FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

VALIDATION OF A SHORTENED, ONLINE-IMPLEMENTABLE COGNITIVE LOAD TASK

COGNITIVE FATIGUE IN THE TLOADDBACK TASK

Word count: 11,423

Sofie Raeymakers
Student number: 0170 4474

Supervisor(s): Dr. Eleonore Smalle

A dissertation submitted to Ghent University in partial fulfillment of the requirements for the degree of Master of Theoretical and Experimental Psychology

Academic year: 2021 - 2022

GHENT
UNIVERSITY

Voorwoord

Deze masterproef betekend het einde van mijn opleiding Theoretische en Experimentele Psychologie. Ik zou graag mijn promotor Dr. Eleonore Smalle bedanken voor haar begeleiding van deze masterproef. Ze gaf mij steeds zeer gerichte en bruikbare feedback en ideeën, en was zowel geduldig als aanmoedigend.

Verder wil ik ook de rest van de docenten en begeleiders bedanken, want ik vond dit een hele fijne, diepgaande opleiding waar ik enorm van heb genoten.

Tenslotte wil ik ook mijn ouders en zus bedanken voor hun aanmoedigingen, steun en vertrouwen in mij.

Samenvatting

De TloadDback task is een cognitieve vermoeidheidstaak die gebruikt kan worden om proceduraal leren te verbeteren in volwassenen, zoals bij het leren van bewegingssequenties en taal (Borragán et al., 2016; Smalle et al., 2021, 2022). Doordat cognitieve vermoeidheid (CV) relatief gemakkelijk toe te brengen is, opent dit de deur voor toepassingen op grotere schaal. Echter, de huidige TloadDback taak heeft een aantal limitaties wat betreft tijdsefficiëntie, compatibiliteit met andere code, en aantrekkelijkheid. Het doel van deze studie was dan ook om een kortere versie van de TloadDback te maken, gecodeerd in Javascript, die beter geschikt is voor online toepassingen op grote schaal. Data werd verzameld van 96 participanten tijdens twee opeenvolgende dagen, in ofwel een hoge cognitieve lading (HCL) of een lage cognitieve lading (LCL) conditie. We vonden geen effect op oplettendheid, waarschijnlijk doordat dat deze versie kort en innemend is. We vonden geen consistente effecten van de taak op prestatie (gewogen accuraatheid), maar wel een verhoging in subjectieve CV. Veranderingen in subjectieve CV en prestatie waren niet geassocieerd, wat er op duidt dat prestatie mogelijks geen valide meting was van objectieve CV in deze studie. Subjectieve CV was minder groot op de tweede dag, wat aantoont dat deze taak mogelijks niet geschikt is voor herhaaldelijk gebruik. Er is nog meer onderzoek nodig om de effecten van stimulus tijdsduur en tijd-op-de-taak te onderscheiden in hun effect op de inductie van CV in de TloadDback taak. Ondanks in staat te zijn in het opwekken van subjectieve CV, is er dus nog ruimte voor verbetering in het huidige paradigma.

Trefwoorden: *TloadDback, cognitieve vermoeidheid, oplettendheid, online, stimulus tijdsduur*

Abstract

The TloadDback task has been shown to increase measures of cognitive fatigue (CF), with CF being an effective way to improve procedural learning in adults, including motor sequences and language learning (Borragán et al., 2016; Smalle et al., 2021, 2022). Since CF is relatively easy to administer, this opens the door for large-scale applications. However, the current TloadDback task has some limitations in terms of time efficiency, compatibility with other codes, and attractiveness. The goal of this study was to make a shorter version of the TloadDback task, coded in JavaScript, that is more suitable for large-scale online applications. Data was collected from 96 participants over two consecutive days, in either a high cognitive load or low cognitive load condition. We found no effect on vigilance, likely due to this version being shorter and more engaging. The task had unconsistent effects on performance (weighted accuracy in the TloadDback task itself) but was effective at inducing subjective CF. Changes in subjective CF and performance were not related, indicating that performance might not have been a valid measure of objective CF in this study. Subjective CF induction was weaker on the second day, which suggests that this task might not be suited for repeated uses. More research is needed to disentangle the effects of stimulus time duration and time-on-task in its effects on the induction of CF in the TloadDback task. While effective at inducing subjective CF, there is still room to improve the current paradigm.

Key Words: TloadDback, cognitive fatigue, vigilance, online, stimulus time duration

Table of Content

Introduction

The TloadDback task has been proven to be effective at inducing cognitive fatigue, which can be used to unlock procedural learning in adults (Borragàn et al., 2017; Borragán et al., 2016; Smalle et al., 2022). Below we will explain the mechanisms behind the induction of cognitive fatigue in the TloadDback task, the mechanisms behind and relevancy of the unlocking of procedural learning through CF, the potential and limits of the current TloadDback task, and the goals and predictions of this study.

Cognitive Fatigue

When engaging with a cognitively challenging task for a long enough time, one will inevitably get mentally exhausted. This feeling of exhaustion, called Cognitive Fatigue (CF) can be understood as an executive failure to maintain and optimize performance, independently of sleepiness (Holtzer et al., 2010). CF depletes frontoparietal resources, impairing cognitive control, high-level information processing, and the use of controlled top-down memory systems (Davranche et al., 2018; Ishii et al., 2014; Lorist et al., 2005; Tanaka et al., 2012). The result is that performance declines and becomes more variable (Trejo et al., 2005). Besides individual factors such as a person's interests, motivation and personal cost-reward predictions (Ackerman & Kanfer, 2009; Boksem et al., 2005; Boksem & Tops, 2008), two important factors in inducing CF are 1) sustained effort over time and 2) cognitive load, with higher cognitive loads inducing CF faster (Borragàn et al., 2017; Ishii et al., 2013; Tanaka et al., 2012). Cognitive load can be influenced by task complexity, such as the number of elements in an N-back paradigm (where participants need to remember if an item is the same as an n-back item in a series), forcing the working memory to keep more elements online (Kirchner, 1958). However, according to the Time Based Resource Sharing model (TBRS), the most important factor in cognitive load is the available time to process (Barrouillet et al., 2007).

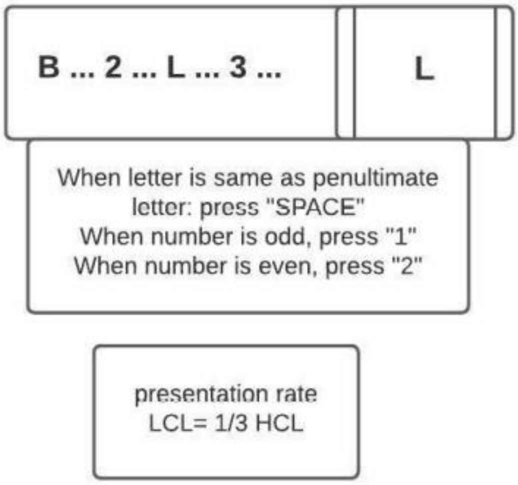
Time Load Dual Back task

To test this assumption, Borragán et al. (2017) used the Time load Dual-back (TloadDback) task, a working memory task that allows tailoring cognitive load levels to an individual's optimal performance capacity. It features two tasks with different cognitive loads: a classical dual-back working-memory updating task, which requires the working memory to keep elements online, and an odd/even task, which does not. On the screen, one of 8 types of letters or 8 types of digits is presented, with the letters for the dual-back task and the numbers for the forced-choice odd/even task. Participants have to press the space bar every time the letter is the same as the letter they saw just before or indicate whether the digit is odd or even by pressing '1' or '2', see Figure 1. To test the effect of time restraints on the induction of CF Borragán et al. (2016) created a High Cognitive Load (HCL) and a Low

Cognitive Load (LCL) condition. During a pre-task session, the maximal Stimulus Time Duration (STD) was calculated for each participant that keeps performance above 85%, with performance being a weighted formula (with the dual-back task representing 65% and the odd/even task representing 35% of the score). The presentation rate of stimuli in the LCL was then a third slower than the HCL rate: STD(LCL) = STD(HCL) + 1/3 STD(HCL).

Figure 1

TloadDback paradigm



Time pressure and under-load.

To test the effect of the cognitive load conditions on cognitive fatigue, they measured objective CF by looking at the evolution of performance (weighted accuracy) throughout the TloadDback task, and subjective CF with the Visual Analogue Scale for fatigue (VAS-f; Lee et al., 1991) before and after the TloadDback task. When time pressure was kept the same (no difference between the HCL and LCL in STD), with only task complexity differing, measures of sleepiness (as measured by the VAS-s; Ishii et al., 2014) would behave similarly to subjective CF and performance. However, when shortening the available time to process while keeping the level of complexity the same, subjective and objective CF got higher while sleepiness was not influenced. This shows a dissociation between sleepiness with subjective CF and objective CF or performance, with higher levels of CF being associated with lower levels of sleepiness dependent on the time available to process. One way to explain this dissociation is that higher cognitive demands (less time to process stimuli) trigger a higher state of arousal because more cognitive resources need to be activated. Low cognitive load might also cause sleepiness, boredom, or decreases in vigilance due to an ‘under-load’, where a low-demanding task causes cognitive resources to shut down because they are not needed (Shigihara et al., 2013). Indeed, past research has shown that under-load can lead to increased sleepiness and boredom, and decreased vigilance (Körber et al., 2015; Pattyn et al., 2008b; Richter et al., 2005), with boredom being strongly connected to vigilance (Sawin & Scerbo, 1995). Borragán et al. (2017) did not investigate post-TloadDback measures of vigilance, but did find that there were no baseline differences between days and cognitive load conditions in terms of vigilance. Vigilance in this case was measured by the Psychomotor Vigilance Task (PVT; Dinges & Powell, 1985), which is a sensitive measure of sustained attention (Sinclair et al., 2013). One important element causing under-arousal and the withdrawal of the supervisory attentional system seems to be time-on-task (the total task being longer) (Pattyn et al., 2008a) and physical fatigue (Angius et al., 2022; Lee et al., 2010). In conclusion, Borragán et al. (2017) demonstrated that the TloadDback task was effective at inducing subjective CF and objective CF (decreases in performance), with time available to process being a crucial element influencing CF, but not sleepiness, which seemed more influenced by time-on-task. Measuring post-task vigilance levels might be interesting to dissociate CF from time-on-task effects of physical fatigue and boredom.

Declarative/procedural learning

The relevancy of inducing CF with the TloadDback task comes from the framework of the declarative/procedural model (Ulman, 2015). According to this dual process theory, there are two types of learning. The first, declarative learning, is a form of explicit or rule-based learning, that requires conscious effort, and is based on prefrontal and medial temporal lobe structures (Smalle et al., 2017; Ullman, 2004). The second, procedural learning, is a form of implicit or statistical learning, is automatic and effortless, and is based on cerebellar and neocortical structures, along with the basal ganglia. (Batterink et al., 2015; Batterink & Paller, 2017; Toro et al., 2005). The declarative/procedural model can help us understand why it is that children, despite adults being stronger in most other cognitive skills, outperform adults when it comes to learning a new language (Finn, 2010, p. 20; Lenneberg, 1967; Newport, 1990). More specifically, children tend to be better at learning aspects of a language that depend on extracting patterns and regularities, such as grammar, syntax, morphology, and phonology (Gomez & Gerken, 1999; Newport et al., 2001; Saffran et al., 1996). They can learn these complex patterns without needing to be explicitly aware of them (Batterink & Paller, 2017; Smalle et al., 2018; Ulman, 2015), effortlessly (Finn et al., 2014), and without the need for consolidation and sleep (Smalle et al., 2017). This ability to extract patterns through mere passive exposure helps children with learning their mother tongue, but it also helps them in learning a second language at a young age (Johnson & Newport, 1989;

Patkowski, 1980). Considering the procedural/declarative framework, it seems like infants rely on statistical information (information about the probability of something to occur), whereas older teenagers and adults have a shift in strategy and use already learned information and rules for word segmentation (Thiessen et al., 2016). This effect can also be found in other areas of procedural learning, such as motor sequences (Adi-Japha et al., 2010; Ashtamker & Karni, 2013; Borragán et al., 2016; Janacsek et al., 2012; Wilhelm et al., 2012). We can understand this talent of children at procedural learning as follows: as children and teenagers grow up, their logic and abstract thinking develop through the maturing of frontal/medial-temporal circuits used for explicit attention (Dumontheil, 2014; Finn, 2010). This declarative system then starts to compete with the implicit learning system for cognitive resources and thus types of procedural learning such as implicit language learning become less evident (Finn et al., 2014; Smalle et al., 2017; Ulman, 2015).

Unlocking procedural learning in adults.

If this is true, this might mean that suppressing the declarative, prefrontal memory system could enhance implicit learning of a new language in adults. Past research has demonstrated that this is indeed possible, through hypnosis (Nemeth et al., 2013), TMS (Galea et al., 2010; Smalle et al., 2014, 2020, 2021, 2022), alcohol consumption (Virag, 2015), benzodiazepines (Frank et al., 2006), distractions (Foerde et al., 2006), and the induction of cognitive fatigue (Borragán et al., 2016; Delpouve et al., 2014; Smalle et al., 2021, 2022). The fact that CF can be used as a way to diminish cognitive control and enhance procedural sequence learning makes it very interesting for use in larger-scale and online studies, as well as for a potential application in language learning programs, since CF is temporary, non-invasive, and relatively easy to elicit. The TloadDback task specifically has been shown to improve procedural motor sequence learning in a serial reaction time task (Borragán et al., 2016), as well as stronger and faster learning of novel phoneme combination constraints (Smalle et al., 2021), and with above-chance distinguishing of word forms learned from a continuous speech stream (a passive listening task) with no acoustic cues for word boundaries (Smalle et al., 2022).

Limitations of the TloadDback task

The TloadDback task has thus proven to be an effective way of suppressing the declarative system to unlock infant-like procedural learning in adults. Compared to methods such as TMS, it seems the most promising method for practical applications to improve procedural learning, such as in a mobile app. However, the TloadDback task as it currently exists has some pitfalls that make it unsuitable for large-scale, online implementations such as in a language learning app. First of all, the original code (Borragán et al., 2017) is programmed in Matlab, a “high-performance language for technical computing, hosted on a programming platform designed for engineers and scientists” (*MATLAB*, 2022). Due to its annual license subscription fee, it is not an accessible programming language, and it tends to be cumbersome to integrate into online applications as it cannot interact with some hardware, browsers, and graphical interfaces (*MATLAB Online, 2022*). It is also around 20 minutes long with the STD calculation included, which is likely more than what most people are willing to spend time on before starting with learning, considering the short attention span of people online (Geri et al., 2017; Lagerstrom et al., 2015). Lastly, visual attractiveness is a key component in whether or not people take an app into use (Chang et al., 2012), and the use of black and white digits and numbers might not be the most attractive.

Creating an online-implementable TloadDback task.

The goal then of this study was then to make and validate a version of the TloadDback task that is suitable for online implementation and large-scale audiences. We thus created a version that is only 5 minutes long. Instead of odd/even digits, pictures of small or big objects are used, and instead of dual-back letters, dual-back colored balls were used. The pre-test condition during which the stimulus time duration was calculated per participant was removed. Instead, two different levels of cognitive load were created, each having a different pre-determined performance percentage. If participants performed better than this percentage, their trials gott 50ms faster. If they performed worse, their trials got 50ms slower, so they would stay around this pre-determined performance percentage. This version of the TloadDback task is thus adaptable to each participant: while the total runtime of the task is always 5 minutes, the speed differs depending on one’s performance between the two cognitive load conditions. This new version was coded in Javascript with JsPsych, which, unlike Matlab, is an open-source library that can be expanded on by the research community itself (de Leeuw, 2015). The fact that this study could be conducted entirely online not only makes it more scalable, but was also the safest option, considering that this study was conducted during the height of the COVID-19 pandemic.

Validating the induction of cognitive fatigue

To test the effectiveness of this new version, the first step was to determine whether or not it was effective at inducing CF. The TloadDback was administered twice, on two consecutive days, to be ablel to look at convergent and discriminant validity. Participants had to conduct the experiment twice and were grouped into either an HCL or LCL condition, making cognitive load a between-subjects factor instead of a within-subjects factor such as was done by Borragán et al. (2017). Because it is important to exclude learning effects or loss of efficiency over time when looking at practical applications, cognitive load was made a between-subjects factor instead of a within-subjects factor, as Borragàn et al. (2017) had done, and because we wanted to make sure that the manipulation of the STD is a crucial factor to elicit CF in this shorter version. This is also why we chose to use a continuously adapting STD, which arguably would be less prone to learning effects compared to a stable STD.

First of all, we wanted to assure baseline vigilance was not a confounding factor. Similar to Borragán et al. (2017), we took pre- TloadDback task vigilance measures with the Psychomotor Vigilance Task (PVT; Basner et al., 2011), with the prediction that we would not find differences in pre-task vigilance between the two days and two conditions. We also took post-task measures to differentiate time-on-task and boredom effects with CF effects. Because this version of the TloadDback task has a short time-on-task of just 5 minutes and is designed to be more engaging and attractive, we predicted to find no effects of the TloadDback task on vigilance. Not finding differences in vigilance is a sign that sustained attention is not affected by the TLoadDback task. Secondly, subjective (self-rated) CF was measured with the Visual Analogue Scale for fatigue (VAS-f; Lee et al., 1991) before and after the TloadDback task, with the prediction that we would find a main effect of cognitive load, a main effect of session (pre- and post-task), no effect of experimental day, and that subjective CF would be higher in the HCL condition. To look at test-retest reliability, we correlated overall subjective CF scores between the two experimental days as well as delta or difference scores (post- minus pre- TloadDback task) between the two experimental days. We predicted to find positive correlations in at least the HCL condition. Thirdly, performance, which can be seen as a measure of objective CF, was measured with a weighted accuracy score (with the dual-back task representing 65%). Our prediction here was that there would be a main effect of cognitive load, with a higher performance at the beginning of the task compared to the end, and no effect of the experimental day. Similarly to subjective CF, correlations were calculated between overall performance as well as delta or difference scores of performance. We predicted to find positive correlations in at least the HCL condition. To validate the calculation of performance with a weighted accuracy score, which assumes the dual-back task has a higher cognitive load, unweighted accuracy scores were calculated. We predicted that Components (colored dual-back balls or big/small pictures) would affect this unweighted accuracy score, with performance being lower by the end of the task, and with the colored balls having a lower mean accuracy. Lastly, to look at convergent validity, delta performance was correlated with delta subjective CF in each condition and on each day. We expected to find negative correlations here (because a positive difference score indicates subjective CF and a negative difference score indicates a drop in performance or objective CF).

Methods

Participants

A total of 96 participants were recruited for this experiment, 48 in the HCL condition (*M* Age = 19.8, *SD* = 4.8), and 48 in the LCL condition (*M* Age = 19.1, *SD* = 3.5). All participants spoke Dutch fluently, and were informed that the study was about language improvement, but did not know what condition they were in. Informed consent was obtained at the start, and a debriefing was given after the completion of the experiment. All participants were first-year Psychology students at the University of Ghent, who obtained a credit for their participation in this experiment and were recruited via Sona (FPPW Research Participation System).

Procedure

The experiment took place over two days and was conducted online on people's computers at home, between the 7th and 19th of March 2022, in the midst of the COVID-19 pandemic. Participants would receive a link to the second experimental session via email, 23 hours after completing the first session. The experiment itself was hosted by Gorilla (Anwyl-Irvine et al., 2020), in JsPsych code (de Leeuw, 2015), and all instructions and tasks were in Dutch. The order of tasks was as follows: a PVT task, a VAS-f questionnaire, two practice rounds for Colors and Pictures respectively, the shortened TloadDback task itself, and a second PVT task and VAS-f questionnaire, making for a total of around 20 minutes. Each of the two days, the order of the tasks was the same.

Measures

Vigilance.

First, each participant got to complete the psychomotor vigilance task (PVT; Dinges & Powell, 1985). Reciprocal reaction times (1/RT; Basner et al., 2011) were recorded to measure vigilance. The PVT has a high internal consistency with a good internal consistency (Cronbach's alpha =.98), and has been shown to have high reliability and validity in assessing sleep loss-induced impairments in cognitive performance (Benderoth et al., 2021). In the PVT task, participants need to press the spacebar as soon as the numbers on the screen start counting up. The start of this counting differs between trials with intervals between 1-10 seconds in a random order, each interval happening five times. If participants press the error bar too early, they get a warning. After the TloadDback task, a second, identical PVT task was administered. The Gorilla-compatible code of the PVT task which was used in this study is made available on Github (Raeymakers, 2022a).

Subjective Cognitive Fatigue.

After the PVT, the Visual Analogic Scale of Fatigue (VAS-f; Lee et al., 1991) was used to measure subjective CF pre- and post-TloadDback on both consecutive days. In this scale, participants need to indicate on a 0-10 scale (not at all – extremely) whether they feel, for example, 'tired'. It has a good internal consistency (Cronbach's alpha =.91 - .96) for healthy participants (Lee et al., 1991), and tends to perform as well or better than other similar and longer scales with respect to sensitivity to change, and correlates well with clinical variables (Wolfe, 2004). After the second PVT task, the second VAS-f questionnaire would follow.

Performance.

Objective CF was computed as the evolution of Performance throughout the TloadDback task. This was done by taking the weighted accuracy, which was computed continuously throughout the task to adjust the level of difficulty (speed) to each participant's performance. In their task, Borragán et al. (2017) divided the data into four successive time periods each including 20% of the total trials, with 20% in between. Because our experiment was much shorter, we used only two time periods, each representing the first and last 40% of the task.

Stimulus Time Duration.

Because Borragán et al. (2017) argued that STD was crucial in creating a higher cognitive load, we needed to assure that our two cognitive load conditions were indeed different in terms of STD, with STD being lower in the HCL condition. We did not look at reaction times, because participants had to react within the STD to not get an error message.

Shortened TloadDback task

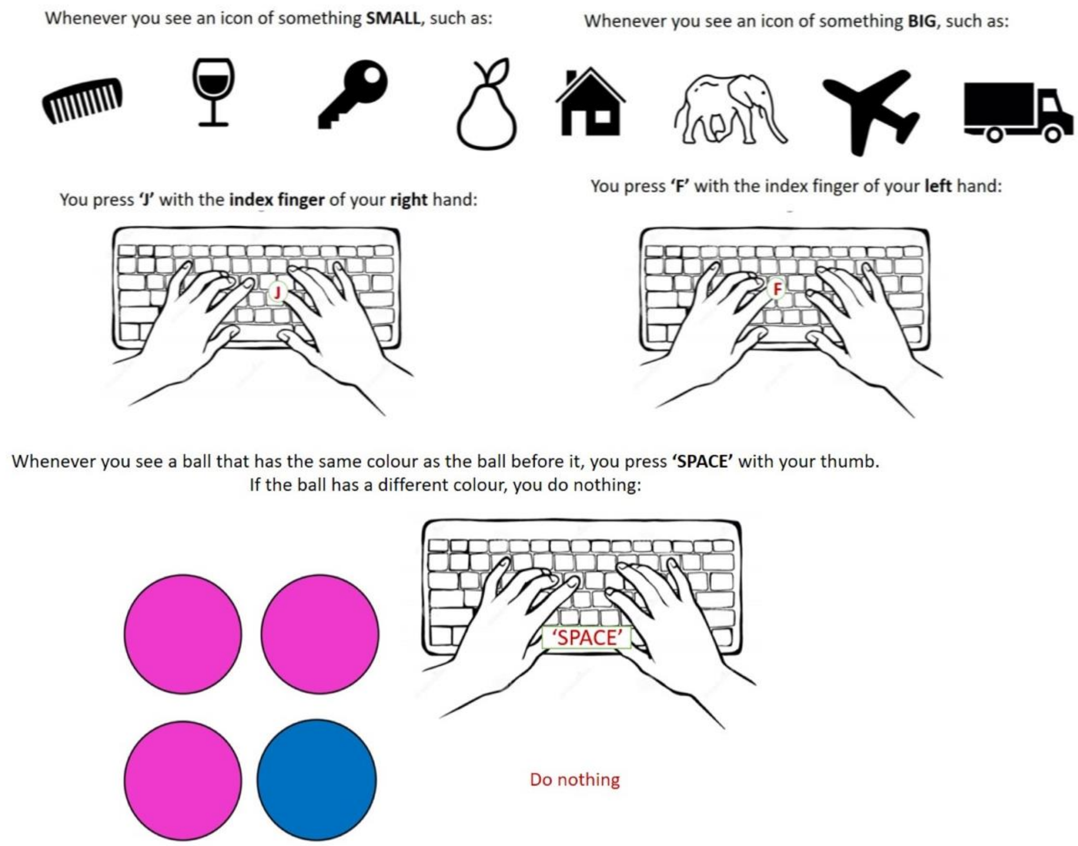
The dual-back task itself, consisting of 8 letters in the original TloadDtask by Borragán et al. (2017), was replaced by 8 different colored balls (blue, red, yellow, green, pink, purple, black ,and white), while the odd/even numbers task was replaced by a task with pictures of 'big' and 'small' objects (elephant, house, truck, plane, comb, glass, pear and key), see Figure 2. During the Colour trials, participants had to press the spacebar if the 1-back ball had the same color, and do nothing otherwise. During the Picture trials, participants had to press 'J' if the picture was of something big, and 'F' if the picture was of something small, see Figure 3. Participants would first get some instructions encouraging them to react as fast and accurately as possible. Participants then got two short (12 seconds) practice rounds for the Colors and Pictures components respectively. After the practice rounds, the trial duration would start at 1500ms. In the HCL, the pre-determined performance percentage would be set to 55%, and in the LCL to 100%. If participants reached a weighted accuracy (with Colors making up 65 % and Pictures 35%) that was higher than this pre-determined performance percentage, the next trial duration would get 50ms faster. If they went below the accuracy goal, the next trial duration would get 50ms slower. If trial duration dropped below 500ms, 50ms would be added to prevent the trials from getting too fast. If participants made a mistake, they would get an error message ('Wrong!'). These error messages were included because this version of the TloadDback task is one-fourth of the original length and we wanted to make sure participants would be maximally engaged during this shorter time frame. The task ended after 5 minutes in both cognitive load conditions. An applicable Javascript code is made available on Github (Raeymakers, 2022a). Since the experiment was hosted via Gorilla, another Gorilla-compatible code was made as well, which can be found in the same repository, along with the Color and Picture training sessions.

Figure 2

Depiction of a typical order of trials

Figure 3

Instructions for the Short TloadDback task



Data Analyses

The data in this experiment were pre-processed and analyzed using R4.2.1 (R Core Team, 2021). The pre-processing code, functions, analysis, and R markdown HTML file (Allaire et al., 2020) are all made available on Github (Raeymakers, 2022b). all ANOVA tests were conducted with linear models using the lm() function from the 'stats' package (R Core Team, 2021). To test for homogeneity of variances, Levene's test for equality of variances were conducted with the 'car' package (Fox & Weisberg, 2019). All conditions in this study had only two levels, making a violation of sphericity impossible. Alpha was set at 5%. Given our sample of 96 participants, a two-tailed t-test with 3 predictors could detect an effect size of .15 with 96% power (calculated with G*Power; Faul et al., 2007). If main effects were detected, follow-up pairwise comparisons were made between the estimated marginal means, using the 'emmeans' package (Lenth et al., 2021) with a false discovery rate to correct for multiple comparisons.

First, Vigilance was measured as the reciprocal reaction times (mean 1/RT) on the PVT task. A three-way repeated-measures ANOVA was conducted with Cognitive Load (HCL/LCL) as a between-subjects factor, and two within-subject factors, namely Session (1 and 2; pre- and post- TloadDback task), and Day (1 and 2), with Vigilance as the dependent variable. We expected to find no main effects.

We then looked at Components, meaning the stimuli in this task: Pictures (8 different pictures of small and big objects and animals) and Colors (8 different dual-back coloured balls). A two-way repeated-measures ANOVA was conducted with Time (1 and 2; the first and last 40% of the trials) and Components (Colors and Pictures) as the two within-subject factors, and Performance (the unweighted accuracy scores in the TloadDback task) as the dependent variable. We expected to find an effect of Colors and Time, with performance being lower in the Color trials.

To see if there were differences in Stimulus Time Duration (STD) between cognitive load conditions and Components, a repeated-measures ANOVA was conducted with Cognitive Load (HCL/LCL) as the between-subjects factor and STD as the dependent variable. We expected to find a main effect of Cognitive Load and that the STD would be higher in the HCL condition.

Subjective CF was measured as the mean scores on the VAS-f questionnaire. A three-way repeated-measures ANOVA was conducted, with Cognitive Load (HCL/LCL) as a between-subjects, and Day (1 and 2) and Session (1 and 2; pre- and post- TloadDback task) as the within-subjects factors, and Subjective CF as the dependent variable. We expected to find a main effect of Session and Cognitive Load with an interaction effect between Session and Cognitive Load, and no effect of Day, as well as significant differences in Subjective CF between the two cognitive load conditions, with higher subjective CF in post-task subjective CF. To look at test-retest reliability, Pearson correlation coefficients were calculated between Subjective CF scores on the two experimental days in both cognitive load conditions, and between delta or difference scores (Session 2 minus Session 1) between the two experimental days in both cognitive load conditions. We expected to find positive correlations.

To measure objective CF, Performance was calculated as the mean weighted accuracy scores 5with Colors representing 65% of the score and Pictures 35%). A three-way repeated-measures ANOVA was conducted with Cognitive Load (HCL/LCL) as a between-subjects factor, and Day (1 and 2) and Time (1 and 2; the first and last 40% of the trials) as within-subjects factors. We expected to find a main effect of Cognitive Load and Time and no effect of Day, as well as an interaction effect of Day with Time, as well as lower

Performance scores in Time 2 in the HCL condition on both days. To look at test-retest reliability, Pearson correlation coefficients were calculated between Performance scores on the two experimental days in both cognitive load conditions, and between delta or difference scores (Time 2 minus Time 1) between the two experimental days in both cognitive load conditions. We expected to find positive correlations.

Lastly, to look at convergent validity, delta Performance was correlated with Delta Subjective CF in the two experimental days and cognitive load conditions. We expected to find negative correlations (because performance being lower is an indication for objective CF, whereas the VAS-f scores being higher is an indication for subjective CF).

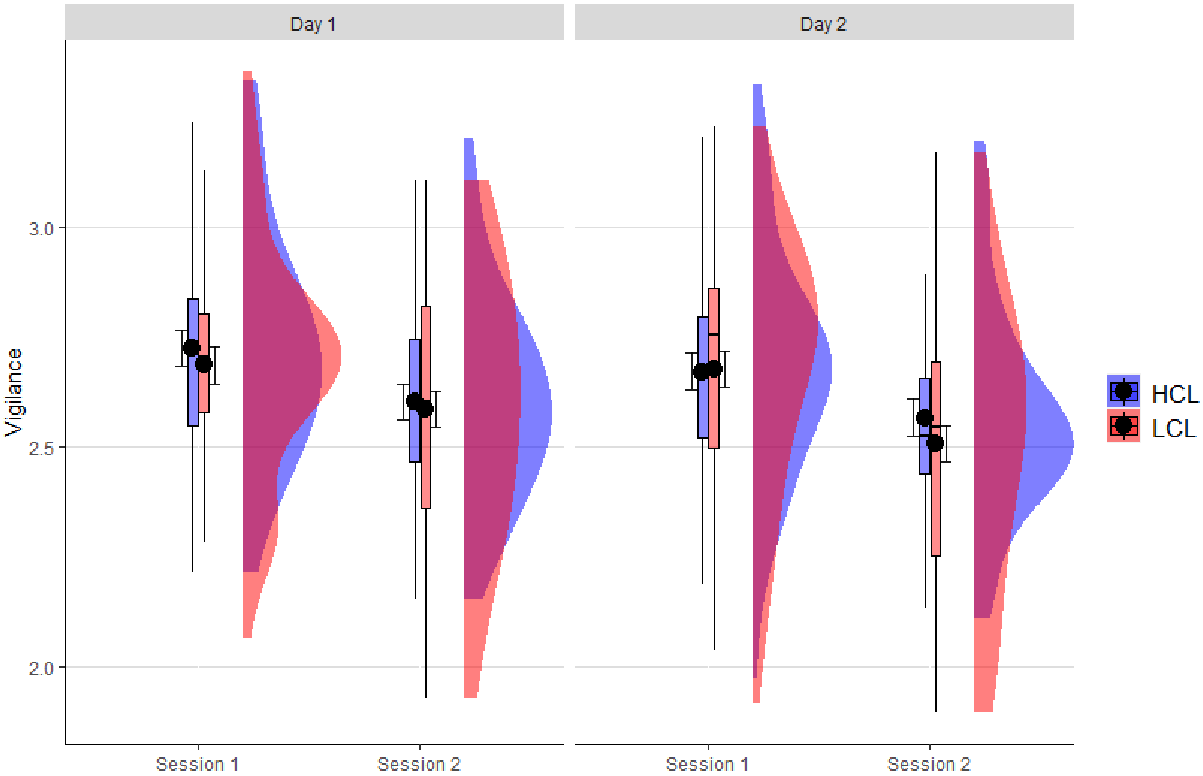
Results

Vigilance

A three-way repeated-measures ANOVA (formula: Mean Vigilance ~ Session * Cognitive Load * Day) showed no statistically significant main effect for Session (*F*(1, 376) = 1.7, *p* = .193, 95% CI [0.00, 1.00]), no significant main effect of Cognitive Load (*F(*1, 376) = .03, *p* = .873, 95% CI [0.00, 1.00]), and no significant main effect of Day (*F*(1, 376) = 1.02, *p* = .873, 95% CI [0.00, 1.00]). There were no significant interaction effects between Session and Cognitive load (*F*(1, 376) = 1.74, *p* = .188, 95% CI [0.00, 1.00]); between Session and Day (*F*(1, 376) = .13, *p* = .723, 95% CI [0.00, 1.00]); between Session and Day (*F*(1, 376) = 1.08, *p* = .298, 95% CI [0.00, 1.00]), or between Cognitive Load, Time and Day (*F*(1, 376) = .03, *p* = .874, 95% CI [0.00, 1.00]). See Figure 4 for a visualisation of the results.

Figure 4

Vigilance between the two conditions on the two experimental days



Note. Descriptive data in the form of distribution plots and boxplots of Vigilance (Mean 1/RT on the PVT task) is visualised for Day 1 and 2 in the HCL (High Cognitive Load) and LCL (Low Cognitive Load) condition, in Session 1 (pre TloadDback task) and Session 2 (post TloadDback task). The black error bars are based on the standard deviations, and the black dots represent the estimated marginal means based on the linear model.

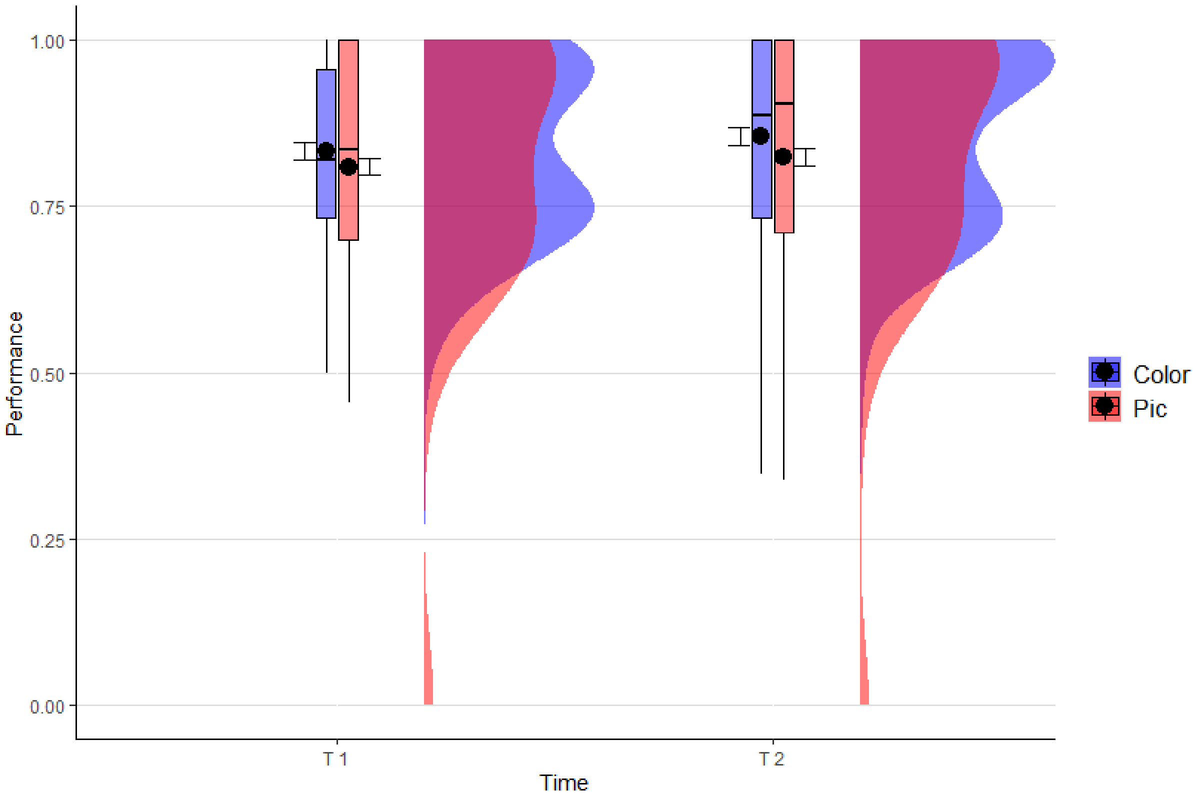
Components

A two-way repeated- measures ANOVA (formula: Performance ~ Time * Components) showed no significant main effect of Time (*F*(1, 772) = 1.98, *p* = .159, 95% CI [0.00, 1.00]), and a significant main effect of Components (*F*(1, 772) = 4.48, *p* = .035, 95% CI [2.17e-04, 1.00]). There was no significant interaction between Time and Components (*F*(1, 772) = .09, *p* = 0.764, 95% CI [0.00, 1.00]) , see Figure 5 for a visualisation.

Follow-up comparisons between the estimated marginal means showed a higher mean Performance on Color trials (*M* = .84, *SD* = .13) compared to Picture trials (*M* = .82, *SD* = .21): *t*(772) = 3.07, *p* = .002.

Figure 5

Performance between Components



Note. Descriptive data in the form of distribution plots and boxplots of Performance (mean unweighted accuracy scores) are visualised in the Color (Colour) and Pic (Picture) Condition, at Time 1 (first 40% of the trials) and Time 2 (last 40% of the trials). The black error bars are based on the standard deviations, and the black dots represent the estimated marginal means based on the linear model.

Stimulus Time Duration

A repeated-measures ANOVA (formula: STD ~ Cognitive Load) was performed, showing a statistically significant main effect of Cognitive Load ($F(1, 190) = 1161.32$, $p < .001$, 95% CI [0.87, 1.00]).

Follow-up pairwise comparisons of the estimated mean STD showed that STD was significantly lower in the HCL condition ($M = 1007.14$, $SD = 511$) compared to the LCL condition ($M = 2955$, $SD = 241$) : $t(192) = -34.12$, $p < .0001$.

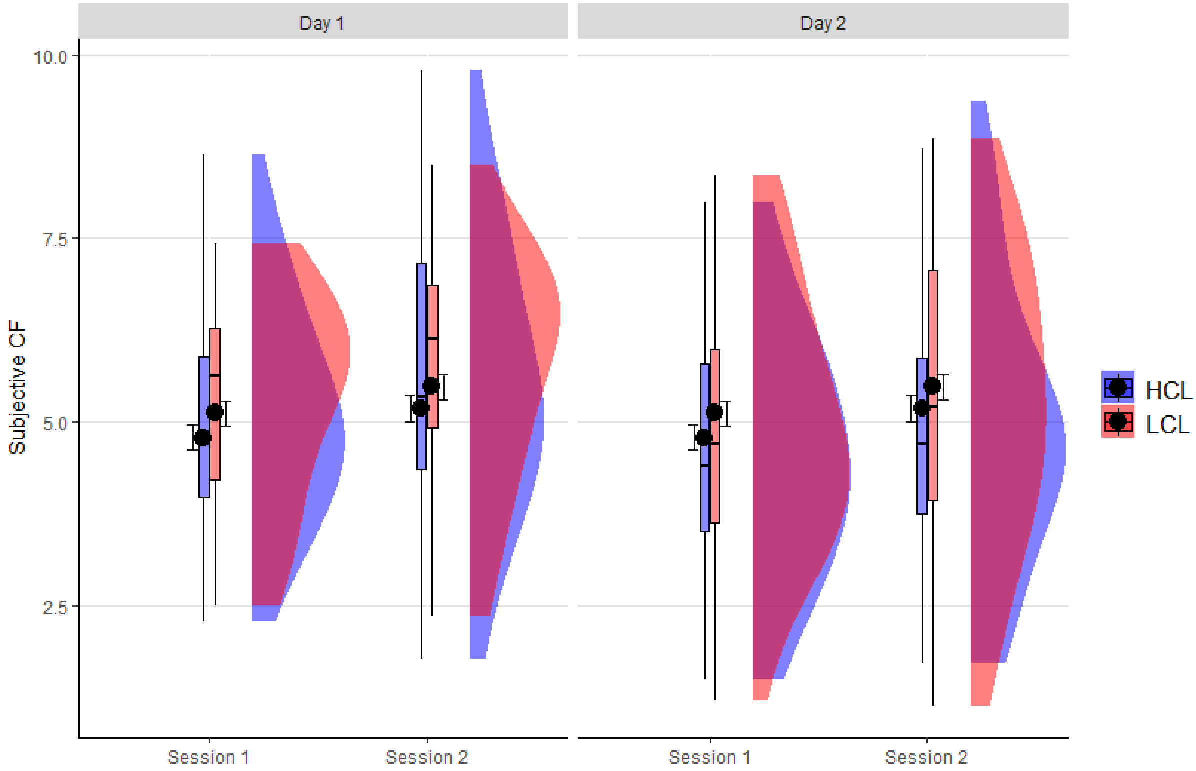
Subjective Cognitive Fatigue

A three-way repeated-measures ANOVA test (formula: Mean VAS-f ~ Session * Cognitive Load * Day) showed a significant main effect of Session ($F(1, 380) = 4.91$, $p = 0.027$, 95% CI [7.50e-04, 1.00]) , a significant main effect of Day ($F(1, 380) = 9.26$, $p = 0.003$, 95% CI [5.02e-03, 1.00]), and no significant main effect of Cognitive Load ($F(1, 380) = 3.31$, $p = 0.070$, 95% CI [0.00, 1.00]). There were no significant interactions between Session and Cognitive Load ($F(1, 380) = 0.01$, p = 0.913, 95% CI [0.00, 1.00]), between Session and Day ($F(1, 380) = 0.01$, p = 0.918, 95% CI [0.00, 1.00]), between Cognitive Load and Day ($F(1, 380) = 0.18$, p = 0.676, 95% CI [0.00, 1.00]), or between Cognitive Load, Session, and Day ($F(1, 380) = 8.36e-03$, p = 0.927, 95% CI [0.00, 1.00]). See Figure 6 for visualisations.

Because Day and Session showed significant main effects without interactions, follow-up comparisons were made between the estimated marginal means of Day and Session. Session 2 ($M = 5.33$, $SD = 1.84$) scored significantly higher on mean Subjective CF compared to Session 1 ($M = 4.96$, $SD = 1.55$), $t(380) = -5.88$, $p < .0001$. Day 1 ($M = 5.4$, $SD = 1.56$) scored significantly higher on mean Subjective CF than Day 2 ($M = 4.89$, $SD = 1.82$), $t(380) = 4.32$, $p < .0001$.

Figure 6

Subjective CF between the two conditions on the two experimental days



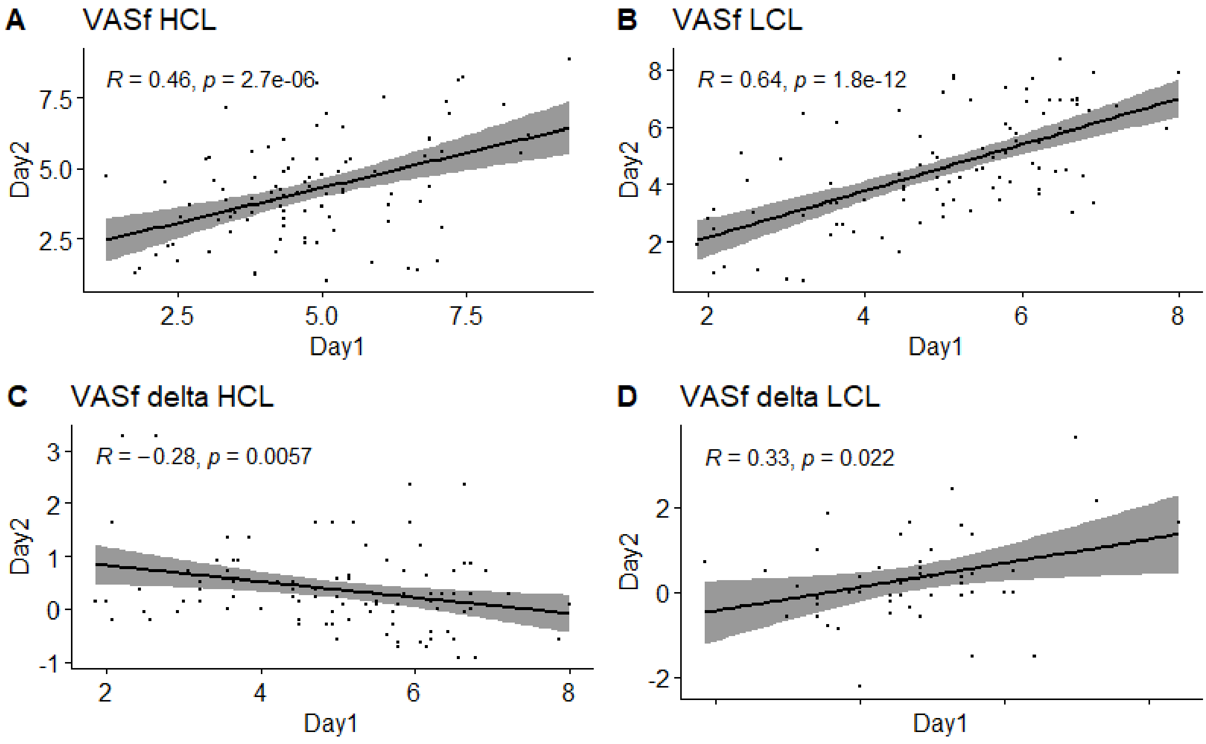
Note. Descriptive data in the form of distribution plots and boxplots of Subjective CF (mean VAS-f scores) are visualised for Day 1 and 2 in the HCL (High Cognitive Load) and LCL (Low Cognitive Load) condition, in Session 1 (pre TloadDback task) and Session 2 (post TloadDback task). The black error bars are based on the standard deviations, and the black dots represent the estimated marginal means based on the linear model.

Test-retest reliability.

There was a significant and positive correlation of Subjective CF between Day 1 and Day 2 in both the HCL condition ($r(46) = .46$, $p < .001$) and the LCL condition ($r(46) = .64$, $p < .001$), see Figure 7, A and B. There was a significant negative correlation of delta subjective CF between Day 1 and Day 2 in the HCL condition ($r(46) = -.28$, $p = .006$) and a significant positive correlation of delta Subjective CF between Day 1 and Day 2 condition ($r(46) = .33$, $p = .022$), see Figure 7, C and D.

Figure 7

Subjective CF and delta Subjective CF correlations



Performance

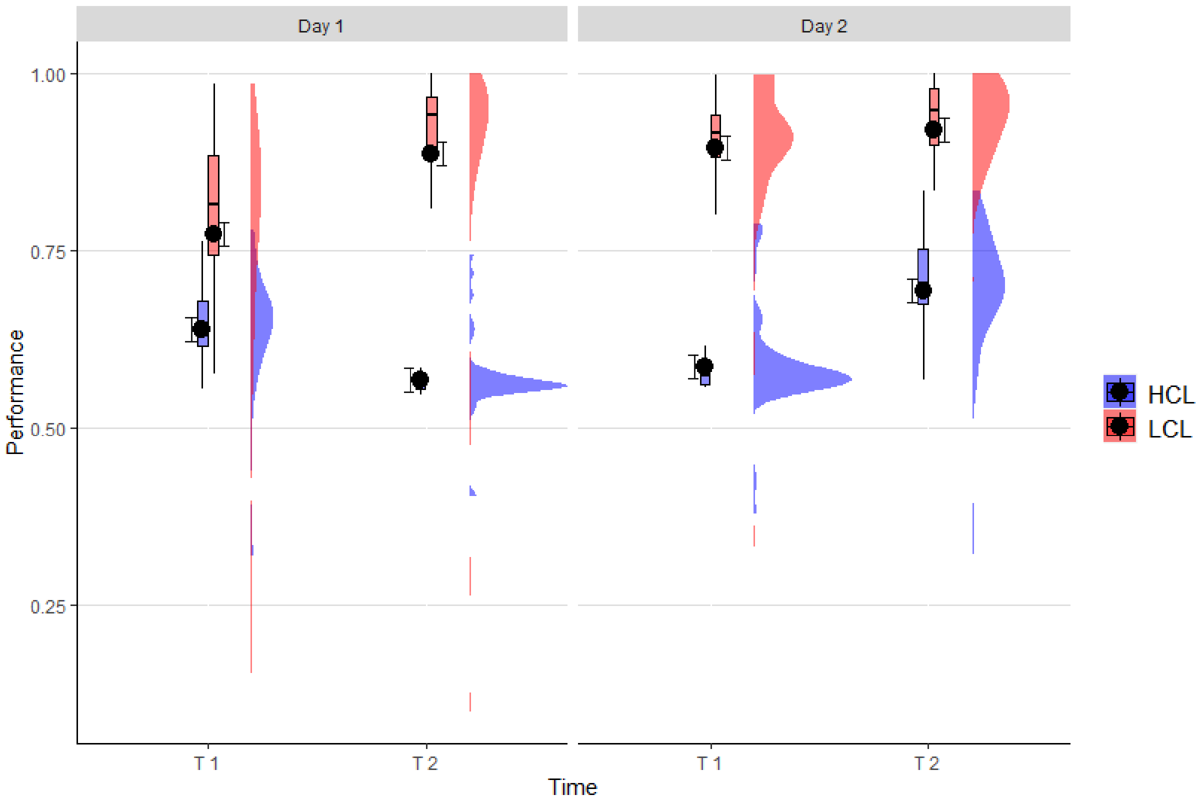
A repeated-measures ANOVA (formula: Performance ~ Time * Cognitive Load * Day) showed a statistically significant main effect of Time ($F(1, 376) = 13.38$, $p < .001$, 95% CI [0.01, 1.00]), a significant main effect of Cognitive Load ($F(1, 376) = 434.16$, $p < .001$, 95% CI [0.48, 1.00]) and a significant main effect of Day ($F(1, 376) = 23.65$, $p < .001$, 95% CI [0.03, 1.00]). There was a significant interaction between Time and Cognitive Load ($F(1, 376) = 4.85$, $p = 0.028$, 95% CI [7.08e-04, 1.00]), no significant interaction between Time and Day ($F(1, 376) = 3.58$, $p = 0.059$, 95% CI [0.00, 1.00]) no significant interaction between Cognitive Load and Day ($F(1, 376) = 2.80$, p = 0.095, 95% CI [0.00, 1.00]) and a significant interaction between Time, Cognitive Load and Day ($F(1, 376) = 31.43$, $p < .001$, 95% CI [0.04, 1.00]), see Figure 8 for a visualisation.

Because Time, Cognitive Load and Day all showed significant main effects and because there was a significant interaction between Time and Cognitive Load and between Time, Cognitive Load, and Day, follow-up comparisons were conducted between the estimated marginal means of Performance between Time 1 and Time 2 in both Cognitive Load conditions and on both Days. There was a significant decrease in Performance between Time 1 ($M = .65$, $SD = .10$) and Time 2 ($M = .56$, $SD = .05$) in the HCL condition on

Day 1: $t(376) = 3.8$, $p < .001$, a significant increase in Performance between Time 1 ($M = .58$, $SD = .07$) and Time 2 ($M = .69$, $SD = .09$) in the HCL condition on Day 2: $t(376) = -4.4$, $p < .0001$, a significant increase in Performance between Time 1 ($M = .75$, $SD = .17$) and Time 2 ($M = .89$, $SD = .17$) in the LCL condition on Day 1: $t(376) = -5.56$, $p < .001$, and no significant difference between Time 1 ($M = .89$, $SD = .12$) and Time 2 ($M = .92$, $SD = .1$) in the LCL condition on Day 2: $t(376) = -1.4$, $p = .178$.

Figure 8

Objective CF between the two conditions on the two experimental days



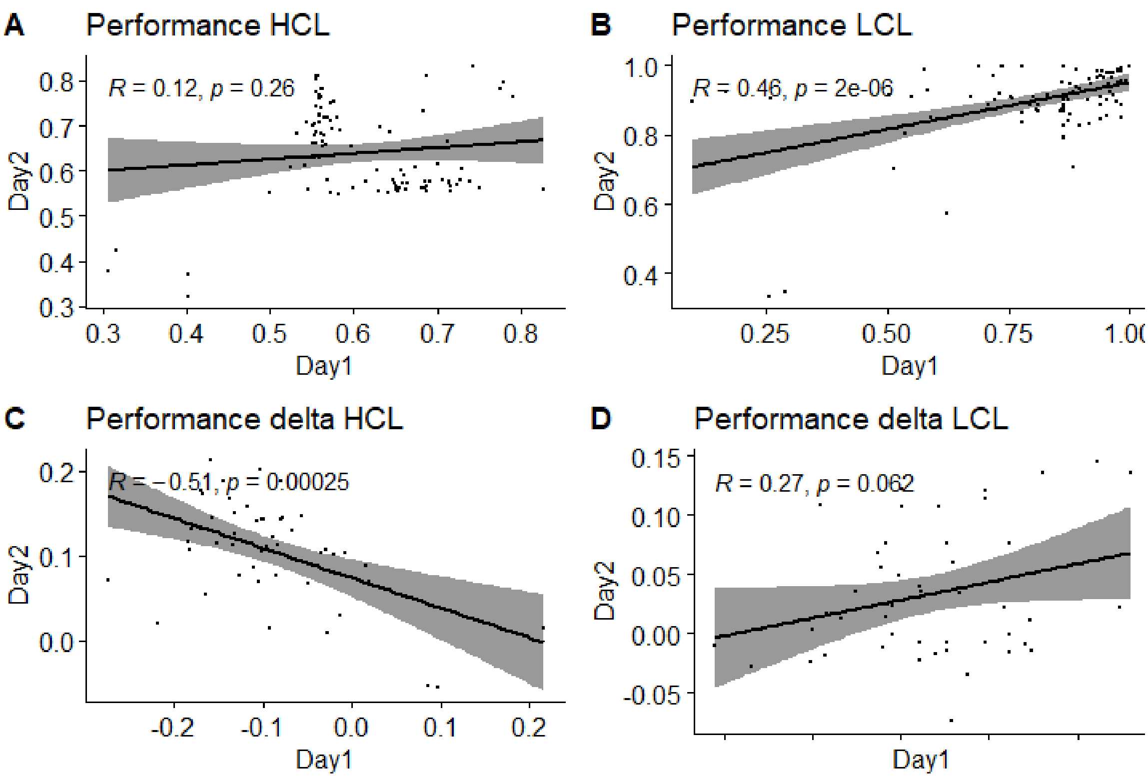
Note. Descriptive data in the form of distribution plots and boxplots of Objective CF, measured as Performance (weighted accuracy scores on the TloadDback task) are visualised for Day 1 and 2 in the HCL (High Cognitive Load) and LCL (Low Cognitive Load) condition, and at T1 (first 40% of the trials) and T2 (last 40% of the trials). The black error bars are based on the standard deviations, and the black dots represent the estimated marginal means based on the linear model.

Test-retest reliability.

There was no correlation of Performance between Day 1 and Day 2 in the HCL condition ($r(46) = .12$, $p = .26$), and a significant positive correlation of Performance between Day 1 and Day 2 in the HCL condition ($r(46) = .46$, $p < .001$), see Figure 9, A and B. There was a significant negative correlation of delta Performance between Day 1 and Day 2 in the HCL condition ($r(46) = -.51$, $p < .001$), and no correlation of delta Performance between Day 1 and Day 2 in the LCL condition ($r(46) = .27$, $p = .062$), see Figure 9, C and D.

Figure 9

Performance and delta Performance correlations

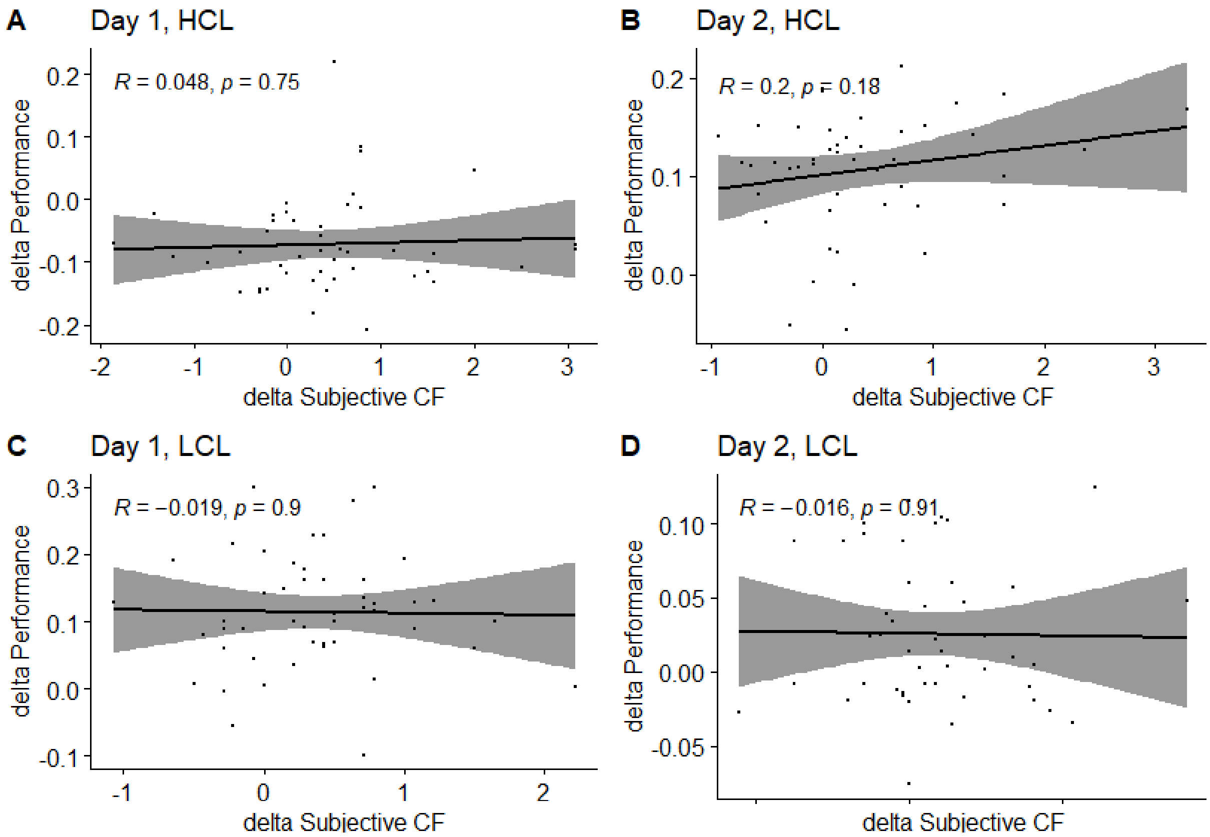


Convergent validity

Correlations were calculated between delta Subjective and delta Performance. There was no significant correlation of delta Performance with delta Subjective CF on Day 1 in the HCL condition ($r(46) = .05$, $p = .75$), no significant correlation of delta Performance with delta Subjective CF on Day 2 in the HCL condition ($r(46) = .2$, $p = .18$), no significant correlation of delta Performance with delta Subjective CF on Day 1 in the LCL condition ($r(46) = -.02$, $p = .9$), and no significant correlation of delta Performance with delta Subjective CF on Day 2 in the LCL condition ($r(46) = -.02$, $p = .91$), see Figure *10.*

Figure 10

Delta Performance and delta Subjective CF correlations



Discussion

By inducing cognitive fatigue with the TloadDback task, procedural learning can be unlocked in adults to facilitate infant-like, implicit learning of certain aspects of language, such as word segmentation (Borragán et al., 2016; Smalle et al., 2021, 2022). Because the TloadDback task is relatively easy to administer, this opens the door for practical applications to help with language learning, such as mobile apps. However, the TloadDback task as it is currently designed takes more than 20 minutes to complete, and uses black and white letters and digits as stimuli. It is thus not attractive to use for a wide audience in terms of practicality, time efficiency and visuals. In this study, a new version of the TloadDback task was created that takes only 5 minutes to complete, and uses more attractive stimuli. In the original version (Borragàn et al., 2017), the available time to process stimuli was considered a crucial element in eliciting CF. For this reason, Stimilus Time Duration (STD) was calculated before the administering of the TloadDback task as the time an individual participant needed to get at least 85% of trials right. In the shorter TloadDback task used in this study, STD was continuously adjusted according to whether or not the performance on the TloadDback task reached a certain pre-determined percentage. This way, two conditions were created: a High Cognitive Load (HCL) condition during which the task would get faster if participants weighted accuracy got above 55%, and slower if they under 55%, and a Low Cogntive Load (LCL) condition during which this percentage was set at 100%. These conditions were created to find out if differences in STD were really crucial to elicit CF in this shorter version. To make the task more appealing to a wider audience, the dual-back letters from the original version were replaced by dual-back coloured balls, and the odd/even digits forced-choice task was replaced by a big/small object picture task. Error messages were added to keep participants maximally engaged with the task. The shorter TloadDback was then completed twice by 96 participants, all first-year students studying Psychology at the University of Ghent, during two consecutive days in March, 2021. Pre- and Post- measures were taken of subjective CF and vigilance. The code was written with JavaScript, which made it possible to conduct the whole experiment online, which was safer in the midst of the COVID-19 pandemic.

Discussion of the findings

Vigilance.

Vigilance was measured as the reaction times on the Psychomotor Vigilance Task task (Dinges & Powell, 1985). The PVT task can be used to measure sleep deprivation (Basner et al., 2011; Benderoth et al., 2021) as well as boredom (Alikonis et al., 2002; Körber et al., 2015; Pattyn et al., 2008b). To look at baseline differences in vigilance between the two experimental days and cognitive load conditions before the completion of the TloadDback task, was well as effects of the TloadDback task on vigilance itself, we tested for effects of cognitive load, day and session (pre- and post-task). We found, in accordance with our prediction and similar to results found by Borragàn et al. (2017), no effects of cognitive load condition or experimental day. This excludes differences in vigilance as a confounding factor between the HCL and LCL conditions on other measures such as performance. We also found no effects of session. Because Borragàn et al. (2017) did not take post-task measures of vigilance, we cannot conclude that our version was more engaging than theirs, but these results do indicate that participants were still just as attentive and engaged by the end of the TloadDback task as they were before (Pattyn et al., 2008a).

Stimulus Time Duration.

Because STD has been argued by Borragàn et al. (2017) to be a crucial factor in eliciting CF, we wanted to make sure our task was succesfully manipulating STD between conditions. Borragàn et al. (2017) found that the mean STD (calculated in a pre-test session) in their original TloadDback task was ± 965ms. They used this STD for the HCL condition, whereas the mean STD in the LCL condition was a third slower (± 1190ms). In our experiment, the mean STD in the HCL condition was ± 890ms, and significantly lower compared to the LCL condition, which was ± 2300ms. This makes the difference in STD between our two conditions greater than the difference in the original TloadDback task, whereas our HCL STD is similar to that in the HCL condition of the experiment by Borragàn et al. (2017).

Components.

To see if the dual-back task really was more difficult, which is the reason behind the calculation of the weighted accuracy used in the original TloadDback task (Borragàn et al., 2017), we also looked at the effect of Components (dual-back Colour trials versus forced-choice Picture trials) in interaction with Time (first versus last 40% of the trials) on unweighted accuracy. We expected to find an effect on Components, with Colors having a lower average performance, and with performance getting lower during the last 40% of the trials. Components showed a significant main effect, with no interaction effect with Time, and no main effect of Time, indicating that unweighted accuracy was relatively stable across the duration of the TloadDback task. Unexpectedly, Colors showed a higher mean unweighted accuracy than Picture trials. This went against our prediction that the dual-back Color trials would be more difficult and thus have a lower unweighted accuracy, and is also in contrast with the findings of Borragàn et al. (2017), who found a main effect of both Time and Components and an interaction effect between Time and Components, with lower performance scores for the dual-back digits task compared to the odd/even letters task, indicating a higher demand on the working memory.

Subjective Cognitive Fatigue.

Measures of Subjective CF were taken before and after each of the two TloadDback tasks with the Visual Analogue Scale fore for fatigue (VAS-f; Lee et al., 1991). Our hypothesis was that the short TloadDback task would influence subjective CF, and that this effect would be higher in the HCL condition, with no significant difference between the two experimental days. We found a main effect of experimental on subjective CF, no main effect of cognitive load, and no interactions. This contrasted our hypothesis as well as the findings by Borragàn et al. (2017), since we expected an effect of cognitive load and no effect of day. It is in line with findings by Smalle et al. (2022) though, who also found no difference between the HCL and LCL in subjective CF. Moreover, participants scored significantly lower on subjective CF on the second day. This indicates that the task gets less efficient at inducing subjective CF upon repetition, which is problematic when wanting to create a CF manipulation that can be used in practical applications for learning. Overall, the TloadDback is able to induce subjective CF, with no differences between the cognitive load conditions, but this effect gets lower on the second day.

Positive correlations were found in subjective CF between both days in both cognitive load conditions. This means that if subjective CF was high overall (over the two sessions) on one day, it was more likely to also be high on the other day, and vice versa, indicating that subjective CF was rather stable throughout the experiment. However, delta or difference scores of subjective CF (post- minus pre-test) correlated negatively between the two days in the HCL condition, and positively in the LCL condition. In other words, an increase in (post-task) CF on the first day was more likely to followed by an increase in (post-task) CF on the second day in the LCL condition, and more likely to be followed by a decrease in the HCL condition, and vice versa. It thus seems that, while the overall subjective CF scores correlate positively between days over the two conditions, the increase in subjective CF over the two days is more consistent in the LCL condition.

Performance.

Performance was measured as the weighted accuracy of participants on the TloadDback task, with the dual-back task representing 65% of the score and the forced-choice task 35% in order to look at objective CF. The hypothesis was that mean performance would be higher for the last 40% of the trials compared to the first 40%, and that this difference would be more pronounced in the HCL condition, with no differences between the experimental days. We found a main effect of Time (first versus last 40% of the trials), Cognitive Load and Day, with interactions between Time and Cognitive Load and between Time, Cognitive Load an Day. The main effects of Time and Cognitive Load were expected, as well as their interaction, but the main effect of and interaction with Day was not.

Besides that, some unexpected directions of differences were found between the means: performance increased between the first and last 40% of the trials in the LCL condition during the first day, but showed no difference during the second day. In the HCL condition, performance decreased during the first day and increased during the second day. This makes it seem like there is only an effect of objective CF in the HCL condition during the first day, and a seemingly learning effect or improvement in performance in the LCL condition on the first day and in the HCL condition on the second day.

Performance did not correlate between the two experimental days in the HCL condition, and had a positive correlation in the LCL condition. This means that if performance was high over the whole TloadDback session on the first day, it would also be high on the second day and vice versa in the LCL condition. Delta or change (mean performance in the last 40% of the trials minus the first 40%) correlated negatively between the two experimental days in the HCL condition while having a positive correlation in the LCL condition. This can be explained by the effect discussed above, which is that the performance difference is opposite between the first and the second day in the HCL condition.

Convergent validity.

Because participants might not always be (subjectively) aware of their CF, Borragàn et al. (2017) compared post-task differences or delta scores of subjective CF with the within-task evolution or delta scores of performance. They found correlations indicating that higher subjective CF post-task was related with higher performance differences between the HCL and LCL conditions, but only in the last 40% of the TloadDback experiment. Because, unlike Borragàn et al. (2017), cognitive load was a between-subjects factor and not a within-subjects factor in this experiment, we conducted slightly different correlations, since we cannot match participants in both the HCL with the LCL condition. We found no significant correlations between delta subjective CF and delta performance in either of the cognitive load conditions or days. This indicates that subjective CF and performance increases or decreases are not related in this experiment.

Effectiveness of the TloadDback task

The shorter version of the TloadDback task tested in this experiment was effective at inducing subjective CF, but showed no difference between the HCL and LCL condition. Subjective CF was lower on the second day, indicating that the CF effect gets lower with repetition. Performance only showed a decrease in the HCL condition on the first day. Overall, the LCL condition, surprisingly, seemed to induce CF more reliably than the HCL condition, and we found no differences in unweighted performance between Colour and Picture trials. Changes in subjective CF were not associated with performance, which might mean they are measuring something different in this experiment. Some prior studies have also found an absence of correlations between subjective CF and objective behavioural measures, showing that fatigue can influence performance without affecting subjective CF and vice versa (Ackerman & Kanfer, 2009; Bailey et al., 2007; Bryant et al., 2004; Krupp & Elkins, 2000; Lim et al., 2010; Schwid et al., 2003). This is possible if participants are not aware of their performance impairment, or when they feel tired in ways that don't affect their performance. However, the nature of this adjusted TloadDback task might also be an explanation. More specifically, performance might not have been a valid measure of objective CF. Below we discuss some limitations of this study and their relation to our findings.

Limitations of this study.

First of all, and probably most importantly, the duration of our task (5 minutes) was much shorter than the original task (20 minutes with STD calculation included). By design, our task might not be long enough to induce objective CF in the form of a lower performance (Gui et al., 2015; Lim et al., 2010). It could be argued that a shorter version could be made more exhausting by speeding up the task even further. In this study, there was a minimum STD of 500ms to avoid the task speeding up so much that participants wouldn't be able to follow. In the original version (Borragàn et al., 2017), the STD ranged from 600 to 1400ms. When making the STD shorter however, there should be thought well about how much time will be allowed for participants to react. Some other studies do not include the timeframe during which participants are allowed to react in the STD (Gajewski et al., 2018). Making the STD might make the task too fast which might bring some new problems, such as participants giving up.

Our sample of first year students studying psychology, which are predominantly female and young (Schelfhout et al., 2021), could be argued to be biased and thus influencing our findings. However, it cannot explain the differences between our findings and those of Borragàn et al. (2017), since they sampled from a similar population of psychology students.

This experiment was conducted online, which adds ssome variability to the environments in which participants conduct the experiment. On the other hand, it also makes the experiment more ecologically valid, which is important when looking testing its effectiveness in the context of practical applications.

Another limitation might be the inclusion of error messages, which were added to keep this shorter version maximally engaging. These might have worked against the cognitive load effect that was intended to be induced. Because the HCL is faster and participants will make more mistakes by design (which, as shown by the overall lower mean performance, was an effective manipulation), the HCL also includes more error messages. These error messages take away from total the amount of trials, and they might also lower time pressure, as they add a little pause after every mistake. This might partially explain why we didn't find an effect of cognitive load on subjective fatigue, and why the HCL condition overall had unreliable results over the two days.

We also found that there were no differences in terms of unweighted performance between the small/big forced-choice Picture trials and the dual-back Color trials. It is possible that in our version, Pictures were more difficult because they consist of different categories (objects, buildings, animals) compared to the odd/even letters. Perhaps the fact that participants had to press 'F' if the picture was of something big and 'J' if it was of something small might also have been counter-intuitive due to people having the tendency to place size on an imaginary left-to-right horizontal line in left-to-right reading cultures (Afsari et al., 2016; Hesse & Bremmer, 2017). Alternatively, the colored balls might be easier to remember than the black and white dual-back letters, because colors can enhance memory (Dzulkifli & Mustafar, 2013). Either way, it seems like performance does not need to be calculated by weighted accuracy where the dual-back task gets a bigger weight, since there is no evidence in this experiment that the dual-back task has a higher cognitive load in this version of the TloadDback task. Our replacement of the original dual-back letters and odd/even numbers by dual-back colored balls and big/small pictures might deviate too much from the original task. Because this shorter version TloadDback task has more attractive stimuli (coloured balls and pictures), it might be more motivating. According to the motivational control theory (Hockey, 2011), cognitive fatigue is a problem of management rather than of energy. It is possible that the lack of consistent performance differences in this experiment is due to the task being too engaging, as indicated by the lack of effects found on vigilance. The lack of effects on vigilance might also be explained as a result of a too short total duration of the TloadDback task, since task duration tends to influence vigilance as measured by the PVT task in similar ways to sleep deprivation. However, even with longer task durations, motivation is still an important factor influencing the amount of decrease in vigilance (Möckel et al., 2015).

Lastly, he calculation of the STD was done before the actual TloadDback task in the version by Borragàn et al. (2017), and kept constant throughout the task itself. In our experiment, STD was continuously adjusted to the performance of participants, meaning that the timing would sometimes get faster, sometimes slower. This might then not only add more variability in terms of STD, but also keeps the performance relatively stable by design. We still see significant differences between the first 40% and last 40% of the trials, but these differences might be less meaningful due to the setup of this experiment. This could also explain the unexpected findings in terms of increases/decreases of performance between the different conditions and experimental days. Because STD differences were even bigger in this version of the TloadDback task compared to the original version, average STD differences seem unlikely to be the reason for differences in findings. In conclusion, it is possible that performance might not have been a valid measure of objective CF for this shorter version of the TloadDback task.

Recommendations for future research

The problem of how to apply research findings that procedural learning can be unlocked in adults to help them learn certain aspects of language in infant-like ways into large-scale applications, has yet to be resolved. This study demonstrates the difficulty of creating ways to implement fundamental research to designs and paradigms that can be used in a real-life context, because doing so often adds extra variables. Our results suggest that not just time pressure, but also the length of the task is important in inducing CF in the TloadDback task, which has been shown in the past to be able to unlock implicit learning in adults (Borragán et al., 2016; Smalle et al., 2022). Future research in this direction should also try and determine in more detail which elements of the TloadDback are crucial to induce which types of CF, as well as the minimum task duration necessary to induce changes in performance. We did not find reliable evidence for differences in performance induced by this shorter version of the TloadDback task, and no differences in subjective CF between different cognitive load conditions. We did however find an effect of the TloadDback task on subjective CF, although the HCL condition gave less reliable results. As discussed above, the lack of consistent findings for performance might be explained due to some limitations in the setup of this shorter version, specifically, the continuous adjustment of STD to performance, making performance a less valid measure of objective CF here, whereas the inclusion of error messages might explain the less reliable results in the HCL condition by effectively evening out the time pressure between the two cognitive load condition by adding small pauses after every mistake. In this light, it might be interesting to look into using the original calculation of a STD per participant beforehand and to keep STD constant throughout the task in order to be able to use performance as an indication of objective CF more reliably. This might make the task less suitable for repeated uses however, which is why it is probably wiser to use other measures of objective CF. The error messages might need to be removed, or the cognitive load conditions should be evened out by having every correct trial be followed by a ‘Correct!’ message. Future research should also look into whether or not the induction of subjective CF alone is enough to unlock implicit learning, or whether changes in performance or other objective measures of CF are necessary as well, and more closely disentangle effects on vigilance with effects on performance and subjective CF. Specifically, it should be studied whether or not making the dual-back and forced-choice tasks more engaging influences the induction of CF or not, since motivation could be a factor in the induction of CF. We should also consider the possibility that perhaps the TloadDback task and the induction of CF are not the most convenient ways to put the unlocking of procedural learning into practice after all. Other options that might be appropriate to investigate for uses in large-scale practical applications such as mobile apps to unlock procedural learning should be considered, such as distractions (Foerde et al., 2006; Lavie, 2010), and hypnosis (Daltrozzo & Valdez, 2018; Nemeth et al., 2013).

Conclusion

In this study, a shorter version of the TloadDback was designed with the goal of making a more attractive, engaging, and time-efficient paradigm to induce CF, suitable for large-scale online applications. We found that vigilance was not influenced by cognitive load, day or the TloadDback task itself. This could be because the current TloadDback task is shorter and more engaging, and vigilance is tied to time-on-task and boredom. The TloadDback task increased subjective CF but this effect got less on the second day, and there were no differences in subjective CF between the HCL condition and LCL condition. Correlations between changes in subjective CF were only positive between the two days in the LCL condition. The performance got lower on the first day in the HCL condition but increased or stayed the same in the LCL condition and/or on the second day. Changes in subjective CF were not correlated to changes in performance. The lack of consistent findings of the TloadDback task on performance might be due to the change from the original TloadDback paradigm to this shorter version, which uses a continuous adjustment of the stimulus time duration to performance, making performance a less valid measure of objective CF in this study. The lack of reliable results in the HCL condition might be explained by the inclusion of error messages in this version of the TloadDback task, which might even out the time pressure between the two cognitive load conditions by adding small pauses after every mistake. Overall, this version of the TloadDback task was effective at inducing subjective CF, with no consistent effects on performance, and with the effect on subjective CF getting less on the second day, which could be problematic for repeated uses of this shorter TloadDback task. More research is needed on whether subjective CF is enough to unlock procedural learning, and on different paradigm variations for a shorter TloadDback task.

References

Ackerman, P. L., & Kanfer, R. (2009). Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions. *Journal of Experimental Psychology: Applied*, *15*(2), 163–181. https://doi.org/10.1037/a0015719

Adi-Japha, E., Berberich-Artzi, J., & Libnawi, A. (2010). Cognitive Flexibility in Drawings of Bilingual Children. *Child Development*, *81*, 1356–1366. https://doi.org/10.1111/j.1467-8624.2010.01477.x

Afsari, Z., Ossandón, J. P., & König, P. (2016). The dynamic effect of reading direction habit on spatial asymmetry of image perception. *Journal of Vision*, *16*(11), 8. https://doi.org/10.1167/16.11.8

Alikonis, C., Warm, J., Matthews, G., Dember, W., Hitchcock, E., & Kellaris, J. (2002). Vigilance, Workload, and Boredom: Two Competing Models. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*, 1531–1535. https://doi.org/10.1177/154193120204601701

Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., & Iannone, R. (2020). *rmarkdown: Dynamic Documents for R*. https://rmarkdown.rstudio.com.

Angius, L., Merlini, M., Hopker, J., Bianchi, M., Fois, F., Piras, F., Cugia, P., Russell, J., & Marcora, S. M. (2022). Physical and Mental Fatigue Reduce Psychomotor Vigilance in Professional Football Players. *International Journal of Sports Physiology and Performance*, *1*(aop), 1–8. https://doi.org/10.1123/ijspp.2021-0387

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Ashtamker, L., & Karni, A. (2013). Motor memory in childhood: Early expression of consolidation phase gains. *Neurobiology of Learning and Memory*, *106*, 26–30. https://doi.org/10.1016/j.nlm.2013.07.003

Bailey, A., Channon, S., & Beaumont, J. G. (2007). The relationship between subjective fatigue and cognitive fatigue in advanced multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *13*(1), 73–80. https://doi.org/10.1177/1352458506071162

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Basner, M., Mollicone, D., & Dinges, D. F. (2011). Validity and Sensitivity of a Brief Psychomotor Vigilance Test (PVT-B) to Total and Partial Sleep Deprivation. *Acta Astronautica*, *69*(11–12), 949–959. https://doi.org/10.1016/j.actaastro.2011.07.015

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory*, *22*(11), 544–556. https://doi.org/10.1101/lm.037986.114

Batterink, L., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. https://doi.org/10.1016/j.cortex.2017.02.004

Benderoth, S., Hörmann, H.-J., Schießl, C., & Elmenhorst, E.-M. (2021). Reliability and validity of a 3-min psychomotor vigilance task in assessing sensitivity to sleep loss and alcohol: Fitness for duty in aviation and transportation. *Sleep*, *44*(11), zsab151. https://doi.org/10.1093/sleep/zsab151

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Brain Research. Cognitive Brain Research*, *25*(1), 107–116. https://doi.org/10.1016/j.cogbrainres.2005.04.011

Boksem, M. A. S., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, *59*(1), 125–139. https://doi.org/10.1016/j.brainresrev.2008.07.001

Borragán, G., Peigneux, P., & Slama, H. (2017). *Time load dual back task (TloadDback)*. https://osf.io/ay6er/

Borragàn, G., Slama, H., Bartolomei, M., & Peigneux, P. (2017). Cognitive fatigue: A Time-based Resource-sharing account. *Cortex*, *89*, 71–84. https://doi.org/10.1016/j.cortex.2017.01.023

Borragán, G., Slama, H., Destrebecqz, A., & Peigneux, P. (2016). Cognitive Fatigue Facilitates Procedural Sequence Learning. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00086

Bryant, D., Chiaravalloti, N. D., & DeLuca, J. (2004). Objective Measurement of Cognitive Fatigue in Multiple Sclerosis. *Rehabilitation Psychology*, *49*(2), 114–122. https://doi.org/10.1037/0090-5550.49.2.114

Chang, T.-R., Kaasinen, E., & Kaipainen, K. (2012). What influences users' decisions to take apps into use? A framework for evaluating persuasive and engaging design in mobile Apps for well-being. *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, 1–10. https://doi.org/10.1145/2406367.2406370

Daltrozzo, J., & Valdez, G. E. (2018). ENHANCING IMPLICIT LEARNING WITH POSTHYPNOTIC SUGGESTION: AN ERP STUDY. *The International Journal of Clinical and Experimental Hypnosis*, *66*(2), 174–210. https://doi.org/10.1080/00207144.2018.1421358

Davranche, K., Tempest, G. D., Gajdos, T., & Radel, R. (2018). Impact of Physical and Cognitive Exertion on Cognitive Control. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02369

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Delpouve, J., Schmitz, R., & Peigneux, P. (2014). Implicit learning is better at subjectively defined non-optimal time of day. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *58*, 18–22. https://doi.org/10.1016/j.cortex.2014.05.006

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652–655. https://doi.org/10.3758/BF03200977

Dumontheil, I. (2014). Development of abstract thinking during childhood and adolescence: The role of rostrolateral prefrontal cortex. *Developmental Cognitive Neuroscience*, *10*, 57–76. https://doi.org/10.1016/j.dcn.2014.07.009

Dzulkifli, M. A., & Mustafar, M. F. (2013). The Influence of Colour on Memory Performance: A Review. *The Malaysian Journal of Medical Sciences : MJMS*, *20*(2), 3–9.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Finn, A. S. (2010). *The sensitive period for language acquisition: The role of age related differences in cognitive and neural function* [UC Berkeley]. https://escholarship.org/uc/item/44j0q2fv

Finn, A. S., Lee, T., Kraus, A., & Hudson Kam, C. L. (2014). When it hurts (and helps) to try: The role of effort in language learning. *PLoS One*, *9*(7). https://doi.org/10.1371/journal.pone.0101806

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the*

United States of America, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Fox, J., & Weisberg, S. (2019). *An R Commpanion to Applied Regression*.

Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychological Science*, *17*(8), 700–707. https://doi.org/10.1111/j.1467-9280.2006.01769.x

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What Does the n-Back Task Measure as We Get Older? Relations Between Working-Memory Measures and Other Cognitive Functions Across the Lifespan. *Frontiers in Psychology*, *9*. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02208

Galea, J. M., Albert, N. B., Ditye, T., & Miall, R. C. (2010). Disruption of the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills. *Journal of Cognitive Neuroscience*, *22*(6), 1158–1164. https://doi.org/10.1162/jocn.2009.21259

Geri, N., Winer, A., & Zaks, B. (2017). Challenging the six-minute myth of online video lectures: Can interactivity expand the attention span of learners? *Online Journal of Applied Knowledge Management (OJAKM)*, *5*(1), 101–111. https://doi.org/10.36965/OJAKM.2017.5(1)101-111

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135. https://doi.org/10.1016/S0010-0277(99)00003-7

Gui, D., Xu, S., Zhu, S., Fang, Z., Spaeth, A. M., Xin, Y., Feng, T., & Rao, H. (2015). Resting Spontaneous Activity in the Default Mode Network Predicts Performance Decline during Prolonged Attention Workload. *NeuroImage*, *120*, 323–330. https://doi.org/10.1016/j.neuroimage.2015.07.030

Hesse, P. N., & Bremmer, F. (2017). The SNARC effect in two dimensions: Evidence for a frontoparallel mental number plane. *Vision Research*, *130*, 85–96. https://doi.org/10.1016/j.visres.2016.10.007

Hockey, G. R. J. (2011). A motivational control theory of cognitive fatigue. In *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 167–187). American Psychological Association. https://doi.org/10.1037/12343-008

Holtzer, R., Shuman, M., Mahoney, J. R., Lipton, R., & Verghese, J. (2010). Cognitive Fatigue Defined in the Context of Attention Networks. *Aging, Neuropsychology, and Cognition*, *18*(1), 108–128. https://doi.org/10.1080/13825585.2010.517826

Ishii, A., Tanaka, M., Shigihara, Y., Kanai, E., Funakura, M., & Watanabe, Y. (2013). Neural effects of prolonged mental fatigue: A magnetoencephalography study. *Brain Research*, *1529*, 105–112. https://doi.org/10.1016/j.brainres.2013.07.022

Ishii, A., Tanaka, M., & Watanabe, Y. (2014). Neural mechanisms of mental fatigue. *Reviews in the Neurosciences*, *25*(4), 469–479. https://doi.org/10.1515/revneuro-2014-0028

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Körber, M., Schneider, W., & Zimmermann, M. (2015). Vigilance, boredom proneness and detection time of a malfunction in partially automated driving. *2015 International Conference on Collaboration Technologies and Systems (CTS)*, 70–76. https://doi.org/10.1109/CTS.2015.7210402

Krupp, L. B., & Elkins, L. E. (2000). Fatigue and declines in cognitive functioning in multiple sclerosis. *Neurology*, *55*(7), 934–939. https://doi.org/10.1212/wnl.55.7.934

Lagerstrom, L., Johanes, P., & Ponsukcharoen, U. (2015). *The Myth of the Six-Minute Rule: Student Engagement with Online Videos*. 26.1558.1-26.1558.17. https://peer.asee.org/the-myth-of-the-six-minute-rule-student-engagement-with-online-videos

Lavie, N. (2010). Attention, Distraction, and Cognitive Control Under Load. *Current Directions in Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lee, I.-S., Bardwell, W. A., Ancoli-Israel, S., & Dimsdale, J. E. (2010). Number of lapses during the psychomotor vigilance task as an objective measure of fatigue. *Journal of Clinical Sleep Medicine: JCSM: Official Publication of the American Academy of Sleep Medicine*, *6*(2), 163–168.

Lee, K., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry Research*, *36*, 291–298. https://doi.org/10.1016/0165-1781(91)90027-M

Lenneberg, E. H. (1967). The Biological Foundations of Language. *Hospital Practice*, *2*(12), 59–67. https://doi.org/10.1080/21548331.1967.11707799

Lenth, R., Buerkner, P., Herve, M., Love, J., Riebl, H., & Singmann, H. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means (1.7.0)* (1.7.0). https://CRAN.R-project.org/package=emmeans

Lim, N., Kim, E., Kim, H., Yang, E., & Lee, S. (2010). Individual and work-related factors influencing burnout of mental health professionals: A meta-analysis. *Journal of Employment Counseling*, *47*. https://doi.org/10.1002/j.2161-1920.2010.tb00093.x

Lorist, M. M., Boksem, M. A. S., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Brain Research. Cognitive Brain Research*, *24*(2), 199–205. https://doi.org/10.1016/j.cogbrainres.2005.01.018

MATLAB (a). (2022). The MathWorks Inc.

MATLAB Online—MATLAB. (2022). Retrieved August 11, 2022, from https://www.mathworks.com/products/matlab-online/limitations.html

Möckel, T., Beste, C., & Wascher, E. (2015). The Effects of Time on Task in Response Selection—An ERP Study of Mental Fatigue. *Scientific Reports*, *5*. https://doi.org/10.1038/srep10113

Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting Human Learning by Hypnosis. *Cerebral Cortex*, *23*. https://doi.org/doi.org/10.1093/cercor/bhs068

Newport, E. L. (1990). Maturational Constraints on Language Learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1207/s15516709cog1401_2

Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 481–502). The MIT Press.

Patkowski, M. S. (1980). The Sensitive Period for the Acquisition of Syntax in a Second Language 1. *Language Learning*, *30*(2), 449–468. https://doi.org/10.1111/j.1467-1770.1980.tb00328.x

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008a). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology and Behavior*, *93*(1–2), 369–378.

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008b). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, *93*, 369–378. https://doi.org/10.1016/j.physbeh.2007.09.016

R Core Team. (2021). *R: A Language and Environment for Statistical Computing* [R]. R Foundation for Statistical Computin. https://www.R-project.org/

Raeymakers, S. (2022a). *Pilot TloadDback*. https://github.com/Raeymakers/PilotTloadDback

Raeymakers, S. (2022b). *Short TloadDback Analysis*. https://github.com/Raeymakers/Short_TLoadDBack_Analysis

Richter, S., Marsalek, K., Glatz, C., & Gundel, A. (2005). Task-dependent differences in subjective fatigue scores. *Journal of Sleep Research*, *14*(4), 393–400. https://doi.org/10.1111/j.1365-2869.2005.00473.x

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Sawin, D. A., & Scerbo, M. W. (1995). Effects of Instruction Type and Boredom Proneness in Vigilance: Implications for Boredom and Workload. *Human Factors*, *37*(4), 752–765. https://doi.org/10.1518/001872095778995616

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Derous, E., De Fruyt, F., & Duyck, W. (2021). How interest fit relates to STEM study choice: Female students fit their choices better. *Journal of Vocational Behavior*, *129*, 103614. https://doi.org/10.1016/j.jvb.2021.103614

Schwid, S., Tyler, C., Scheid, E., Weinstein, A., Goodman, A., & McDermott, M. (2003). Cognitive fatigue during a test requiring sustained attention: A pilot study. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *9*, 503–508. https://doi.org/10.1191/1352458503ms946oa

Shigihara, Y., Tanaka, M., Ishii, A., Tajima, S., Kanai, E., Funakura, M., & Watanabe, Y. (2013). Two different types of mental fatigue produce different styles of task performance. *Neurology, Psychiatry and Brain Research*, *19*(1), 5–11. https://doi.org/10.1016/j.npbr.2012.07.002

Sinclair, K. L., Ponsford, J. L., Rajaratnam, S. M. W., & Anderson, C. (2013). Sustained attention following traumatic brain injury: Use of the Psychomotor Vigilance Task. *Journal of Clinical and Experimental Neuropsychology*, *35*(2), 210–224. https://doi.org/10.1080/13803395.2012.762340

Smalle, E., Daikoku, T., Duyck, W., Szmalec, A., & Möttönen, R. (2020). P80 Language learning in the adult brain: TMS-induced disruption of the left dorsolateral prefrontal cortex enhances neural entrainment indexes to statistical language learning. *Clinical Neurophysiology*, *131*, e56. https://doi.org/10.1016/j.clinph.2019.12.191

Smalle, E., Daikoku, T., Szmalec, A., Duyck, W., & Möttönen, R. (2022). Unlocking adults' implicit statistical learning by cognitive depletion. *Proceedings of the National Academy of Sciences*, *119*(2), e2026011119. https://doi.org/10.1073/pnas.2026011119

Smalle, E. H. M., Panouilleres, M., Szmalec, A., & Möttönen, R. (2017). Language learning in the adult brain: Disrupting the dorsolateral prefrontal cortex facilitates word-form learning. *Scientific Reports*, *7*(1), 13966. https://doi.org/10.1038/s41598-017-14547-x

Smalle, E., Muylle, M., Szmalec, A., & Duyck, W. (2021). Less is More: Depleting cognitive resources enhances language learning abilities in adults. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001058

Smalle, E., Page, M. P. A., Duyck, W., Edwards, M., & Szmalec, A. (2018). Children retain implicitly learned phonological sequences better than adults: A longitudinal study. *Developmental Science*, *21*(5), e12634. https://doi.org/10.1111/desc.12634

Smalle, E., Rogers, J., & Rikka, M. (2014). Dissociating Contributions of the Motor Cortex to Speech Perception and Response Bias by Using Transcranial Magnetic Stimulation. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhu218

Tanaka, M., Shigihara, Y., Funakura, M., Kanai, E., & Watanabe, Y. (2012). Fatigue-associated alterations of cognitive function and electroencephalographic power densities. *PloS One*, *7*(4), e34774. https://doi.org/10.1371/journal.pone.0034774

Thiessen, E. D., Girard, S., & Erickson, L. C. (2016). Statistical learning and the critical period: How a continuous learning mechanism can give rise to discontinuous learning. *Cognitive Science*, *7*, 276–288. https://doi.org/10.1002/wcs.1394

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cogntition*, *97*(2), B25-34. https://doi.org/10.1016/j.cognition.2005.01.006

Trejo, L., Kochavi, R., Kubitz, K., Montgomery, L., Rosipal, R., & Matthews, B. (2005). Measures and models for predicting cognitive fatigue. *SPIE Defense + Commercial Sensing*. https://doi.org/10.1117/12.604286

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ulman, M. T. (2015). The Declarative/Procedural Model: A Neurobiologically Motivated Theory of First and Second Language. In *Theories in Second Language Acquisition:*

An Introduction (2nd ed.). Routledge. file:///C:/Users/SOFIER~1/AppData/Local/Temp/Ullman2015-1.pdf

Virag, M. (2015). Competition between frontal lobe functions and implicit sequence learning: Evidence from the long-term effects of alcohol. *Experimental Brain Research*, *233*, 2081–2089. https://doi.org/10.1007/s00221-015-4279-8

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation—What can be learned from children? *Neuroscience and Biobehavioral Reviews*, *36*(7), 1718–1728. https://doi.org/10.1016/j.neubiorev.2012.03.002

Wolfe, F. (2004). Fatigue assessments in rheumatoid arthritis: Comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. *The Journal of Rheumatology*, *31*(10), 1896–1902.