoring on performance. Employing self-monitoring throughout the session did not result in any significant main effects or interactions while taking inattention or hyperactivity/impulsivity score into account for both log-transformed reaction times and accuracy. This is contrast with previous literature that did find positive effects on performance (Harris et al., 2005; Scheithauer & Kelley, 2017; Reid et al., 2005). However, previous literature also investigated self-monitoring on on-task behavior on children (Harrison et al., 2022; Mathes & Bender, 1997; Wills & Mason, 2014; Vogelgesang et al., 2016; Reid et al., 2005) over multiple sessions and/or with fading.

Different session durations were employed in previous research on children, such as 20 minutes with 5 minute interval cues to self-monitor (Harrison et al., 2022) 55 minutes and 15 minutes of observation, with 5 minute interval cues (Wills & Mason, 2014), 45 to 60 minutes with 10 minute interval cues (Vogelgesang et al., 2016), 10 minutes of observation with review of self-monitoring each 2 days and with cues at random, but on average with a 45 second interval (Mathes & Bender, 1997) 15 minutes with 10 minute observations and also cues at random with minimum 10 seconds and maximum 90 and average 45 (Harris et al., 2005). These intervals and session duration vary in literature. However, our self-monitoring of attention session of at least 8 minutes is on the short end and only occurred once instead of multiple sessions, with a self-monitoring interval of approximately 90 seconds. Moreover, these design choices of session duration and intervals are based on children, not adults. Perhaps adults might be able to focus longer than children and therefore these design choices in duration might not translate to adults as much.

Furthermore, none of the participants omitted an answer to the self-monitoring question and all responded before the maximum response time of fifteen seconds. We have seen from Figure 10 that most participants were quite fast in their response. Perhaps participants did not fully process their answer to become more conscious of their attention. On the other hand, as explained above, participants might have stayed on-task for the whole duration of the session and might have simply not needed the self-monitoring prompts. Another explanation could be quite a portion of the fourteen participants in the self-monitoring group scored on the low end or even zero on the inattention and hyperactivity/impulsivity scores. These participants might therefore not struggle with losing focus or continuing with the task as much or at all.

Additionally, the same explanations of this sample of participants in the breaks group can be applied to the self-monitoring group. The sample size is small, and comorbidity or other disorders could have been at play and influences results that we did not ask.

5.3 Interview

We have seen from the interview responses that a lot of participants were not distracted in the sessions and felt that breaks or self-monitoring prompts distracted them from the task or were unnecessary. We have seen from Figure 13 and Figure 12 that those with scores on the low end of especially inattention seem to opt more for uninterrupted learning instead of breaks or self-monitoring. Alternatively, those that score more on the high-end seem to prefer either with self-monitoring or breaks or depending on factors. This sketches an interesting image of participants' preferences, which is not visible from behavioral data. However, those that opted for self-monitoring or breaks were a minority, nor were the differences of inattention and hyperactivity/impulsivity scores significant between condition preferences.

Another notable point is that one participant with ADHD characteristic scores on the high end, states to have felt distracted in both sessions. However, they report in the interview that the self-monitoring prompts did provide a difference for them and were easy to click away. This also indicates that more work is needed to investigate these approaches for adults (with ADHD characteristics) in an adaptive learning system.

5.4 General

Another possible explanation for the results of both approaches could be the fast-paced computerized learning system, where it is not required of the learner to continue to learn the next ”flashcard”, instead the next item is presented to them. Moreover, participants were instructed to respond as fast and as accurate as possible, which perhaps could have induced a motivation to stay focused. Even though a spike in accuracy and a decrease in reaction time is seen after each intervention in Figure 8a and Figure 8b, this did not seem to transfer to next trials. A possible reasoning could be that part of the participants mentioned skipping the break or only taking a break for a few seconds, which might have shortend the effect of the break.
According to American Psychiatric Association (2013), symptoms might present itself in a lesser form or not at all when the learning is in a new environment, has constant outside inputs or stimuli, such as a digital display or is doing an engaging task among other factors. These factors mentioned above were all in play during the experiment. It might be argued that the novelty of the setting and the online learning system itself might have limited the display of symptoms. Furthermore, the task itself, learning flags could subjectively be argued as a more fun learning tasks, as mentioned by a participant. Instead of tasks where the participant is doing their homework (Harrison et al., 2022).

5.5 Limitations

There were several other limitations during the experiments.
Another limitation during the experiment was the transcription of interview responses instead of audio recordings. The interviews were held by one experimenter and transcribed at the same time, which made it harder to focus on answers as well as transcribing them properly. Not using audio recording might have also meant loss of data in the form of intonation, nuances and exact spacing participants use between answers, where a difference might lie between confident and doubtful answers. Additionally, the annotating of interview responses was done in solo and quite difficult with nuances provided. For instance, a participant that says they have no preference, but if they must choose then they would opt for including self-monitoring.
The self-monitoring sheet was based on literature on children, such as the sheet in Blood et al. (2011). This might not have translated as well too adults. Perhaps adults are less reinforcement by smileys than children. The specific parameters, such as time between approaches, and break duration were not clear from literature or not as transferrable to this task. Perhaps the certain design choices made, such as four approaches in a session of at least 8 minutes might have been too many for this sample. The sessions itself were in turn quite short.
Two participants took part in the experiment in a different location outside of the lab. Furthermore, spelling was also tested, but instead a calculation could have been utilized in order for participants that recognize the flag of ”Antarctica” to be still correct when answering ”Antartica”. Therefore in this project, spelling has been taken into account as a ”target skill”, but it might have actually been an ”access skill”. Another limitation is comorbidity in this project. We did not record whether participants had any other diagnoses besides ADHD or without ADHD. However, comorbidity within individuals with ADHD is not a rarity (American Psychiatric Association, 2013). This might have an impact on the performance with these specific strategies employed.
Lastly, the aim was also to observe participants during the sessions, but this turned out to be quite invasive with the way the lab was set up. Therefore, we only observed participants through their screens in the lab. It would be interesting if future work would also analyze participants’ behavior and their on-task attention, besides solely performance and experience. This meant that we could not

compare on-task behavior to other literature.
Future work is needed to investigate breaks and self-monitoring on on-task attention besides performance in an adaptive learning task on longer durations and with various fact-learning tasks. It might in turn be interesting if future work takes comorbidity into account as well. Furthermore, it might also be interesting to apply more choice in self-monitoring and breaks through opting out or in and its number of duration. In that way, participants that do not need approaches might not be distracted by them and those that do can opt to do so. Participants are represented in this project by their ADHD characteristics score, but it is important to keep in mind that each individual is unique and might have different preferences and needs and perhaps already their own strategies.

6 Conclusion

We have investigated the effect of breaks and self-monitoring on performance through reaction time and accuracy, while taking ADHD characteristics scores into account. We investigated this effect both on the whole session and in intermediate sections of time between each break or self-monitoring prompt. No significant main effects or interaction effects were found on the complete sessions for both conditions and for both inattention and hyperactivity/impulsivity characteristics scores.
Significant main effects and an a significant interaction were found on accuracy for the time right after a break or self-monitoring prompt in comparison to the time just before. However, the task, of at least 8 minutes per session is quite short in comparison to other literature on self-monitoring for children (e.g. Wills and Mason (2014); Vogelgesang et al. (2016)). Learning flags of nations in a paced and computerized task might have been engaging and might have already reduced symptoms of those with ADHD characteristics (American Psychiatric Association, 2013).
Perhaps adaptive online learning systems are already beneficial for learners with ADHD characteristics. More work is needed to investigate these approaches on attention and performance in different and perhaps longer tasks.

References

Adamou, M., Asherson, P., Arif, M., Buckenham, L., Cubbin, S., Dancza, K., . . . Young, S. (2021). Recommendations for occupational therapy interventions for adults with ADHD: a consensus statement from the UK adult ADHD network. *BMC psychiatry*, *21*, 72. doi: https://doi.org/10.1186/s12888-021-03070-z

Advokat, C., Lane, S. M., & Luo, C. (2011). College students with and without ADHD: Comparison of self-report of medication usage, study habits, and academic achievement. *Journal of Attention Disorders*, *15*(8), 656-666. doi: https://doi.org/10.1177/1087054710371168

American Psychiatric Association. (2013). Attention-deficit/hyperactivity disorder. In *Diagnostic and statistical manual of mental disorders* (5th ed.). doi: https://doi.org/10.1176/appi.books.9780890425596

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press. doi: https://doi.org/10.1093/acprof:oso/9780195324259.003.0003

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: https://doi.org/10.18637/jss.v067.i01

Biederman, J., Petty, C. R., Fried, R., Kaiser, R., Dolan, C. R., Schoenfeld, S., . . . Faraone, S. V. (2008). Educational and occupational underattainment in adults with attention-deficit/hyperactivity disorder: A controlled study. *The Journal of Clinical Psychiatry*, *69*(8), 1217-1222. doi: https://doi.org/10.4088/jcp.v69n0803

Blood, E., Johnson, J. W., Ridenour, L., Simmons, K., & Crouch, S. (2011). Using an iPod touch to teach social and self-management skills to an elementary student with emotional/behavioral disorders. *Education and Treatment of Children*, *34*(3), 299–321. Retrieved from http://www.jstor.org/stable/42900117

Children and Adults with Attention-Deficit/Hyperactivity Disorder (CHADD). (n.d.). *Classroom accommodations.* Retrieved July 15, 2025, from https://chadd.org/for-educators/classroom-accommodations/

Conners, C., Epstein, J., Angold, A., & Klaric, J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of Abnormal Child Psychology*, *31*, 555-562. doi: https://doi.org/10.1023/A:1025457300409

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. doi: https://doi.org/10.1207/S15327965PLI1104_01

Elliott, S. N., Braden, J. P., & White, J. L. (2001). *Assessing one and all: Educational accountability for students with disabilities*. Arlington, VA: Council for Exceptional Children. Retrieved from https://eric.ed.gov/?id=ED458746

Gaastra, G. F., Groen, Y., Tucha, L., & Tucha, O. (2016, 02). The effects of classroom interventions on off-task and disruptive classroom behavior in children with symptoms of attention-deficit/hyperactivity disorder: A meta-analytic review. *PLOS ONE*, *11*(2), e0148841. doi: https://doi.org/10.1371/journal.pone.0148841

Graham, S., Harris, K. R., & Reid, R. (1992). Developing self-regulated learners. *Focus on Exceptional children*, *24*(6), 1-16.

Hall-Ruiz, A. N. (2016). *Exploring the wandering mind in adhd* (PhD thesis, University of Michigan, University of Michigan). University of Michigan Library. Retrieved from https://deepblue.lib.umich.edu/handle/2027.42/135880

Harris, K. R., Friedlander, B. D., Saddler, B., Frizzelle, R., & Graham, S. (2005). Self-monitoring of attention versus self-monitoring of academic performance: Effects among students with ADHD in the general education classroom. *The Journal of Special Education*, *39*(3), 145-157. doi: https://doi.org/10.1177/00224669050390030201

Harrison, J. R., Bunford, N., Evans, S. W., & Owens, J. S. (2013). Educational accommodations for students with behavioral challenges: A systematic review of the literature. *Review of Educational Research*, *83*(4), 551-597. doi: https://doi.org/10.3102/0034654313497517

Harrison, J. R., Evans, S. W., Zatz, J., Mehta, P., Patel, A., Syed, M., ... Custer, B. A. (2022). Comparison of four classroom-based strategies for middle school students with ADHD: A pilot randomized controlled trial. *Journal of Attention Disorders*, *26*(11), 1507-1519. doi: https://doi.org/10.1177/10870547221081108

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review*, *19*, 429–442. doi: https://doi.org/10.1007/s10648-006-9027-y

Ketterlin-Geller, L. R. (2008). Testing students with special needs: A model for understanding the interaction between assessment and student characteristics in a universally designed environment. *Educational Measurement: Issues and Practice*, *27*, 3-16. doi: https://doi.org/10.1111/j.1745-3992.2008.00124.x

Ketterlin-Geller, L. R., Jamgochian, E. M., Nelson-Walker, N. J., & Geller, J. P. (2012). Disentangling mathematics target and access skills: Implications for accommodation assignment practices. *Learning Disabilities Research & Practice*, *27*(4), 178-188. doi: https://doi.org/10.1111/j.1540-5826.2012.00365.x

Knouse, L. E., Anastopoulos, A. D., & Dunlosky, J. (2012). Isolating metamemory deficits in the self-regulated learning of adults with ADHD. *Journal of Attention Disorders*, *16*(8), 650-660. doi: https://doi.org/10.1177/1087054711417231

Kooij, J. J. S. (2009). *Adhd bij volwassenen: diagnostiek en behandeling* (3rd ed.). Amsterdam: Pearson.

Kooij, J. J. S. (2022). *Adult adhd: Diagnostic assessment and treatment* (4th ed.). Springer Nature Switzerland AG.

Kooij, J. J. S., Buitelaar, J. K., Van den Oord, E. J., Furer, J. W., TH. Rijnders, C. A., & Hodiamont, P. P. G. (2005). Internal and external validity of attention-deficit hyperactivity disorder in a population-based sample of adults. *Psychological Medicine*, *35*(6), 817–827. doi: https://doi.org/10.1017/s003329170400337x

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. doi: https://doi.org/10.1080/01621459.1952.10483441

Lovett, B. J., & Fienup, D. M. (2023). Using interventions to increase test access: An alternative to accommodations? *Assessment for Effective Intervention*, *49*(1), 41-49. doi: https://doi.org/10.1177/15345084231178794

Mathes, M. Y., & Bender, W. N. (1997). The effects of self-monitoring on children with attention-deficit/hyperactivity disorder who are receiving pharmacological interventions. *Remedial and Special Education*, *18*(2), 121-128. doi: https://doi.org/10.1177/074193259701800206

McKnight, L. (2010). Designing for adhd in search of guidelines. In *IDC 2010 digital technologies and marginalized youth workshop* (Vol. 30).

MemoryLab. (n.d.). *Onze wetenschap — MemoryLab.* Retrieved August 9, 2024, from `https://www.memorylab.nl/en/wetenschap/`

Microsoft Corporation. (2024, 3 6). Visual Studio Code. [Computer software]. Retrieved from `https://code.visualstudio.com/`

Ministerie van Onderwijs, Cultuur en Wetenschap. (n.d.). *Wanneer kan ik een aangepast examen doen in het voortgezet onderwijs?* Rijksoverheid. Retrieved December 5, 2024, from https://www.rijksoverheid.nl/onderwerpen/eindexamens/aangepast-examen/aangepast-examen-voortgezet-onderwijs

Nationale Drug Monitor. (2025, 04 29). *ADHD-medicatie 15.2 Gebruik: volwassenen - Nationale Drug Monitor.* Trimbos-instituut & WODC. Retrieved from https://www.nationaledrugmonitor.nl/ritalin-gebruik-algemene-bevolking/

Pavlik Jr., P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, *29*(4), 559-586. doi: https://doi.org/10.1207/s15516709cog0000_14

Posit team. (2025). Rstudio: Integrated development environment for r [Computer software manual]. Boston, MA. Retrieved from http://www.posit.co/

Pritchard, A. E., Koriakin, T., Carey, L., Bellows, A., Jacobson, L., & Mahone, E. M. (2016). Academic testing accommodations for ADHD: Do they help? *Learning Disabilities: A Multidisciplinary Journal*, *21*(2), 67-78. doi: https://doi.org/10.18666/LDMJ-2016-V21-I2-7414

R Core Team. (2024). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

R Core Team. (2025). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Ramaekers, M., & van der Mooren, F. (2022, 4 4). *Aansluiting onderwijs arbeidsmarkt van afgestudeerden in het hoger onderwijs.* Centraal Bureau voor de Statistiek. Retrieved from https://www.cbs.nl/nl-nl/longread/rapportages/2022/aansluiting-onderwijs-arbeidsmarkt-van-afgestudeerden-in-het-hoger-onderwijs?onepage=true

Reeve, J., Nix, G., & Hamm, D. (2003). Testing models of the experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of educational psychology*, *95*(2), 375-392. doi: https://doi.org/10.1037/0022-0663.95.2.375

Reid, R., Trout, A., & Schartz, M. (2005). Self-regulation interventions for children with attention deficit/hyperactivity disorder. *Exceptional Children*, *71*(4), 361–377.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249-255. doi: https://doi.org/10.1111/j.1467-9280.2006.01693.x

Scheithauer, M. C., & Kelley, M. L. (2017). Self-monitoring by college students with adhd: The impact on academic performance. *Journal of Attention Disorders*, *21*(12), 1030-1039. doi: https://doi.org/10.1177/1087054714553050

Sense, F., Behrens, F., Meijer, R., & van Rijn, H. (2016). An individual's rate of forgetting is stable over time but differs across materials. *Topics in Cognitive Science*, *8*(1), 305–321. doi: https://doi.org/10.1111/tops.12183

ten Have, M., Tuithof, M., van Dorsselaer, S., Schouten, F., Luik, A. I., & de Graaf, R. (2023). Prevalence and trends of common mental disorders from 2007-2009 to 2019-2022: results from the Netherlands Mental Health Survey and Incidence Studies (NEMESIS), including comparison of prevalence rates before vs. during the COVID-19 pandemic. *World Psychiatry*, *22*(2), 275-285. doi: https://doi.org/10.1002/wps.21087

van der Velde, M., Sense, F., Borst, J., & van Rijn, H. (2021). Alleviating the cold start problem in adaptive learning using data-driven difficulty estimates. *Computational Brain & Behavior*, *4*(2), 231–249. doi: https://doi.org/10.1007/s42113-021-00101-6

van Rijn, H., van Maanen, L., & van Woudenberg, M. (2009). Passing the test: Improving learning gains by balancing spacing and testing effects. In A. Hoses, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 9th international conference of cognitive modeling* (pp. 110–115). Retrieved from https://iccm-conference.neocities.org/2009/proceedings/

cd/papers/200/paper200.pdf

Vogelgesang, K. L., Bruhn, A. L., Coghill-Behrends, W. L., Kern, A. M., & Troughton, L. C. (2016). A single-subject study of a technology-based self-monitoring intervention. *Journal of Behavioral Education*, *25*, 478–497. doi: https://doi.org/10.1007/s10864-016-9253-4

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2023). stringr: Simple, consistent wrappers for common string operations [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=stringr (R package version 1.5.1)

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). dplyr: A grammar of data manipulation [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=dplyr (R package version 1.1.4)

Wickham, H., Vaughan, D., & Girlich, M. (2024). tidyr: Tidy messy data [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=tidyr (R package version 1.3.1)

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80-83. doi: https://doi.org/10.2307/3001968

Wilens, T. E., Biederman, J., Faraone, S. V., Martelon, M., Westerberg, D., & Spencer, T. J. (2009). Presenting adhd symptoms, subtypes, and comorbid disorders in clinically referred adults with ADHD. *The Journal of clinical psychiatry*, *70*(11), 1557-1562. doi: https://doi.org/10.4088/JCP.08m04785pur

Wills, H. P., & Mason, B. A. (2014). Implementation of a self-monitoring application to improve on-task behavior: A high-school pilot study. *Journal of Behavioral Education*, *23*, 421–434. doi: https://doi.org/10.1007/s10864-014-9204-x

Appendices

A Interview questions

A.1 Breaks

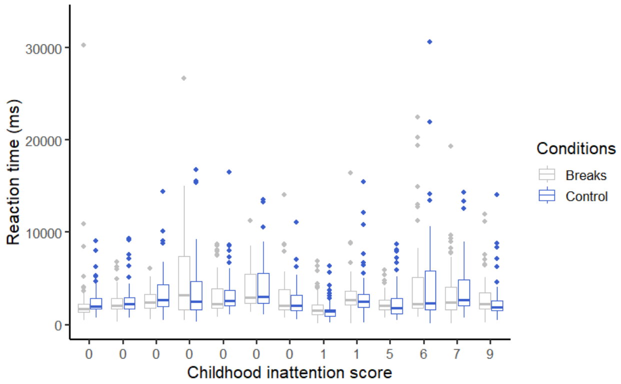
1. Welke oefening vind je fijner werken? [1 of 2, allebei/weet niet] Waarom? [Toelichting]
2. Werd je afgeleid in een van de oefeningen? [Ja, nee, weet niet] Zo ja, welke oefening en waarom? [1 of 2, allebei, weet niet, toelichting]
3. Hielpen de pauzes om je aandacht bij het leren te houden? [Ja, nee, weet niet] Waarom wel/niet? [Toelichting]
4. Wat vond je van de lengte van de pauzes? [Goed, te kort, te lang, weet niet]
5. Wat vond je van het aantal pauzes? [Goed, te weinig, te veel, weet niet]
6. Heb je van alle pauzes gebruik gemaakt (of ben je eerder doorgegaan met de taak)? [Ja, nee, weet niet] Waarom? [Toelichting]
7. Wat vond je van hoe het scherm eruit zag tijdens de pauze? [Goed, niet goed, weet niet, toelichting]
8. Wat deed je tijdens de pauze? [Nadenken over de taak, nadenken over iets anders, weggaan uit de stoel, niks, weet niet]
9. Als je de taak nog een keer zou doen, zou je die dan met of zonder pauzes willen doen? [Met, zonder, weet niet]
10. Is er nog iets anders wat je misschien kwijt zou willen dat nog niet is besproken? [Ja, nee, weet niet, toelichting]

A.2 Self-monitoring

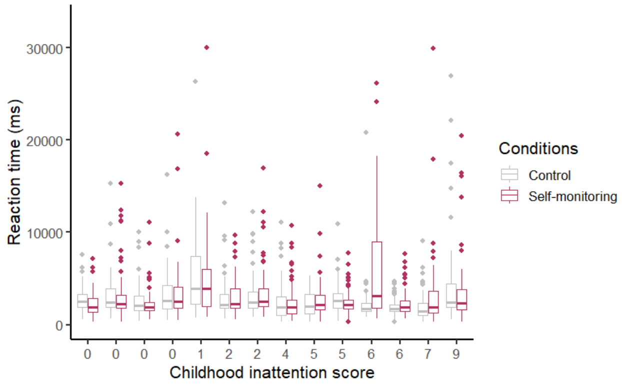
1. Welke oefening vind je fijner werken? [1 of 2, allebei/weet niet] Waarom? [Toelichting]
2. Werd je afgeleid in een van de oefeningen? [Ja, nee, weet niet] Zo ja, welke oefening en waarom? [1 of 2, allebei, weet niet, toelichting]
3. Hielpen de aandachtsvragen om je aandacht bij het leren te houden? [Ja, nee, weet niet] Waarom wel/niet? [Toelichting]
4. Er stond een tijdslimiet op het beantwoorden van de aandachtsvraag. Heb je deze gemerkt? [Ja, nee, weet niet] Zo ja, was dit [goed, te kort, te lang, weet niet]
5. Wat vond je van de manier waarop de aandachtsvraag werd gesteld? [Duidelijk, niet duidelijk, weet niet] Waarom? [Toelichting]
6. Op de pagina waarop je kon beantwoorden of je je aandacht erbij had kon je op de achtergrond je vorige antwoorden zien. Heb je dit opgemerkt? [Ja, nee, weet niet] Zo ja, wat vond je ervan dat je je vorige antwoorden kon zien? [Handig, niet handig, toelichting]

7. Wat vond je van het aantal aandachtsvragen? [Goed, te weinig, te veel, weet niet]
8. Heb je alle aandachtsvragen beantwoord? [Ja, nee, weet niet] Waarom wel, waarom niet? [Toelichting]
9. Als je de taak nog een keer zou doen, zou je die dan met of zonder de aandachtsvragen willen doen? [Met, zonder, weet niet]
10. Is er nog iets anders wat je misschien kwijt zou willen over de taak dat nog niet is besproken? [Ja, nee, weet niet, toelichting]

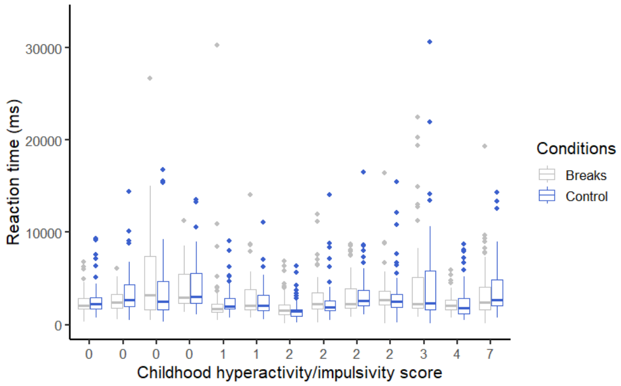
B Childhood inattention and hyperactivity/impulsivity



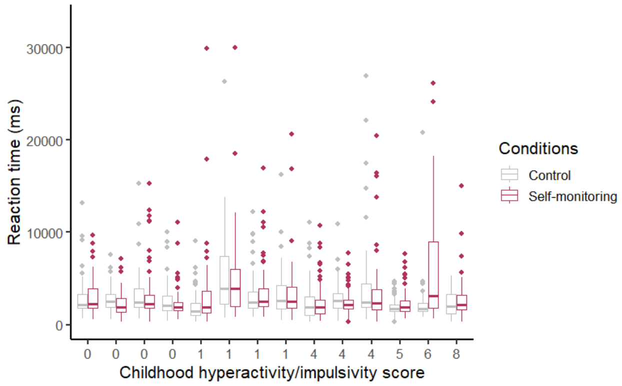
(a) Childhood inattention scores.



(b) Childhood inattention scores.

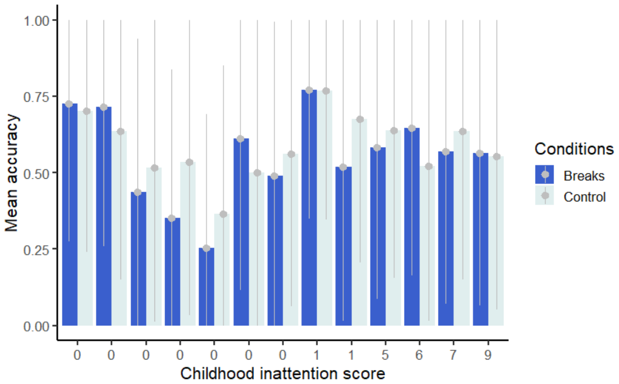


(c) Childhood impulsivity/hyperactivity scores.

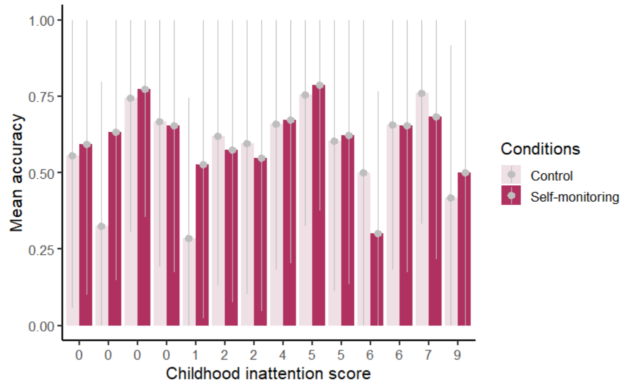


(d) Childhood impulsivity/hyperactivity scores.

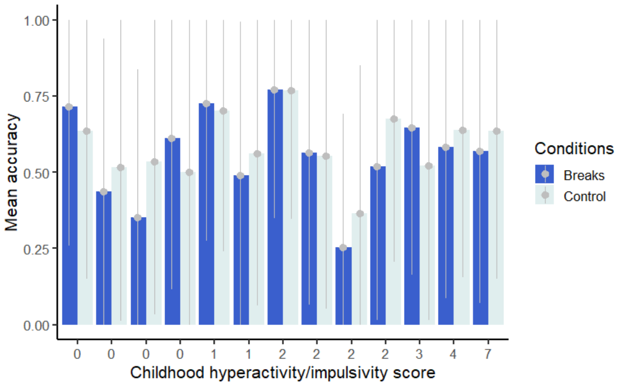
Figure 14: The reaction time for both conditions and their symptom score for each participant. Blue denotes the break condition and red denotes the self-monitoring condition.



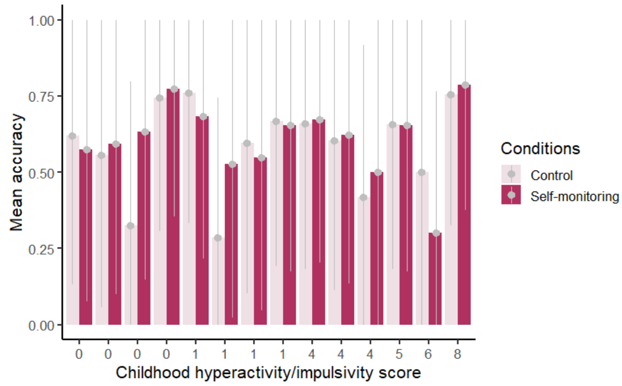
(a) Childhood inattention scores.



(b) Childhood inattention scores.



(c) Childhood impulsivity/hyperactivity scores.



(d) Childhood impulsivity/hyperactivity scores.

Figure 15: The mean accuracies and their ADHD characteristics score for each participant. Blue denotes the break condition and red denotes the self-monitoring condition. The lighter colors represent the control condition for the participant. Standard deviation has a maximum of 1 in the plot for readability, but is often larger.

C Model results adulthood

	Estimates	SE	t-value
Intercept	7.82	0.07	105.24
Break	−0.04	0.08	−0.50
Self-monitoring	−0.04	0.08	−0.44
Inattention	−0.01	0.02	−0.79
Bin 1	−0.06	0.06	−1.02
Bin 2	−0.04	0.09	−0.43
Bin 3	−0.11	0.09	−1.21
Break x inattention	0.03	0.02	1.68
Self-monitoring x inattention	0.01	0.01	0.93
Break x Bin 1	−0.03	0.08	−0.32
Self-monitoring x Bin 1	0.01	0.07	0.14
Break x Bin 2	−0.03	0.11	−0.25
Self-monitoring x Bin 2	0.01	0.10	0.09
Break x Bin 3	−0.01	0.11	−0.07
Self-monitoring x Bin 3	−0.02	0.10	−0.23
Inattention x Bin 1	−0.00	0.01	−0.19
Inattention x Bin 2	−0.02	0.01	−1.16
Inattention x Bin 3	0.01	0.01	1.04

Table 8: The estimates from the linear mixed effects model on log-transformed reaction time with inattention characteristics score, the standard errors and the t-values. The reference was the control condition.

	Estimates	SE	t-value
Intercept	7.82	0.09	87.13
Break	−0.01	0.10	−0.07
Self-monitoring	−0.01	0.09	−0.12
Hyperactivity/impulsivity	−0.01	0.02	−0.48
Bin 1	−0.02	0.06	−0.37
Bin 2	−0.08	0.10	−0.81
Bin 3	−0.10	0.10	−1.03
Break x hyperactivity/impulsivity	0.01	0.03	0.49
Self-monitoring x hyperactivity/impulsivity	0.01	0.02	0.34
Break x Bin 1	−0.03	0.08	−0.37
Self-monitoring x Bin 1	0.01	0.07	0.14
Break x Bin 2	−0.01	0.11	−0.14
Self-monitoring x Bin 2	−0.00	0.10	−0.01
Break x Bin 3	−0.02	0.11	−0.17
Self-monitoring x Bin 3	−0.02	0.10	−0.17
Hyperactivity/impulsivity x Bin 1	−0.01	0.02	−0.96
Hyperactivity/impulsivity x Bin 2	−0.00	0.02	−0.04
Hyperactivity/impulsivity x Bin 3	0.01	0.02	0.56

Table 9: The estimates from the linear mixed effects model on log-transformed reaction time with hyperactivity/impulsivity characteristics score, the standard errors and the t-values. The reference was the control condition.

	Estimates	SE	t-value
Intercept	7.79	0.06	128.45
Break	0.02	0.07	0.30
Self-monitoring	0.01	0.07	0.10
Bin 1	−0.06	0.05	−1.33
Bin 2	−0.08	0.08	−0.98
Bin 3	−0.07	0.08	−0.85
Break x Bin 1	−0.02	0.08	−0.31
Self-monitoring x Bin 1	0.01	0.07	0.12
Break x Bin 2	−0.01	0.10	−0.13
Self-monitoring x Bin 2	−0.00	0.10	−0.01
Break x Bin 3	−0.02	0.11	−0.19
Self-monitoring x Bin 3	−0.02	0.10	−0.16

Table 10: The estimates from the linear mixed effects model on log-transformed reaction time with bins, condition and their interaction the standard errors and the t-values. The reference was the control condition.

	Estimates	SE	t-value
Intercept	7.79	0.06	124.71
Break	−0.06	0.06	−1.07
Self-monitoring	−0.04	0.07	−0.58
Inattention	−0.02	0.01	−1.45
Break x inattention	0.04	0.02	1.88
Self-monitoring x inattention	0.02	0.01	1.48

(a) The estimates with inattention characteristics scores.

	Estimates	SE	t-value
Intercept	7.82	0.07	105.51
Break	−0.07	0.08	−0.79
Self-monitoring	−0.05	0.08	−0.72
Hyperactivity/impulsivity	−0.03	0.02	−1.48
Break x hyperactivity/impulsivity	0.03	0.03	1.04
Self-monitoring x Hyperactivity/impulsivity	0.03	0.02	1.48

(b) The estimates with hyperactivity/impulsivity characteristics scores.

Table 11: The estimates from the linear mixed effects models on log-transformed reaction time, the standard errors and t-values with inattention or hyperactivity/impulsivity characteristics scores. The reference was the control condition the dataset is only comprised of bin 1.