Applicability of the Sustained Attention to Response Task (SART) in hypersomnolence: experience and results from a tertiary referral center

Hoeven, A.E. van der; Bijlenga, D.; Bouhuijs, P.; Schie, M.K.M. van; Lammers, G.J.; Fronczek, R.

Sleep Medicine 108 (2023) 105–113

Contents lists available at ScienceDirect

Sleep Medicine

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

Applicability of the Sustained Attention to Response Task (SART) in hypersomnolence: Experience and results from a tertiary referral center

Adrienne Elisabeth van der Hoeven [a, b], Denise Bijlenga [a, b], Puck Bouhuijs [b], Mojca Kristina Maria van Schie [a, b], Gert Jan Lammers [a, b], Rolf Fronczek [a, b, *]

[a] Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands
[b] Sleep-Wake Center, Stichting Epilepsie Instellingen Nederlands (SEIN), Heemstede, the Netherlands

ARTICLE INFO

Article history:
Received 12 July 2022
Received in revised form
4 May 2023
Accepted 8 June 2023
Available online 10 June 2023

Keywords:
narcolepsy Type 1
Narcolepsy type 2
Idiopathic hypersomnia
Vigilance

ABSTRACT

Objective/background: Evaluation of hypersomnolence disorders ideally includes an assessment of vigilance using the short Sustained Attention to Response Task (SART). We evaluated whether this task can differentiate between hypersomnolence disorders, whether it correlates with subjective and objective sleepiness, whether it is affected by the time of day, and symptoms of anxiety and depression.
Patients/methods: We analyzed diagnostic data of 306 individuals with hypersomnolence complaints diagnosed with narcolepsy type 1 (n=100), narcolepsy type 2 (n=20), idiopathic hypersomnia (n=49), obstructive sleep apnea (n=27) and other causes or without explanatory diagnosis (n=110). We included the Multiple Sleep Latency Test (MSLT), polysomnography, Epworth Sleepiness Scale (ESS), Hospital Anxiety and Depression Scale and SART, which were administered five times during the day (outcomes: reaction time, total, commission and omission errors).
Results: The SART outcomes did not differ between groups when adjusted for relevant covariates. Higher ESS scores were associated with longer reaction times and more commission errors ($p<.01$). The main outcome, total errors, did not differ between times of the day. Reaction times and omission errors were impacted ($p<.05$).
Conclusions: The SART quantifies disturbed vigilance, an important dimension of disorders of hypersomnolence. Results do not suggest that depressive symptoms influence SART outcomes. A practice session is advised. Testing time should be taken into account when interpreting results. We conclude that the SART does not differentiate between central disorders of hypersomnolence. It may be a helpful addition to the standard diagnostic workup and monitoring of these disorders.

Statement of significance

The Sustained Attention to Response Task (SART) is suitable for assessing vigilance in individuals with hypersomnolence. It is easy to implement and requires little time and resources when combined with the Multiple Sleep Latency Test (MSLT) in the routine workup. Our study assessed the association between SART and polysomnography (PSG) outcomes. We also evaluated the implementation of the SART in a large sample in clinical practice. This sample showed that the SART measures a specific aspect, namely vigilance, that is often overlooked despite having a significant impact on people's daily lives.

1. Introduction

Individuals suffering from disorders of hypersomnolence often complain of disturbed vigilance, resulting in disturbed sustained attention [1–3], which has the potential to have a profound, negative impact on daily functioning [4–7]. The most commonly used tests to evaluate these disorders focus solely on sleep and wakefulness. These tests are either subjective, e.g. the Epworth Sleepiness Scale (ESS), or objective, using polysomnography (PSG) and/or the Multiple Sleep Latency Test (MSLT) [8–10]. Daytime vigilance is still rarely assessed in clinical practice of

* Corresponding author. Achterweg 5, 2103 SW, Heemstede, the Netherlands.
E-mail address: r.fronczek@lumc.nl (R. Fronczek).

https://doi.org/10.1016/j.sleep.2023.06.007

hypersomnolence disorders. Vigilance tests such as the Sustained Attention to Response Task (SART) generally take little time and provide important insight into an often neglected disease aspect [11].

The SART is a short (<5 min), and inexpensive computer task in which the subject should click a button when a target appears on the screen and inhibit to a non-target. It can be easily used in combination with the MSLT, as it is assessed five times over the day, e.g. just before each MSLT session [12,13]. It was initially developed in 1997 to measure vigilance in individuals with traumatic brain injury and has since been used in various disorders, including central disorders of hypersomnolence: narcolepsy types 1 and 2 (NT1, NT2) and idiopathic hypersomnia (IH) [12–15]. The Psychomotor Vigilance Task (PVT) [16] and Oxford Sleep Resistance test (OSLER) [17] have also been used to measure vigilance impairment in sleep-wake disorders. However the SART is more sensitive to measure treatment effects in individuals with narcolepsy type 1 than the PVT [18], and while the OSLER is also suitable to measure vigilance, it takes considerably more time to perform [17].

We previously showed that the SART could differentiate between healthy controls and individuals with NT1 [12], but not between NT1, NT2 and IH [13]. The study also showed that the NT1 group was relatively slower in their reaction to the presented stimuli and that they made more mistakes on the SART in morning sessions than in the afternoon sessions [13]. This was either due to a learning or a time-of-the-day effect. Follow-up research among healthy subjects [15] showed that this was presumably due to a learning effect, but that applied to the specific test instruction that accuracy was more important than speed, whereas in our earlier research among patients with sleep disorders a different instruction was used, namely that accuracy and speed were equally important. No correlations were found between the SART scores and the MSLT or ESS outcomes. Thus, SART scores reflected a different aspect of disorders of hypersomnolence [12,13]. However, the relationship between SART and PSG results has not yet been assessed.

This report provides clinical data from a tertiary sleep center, collected over a period when the SART was routinely assessed in the diagnostic workup of suspected hypersomnolence. In addition to providing data regarding the daily clinical practice of implementing the SART, we explored whether the SART (1) can be used to distinguish between different disorders of hypersomnolence, (2) measures a different disease aspect than the PSG, MSLT or ESS, (3) is affected by testing time in clinical practice, and (4) is affected by anxiety or depression (as measured using the Hospital Anxiety and Depression Scale; HADS).

2. Methods

2.1. Data collection

Data were collected between March 2014 and October 2021 for clinical purposes in a tertiary Sleep-Wake center (SEIN, Heemstede, the Netherlands). Included were data of people (≥16 years old) who completed the SART, MSLT, PSG and/or ESS as part of routine diagnostic workup (the diagnostic sample), or completed the SART and ESS for their driver's license evaluation (the driver's license sample). Individuals with multiple sleep diagnoses were excluded, except for mild obstructive sleep apnea (OSA; i.e. apnea-hypopnea index (AHI) <15). This secondary sleep diagnosis is often an incidental finding after a diagnostic polysomnography [19,20].

Experienced neurologists-somnologists made the clinical diagnoses of people in the diagnostic sample. The diagnoses of people in the driver's license sample were verified based upon clinical information obtained from the electronic health records from the various sleep-wake clinics or (if this was not possible) extracted from referral letters. Individuals were classified into the following primary diagnostic groups: NT1, NT2, IH, OSA or complaints of excessive daytime sleepiness without explanatory diagnosis (CEDS). The CEDS group consisted of the following categories: insomnia (20.0%), no primary sleep diagnosis (18.2%), unclear diagnosis but no central hypersomnolence disorder (16.4%), behaviorally induced insufficient sleep syndrome (10.0%), restless legs syndrome (6.4%), suboptimal sleep hygiene (6.4%), mood issues (3.6%), circadian rhythm disorder (3.6%), psychogenic non-epileptic seizures (2.7%), hypersomnolence due to a medical disorder (2.7%), and 10.0% other. A secondary analysis was performed on the individuals strictly meeting the International Classification of Sleep Disorders (third edition, ICSD3) criteria (taking into account the results of any previous diagnostic testing performed) and the CEDS group and can be found in the supplementary material.

2.2. Diagnostic workup

The diagnostic workup consisted of the ESS, SART, and MSLT performed in a clinical setting on a single day. A PSG was performed either in the person's home or a clinical setting on a preceding night (see Fig. 1). Before the first SART session, individuals had the opportunity to practice the SART for 30 s. A SART session preceded every MSLT nap opportunity.

2.3. Materials

2.3.1. Sustained Attention to Response Task (SART)

The SART [21] is a short (4 min and 20 s) computerized go/no-go task and was administered on a laptop model HP ProBook 6570 b with a monitor refresh of 60 Hz and utilizing the operating system Windows 7 Professional. To improve the reliability of measured reaction times, the SART computer program was executed in priority mode, with minimal background programs running. It is performed in a quiet room with dimmed lights, which is the same room where the MSLT is also conducted. Numbers 1 to 9 are consecutively presented in a random sequence on a computer screen. Participants are instructed to press a button whenever a number (the target) appears on the screen, except for number 3 (the non-target). Each number is shown 25 times (225 in total), for 250 ms each, as white numbers on a black screen, followed by a black screen for 900 ms. Individuals have to respond before the following number appears. Participants are instructed to aim for accuracy over speed [13]. The main outcome is the mean total error score over the five SART sessions, consisting of the sum of commission errors (pressing a key after a non-target) and omission errors (not pressing a key after a target). Another SART outcome was the mean reaction time (SART RT) which is estimated per session over correct responses.

2.3.2. Multiple Sleep Latency Test (MSLT)

The MSLT is an objective test that can provide EDS or daytime sleep pressure information. The MSLT is performed under standard conditions with dimmed light. Individuals undergoing this test lie in bed and are instructed to nap for up to 20 min on five occasions on a single day, with electro-encephalographic recording of sleep [22]. The outcome parameters of the MSLT are sleep latency (SL, with a cut-off of ≤8 min), and the presence of sleep-onset REM periods (SOREMP, with a cut-off of ≥2 SOREMPs), which are part of the diagnostic criteria of NT1 and NT2 [9].

2.3.3. Polysomnography (PSG)

From the PSG performed the preceding night, the parameters relevant to the diagnostic process based on the official ICSD3

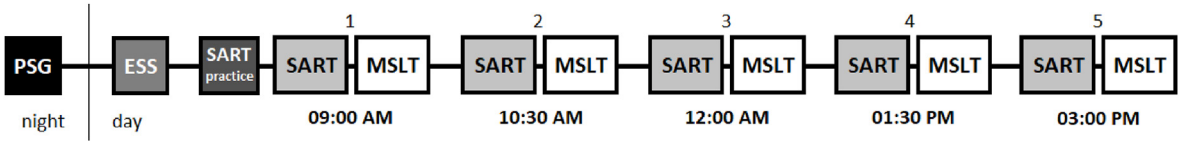

Fig. 1. Representation of our routine diagnostic work-up in a clinical setting. During the night a polysomnography (PSG; black) is performed, followed by the Epworth Sleepiness Scale (ESS; dark grey), five sessions of the Sustained Attention to Response Task (SART; light grey) the first preceded by a SART practice round and the Multiple Sleep Latency Test (MSLT; white).

criteria were used: sleep latency (PSG SL), duration of the time in bed (TIB), total sleep time (TST), sleep efficiency (SE, i.e. TST/TIB*100%) and the REM-sleep latency, which was used to determine whether a SOREMP was present [23]. Furthermore, the apnea hypopnea index was collected, representing the average number of apneas and hypopneas per hour individuals experience during the night. The hypopnea definition used in this study was a decrease of at least 30% in airflow, coupled with a desaturation of at least 3%, as is recommended by the American Academy of Sleep Medicine (AASM) [24].

2.3.4. Epworth Sleepiness Scale (ESS)

The ESS [10] is a self-report questionnaire developed to measure subjective daytime sleepiness. Individuals are asked to estimate their likelihood of falling asleep in certain situations. Total scores range from 0 to 24; the higher the score, the higher the subjective sleepiness during the day. A cut-off of $\geq$10 is used to indicate EDS [12,25].

2.3.5. Hospital Anxiety and Depression Scale (HADS)

The HADS is a questionnaire screening tool for anxiety and depression [26]. A cut-off of $\geq$8 (out of 21) indicates increased anxiety or depression symptoms [27].

2.4. Ethics statement

The study was conducted following the Helsinki Declaration as revised in 2013. Due to the historical nature of data, the Medical Ethical Committee of Leiden-Den Haag-Delft (registration number: G20.044) allowed a waiver of the requirement for informed consent.

2.5. Data availability statement

The data are available from the corresponding author upon reasonable request.

2.6. Data analyses

Data were analyzed using SPSS version 24 (Chicago, IL). Unless mentioned otherwise, a statistical significance level of α=0.05 (2-tailed) was used.

Prevalences (frequencies and percentages) were used to describe categorical variables. Continuous data were presented using means and standard deviations or median and interquartile range (IQR) depending on the distribution. Pearson's Chi-Square test compared categorical data except where an expected count was below 5, then Fisher's exact test was used. Depending on distributions, one-way ANOVA or the Kruskal-Wallis test was used to compare multiple groups. In case of significant differences across three or more groups, post hoc analyses were performed to examine which groups differed significantly, using pairwise comparisons (Dunn's test) in the case of the Kruskal-Wallis test and multiple comparisons (Tukey's test) in the case of one-way ANOVA analysis. Pairwise Chi-Square tests were performed in case of categorical variables.

We first used an ANOVA or Kruskal-Wallis test, depending on the distribution, to determine whether SART outcomes differed significantly between groups (aim 1). We used multiple linear regression models for the significant outcomes of the SART, corrected for age, BMI and sex.

Regression analyses were used to determine whether SART outcomes were related to MSLT, ESS, PSG (aim 2) or HADS (aim 4) outcomes after verifying that the relevant key assumptions were met: linearity, multivariate normality, no multi-collinearity, homoscedasticity and independence of observations. Univariate (with reaction time, commission and omission errors as independent variables) and multivariate linear regression analyses (with additional variables age, sex and BMI) were performed. The dependent variables used in separate models were: MSLT SL, PSG TST, ESS, HADS total, anxiety and depression scores. The variables were entered simultaneously. One outlier (defined as a standardized residual more prominent than 3 in absolute value) was excluded in the association between SART outcomes and the MSLT SL and the PSG TST. As the distributions of the PSG SL, AHI and SE residuals were skewed, Spearman's Rank correlation was used. Where there were significant associations, log transformations were applied to the PSG SL, AHI and SE outcomes and further analyzed in a regression analysis with correction for age, BMI and sex.

The Friedman test and post hoc testing Wilcoxon Signed Ranks Test were used to determine any impact of the five SART session times on SART outcomes (aim 3).

A secondary analysis was performed exclusively on the outcomes of individuals strictly meeting the ICSD3 criteria and those of the CEDS group [28]. The outcomes of this secondary analysis were similar to the primary analysis and can be found in the accompanying supplementary materials.

3. Results

3.1. Study population

Data of 306 individuals were collected, of whom 233 (76%) were tested as part of the normal diagnostic procedure and 73 (24%) were tested as part of their driver's license medical evaluation. Individuals were categorized according to their final clinical diagnosis: NT1 (diagnostic sample: n=43, driver's license sample: n=57), NT2 (diagnostic sample: n=10, driver's license sample=10), IH (diagnostic sample: n=43, driver's license sample 6), OSA (n=27) or CEDS (n=110). Of these, 77.0% of individuals with NT1, 70.0% with NT2, 42.9% with IH and 100% with OSA strictly met the ICSD3 criteria. 3 out of 27 individuals from the OSA group had received CPAP during the diagnostic procedures, which they underwent due to remaining hypersomnolence complaints in spite of treatment.

Table 1
Descriptive statistics of characteristics and diagnostic outcomes for all subgroups of the diagnostic sample.

Diagnostic sample (n=233)	1 NT1 (n=43)	2 NT2 (n=10)	3 IH (n=43)	4 OSA (n=27)	5 CEDS (n=110)	Test statistic	p-value
Age in years	27 (21–44) *2,4,5	24 (19–27) *1,4,5	29 (24–41) *4,5	58 (47–65) *1,2,3,5	47 (29–54) *1,2,3,4	H=64.942	<.001*
Sex: count (% male)	21 (48.8) *2,4	0 (0.0) *1,3,4,5	16 (37.2) *2,4	24 (88.9) *1,2,3,5	48 (43.6) *2,4	FET	<.001*
BMI	25.0 (23–28) *4	23.1 (21–27) *4	24.6 (21–28) *4	29.9 (26–33) *1,2,3,5	25.2 (22–29) *4	H=18.899	.001*
Use of antidepressants: count (%)	0 (0.0) *3,5	0 (0.0)	7 (16.3) *1	3 (11.1)	23 (20.9) *1	FET	.003*
PSG:	n=42	n=10	n=41	n=27	n=108		
SL in minutes	5.4 (2–9) *3,4,5	4.7 (2–10) *3,4,5	10.4 (7–15) *1,2,5	10.8 (7–16) *1,2	15.4 (7–26) *1,2,3	H=33.400	<.001*
TST in hours: mean ±SD	6.4 ±1.1	7.3 ±1.4	7.1 ±0.7 *4,5	6.2 ±1.2 *3	6.4 ±1.2 *3	F=5.038	.001*
TIB, hours, mean ±SD	7.6 ±1.0	7.8 ±1.4	7.9 ±0.8	7.5 ±1.0	7.5 ±1.1	F=1.167	.326
SE	88.3 (80–92) *2,3	94.7 (91–96) *1,4,5	92.5 (89–94) *1,4,5	84.2 (79–93) *2,3	87.5 (82–91) *2,3	H=28.808	<.001*
SOREMP present, count (%)	20 (47.6) *3,4,5	5 (50.0) *3,4,5	1 (2.4) *1,2	0 (0.0) *1,2	1 (0.9) *1,2	FET	<.001*
AHI	1.4 (0.1–4.6) *3,4	0.4 (0.2–1.8) *4	0.4 (0.1–1.3) *1,4,5	18.4 (13.7–31.8) *1,2,3,5	1.6 (0.2–4.2) *3,4	H=77.891	<.001*
Location: count (% ambulatory)	38 (90.5)	10 (100.0)	34 (82.9)	24 (88.9)	99 (91.7)	FET	.501
MSLT:	n=42	n=10	n=43	n=26	n=109		
SL, minutes	4.6 (3–7) *3,4,5	3.7 (2–5) *3,4,5	8.5 (6–11) *1,2,5	10.0 (6–13) *1,2,5	13.5 (11–16) *1,2,3,4	H=91.164	<.001*
Number of SOREMPs	3 (2–4) *3,4,5	3 (1–4) *3,4,5	0 (0–0) *1,2	0 (0–0) *1,2	0 (0–0) *1,2	H=144.557	<.001*
SOREMP present, count (%)	36 (85.7) *3,4,5	8 (80.0) *3,4,5	5 (11.6) *1,2	3 (11.5) *1,2	5 (4.6) *1,2	FET	<.001*
Questionnaires:	n=36	n=7	n=34	n=21	n=81		
ESS, n	16 (14–19), 36 *5	16 (15–18), 7	15 (12–17), 34	15 (10–19), 21	13 (10–16), 76 *1	H=18.315	.001*
HADS, n	11.5 (7–14), 32	5.5 (3–11), 6	11.5 (8–15), 28	12.0 (5–21), 13	11.0 (7–15), 81	H=6.263	.180
-Anxiety	7.0 (4–9)	2.5 (1–4)	6.0 (3–8)	7.0 (3–11)	6.0 (4–9)	H=8.518	.074
-Anxiety score ≥8, count (% ≥8)	11 (34.4)	0 (0.0)	9 (32.1)	4 (30.8)	27 (33.3)	FET	.603
-Depression	5.0 (3–7)	3.0 (2–6)	6.0 (3–9)	5.0 (2–10)	5.0 (3–9)	H=3.010	.556
-Depression score ≥8, count (% ≥8)	7 (21.9)	1 (16.7)	9 (32.1)	6 (46.2)	24 (29.6)	FET	.554

Abbreviations: NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; OSA, obstructive sleep apnea; CEDS, complaints of excessive daytime sleepiness; FET, Fisher's exact test; BMI, body mass index; PSG, polysomnography; SL, sleep latency; TST, total sleep time; TIB, time in bed; SE, sleep efficiency; SOREMp, sleep-onset rapid eye movement period; AHI, apnea-hypopnea index; MSLT, Multiple Sleep Latency Test; ESS, Epworth Sleepiness Scale; HADS, Hospital Anxiety and Depression Scale.
Notes: Median (IQR) was used unless specified otherwise. *p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH, (4) OSA, (5) CEDS.

Table 2
Descriptive statistics and diagnostic outcomes for all subgroups of the driver's license sample.

Driver's license sample (n=73)	1 NT1 (n=57)	2 NT2 (n=10)	3 IH (n=6)	Test statistic	p-value
Age in years	37 (25–48)	41 (30–58)	29 (27–44)	H=1.469	.480
Sex, count (% male)	32 (56.1)	7 (70.0)	2 (33.3)	FET	.373
BMI, n	26.8 (23–30), 53	25.1 (24–33), 10	21.7 (19–26), 6	H=4.722	.094
Medication:	n=57	n=10	n=6		
Use of stimulants, SXB and/or AD: count/n (%)	52 (91.2)	9 (90.0)	4 (66.7)	FET	.195
Stimulants: count (%)	42 (73.7)	8 (80.0)	4 (66.7)	FET	.802
Sodium oxybate: count (%)	27 (47.4) *2,3	1 (10.0) *1	0 (0.0) *1	FET	.008*
Antidepressant: count (%)	14 (24.6)	0 (0.0)	0 (0.0)	FET	.088
Questionnaires:	n=56	n=10	n=6		
ESS	7.0 (6–9) *3	8.5 (7–9) *3	5.5 (3–7) *1,2	H=6.253	.044*
SART outcomes:	n=57	n=10	n=6		
Commission errors	4.8 (2–7)	3.5 (2–5)	2.0 (2–4)	H=3.154	.207
Omission errors	0.3 (0–1)	0.6 (0–2)	0.4 (0–1)	H=.514	.774
Total errors	5.8 (3–9)	3.9 (2–7)	2.8 (2–5)	H=3.663	.160
Reaction time, ms	413 (359–468)	384 (350–434)	434 (402–515)	H=3.417	.181

Abbreviations: NT1, narcolepsy type 1; NT2, narcolepsy type 2; IH, idiopathic hypersomnia; FET, Fisher's exact test; BMI, body mass index; SXB, sodium oxybate; AD, antidepressant; ESS, Epworth Sleepiness Scale; SART, sustained attention to response task; Driver's license sample, individuals with treated sleep disorders as concluded by the Dutch central driving license office.
Notes: Median (IQR) was used unless specified otherwise. *p<.05. Significant pairwise difference with: (1) NT1, (2) NT2, (3) IH.

Descriptive statistics are presented in Tables 1 and 2. The outcomes of the secondary analysis were congruent with those of the primary

analysis and can be seen in the supplementary material (tables S1, S2 and S4).

3.2. Individual characteristics and diagnostic outcomes

In the diagnostic sample, the OSA group was significantly older and had a significantly higher BMI than the other groups. The OSA group also had substantially more males ($p<.05$). The NT2 group was younger and had a lower proportion of males (0%) than the NT1, CEDS and OSA groups ($p<.05$).

The NT1 and NT2 groups had a shorter MSLT and PSG sleep latency and more SOREMPs than the other groups (all $p<.05$). SOREMPs were not limited to narcolepsy patient groups: 5 out of 109 individuals with CEDS, 3 out of 26 individuals with OSA and 6 out of 43 individuals with IH had at least one SOREMP during the MSLT and PSG. ESS scores were higher in the NT1 group than CEDS group ($p<.05$). The percentage of antidepressant use was lower in the narcolepsy type 1 group than the IH and CEDS groups ($p<.05$). Sodium oxybate and stimulants were not used in any of the diagnostic sample subgroups. HADS scores did not differ significantly between groups.

The characteristics of the driver's license group (see Table 2) did not differ between diagnostic groups (NT1, NT2 and IH).

3.3. Vigilance across diagnostic groups

Differences in SART outcomes were analyzed between diagnostic groups (see Table S3). Fig. 2 illustrates the median outcomes, the IQR, the minimum and maximum for each group and significance levels corrected for age, sex and BMI. No significant differences were found. Additional correction for antidepressant use did not substantially impact the outcomes.

3.4. Confounders

Age was negatively associated to SART commission errors ($B=-.058$, $p=.013$, $R^2=.024$), and positively associated to SART reaction time ($B=1.468$, $p<.001$, $R^2=.092$). Age was not associated to other SART outcomes. Correction for diagnosis did not change these outcomes. Sex and BMI were not associated with any SART outcomes.

3.5. Associations between SART and PSG, MSLT and ESS

There were no associations between MSLT sleep latency and any SART outcomes in both the univariate and multivariate regression analysis. PSG TST, SE and SL were also not significantly associated with SART RT, commission or omission errors. A higher AHI was associated with a longer reaction time when not adjusted for confounders ($p<.001$), however after adjusting for age, sex and BMI the association was no longer significant ($p=.259$).

There were significant positive associations between the ESS score and SART RT and commission errors, in the univariate and multivariate regression analyses. I.e. subjects with higher sleepiness scores tend to react slower and make more commission errors (see Table 3). These significant associations remained when only

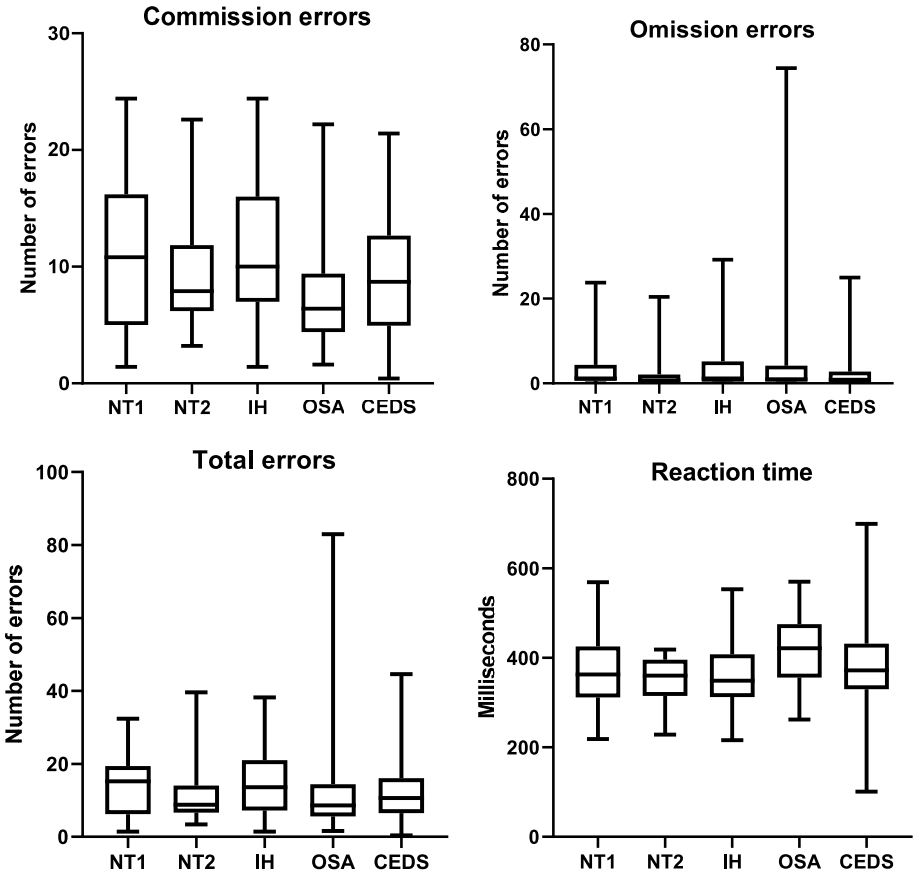

Fig. 2. Differences in Sustained Attention to Response Task outcomes between diagnostic sample groups: narcolepsy type 1 (NT1, n=43), narcolepsy type 2 (NT2, n=10), idiopathic hypersomnia (IH, n=43), obstructive sleep apnea (OSA, n=27), individuals with complaints of excessive daytime sleepiness (CEDS, n=110). Boxplots represent median, IQR, minimum and maximum.

Table 3
Associations between ESS as dependent variable, and: SART outcomes (univariate) or SART outcomes and general characteristics (multivariate).

	Dependent variable: ESS score			
	Univariate (n=245)		Multivariate (n=241)	
	Estimate	p-value	B	p-value
SART Reaction time	.018	.006*	.021	.002*
SART Commission errors	.448	<.001*	.458	<.001*
SART Omission errors	−.089	.295	−.072	.417
Age	[a]	[a]	−.028	.228
Sex	[a]	[a]	1.023	.110
BMI	[a]	[a]	.071	.262
Model information	R^2= .093, p<.001*		R^2=.108, p<.001*	

Abbreviations: ESS, Epworth Sleepiness Scale; SART, sustained attention to response task.
Notes: *p<.05.
[a] Unused variables in univariate regression analysis.

including individuals from the diagnostic sample (see Table S5) and adding antidepressant use as independent variable in the multivariate analysis.

3.6. Relationship between depression, anxiety, and vigilance

There was no significant association between HADS total score (or the anxiety and depression sub-scores) and any SART outcome.

Comparisons of SART outcomes between individuals from the diagnostic sample group based on HADS depression and anxiety cut-off scores are shown in Table 4. People with a higher anxiety score (≥8) made near-significantly more omission errors than those with a lower score (<8, p=.055).

3.7. Time of day effects on vigilance outcomes

Commission and total errors did not differ significantly between SART sessions (see Fig. 3).

Reaction time was significantly longer during the first 3 sessions than in the last 2 (median and IQR: session 1, 382 (343–439); session 2, 381 (332–446); session 3, 380 (328–435); session 4, 373 (326–422); session 5, 355 (314–414), in all comparisons p<.05).

Omission errors, although scarce, differed significantly between sessions 1 and 2 (median and IQR 1 (0–3) vs 0 (0–2), p<.005), 1 and 3 (median and IQR 0 (0–2), p<.001), 2 and 3 (0 (0–2)), 3 and 4 (0 (0–2) vs 1 (0–2), p<.05) and 3 and 5 (1 (0–3), p<.05).

3.8. Driver's license subgroup

The number of SART errors (omission, commission and total errors) was lower in the driver's license group than in the diagnostic sample (in the NT1 group 0.3 vs 1.1, 4.8 vs 10.6 and 5.8 vs 13.8, see Supplemental Table S2). On the other hand, reaction times were longer in the driver's license sample (in the NT1 group 413 vs 361 ms).

4. Discussion

We evaluated the applicability of the SART in the diagnostic workup and driver's license evaluations of a tertiary referral center for disorders of hypersomnolence. We assessed the association between SART and PSG outcomes. Additionally, while associations with ESS and MSLT outcomes were previously reported, we examined them in larger samples relevant to clinical practice, including a sample consisting of people with EDS complaints without a diagnosis of a disorder of hypersomnolence. While associations between ESS and MSLT and vigilance outcomes have been assessed before in individuals without explanatory diagnosis [29,30], the SART was not explicitly evaluated. We conclude that the SART (1) can detect disturbed vigilance in hypersomnolence disorders but cannot differentiate between different conditions, (2) measures a different EDS disease aspect than the MSLT and PSG, (3) is influenced by the time of day and (4) is not influenced by depression. Notably, individuals with increased anxiety (sub-score ≥8) did make near-significantly more omission errors. Additionally, (5) total, commission and omission errors are lower in individuals undergoing a driver's license evaluation, probably because of treatment and/or motivational factors.

It is worth noting that there is currently no universally accepted cut-off for the SART total error score in the context of hypersomnolence disorders. While a cut-off of 5 total errors was established by Fronczek et al. (2006) in a small study with 15 untreated narcoleptics and 15 matched controls, it has not been further validated in larger or more diverse populations [12]. Despite this, our assessment of different hypersomnolence disorders demonstrated impaired vigilance, as evidenced by all groups exhibiting a median total error score above this previously determined cut-off. Previous research [15], conducted with healthy participants, showed that this cut-off was only applicable when using the SART with the instruction that accuracy is more important than response speed. Therefore, the finding of our previous descriptive study [13] that individuals with hypersomnolence disorders have abnormal SART outcome measures became questionable. However, abnormal SART results have now also been reproduced in individuals who were instructed to prioritize accuracy, thereby providing further support to previous findings. Nonetheless, to arrive at definitive

Table 4
Comparisons in SART outcomes based on HADS depression and anxiety cut-off scores.

Diagnostic sample (n=160)	Depression<8 (n=113)	Depression≥8 (n=47)	Test statistic	p-value
Commission errors	8.6 (5–14)	9.4 (5–14)	H=2862	.440
Omission errors	0.8 (0–3)	1.0 (0–3)	H=2779	.644
Total errors	9.8 (6–18)	11.0 (6–21)	H=2895	.371
Reaction time, ms	372 (326–427)	346 (316–446)	H=2570	.749
	Anxiety<8 (n=109)	Anxiety≥8 (n=51)		
Commission errors	8.8 (5–14)	9.0 (5–14)	H=2879	.591
Omission errors	0.8 (0–3)	1.4 (0–3)	H=3302	.055
Total errors	9.8 (6–18)	11.0 (7–19)	H=3136	.192
Reaction time, ms	370 (323–423)	367 (316–455)	H=2817	.892

Abbreviations: SART, sustained attention to response task; HADS, Hospital Anxiety and Depression Scale.
Notes: Median (IQR) was used unless specified otherwise. *p<.05.

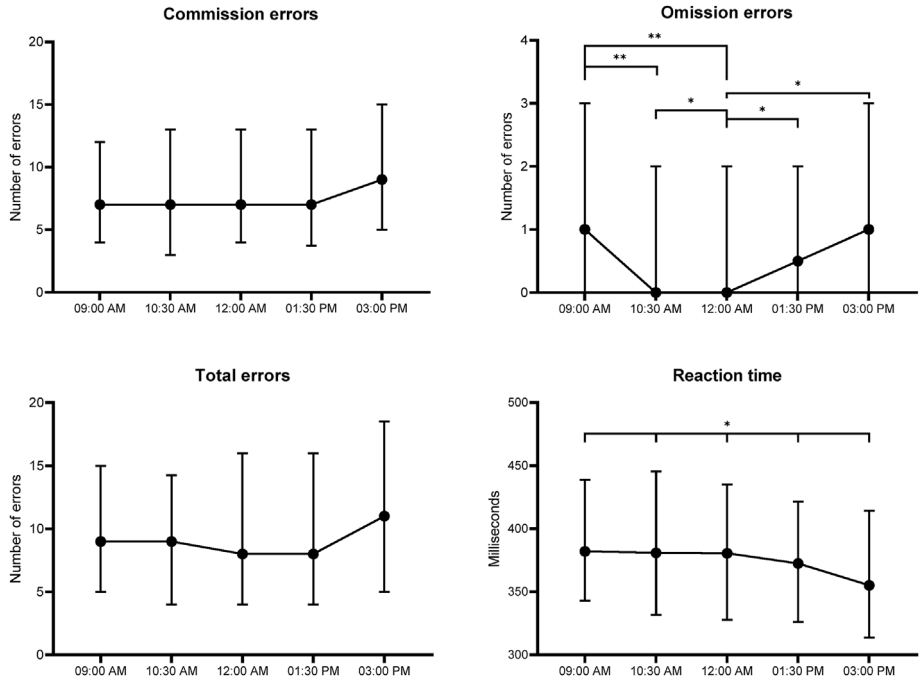

Fig. 3. Differences in Sustained Attention to Response Task outcomes between five sessions starting at 09:00 a.m., 10:30 a.m., 12:00 a.m., 01:30 p.m. and 03:00 p.m. Median and IQR are shown. *p<.05; **p<.005. Reaction time differed significantly between all sessions.

conclusions, further research is required, which should involve both individuals with hypersomnolence disorders and a healthy control group, as well as testing the effects of differential instructions.

In line with previous research [13], we conclude that the SART is not a diagnostic tool for a specific sleep disorder. However, it may have potential as a monitoring tool during follow-up. The SART has already been proven capable of measuring treatment effects in individuals with narcolepsy [31,32].

There was an overlap between subjective EDS as measured by the ESS and vigilance impairment as measured by the SART. To rule out a confounding effect of treatment on these associations, a post hoc analysis was performed using only the diagnostic sample and adjusting for antidepressant use. This association was not found between the SART on the one hand, and PSG and MSLT outcomes on the other hand. Vigilance impairments are thus associated with subjective sleepiness but not with objective sleepiness. This suggests that the SART measures a different disease aspect, which is supported by multiple, previously conducted studies [33–35].

We found a clear association between SART outcomes and age, where older individuals tend to have longer reaction times. Increased age was associated with longer reaction times and fewer commission errors. This may be related to the speed-accuracy trade-off [15], where longer reaction times result in fewer errors. These results are in line with expectations, given the known relationship between decreasing reaction times with increasing age [36,37]. It is possible that age needs to be taken into account when interpreting SART results. Notably, the primary SART outcome measure, total errors, was not significantly associated with age, making this association less relevant in clinical practice.

The number of total errors also did not differ significantly between different times of the day. The time of the day significantly impacted reaction times and the number of omission errors, with reaction time being higher in the morning and fewer omission errors around noon. The differences in omission errors were marginal and not clinically relevant. Based on previous research we suspect these differences are more likely to be caused by a time of day effect than a learning effect [15]. Unlike two previous studies by our research group, we did not find that total error scores were significantly higher in the first session [13,15]. The longer reaction times in the early morning confirmed past results [15]. We found no evidence of a clinically relevant time-of-day effect on total errors scores. However the time of the day should be taken into account when interpreting the reaction times.

There were no associations between subjective depression symptoms and any SART outcomes. This may seem counterintuitive as multiple studies have shown attention deficits in people with depression [38,39]. In our study population only two individuals had a cut-off score $\geq$16 on the HADS (both from the CEDS group), indicating severe depression symptoms [27]. Therefore we cannot reliably examine the effect of depression on SART outcomes. It was, however, found that individuals with increased anxiety (sub-score $\geq$8) made near-significantly more omission errors. This finding aligns with a previous report that anxiety is associated with response inhibition [40]. As with depression, there were few people in this study with severe anxiety complaints (the maximum anxiety sub-score in the sample was 13/21).

The impact of motivation on SART performance is an interesting topic that needs further examination. The driver's license group made fewer errors during the SART than the diagnostic group. There were substantial differences between these groups: the driver's license group may have been more motivated to stay awake and had received treatment. Despite treatment, individuals with severe complaints would not be considered for a driver's license testing, resulting in selection bias. These factors make it difficult to determine the cause of differences between the groups.

4.1. Strengths and limitations

Our study population reflected the daily clinical practice of a highly specialized Sleep-Wake center. Therefore, our OSA group is not generalizable to the general population with OSA, as not all

people with OSA will experience complaints of hypersomnolence severe enough to result in referral to a Sleep-Wake center.

Our diagnostic groups were relatively large, but the NT2 group consisted of only 10 individuals, thus making it difficult to draw definite conclusions regarding this sub-group.

In addition, it should be noted that the SART program used in our study was run in priority mode and was not performed on a calibrated platform, which may have introduced some variability in the results. However, the impact on the results is expected to be small given the large sample size used.

Another limitation is related to the setup of the diagnostic procedures. The PSG is always followed by tests during the following day starting at 09:00 a.m., this has probably shortened the natural sleep duration of at least some of those included. This means that the PSG TST and TIB results are probably underestimations and need to be interpreted with care.

Finally, no healthy control group was included in this study with which the sub-groups could be compared.

4.2. Conclusion

Impaired vigilance can significantly affect quality of life, as vigilance is needed for everyday tasks at work, school, parenting tasks, household, and social interactions such as a conversation [41,42]. Thus, not only EDS complaints but also vigilance impairment should be regularly monitored in individuals with central disorders of hypersomnolence. The SART can be used to assess vigilance in these people. Unlike the MSLT and PSG, the SART is a short and easy to administer task and a good addition to the standard diagnostic workup and monitoring of central disorders of hypersomnolence.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Adrienne Elisabeth van der Hoeven: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Denise Bijlenga:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Puck Bouhuijs:** Investigation, Formal analysis. **Mojca Kristina Maria van Schie:** Conceptualization, Writing – review & editing. **Gert Jan Lammers:** Conceptualization, Methodology, Writing – review & editing. **Rolf Fronczek:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of competing interest

Authors declare none.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sleep.2023.06.007.

References

[1] Davies DR, Parasuraman R. The psychology of vigilance. Academic Press; 1982.
[2] Oken BS, Salinsky MC, Elsas SM. Vigilance, alertness, or sustained attention: physiological basis and measurement. Clin Neurophysiol : official journal of the International Federation of Clinical Neurophysiology 2006;117(9):1885–901.
[3] van Schie MKM, et al. Vigilance: discussion of related concepts and proposal for a definition. Sleep Med 2021;83:175–81.
[4] Khan Z, Trotti LM. Central disorders of hypersomnolence: focus on the narcolepsies and idiopathic hypersomnia. Chest 2015;148(1):262–73.
[5] Al-Shargie F, et al. Vigilance decrement and enhancement techniques: a review. Brain Sci 2019;9(8):178.
[6] Mathis J, Hess CW. Sleepiness and vigilance tests. Swiss Med Wkly 2009;139(15–16):214–9.
[7] Lyznicki JM, et al. Sleepiness, driving, and motor vehicle crashes. Council on scientific affairs, American medical association. JAMA 1998;279(23):1908–13.
[8] Komada Y, et al. Difference in the characteristics of subjective and objective sleepiness between narcolepsy and essential hypersomnia. Psychiatr Clin Neurosci 2005;59(2):194–9.
[9] Carskadon MA, et al. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. Sleep 1986;9(4):519–24.
[10] Johns MW. Reliability and factor analysis of the Epworth sleepiness scale. Sleep 1992;15(4):376–81.
[11] Jung CM, et al. Comparison of sustained attention assessed by auditory and visual psychomotor vigilance tasks prior to and during sleep deprivation. J Sleep Res 2011;20(2):348–55.
[12] Fronczek R, et al. Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART). Sleep 2006;29(2):187–91.
[13] Van Schie MK, et al. Sustained attention to response task (SART) shows impaired vigilance in a spectrum of disorders of excessive daytime sleepiness. J Sleep Res 2012;21(4):390–5.
[14] Whyte J, et al. Measuring sustained attention after traumatic brain injury: differences in key findings from the sustained attention to response task (SART). Neuropsychologia 2006;44(10):2007–14.
[15] van Schie MK, et al. The influences of task repetition, napping, time of day, and instruction on the Sustained Attention to Response Task. J Clin Exp Neuropsychol 2014;36(10):1055–65.
[16] Wilkinson RT, Houghton D. Field test of arousal: a portable reaction timer with data storage. Hum Factors 1982;24(4):487–93.
[17] Bennett LS, Stradling JR, Davies RJ. A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. J Sleep Res 1997;6(2):142–5.
[18] van Schie MK, et al. Improved vigilance after sodium oxybate treatment in narcolepsy: a comparison between in-field and in-laboratory measurements. J Sleep Res 2016;25(4):486–96.
[19] Chokroverty S. Sleep apnea in narcolepsy. Sleep 1986;9(1 Pt 2):250–3.
[20] Pataka AD, et al. Narcolepsy and sleep-disordered breathing. Eur J Neurol 2012;19(5):696–702.
[21] Robertson IH, et al. 'Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia 1997;35(6):747–58.
[22] Baiardi S, Mondini S. Inside the clinical evaluation of sleepiness: subjective and objective tools. Sleep Breath 2020;24(1):369–77.
[23] Ibanez V, Silva J, Cauli O. A survey on sleep assessment methods. PeerJ 2018;6:e4849.
[24] Iber C, et al. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. Westchester, IL: American Academy of Sleep Medicine; 2007.
[25] Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14(6):540–5.
[26] Brennan C, et al. The Hospital Anxiety and Depression Scale: a diagnostic meta-analysis of case-finding ability. J Psychosom Res 2010;69(4):371–8.
[27] Bjelland I, et al. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. J Psychosom Res 2002;52(2):69–77.
[28] American Academy of Sleep M. The international classification of sleep disorders. 2014.
[29] Ramm M, et al. Disease-specific attention impairment in disorders of chronic excessive daytime sleepiness. Sleep Med 2019;53:133–40.
[30] Mathis J, et al. The diagnostic value of sleep and vigilance tests in central disorders of hypersomnolence. Sleep 2022;45(3).
[31] Dauvilliers Y, et al. Pitolisant versus placebo or modafinil in patients with narcolepsy: a double-blind, randomised trial. Lancet Neurol 2013;12(11):1068–75.
[32] van der Heide A, et al. Comparing treatment effect measurements in narcolepsy: the sustained attention to response task, Epworth sleepiness scale and maintenance of wakefulness test. Sleep 2015;38(7):1051–8.
[33] Chervin RD, et al. Comparison of the results of the Epworth sleepiness scale and the multiple sleep latency test. J Psychosom Res 1997;42(2):145–55.
[34] Sangal RB, Mitler MM, Sangal JM. Subjective sleepiness ratings (Epworth sleepiness scale) do not reflect the same parameter of sleepiness as objective sleepiness (maintenance of wakefulness test) in patients with narcolepsy. Clin Neurophysiol 1999;110(12):2131–5.
[35] Furuta H, et al. Epworth Sleepiness Scale and sleep studies in patients with obstructive sleep apnea syndrome. Psychiatr Clin Neurosci 1999;53(2):301–2.
[36] Fozard JL, et al. Age differences and changes in reaction time: the baltimore longitudinal study of aging. J Gerontol 1994;49(4):P179–89.
[37] Bento-Torres NV, et al. Influence of schooling and age on cognitive performance in healthy older adults. Braz J Med Biol Res 2017;50(4):e5892.
[38] Li X, et al. Study on the executive function of attention in depression patients based on SPECT technology. Int J Clin Exp Med 2014;7(4):1110–5.
[39] Keller AS, et al. Paying attention to attention in depression. Transl Psychiatry 2019;9(1):279.

[40] Grillon C, et al. Clinical anxiety promotes excessive response inhibition. Psychol Med 2017;47(3):484–94.
[41] Valley V, Broughton R. Daytime performance deficits and physiological vigilance in untreated patients with narcolepsy-cataplexy compared to controls. Rev Electroencephalogr Neurophysiol Clin 1981;11(1):133–9.
[42] Lee I-S, et al. The Relationship between psychomotor vigilance performance and quality of life in obstructive sleep apnea. J Clin Sleep Med : JCSM : official publication of the American Academy of Sleep Medicine 2011;7(3):254–60.