ELSEVIER

Contents lists available at ScienceDirect

## Consciousness and Cognition



journal homepage: www.elsevier.com/locate/concog

# Revisiting the link between the sustained attention to response task (SART) and daily-life cognitive failures<sup> $\star$ </sup>



Annika M. Schepers <sup>a,1</sup>, Leonie Schorrlepp <sup>a,1,\*</sup>, Juriena D. de Vries <sup>b</sup>, Tamara de Kloe <sup>c,d</sup>, Dimitri van der Linden <sup>e</sup>, Erik Bijleveld <sup>a</sup>

<sup>a</sup> Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands

<sup>b</sup> Department of Health, Medical and Neuropsychology, Leiden University, Leiden, the Netherlands

<sup>c</sup> Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Department of Cognitive Neuroscience, Nijmegen, the Netherlands

<sup>d</sup> Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Nijmegen, the Netherlands

<sup>e</sup> Department of Psychology, Education, and Child Studies, Erasmus University, Rotterdam, the Netherlands

#### ARTICLE INFO

Keywords: Concentration Distractibility Mental fatigue Cognitive fatigue Burnout External validity

#### ABSTRACT

In this study, we examined the relationship between errors of commission on the Sustained Attention to Response Task (SART) and scores on the Cognitive Failures Questionnaire (CFQ). The goal was to assess the ecological validity of the SART in a sample of people scoring high on fatigue complaints. SART errors of commission were positively associated with CFQ scores and this finding remained after controlling for fatigue level, age, and SART reaction times. Thus, our results generally supported the ecological validity of the SART. However, when examining sub-samples separately, we found the association between SART and CFQ only in our subsample of employees, not in our subsample of university students. The three subscales of the CFQ showed the same pattern of findings. Our results imply that, when using the SART to draw conclusions about everyday life, it is crucial to consider the characteristics of one's sample and control for relevant confounding variables.

#### 1. Introduction

The human mind often wanders off while doing mental work, while driving, or while engaging in leisure activities, such as reading. One of the most-used neuropsychological tasks that aims to quantify people's capacity for sustained attention is the Sustained Attention to Response Task (SART; Robertson et al., 1997). In this task, participants are instructed to quickly press a button every time a number appears on a screen, except the digit "3", for which their response must be withheld. Initially designed for the use in patients with traumatic brain injuries (Robertson et al., 1997), the SART has been employed in a variety of clinical samples, such as those with ADHD and depression (Smilek et al., 2010). In this setting, the SART is a promising instrument. For example, it demonstrated meaningful differences between military personnel with and without depression (Farrin et al., 2003) and it contributed to new insights

#### https://doi.org/10.1016/j.concog.2023.103558

Received 12 January 2023; Received in revised form 12 July 2023; Accepted 9 August 2023

Available online 30 August 2023

<sup>\*</sup> This research was not funded by any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

<sup>\*</sup> Corresponding author at: Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands.

*E-mail addresses*: annika.schepers@ru.nl (A.M. Schepers), leonie.schorrlepp@ru.nl (L. Schorrlepp), j.d.de.vries@fsw.leidenuniv.nl (J.D. de Vries), tamara.dekloe@donders.ru.nl (T. de Kloe), vanderlinden@essb.eur.nl (D. van der Linden), erik.bijleveld@ru.nl (E. Bijleveld).

<sup>&</sup>lt;sup>1</sup> Contributed equally to the manuscript and thus share first authorship.

<sup>1053-8100/© 2023</sup> The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

into the cognitive effects associated with burnout (Van der Linden et al., 2005).

Beyond its initial application in clinical settings, there is evidence for the general ecological validity of the SART. SART scores are associated with scores on the Cognitive Failures Questionnaire (CFQ, Broadbent et al., 1982). The CFQ assesses failures in everyday life that many people can relate to, for example, "Do you find you forget people's names?" (Broadbent et al., 1982). These failures represent situations in which individuals' intentions deviate from their actions, not due to failed self-control, but rather due to a cognitive failure, such as a failure of sustained attention. The existence and strength of this association points to the extent to which the SART, which is used in lab settings, explains and relates to cognitive failures in everyday life. Initial data suggests that the association between the SART and the CFQ is stable, with a meta-analysis reporting a correlation since the SART is designed to assess specific cognitive processes such as sustained attention and inhibition, whereas the CFQ rather measures general behaviour related to a wider range of cognitive processes. In this research, we examine the link between the SART and the CFQ, addressing this criticism and going beyond previous work in three ways.

First, prior work on the SART–CFQ relation has typically failed to control for speed–accuracy trade-offs (SATOs). Whilst the SART is useful as an attentional measure, it is also susceptible to SATOs as people might employ different strategies, for example, take longer but be more accurate (Seli et al., 2013). Different response strategies, however, can mask failures of attention as reflected by errors of commission on the SART. After Seli and colleagues (2013) established how these strategies might influence the SART–CFQ relationship, not much work has been done to explore how this influence may play out in different samples. SATOs are known to be particularly relevant in relation to fatigue and age. Indeed, it has been suggested that one of the main behavioural effects of fatigue is that people tend to switch strategies in order to maintain levels of performance (Hockey, 1997). Moreover, a common finding in experimental research is that elderly people prefer a focus on accuracy compared to speed (Smith & Brewer, 1995). Accordingly, in the present study we will control for reaction times, level of fatigue, and age, potential confounders of the SART–CFQ relationship.

Second, we examine two populations, university students and employees, with persistent fatigue complaints. Fatigued individuals are a potential target population for the SART in practice and thus are especially interesting for investigation. Here, we define and operationalise fatigue as a feeling of tiredness (Dora et al., 2021) that is associated with a reduced willingness to invest - more effort into tasks (Hockey, 2013). Cognitive impairment (as reflected in high CFQ scores) and reduced attention (as measured by the SART) are characteristic of fatigue (Cveijc et al., 2016), making a fatigued sample intriguing for investigating the SART-CFQ relationship. Fatigued people may score especially high, perhaps close to the ceiling, on both SART errors and CFQ, which may lead to a weaker SART-CFQ relationship. So, using a sample of people with fatigue complaints, will give more detailed insight into the ecological validity of the SART-CFQ association.

Third, the use of the SART–CFQ association to show ecological validity of the SART has been criticised due to the specificity of the SART compared to the generality of the CFQ (Smilek et al., 2010). To address this criticism, we explore the association between SART and subscales of the CFQ. As the subscales reflect more specific types of cognitive failures, doing so enables us to investigate the relationship between SART and the specific CFQ subscales. This way we can visualise the relationship between SART and CFQ in a more differentiated manner rather than exclusively relying on the full CFQ score. For this we will use the subscales Forgetfulness, Distractibility and False Triggering as identified by Rast and colleagues (2009). Additionally, we use a custom subscale based on Cheyne and colleagues (2006), which only includes items that reflect attention failures. We have no specific hypotheses regarding these subscales.

In sum, we examine the ecological validity of the SART in novel ways, by 1) elaborating on the SART–CFQ relationship, increasing our understanding of what the SART can tell us about everyday cognitive impairment, and 2) exploring the question of how the relationship between SART and CFQ is characterised in a fatigued sample. We hypothesise that SART errors of commission as indicators of sustained attention failures will be associated with CFQ scores (*Hypothesis 1*) and that SART errors of commission will still predict CFQ scores when controlling for fatigue level, age, and SART reaction times (*Hypothesis 2*).

#### 2. Methods

#### 2.1. Data

This project used secondary data, combining baseline data from two studies by de Vries and colleagues (2016, 2017) which investigated the effectiveness of an exercise intervention for reducing work- and study-related fatigue. In the 2016 study, de Vries and colleagues collected data from a sample of 99 students who scored above cut-off values on both the Emotional Exhaustion Scale of the Utrecht Burnout Scale for Students (Cronbach's  $\alpha = 0.81$ ; Schaufeli et al., 2002) and the 10-item Fatigue Assessment Scale (FAS; Cronbach's  $\alpha = 0.79$ ; De Vries et al., 2004). Participating students did not experience fatigue due to a medical condition and did not receive psychological or pharmacological treatment for fatigue. Forty-four percent of the students had a part-time job next to their study. On average, those who had a job worked 7.8 h per week. In the 2017 study, data were collected from 96 employees from different work environments according to the same criteria as the student sample. In addition to the FAS (Cronbach's  $\alpha = 0.84$ ), fatigue was assessed using the Utrecht Burnout Scale without adjustment for students (Cronbach's  $\alpha = 0.80$ ; Schaufeli & Van Dierendonck, 2001). Furthermore, fatigue was additionally assessed with the Need for Recovery Scale (van Veldhoven & Broersen, 2003) with a Cronbach's  $\alpha$  of 0.85 for employees and 0.75 for students. Both studies measured sleep quality using the Questionnaire on the Experience and Evaluation of Work (van Veldhoven et al., 2015), with Cronbach's  $\alpha = 0.62$  for employees and Cronbach's  $\alpha = 0.61$  for students. In addition, both studies measured sleep quantity (mean hours of sleep per night).

To assess cognitive functioning, before the start of both interventions, participants first self-rated their momentary level of fatigue

by responding to the item "How fatigued do you currently feel?" (on a 1–10 scale; for validation, see Van Hooff et al., 2007). Then, they filled in the 25-item Dutch version of the CFQ and completed the SART. A more detailed study protocol for both studies is available online (de Vries et al., 2015; de Vries, 2014). Combining the data from these two studies and excluding participants with missing or incomplete values on the study variables (removing 12 employees and 2 students), left us with an initial data set of 181 participants with gender recorded as a binary variable (147 women). This data set is made up of 83 employees between 24 and 65 years old (67 women,  $M_{age} = 45.4$ , SD = 10.6) and 98 students between 18 and 30 years old (80 women,  $M_{age} = 20.9$ , SD = 2.3).

#### 2.2. Measures

**2.2.1. Daily Cognitive Failures.** The answer scale of the CFQ ranges from never (1) to very often (5) to indicate the frequency at which participants recall encountering the items of the questionnaires with higher total scores representing more cognitive failures (Broadbent et al., 1982). The collection of items includes everyday occurrences, such as "Do you fail to listen to people's names when you are meeting them?", or "Do you fail to see what you want in a supermarket (although it's there)?". The employee sample answered the CFQ on a scale of 0–4 and the student sample on a scale of 1–5. To harmonise both datasets, the adjusted CFQ scores from the student sample reflect the 0–4 scale. For the employees, the CFQ had a Cronbach's  $\alpha$  of 0.90, for the students 0.83. Four sub-scales of the CFQ, as derived from the literature, were created. The first subscale reflects failures of attention. As the CFQ does not exclusively reflect attention failures, the creation of subscales based on the CFQ enabled us to measure failures related to attention more accurately. This sub-scale was based on the ARCES (Cheyne et al., 2006), a scale that reflects attention-related cognitive errors and shows overlap with CFQ items (1, 6, 13, 19, 21). Furthermore, we included the following subscales into our analysis: Forgetfulness (CFQ items 1, 2, 5, 7, 17, 20, 22, 23), Distractibility (CFQ items 8, 9, 10, 11, 14, 19, 21, 25, and False Triggering (CFQ items 2, 3, 5, 6, 12, 18, 23, 24, the latter describing interruptions of intended behaviours (Rast et al., 2009), for example, throwing away the tomatoes instead of their package, with the initial intention to keep the food and throw away the packaging.

**2.2.2. Sustained Attention.** SART errors-of-commission are indicators of failures of sustained attention. The SART is a reversed GO/NO GO task, in which the GO stimuli appear more frequent than the NOGO stimuli for which the participants have to withhold their response. Participants see random digits between 1 and 9 on a screen and are instructed to press a button as quickly as possible when the number appears on the screen. However, when a 3 is shown, participants have to withhold their response. Participants completed 450 trials, the numbers were shown for 250 ms with intervals between trials fixed at 850 ms (de Vries et al., 2016). In both studies by de Vries et al. (2016, 2017), participants were given the instruction to "click as fast as possible on the button" when they saw a digit, except when the digit was 3. There are two kinds of errors that participants can make: *errors of omission*, which occur when participants fail to press the button at a GO stimulus, and *errors of commission*, which occur when participants press the button at a NOGO stimulus. Though errors of omission can be informative (Cheyne et al., 2009), they were rare in our data (i.e., M = 2.2 out of 400 GO stimuli), and, in line with our a priori analysis plan, we will not consider them further.

#### 2.3. Statistical analyses

The first hypothesis that SART errors of commission are associated with CFQ scores, was tested using correlations. The second hypothesis was tested using a linear regression with SART errors of commission as an independent variable and CFQ scores as a dependent variable, whilst controlling for fatigue level, age and SART reaction times. All independent variables were standardised. To investigate whether estimates of the model including these control variables are different from the initial model we computed, as well as compared, the respective confidence intervals.

Furthermore, exploratory analyses involved testing whether SART errors of commission related to individual subscales of the CFQ, after adjusting for fatigue level, age, and SART reaction times. For this, the same linear model as in hypothesis 2 was used to predict subscale sum scores in four separate regressions.

Then, we investigated whether the relationship between SART errors of commission and CFQ scores is different between the student and the employee sample by using linear regressions as specified above (SART scores predicting CFQ, controlling for fatigue, age, RT) for each group. The models were compared, and independent t-tests conducted to inspect the differences found.

The study was pre-registered after the data was collected, but before it was further pre-processed for the current project. The preregistration can be found online (https://aspredicted.org/u4hx6.pdf). We conducted all analyses in line with this pre-registration, with one exception: In addition to the analyses, we report below, we pre-registered to explore the SART–CFQ association separately for younger and older participants. However, we refrained from carrying out this analysis, as this analysis would overlap strongly with our analysis of the two separate samples (students were all younger than thirty; employees covered a much larger range of ages up to retirement age). All analyses were conducted in RStudio Version 1.4.1717 (RStudio Team, 2021), using R Version 4.1.1. (R Core Team, 2021).

The data used for this analysis and R code of the analysis can be found on the Open Science Framework website (https://osf.io/e7faz/).

#### 3. Results

Univariate distributions of the continuous variables of interest were approximately normal. No extreme values were found in these variables either, except three outliers that deviated  $\geq 3$  SD from the mean were found in the SART reaction times. As preregistered, they were not excluded. However, the results of the analysis did not change significantly when the outliers were excluded. Scores on

the CFQ ranged from 15 to 76 (*M* = 44.0, 95% CI [42.3, 45.7]). SART error-of-commissions ranged from 1 to 45 (*M* = 22.7, 95% CI [21.4, 24.0]).

#### 3.1. Hypothesis I – Correlation between SART and CFQ

SART errors of commission were positively associated with CFQ scores (r = 0.28, 95% CI [0.15, 0.41], p < .001). So, on average, a higher number of SART errors of commission was associated with higher CFQ scores thereby confirming our first hypothesis.

#### 3.2. Hypothesis II – Association between SART and CFQ when controlling for confounding variables

After including the control variables, the model explained 12.0% of the variance in CFQ scores, F(4, 176) = 5.97, p < .001. Supporting the hypothesis, SART errors of commission predicted CFQ sum scores (b = 3.99, SE = 1.26, 95% CI [1.50, 6.48],  $\beta = 0.34$ , SE = 0.11, p = .002), in the same direction as the correlation. Additionally, fatigue scores were positively associated with CFQ scores (b = 2.14, SE = 0.84, 95% CI [0.48, 3.79],  $\beta = 0.18$ , SE = 0.07, p = .012), so higher fatigue levels predicted higher CFQ scores. Reaction time was not a significant predictor of overall CFQ sum scores in this model (b = 1.39, SE = 1.31, p = .289), nor was age (b = -0.31, SE = 0.98, p = .755).

To further examine whether the inclusion of the control variables fatigue, age and SART reaction time made a difference to the model (apart from the higher variance explained) we compared the confidence intervals for the models' parameter estimates (see Table 1). Table 1 and Fig. 1 suggest that the different models' estimate for SART errors of commission is very similar and their confidence intervals overlap, regardless of whether we controlled for fatigue and age, indicating that there is no significant difference of estimates between the two models.

We checked for influential cases by checking Cook's distance. None of the cases surpassed our preregistered cut-off value of Cook's distance values greater than 1. Additionally, the proportion of standardised residuals greater than 2 (3.3%), greater than 2.5 (0.6%), and greater than 3 (0.0%) was calculated with all values suggesting that the model is acceptable.

#### 3.3. Exploratory analyses

To explore whether SART errors of commission specifically relate to the subscales of the CFQ when controlling for fatigue level, age, and SART reaction times, we performed linear regression analyses separately for each subscale (Fig. 2). The independent variables included in the regression analyses together explained 10.2% of the variance of CFQ attention subscale scores (F(4, 176) = 5.04, p < .001), 18.5% of the variance of CFQ forgetfulness subscale scores (F(4, 176) = 9.97, p < .001), 7.2% of the variance of CFQ distractibility subscale scores (F(4, 176) = 3.45, p = .010), and 8.0% of the variance of CFQ false triggering subscale scores (F(4, 176) = 3.84, p = .005). SART errors of commission were positively associated with CFQ attention subscale scores (b = 1.02, SE = 0.33, 95% CI [0.37, 1.67],  $\beta = 0.34$ , SE = 0.11, p = .002), with CFQ distractibility subscale scores (b = 1.42, SE = 0.47, 95% CI [0.49, 2.35],  $\beta = 0.33$ , SE = 0.11, p = .003), and with CFQ false triggering subscale scores (b = 1.31, SE = 0.52, 95% CI [0.29, 2.34],  $\beta = 0.28$ , SE = 0.11, p = .012). This indicates that participants who made more mistakes on the SART also had higher scores on the attention, distractibility, and false triggering subscales. There was no support for the idea that SART errors predicted scores on the forgetfulness subscale (b = 0.87, SE = 0.49, 95% CI [-0.10, 1.83],  $\beta = 0.18$ , SE = 0.10, p = .077; for regression tables, see Appendix A).

Participant age was negatively associated with CFQ forgetfulness subscale scores (b = -1.27, SE = 0.38,  $\beta = -0.43$ , SE = 0.11, 95% CI [-2.02, -0.52], p = .001), meaning that older participants scored lower on forgetfulness subscale scores. Furthermore, age was positively associated with scores on the CFQ false triggering subscale (b = 0.96, SE = 0.40,  $\beta = 0.20$ , SE = 0.09, 95% CI [0.16, 1.76], p = .018), indicating that older participants scored higher on the false triggering subscale.

Lastly, we were interested in whether the subsamples (i.e., students, employees) differed in their SART–CFQ relationship. Therefore, we fit the model from Hypothesis 2 to each of the sub-samples (Fig. 3). For employees, the analysis explained 13.1% of variance in CFQ sum scores (F(4, 78) = 2.94, p = .025). SART errors-of-commission significantly predicted CFQ sum scores (b = 6.26, SE = 2.06, 95% CI [2.16, 10.35],  $\beta = 0.53$ , SE = 0.18, p = .003), meaning that employees who made more mistakes on the SART also had higher CFQ scores. For students, the analysis explained 14.6% of the variance in CFQ sum scores (F(4, 93) = 3.96, p = .005). Importantly, however, SART errors of commission did not significantly predict CFQ sum scores (b = 1.91, SE = 1.65, 95% CI [-1.36,

#### Table 1

Regression table for model predicting CFQ-scores from SART errors of commission with and without control variables.

Predictor variables	Model with control variables		Model with only RT controlled		Model without control variables		
	β	95% CI	β	95% CI	β	95% CI	
Intercept SART errors-of-commission SART RT Age Fatigue	0.00 0.34** 0.12 -0.03 0.18*	-0.14, 0.14 0.13, 0.55 -0.10, 0.34 -0.19, 0.14 0.04, 0.32	0.00 0.36** 0.10	-0.14, 0.14 0.14, 0.57 -0.12, 0.31	0.00 0.28***	-0.14, 0.14 0.14, 0.43	

Note. \*p < .05, \*\*p < .01, \*\*\*p < .001; Beta-values and confidence intervals for the model from hypothesis 2 and for exploratory models controlling only for reaction time and controlling for no other variables.



Fig. 1. Note. Association of SART errors-of-commission and CFQ scores. Predicted values for the linear regression including control variables (continuous line) were not significantly different from regression without control variables (dashed line). Dots are raw data from individual participants, employees in light blue. The distributions of the CFQ scores and the SART errors of commission are shown on the sides. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.18],  $\beta = 0.16$ , SE = 0.14, p = .250). Instead, fatigue levels predicted students' CFQ sum scores (b = 3.85, SE = 1.10, 95% CI [-1.67, 6.02],  $\beta = 0.33$ , SE = 0.09, p < .001), indicating that students who were more fatigued reported more incidences of cognitive failures in their daily lives.

To explore differences between the two subsamples further, t-tests were computed to compare the samples on the key variables. Students made significantly more errors (M = 26.4) than employees (M = 18.3), t(175.9) = 6.74, p <.001). However, students had significantly lower reaction times (M = 262 ms) than employees (M = 306 ms), t(150.6) = -6.52, p <.001). We found no evidence that students had lower or higher CFQ scores (M = 45.6) than employees (M = 42.2), t(167.8) = 1.95, p =.053. Also, we found no evidence that students (M = 6.09) were less or more fatigued than employees (M = 5.60), t(174.3) = 1.76, p =.080. For completeness, we note that students tended to sleep longer (M = 7.4 h) than employees (M = 7.1 h), t(176.7) = 2.13, p =.034, but reported worse sleep quality (M = 2.9) than employees (M = 3.6), t(176.9) = 3.18, p =.002.

#### 4. Discussion

The present study explored the relationship between SART errors of commission and CFQ sum scores in a fatigued sample (students and employees). In line with our expectations, SART errors of commission as indicators of sustained attention failures were associated with more self-reported cognitive failures (i.e., CFQ) in daily life. Further, also in line with our expectations, this association remained significant after controlling for fatigue level, age and reaction times, generally supporting the ecological validity of the SART.

We further found that fatigue uniquely predicted CFQ sum scores, suggesting a link between daily cognitive failures and fatigue symptoms in this sample. The CFQ-SART relationship had approximately the same magnitude for the CFQ subscales attention, distractibility, and false triggering. The relationship was not significant, however, when considering only the forgetfulness subscale. Finally, we explored whether the association between SART errors of commission and daily cognitive failures was different between the student and the employee sample. Indeed, SART errors of commission predicted CFQ sum scores only in the employee sample. For students, only fatigue statistically predicted CFQ sum scores. The association between the CFQ and the SART and was not significant in the student sample. On average, students made more mistakes on the SART and were more fatigued than employees.



Fig. 2. Note. Association of SART errors-of-commission and CFQ subscale scores. The association between SART errors of commission and the CFQ subscale scores was significant for all subscales except the forgetfulness subscale.

#### 4.1. Ecological validity

Even though the present samples differed from the samples Smilek and colleagues (2010) used in their *meta*-analysis, the association we found (r = 0.29) was rather similar to the overall correlation from their *meta*-analysis (r = 0.21). Thus, at first sight, our study's findings are in line with the idea that the SART is a stable predictor of everyday cognitive failures. However, when analysing employees and students separately, we found that the two samples differed in the SART-CFQ association. Whilst SART errors were clearly a significant predictor for employees, SART errors did not significantly predict CFQ scores in students. Instead, only the students' fatigue level predicted their CFQ scores. Consequently, there was no evidence that the SART gives direct insight into everyday cognitive failure in a student population.

One possible explanation for the lack of the SART-CFQ association in students is that their high levels of fatigue caused a ceiling effect that prevented the association from being detectable. The high levels of fatigue in our student sample may have led to higher errors and higher CFQ scores, thus hiding the relationship between sustained attention and daily cognitive failures. Indeed, at least for some individuals, fatigue is associated with lower cognitive performance (Ackerman, 2011; Hopstaken et al., 2015). Similarly, in children and adolescents (Sievertsen et al., 2016) as well as university students (Smith, 2018), cognitive fatigue is generally associated with lower academic performance. In line with this research, we found that cognitive failures were frequent in students' everyday life. So, whilst there might not have been an association between self-reported daily cognitive failures and a more objective measure of attention in this sample, there may well be an influence of fatigue on cognitive performance in general. At least, the lack of association between the SART and the CFQ in the fatigued student sample in contrast to the association found among employees, highlights the importance of considering sample characteristics when using the SART as an indicator of daily cognitive failures.



Fig. 3. Note. Association of SART Errors-of-Commission and CFQ Scores for Employees and Students. CFQ sum scores were predicted by SART errors-of-commission for employees but not for students. Employees visualized in light blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

#### 4.2. Limitations and future directions

In this study we did not consider the time course of attention during the SART in our analysis of the relationship with the CFQ. Sustained attention is not a static concept, it changes over the time that it takes to complete the SART. The loss of information about how individual participants' performance develops throughout the task may have led to a loss of insight into the relationship between SART and CFQ. In future research, a more fine-grained analysis may provide novel insights into individual differences in whether and how people's capacity for sustained attention affects daily life.

Seli and colleagues (2013) argued that it is crucial to control for RTs when using SART errors of commission as a predictor, as SART errors of commission may otherwise mainly reflect response strategies, rather than failures of attention. In our sample, controlling for RTs changed neither our estimates nor our conclusions based on them. Nevertheless, we do not think our finding invalidates Seli and colleagues' (2003) recommendation to continue considering this covariate at all times in order to draw conclusions about people's attention in everyday life. As seen in this study, reaction times and errors of commission are highly correlated and this relationship can be expected to exist in other samples, too.

Building on the previous point, it is worth noting that students were relatively fast (at the cost of accuracy), whereas employees were relatively accurate (at the cost of speed). This finding is intriguing, as both subsamples received identical instructions, which were worded such that they emphasized speed (see Method). Speculatively, this difference may have emerged because students picked up on the task instructions better, perhaps because they are more accustomed to computerized tests. Alternatively, this difference may have emerged due to differences in both subsamples' learning history (e.g., students were younger, and thus, more likely to have grown up around technology; but see Bennett et al., 2008). Finally, students and employees may have differed in their sleep duration and sleep quality, which may also affect speed–accuracy trade-offs (Stawarczyk & D'Argembeau, 2016). Regardless of the cause of students' relative tendency to prioritize speed, this tendency can explain why we did not find a clear SART–CFQ association among students. Perhaps, students' lowered their response threshold, which made the commission errors score from the SART more noisy, and thus, less informative. In further research on the SART (and its relation to other variables), we recommend that researchers take into account not just RTs, but also sample characteristics that are potentially related to how people make speed–accuracy trade-offs.

Another limitation is the generalizability of our findings. We note that the sample was largely composed of women. More importantly, however, it should be noted that the sample was largely white and indicated a high level of education as is the case in most

research into this association. Furthermore, data were collected in an industrialised, rich, **d**emocratic country (WEIRD). The samples in Smilek et al.'s (2010) *meta*-analysis were also all WEIRD, pointing to the larger issue with generalising findings in this line of research. It is crucial to not blindly extend these findings to other populations, especially as the cognitive failures on the CFQ are informed by the lifestyle in a WEIRD country, was validated with a WEIRD sample, and may not apply at all to people who do not fit the WEIRD description, thus reducing the power of the CFQ to give ecological validity to the SART.

#### 5. Conclusion

Our findings generally support the ecological validity of the SART by confirming its relationship with cognitive failures in daily life, highlighting its usefulness as a tool in research and clinical practice. However, the present study also emphasises the importance of sample characteristics and potential confounders (especially reaction times and fatigue). The illustration of the relationship's complexity paves the way for more differentiated approaches toward this topic in the future.

### CRediT authorship contribution statement

Annika M. Schepers: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. Leonie Schorrlepp: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. Juriena D. de Vries: Investigation, Resources, Writing – review & editing. Tamara de Kloe: Conceptualization, Software. Dimitri van der Linden: Writing – review & editing. Erik Bijleveld: Conceptualization, Software, Writing – review & editing, Supervision.

#### **Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Appendix A

Regression table for models of subscales

Predictors	Attention		Forgetfulnes	Forgetfulness		Distractibility		False Triggering	
	β	95% CI	β	95% CI	β	95% CI	β	95% CI	
SART errors-of-commission SART RT Age Fatigue	0.34 ** 0.14 -0.06 0.13	0.12, 0.55 -0.08, 0.36 -0.23, 0.10 -0.01, 0.27	0.18 0.01 -0.27** 0.14*	-0.02, 0.39 -0.19, 0.23 -0.43, -0.11 0.01, 0.28	0.33 ** 0.16 0.11 0.14	0.11, 0.55 -0.07, 0.38 -0.06, 0.28 -0.01, 0.28	0.28* 0.03 0.20* 0.14	0.06, 0.50 -0.20, 0.25 0.04, 0.37 -0.01, 0.28	

*Note.* \*p < .05, \*\*p < .01; The confidence intervals of the SART – CFQ relationship for each of the subscales are overlapping; therefore, we cannot conclude that the relationship was different for one of the subscales. However, we can conclude that the relationship was not significant for the forgetfulness subscale.

#### References

- Ackerman, P. L. (2011). 100 years without resting. In P. L. Ackerman (Ed.), Cognitive fatigue: Multidisciplinary perspectives on current research and future applications (pp. 11–43). American Psychological Association. https://doi.org/10.1037/12343-001.
- Bennett, S., Maton, K., & Kervin, L. (2008). The 'digital natives' debate: A critical review of the evidence. British Journal of Educational Technology, 39(5), 775–786. https://doi.org/10.1111/j.1467-8535.2007.00793.x

de Vries, J. D., van Hooff, M. L., Geurts, S. A., & Kompier, M. A. (2017). Exercise to reduce work-related fatigue among employees: A randomized controlled trial. Scandinavian Journal of Work, Environment & Health, 4, 337–349. https://doi.org/10.5271/sjweh.3634

de Vries, J. D., van Hooff, M. L. M., Geurts, S. A. E., & Kompier, M. A. J. (2015). Efficacy of an exercise intervention for employees with work-related fatigue: Study protocol of a two-arm randomized controlled trial. *BMC Public Health*, 15(1), 1117. https://doi.org/10.1186/s12889-015-2434-6

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, 21(1), 1–16. https://doi.org/10.1111/j.2044-8260.1982.tb01421.x

Cheyne, A. J., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, 111(1), 98–113. https://doi.org/10.1016/j.cognition.2008.12.009

de Vries, J. D. (2014, January 27). Trial NL4201 (NTR4412). Netherlands Trial Register. https://www.trialregister.nl/trial/4201.

- de Vries, J. D., van Hooff, M. L. M., Geurts, S. A. E., & Kompier, M. A. J. (2016). Exercise as an intervention to reduce study-related fatigue among university students: A two-arm parallel randomized controlled trial. *PLoS One1*, 11(3), e0152137.
- de Vries, J., Michielsen, H., Van Heck, G. L., & Drent, M. (2004). Measuring fatigue in sarcoidosis: The Fatigue Assessment Scale (FAS). British Journal of Health Psychology, 9(3), 279–291. https://doi.org/10.1348/1359107041557048
- Farrin, L., Hull, L., Unwin, C., Wykes, T., & David, A. (2003). Effects of depressed mood on objective and subjective measures of attention. The Journal of Neuropsychiatry and Clinical Neurosciences, 15(1), 98–104. https://doi.org/10.1176/jnp.15.1.98

Hockey, R. (2013). The Psychology of Fatigue: Work, Effort and Control. Cambridge University Press. https://doi.org/10.1017/CB09781139015394.

Hopstaken, J. F., Van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. Psychophysiology, 52(3), 305–315. https://doi.org/10.1111/psyp.12339

Van Hooff, M. L. M., Geurts, S. A. E., Kompier, M. A. J., & Taris, T. W. (2007). "How fatigued do you currently feel?" Convergent and discriminant validity of a singleitem fatigue measure. Journal of Occupational Health, 49(3), 224–234. https://doi.org/10.1539/joh.49.224

- Rast, P., Zimprich, D., Van Boxtel, M., & Jolles, J. (2009). Factor structure and measurement invariance of the Cognitive Failures Questionnaire across the adult life span. Assessment, 16(2), 145–158. https://doi.org/10.1177/1073191108324440
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35(6), 747–758. https://doi.org/10.1016/S0028-3932(97)00015-8
- Schaufeli, W. B., Martinez, I. M., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002). Burnout and engagement in university students: A cross-national study. Journal of Cross-Cultural Psychology, 33(5), 464–481.
- Schaufeli, W., & Van Dierendonck, D. (2001). Utrechtse Burnout Schaal (UBOS)[Utrecht Burnout Scale]. De Psycholoog, 36, 9-12.
- Seli, P., Jonker, T., Cheyne, J., & Smilek, D. (2013). Enhancing SART validity by statistically controlling speed-accuracy trade-offs. Frontiers in Psychology, 4, 265. https://doi.org/10.3389/fpsyg.2013.00265
- Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. Proceedings of the National Academy of Sciences, 113(10), 2621–2624.
- Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Failures of sustained attention in life, lab, and brain: Ecological validity of the SART. Neuropsychologia, 48(9), 2564–2570. https://doi.org/10.1016/j.neuropsychologia.2010.05.002

Smith, A. P. (2018). Cognitive fatigue and the well-being and academic attainment of university students. Journal of Education, Society and Behavioral Science.

Smith, G. A., & Brewer, N. (1995). Slowness and age: Speed-accuracy mechanisms. Psychology and Aging, 10(2), 238-247. https://doi.org/10.1037/0882-7974.10.2.238

- Stawarczyk, D., & D'Argembeau, A. (2016). Conjoint influence of mind-wandering and sleepiness on task performance. Journal of Experimental Psychology: Human Perception and Performance, 42(10), 1587–1600. https://doi.org/10.1037/xhp0000254
- Van der Linden, D., Keijsers, G., Eling, P., & Schaijk, R. (2005). Work stress and attentional difficulties: An initial study on burnout and cognitive failures. Work and Stress - WORK STRESS, 19, 23–36. https://doi.org/10.1080/02678370500065275
- van Veldhoven, M., & Broersen, S. (2003). Measurement quality and validity of the "need for recovery scale". Occupational and Environmental Medicine, 60(suppl 1), Article i3. https://doi.org/10.1136/oem.60.suppl 1.i3
- van Veldhoven, M. J. P. M., Prins, J., van der Laken, P. A., & Dijkstra, L. (2015). QEEW2.0: 42 short scales for survey research on work, well-being and performance. *SKB*.