Dissecting the cognitive phenotype of post-stroke fatigue using computerized assessment and computational modeling of sustained attention

Kristine M. Ulrichsen1,2,3,4 | Dag Alnaes1,2 | Knut K. Kolskar1,2,3,4 | Geneviève Richard1,2,3,4 | Anne-Marthe Sanders1,2,3,4 | Erlend S. Dørum1,2,3,4 | Hege Ihle-Hansen5 | Mads L. Pedersen3,6 | Sveinung Tornås4 | Jan E. Nordvik7 | Lars T. Westlye1,2,3,8

1NORMENT, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
2Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway
3Department of Psychology, University of Oslo, Oslo, Norway
4Sunnaas Rehabilitation Hospital HF, Nesodden, Norway
5Department of Neurology, Oslo University Hospital, Oslo, Norway
6Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, RI, USA
7CatoSenteret Rehabilitation Center, Son, Norway
8K.G. Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway

Abstract

Post-stroke fatigue (PSF) is prevalent among stroke patients, but its mechanisms are poorly understood. Many patients with PSF experience cognitive difficulties, but studies aiming to identify cognitive correlates of PSF have been largely inconclusive. With the aim of characterizing the relationship between subjective fatigue and attentional function, we collected behavioral data using the attention network test (ANT) and self-reported fatigue scores using the fatigue severity scale (FSS) from 53 stroke patients. In order to evaluate the utility and added value of computational modeling for delineating specific underpinnings of response time (RT) distributions, we fitted a hierarchical drift diffusion model (hDDM) to the ANT data. Results revealed a relationship between fatigue and RT distributions. Specifically, there was a positive interaction between FSS score and elapsed time on RT. Group analyses suggested that patients without PSF increased speed during the course of the session, while patients with PSF did not. In line with the conventional analyses based on observed RT, the best fitting hDD model identified an interaction between elapsed time and fatigue on non-decision time, suggesting an increase in time needed for stimulus encoding and response execution rather than cognitive information processing and evidence accumulation. These novel results demonstrate the significance of considering the sustained nature of effort when defining the cognitive phenotype of PSF, intuitively indicating that the cognitive phenotype of fatigue entails an increased vulnerability to sustained effort, and suggest that the use of computational approaches offers a further characterization of specific processes underlying behavioral differences.

Abbreviations: ANT, attention network test; CI, confidence interval; DIC, deviance information criteria; FSS, Fatigue Severity Scale; hDDM, hierarchical drift diffusion model; MCMC, Markov chain Monte Carlo; MMSE, Mini-Mental Status Examination; NIHSS, National Institutes of Health stroke scale; PHQ-9, Patient Health Questionnaire; PPC, posterior predictive checks; PSF, post-stroke fatigue; RT, reaction time; SD, standard deviation; TOAST, Trial of ORG 10,172 in Acute Stroke Treatment.

Edited by Susan Rossell

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. European Journal of Neuroscience published by Federation of European Neuroscience Societies and John Wiley & Sons Ltd
1 | INTRODUCTION

Post-stroke fatigue (PSF) is a common complaint among stroke survivors, with an estimated prevalence ranging between 25% and 85% (Cumming, Packer, Kramer, & English, 2016). The symptom burden is often pervasive and persistent (Duncan, Wu, & Mead, 2012; Schepers, Visser-Meily, Ketelaar, & Lindeman, 2006; van der Werf, van den Broek, Anten, & Bleijenberg, 2001) and associated with poorer outcome after rehabilitation, higher mortality (Michael, 2002; Naess, Lunde, Brogger, & Waje-Andreassen, 2012) and increased probability of institutionalization (Glader, Stegmayr, & Asplund, 2002). Post-stroke fatigue has been defined as a highly prioritized future research topic by stroke survivors, family members and healthcare professionals (Pollock, St George, Fenton, & Firkins, 2012).

Although a universally accepted definition is lacking (Deluca, 2005), PSF is generally conceptualized as the feeling of debilitating tiredness and loss of energy (Stulemeijer, Fasotti, & Bleijenberg, 2005). Moreover, many patients suffering from PSF experience cognitive difficulties such as problems concentrating (Johansson & Rönnbäck, 2012; Koopman et al., 2009) and report increased fatigue when engaging in cognitively demanding activities over time, often referred to as mental or cognitive fatigue (Johansson & Ronnback, 2014). To date, identifying robust and objective cognitive correlates of PSF has proven difficult, and the literature has failed to confirm or refute an association between self-reported fatigue and cognitive function (Lagogianni, Thomas, & Lincoln, 2018). However, this may partly be due to the use of multifactorial neuropsychological tests, with varying or low cognitive specificity and which do not account for the temporal aspects during the course of a test session (Holtzer, Shuman, Mahoney, Lipton, & Verghese, 2010). In line with this, many of the studies failing to identify an association use rather general measures of cognitive function such as the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975); van Eijsden, van de Port, Visser-Meily, & Kwakkel, 2012; Kutlabaev et al., 2013) and a recent review on factors associated with PSF concluded that although the evidence does not support a link between general cognitive function and PSF, there may be an association between attentional functioning, processing speed and fatigue (Pihlaja, Uimonen, Mustanoja, Tatlisumak, & Poutiainen, 2014; Ponchel, Bombos, Bordet, & Hémon, 2015).

With the assumption that a critical characteristic of cognitive fatigue is the failure to maintain or sustain cognitive effort over time, monitoring performance over time should increase sensitivity to cognitive manifestations of fatigue (Holtzer et al., 2010) and would also be closer in line with the conceptual definition of cognitive fatigue as “decreased performance during acute but sustained mental effort” (Deluca, 2005). Accordingly, the attentional network task (ANT; (Fan, McCandliss, Sommer, Raz, & Posner, 2002)) appears to be appropriate for examining the relationship between self-reported fatigue and attentional function over time in stroke patients. ANT combines a flanker test (Eriksen & Eriksen, 1974), and a cued reaction time task (Posner, 1980) in a computerized behavioral paradigm requiring sustained attention over time. The full version lasts for about 20 min, where accuracy and response times (RT) are tracked over time in 288 trials with varying cognitive demands. The ANT allows for estimation of individual-level attention network scores such as the alerting, orienting and executive components, defined as relative differences in average RTs between different flanker and cue conditions (Fan et al., 2002).

ANT has been applied in studies of fatigue and attention in other neurological patient groups, such as Parkinson’s disease, where fatigue was associated with reduced efficiency in the executive attentional network (Paulielli et al., 2017) and chronic fatigue syndrome, associated with higher RT in the most cognitively demanding condition (Togo, Lange, Natelson, & Quigley, 2015).

Although representing a widely applied and valuable contribution to theories on attentional function, analytical approaches based on mean RTs are vulnerable to trade-offs between speed and accuracy which are not accounted for in the model (Miller & Ulrich, 2013), and they do not provide information about which underlying mechanisms give rise to observed RT differences. In contrast, computational approaches such as the drift diffusion models (DDM; (Ratcliff, 1978)) simultaneously model the full distribution of RTs and accuracies to estimate parameters reflecting specific theoretical cognitive constituents of the decision process. DDMs are frequently...
applied to simple and speeded decision-making tasks (Ratcliff & McKoon, 2008; Ratcliff, Smith, Brown, & McKoon, 2016), offering both a theoretical framework to understand basic cognitive processes, and a psychometric tool to translate behavioral data into subcomponents of cognitive processing (Ratcliff & McKoon, 2008). DDMs conceptualize decision-making as a noisy process where information is accumulated over time, continuing until a decision threshold is reached and a response is initiated (Ratcliff & McKoon, 2008). Four parameters are postulated in the original model (Ratcliff, 1978): drift rate (ν), describing the rate or the speed of information accumulation, reflecting processing efficiency; non-decision time (t) representing time needed for stimulus encoding and response execution; decision boundary separation (α) indicating how much evidence is needed before a decision is made; and the starting point (z), reflecting any bias toward one of the two responses (Ratcliff & McKoon, 2008). The parameters have been validated in various experimental paradigms (Lerche & Voss, 2017; Voss, Rothermund, & Voss, 2004).

Applying computational models such as the DDM in clinical research may allow for a dissection of specific cognitive processes underlying observed group and individual differences in RT patterns. For example, assessing young and older subjects with a signal detection task, Ratcliff, Thapar, and McKoon (2001) found that the prolonged RTs often observed in older individuals were not explained by slower drift rates but rather longer non-decision times and higher decision thresholds, which provided a relevant adjustment to the long-held notion of a general slowing in cognitive aging (Brinley, 1965; Salthouse, 1985). In the context of stroke patients and PSF, such computational approaches may provide a valuable, supplementary tool to expand our understanding of cognitive function beyond conventional methods of neuropsychological assessment and statistical analysis.

In sum, a large number of stroke patients suffer from PSF, and many experience cognitive difficulties and cognitive fatigue. Attentional deficits may be particularly involved. The ANT paradigm allows us to determine whether and how subjective fatigue manifests cognitively during prolonged effort, and assess associations between subjective fatigue and efficiency of the attentional networks. With the aim of characterizing the relationship between subjective fatigue and attentional function, we collected behavioral data using the ANT and self-reported symptoms of fatigue using the fatigue severity scale (FSS; Krupp, LaRoca, Muir-Nash, & Steinberg, 1989) from 53 chronic stroke patients (>6 months since hospital admission). We hypothesized that self-reported symptoms of fatigue as measured by the fatigue severity scale (FSS; Krupp et al., 1989) would interact with time on task, manifesting in an increase in RT for patients with high fatigue levels relative to patients with low levels of fatigue. Further, we expected to find a negative association between fatigue and executive network functioning, in line with previously mentioned literature. Main analyses were conducted with FSS score as a continuous predictor, and follow-up sensitivity analyses were conducted with PSF group (high/low PSF) as a factor predictor, or separately for patients with high/low PSF to assess manifestation of group differences. Lastly, evaluating whether DDM modeling can elaborate our understanding further by characterizing the specific cognitive processes underlying observed differences in RT patterns, we performed an exploratory analysis where we fitted a hDDM to the ANT behavioral data and tested for associations between the model parameters (drift rate (ν), non-decision time (t) and boundary separation (α)) and fatigue (FSS) score. To account for the temporal aspects of task performance, we specifically tested for interactions between FSS, trial number and performance. In line with our first hypothesis, we hypothesized that any associations between subjective fatigue and model parameters will interact with time, with increasing associations between fatigue and model parameters with more sustained performance.

2 | MATERIALS AND METHODS

2.1 | Sample

Stroke patients who had been previously admitted with acute stroke to the Stroke Unit, Oslo University Hospital, or the Geriatric Department, Diakonhjemmet Hospital, between 2013 and 2016, were invited by letter. Patients had to be in a chronic phase, defined as minimum 6 months post-stroke, with no other severe neurological, psychiatric or neurodevelopmental conditions. Among the approximately 900 invitation letters, 250 patients responded to decline or obtain more information. Seventy-seven were interested and eligible for inclusion and provided informed consent. Nineteen of the 77 patients withdrew during the course of the study and before the data for the current paper were collected. Four additional patients were excluded because of medical conditions. One patient was excluded due to behavioral criteria for the ANT (see below), resulting in a final sample of n = 53 stroke patients.

Table 1 summarizes relevant demographic and clinical information of the patient group, and Figure 1 shows the age distribution. This work was part of an intervention study on cognitive rehabilitation after stroke with a double baseline, randomized controlled design (see Kolskaar et al., 2019) for more details, including a description of overall study design. All data for the current study were collected from the baseline assessments prior to the intervention, starting 6–45 months after the acute stroke.
Mini-Mental Status Examination scores < 24 may indicate cognitive impairment and warrant further examination (Strobel & Engedal, 2008). One patient scored below 24, but further neuropsychological assessments done by a clinical psychologist indicated that cognitive function was sufficient for participation and that the inclusion criteria were not violated. The study was approved by the Regional Committee for Medical and Health Research Ethics, south-east Norway. All participants provided their written informed consent prior to inclusion.

### 2.2 FSS

Fatigue was measured by the FSS (Krupp et al., 1989), which is a one-dimensional, 9-item self-report scale, and one of the most frequently used measures to assess fatigue after stroke and other neurological conditions (Cumming et al., 2016; Lerdal et al., 2009; Whitehead, 2009). The nine items are statements about impact of fatigue on different areas of daily life, and responses are given on a nine-point Likert scale reflecting degree of agreement (minimum mean score 1, maximum mean score 7). A review of 22 fatigue measures concluded that FSS was among the three scales that demonstrated good psychometric properties, as well as sensitivity to change in fatigue over time (Whitehead, 2009). Figure 1 shows the distribution of mean FSS scores by sex. Average FSS score was 3.53 (SD = 1.46), and 35% of the patients reported mean FSS > 4, which is a commonly adapted threshold for clinical fatigue in stroke studies (Krupp et al., 1989; Schepers et al., 2006; Tang et al., 2010). Table S1 shows the mean scores per item for patients with- and without PSF according to this cutoff value, offering a more detailed characterization of fatigue complaints in the sample. The PSF group scored significantly higher on all items.

### 2.3 PHQ-9

Depressive symptoms were measured by the self-report scale Patient Health Questionnaire (PHQ-9; Spitzer, Kroenke, Williams, & Patient Health Questionnaire Primary Care Study, 1999). PHQ-9 consists of nine items based on the DSM-IV criteria for depression. These are scored 0–3, providing severity scores ranging from 0 to 27. Briefly, sum scores of 5, 10, 15 and 20 represent mild, moderate, moderately severe and severe symptom levels. Average PHQ score in the patient sample was 4.79.

### Table 1 Sample characteristics

| Current demographic and clinical information | Mean  | SD   | Min | Max |
|---------------------------------------------|-------|------|-----|-----|
| Age                                         | 69.00 | 7.43 | 47  | 81  |
| Males/females (count)                       | 38/15 |      |     |     |
| Education in years                          | 14.56 | 3.65 | 9   | 30  |
| FSS                                         | 3.53  | 1.46 | 1.11| 6.77|
| PHQ-9                                       | 4.79  | 3.61 | 1   | 14  |
| MMSE                                        | 28.22 | 1.68 | 22  | 30  |

Stroke-related information

| NIHSS at hospital discharge                | 1.14  | 1.23 | 0   | 6   |
| Months since stroke                        | 25.59 | 9.40 | 6.00| 45.00|
| TOAST classification for ischemic stroke   |       |      |     |     |
| Small vessel occlusion                     |       |      |     |     |
| Cardioembolism                             |       |      |     |     |
| Large artery atherosclerosis (19)          |       |      |     |     |
| Other known/not known (10)                |       |      |     |     |
| Right Hemisphere (20)                     |       |      |     |     |
| Left Hemisphere (18)                      |       |      |     |     |
| Brainstem/cerebellum (7)                  |       |      |     |     |
| Both Hemispheres (5)                      |       |      |     |     |
| Not specified (3)                          |       |      |     |     |

*All but one patient suffered ischemic stroke.*

### Figure 1 Distribution of mean FSS scores by gender and distribution of age by gender. Red line denotes the mean [Colour figure can be viewed at wileyonlinelibrary.com]
2.4 Attention network test

A conventional version of the ANT was applied, as previously described (Fan et al., 2002). In the ANT, accuracy and response times (RT) are tracked over time in trials with varying cognitive demands in a computerized paradigm. By combining a flanker test (Eriksen & Eriksen, 1974), and a cued reaction time task (Posner, 1980), the ANT estimates network scores as relative differences in mean RTs between different flanker and cue conditions (Fan et al., 2002). Figure 2 depicts the details of the task.

Briefly, participants were instructed to direct their gaze at fixation cross that was presented with a duration of 400, 800, 1,200 or 1,600 milliseconds. Immediately following the fixation cross, one out of four cue conditions would appear for 100 milliseconds; no cue, a center cue (temporal cue only), a double cue (temporal cue only), or a spatial cue (temporal and spatial cue), alerting the attention toward the stimulus about to appear. Then, five small arrows or lines were presented for 1,700 milliseconds, and the task was to, as quickly and correctly as possible, decide whether the middle arrow (target arrow) was pointed left or right. Participants responded by pressing the left or the right mouse button. The four flanker arrows/lines surrounding the middle, target arrow could point in either the same direction (congruent flankers) or the opposite (incongruent flankers) direction as the middle, target arrow, or they could simply be lines without direction, constituting neutral flankers. The flanker arrows/lines represent the different stimulus conditions associated with different cognitive demands, where incongruent flankers typically result in the highest error rates and RTs (Westlye, Grydeland, Walhovd, & Fjell, 2010).

Starting with a practice run of 24 trials, the full test consisted of 288 trials, divided into three rounds (96 trials per round), lasting about 20 min. Participants were instructed to take a short break between rounds. For setting up the experiment and collecting responses, E-prime software (Psychology Software Tools, Pittsburg, PA) was applied.

2.5 Statistical analyses

Statistical analyses were performed using R version 3.4.0 (2017-04-21; R Core Team, 2017) and the python toolbox HDDM (Wiecki, Sofer, & Frank, 2013). Figures were produced using the ggplot2 package (Wickham, 2009).

2.5.1 Outlier exclusion and data cleaning

Trials with RT < 200 ms, thought to reflect fast guesses, were removed from the analysis, in line with previous ANT reports (Chang, Pesce, Chiang, Kuo, & Fong, 2015; Westlye et al., 2010). 2% of the responses were removed due to this criterion. Participants having more than 50% incorrect responses within any of the flanker conditions were discarded. One participant was removed due to this criterion.

2.5.2 Associations between FSS, time and RT

In order to characterize the relationship between subjective fatigue (FSS_z), time (trial 1–288) and RT, we applied linear mixed-effects models using the lme function from the nlme package in R (Pinheiro, Bates, DebRoy, & Sarkar, 2013). Following the recommendations from Barr, Levy, Scheepers, and Tily (2013), we started with a maximal model, including by-subject random slopes for FSS_z * time at the subject level, in addition to random
intercepts, and all fixed effects or covariates of potential interest. These were z-normalized FSS scores × time, age, sex, flanker condition, stroke topography (left or right hemisphere, brainstem/cerebellum), lesion volume (defined by number of voxels affected), TOAST classification for stroke etiology (large artery atherosclerosis, small vessel occlusion, cardioembolism or “other known or unknown factors”), NIHSS scores and z-normalized PHQ scores. Non-converging models were dealt with by sequentially simplifying the fixed effect structure until reaching convergence. The full model did not converge, and we dropped NIHSS, on the basis that the variability in NIHSS scores was small (mean = 11.4, median = 1, SD = 1.23), reflecting the fairly highly functioning patient sample. Next, we removed TOAST classification of stroke etiology, due to a large number of cases in the “not specified/unknown” category, and then excluded PHQ scores because of high correlations with FSS.

The most complex converging model was specified as follows: lme (RT ≈ FSS_z × time + age + sex + flanker + lesion volume + lesion location, random = 1 + FSS_z × timelid, data = data, method = “REML”). As a formal test of whether random slope effects were warranted, we used the ANOVA function in R to compare model fit between this model and a similar model without a random slope term, and results indicated that random slopes should be included. To further refine the model, we tested whether removing independent variables that did not provide predictive value improved model fit. Model fit improved marginally by removing lesion volume and lesion location. As an indication of FSS effect size, we compared the final model with a model that did not include FSS score. Model formulae and notes on model selection are provided in Table S2.

Assessing whether PSF status (PSF defined by mean FSS score > 4, in line with common practice (Krupp et al., 1989; Schepers et al., 2006)) interacted with the effect of time/trial number, we reran the above-described regression model with PSF status included in the model instead of FSS score as a continuous measure. Additionally, to test whether effects varied between flanker conditions, we estimated the full regression model separately for each flanker condition. In these follow-up models, random slopes were not estimated in order to secure convergence. To explore whether the relationship between time, fatigue and performance manifested differently according to PSF status, we repeated the above-described within-flanker linear mixed-effects models within patients with PSF and patients without PSF.

Importantly, to test whether potential effects were specific for fatigue or could be explained by depressive symptoms, the full linear mixed-effects model was repeated with PHQ instead of FSS, keeping all other model specifications constant.

In all analyses, the time variable refers to trial number (1–288).

### 2.5.3 Associations between conventional ANT network scores and FSS

Based on a previous definition (Westlye et al., 2010), we computed the conventional ANT network scores orienting, alerting and executive control network based on median RTs:

- Executive control = (RT incongruent – RT congruent)/RT congruent
- Alerting = (RT no cue – RT center cue)/RT center cue
- Orienting = (RT center cue – RT spatial cue)/RT spatial cue

To assess the association between estimated attentional networks and subjective fatigue, we ran a linear model for each attentional network and tested for main effects of FSS, covarying for age and sex. We then estimated change in network efficiency over time (network slope) for each network and fitted this to a linear model adding FSS, age and sex as predictors to test for interactions between attentional networks, time and FSS. Network slopes were created in two steps: First, we ran linear models for each patient within each flanker and cue condition separately, predicting RT by trial number. Then, change in network efficiency/network slope was calculated for each patient by subtracting the betas from the first models in the same way as outlined above, that is:

Executive slope = (beta incongruent – beta congruent)

Following the same procedure as in the RT models above, we reran the network analyses replacing FSS with PSF status as independent variable, to investigate whether attentional networks were differently affected by time dependent on PSF status.

As an additional test of potential associations between subjective fatigue and stroke-related variables, we estimated the correlations between FSS score, NIHSS score, lesion volume and months since stroke, respectively.

### 2.5.4 Hierarchical drift diffusion modeling

Cleaned RT and accuracy data were submitted to hierarchical drift diffusion modeling by use of the python toolbox HDDM (Wiecki et al., 2013). HDDM uses hierarchical Bayesian parameter estimation, which provides enhanced statistical power and allows for estimation of both individual and group parameters simultaneously (Wiecki et al., 2013). We applied mildly informative priors and starting points as predefined in the toolbox (Wiecki et al., 2013). We did not estimate any bias in starting point. The data were accuracy-coded (accurate responses = 1, erroneous responses = 0). In addition to the data cleaning described above, an outlier mixture model...
included in the HDDM was applied, which assumes that a fixed proportion (5%) of trials are outliers that come from a uniform distribution not generated by a diffusion process (Wiecki et al., 2013). A mixed-effects model allowing for some outliers has been shown to provide a better fit in likelihood models than models not allowing for any outliers at all (Wiecki et al., 2013).

2.5.5  Model selection/ defining parameters

When parametrizing the hDDM, we tested different cognitively plausible models to identify the model that best explained data, guided by the theoretical assumption that drift rate (v) should be allowed to vary as a function of stimulus difficulty condition (Ratcliff, Smith, & McKoon, 2015). Further, decision threshold (a) was assumed to be constant across stimulus conditions, following the logic that if a varies with stimulus conditions, the participant would have to first identify the condition, before adjusting threshold and then start accumulating information from the stimulus, a sequence of events that does not seem plausible (Thapar, Ratcliff, & McKoon, 2003). Non-decision time (t, stimulus encoding and motor responses) was not expected to be affected by flanker condition, given that the visual stimuli were highly similar across flanker conditions and motor responses were simple and identical across conditions (simple button press).

Building on the above-mentioned assumptions, we estimated different models and tested which combination of parameter fixations provided the best model fit. See Table 2 for an overview of models tested.

Variability estimates were included in the preliminary models, but were discarded as they failed to converge adequately and slightly worsened model fit. Variability parameters are often estimated poorly, and less complex models may improve estimates of the parameters of interest (Lerche & Voss, 2016). To evaluate in which parameter the interaction between time and FSS should be localized, we estimated a

| Model | Samples | DIC     |
|-------|---------|---------|
| (a ~ warningcue) | A1 | 1,500 | −15,424 |
| (t ~ warningcue) | A2 | 1,500 | −15,717 |
| (v ~ warningcue) | A3 | 1,500 | −15,169 |
| (v ~ flanker + time * FSS, t ~ time * FSS + warningtype, a ~ FSS + time), group_only = True) | B1 | 12,000 | −16,608 |
| (a ~ time) | C1 | 1,500 | −14,356 |
| (t ~ time) | C2 | 1,500 | −14,448 |
| (v ~ time) | C3 | 1,500 | −14,219 |
| (a ~ FSS) | D1 | 1,500 | −14,109 |
| (t ~ FSS) | D2 | 1,500 | −14,110 |
| (v ~ FSS) | D3 | 1,500 | −14,109 |
| (a ~ FSS:time) | E1 | 1,500 | −14,353 |
| (t ~ FSS:time) | E2 | 1,500 | −14,454 |
| (v ~ FSS:time) | E3 | 1,500 | −14,222 |
| (v ~ flanker, t ~ warningtype, a ~ FSS + time), group_only = True) | F1 | 6,000 | −16,590 |
| (v ~ flanker, t ~ time * FSS + warningcue), group_only = True) | F2 | 6,000 | −16,602 |
| (v ~ flanker + time * FSS, t ~ warningtype), group_only = True) | F3 | 6,000 | −16,575 |
| (v ~ flanker, t ~ time * FSS + warningcue), group_only = False) | G1 | 6,000 | −17,648 |
| (v ~ flanker, t ~ time * FSS + warningcue), group_only = False) | G2 | 12,000 | −17,648 |
| (v ~ flanker, t ~ time * FSS + warningcue), group_only = True) | H1 | 12,000 | −16,602 |

Note: DIC, deviance information criterion, where lower values indicate a better model fit. DIC values in bold indicate the combination of parameter fixation that provided the best fit for each model comparison (model comparisons between same letter models (i.e., A1–A3)).
regression model where all three parameters \((a, t \text{ and } v)\) were allowed to vary by the interaction term. To further explore which parameter fixations provided the best model fit, we ran nine simple models with (a) the main effect of time on either \(a, t \text{ or } v\); (b) the main effect of FSS on either \(a, t \text{ or } v\); and (c) the FSS*time interaction on \(a, t \text{ or } v\) separately. Finally, we estimated the best model with individual regressors and group only regressors.

Model fit was assessed by comparing the deviance information criteria (relative DIC values) between models. In Bayesian analyses, the DIC provides an estimation of fit of the model to the data, where lower DIC values indicate that the model has better support (François & Laval, 2011). In models where individual regressors were estimated, we simulated data from the respective models and performed posterior predictive checks (PPC) to evaluate whether the model was able to reproduce central patterns in the observed data (Wiecki, 2016). 500 data sets were simulated by drawing 500 samples for each parameter from the estimated posterior distribution. The simulations thus capture the uncertainty in the estimated model and allow for comparisons with the observed data.

Final choice of model was based on a combination of model fit and convergence (see below).

2.5.6 Estimating the posterior distributions and assessing convergence (model diagnostics)

We used a Bayesian framework and Markov chain Monte Carlo sampling (MCMC) to estimate the posterior distributions (Kruschke, 2014). In the preliminary models, when testing and comparing parameter fixations, models were estimated on 1,500 or 6,000 samples. The final model was run on 12,000 samples. To improve convergence, the 4,000 first samples were discarded, and thinning was set to 2 (keeping only every second sample).

A valid model should demonstrate convergence of the MCMC chains (Wiecki, 2016). Convergence was assessed by plotting and visually inspecting traces and autocorrelation plots for each estimated parameter. As a more formal test of convergence, the Gelman–Rubin statistics \((R^\text{\textsuperscript{a}})\) (Gelman & Rubin, 1992) were calculated. These values should be close to 1 and not exceed 1.1 if the chains have converged successfully, that is, if the samples of the different chains are similar (Wiecki et al., 2013).

2.5.7 Hypothesis testing within the hDDM

Effects of task and cue conditions, as well as the effects of time and fatigue status, were determined by Bayesian hypothesis testing, by assessing the degree of overlap between posterior distributions. If less than 5 percent of the posterior distributions of two parameters overlap, the difference is said to be credible, or an effect is credibly different than null when at least 95 percent of the posterior distribution does not contain zero.

3 RESULTS

3.1 ANT behavioral results

Table 3 shows mean RT and error rates for each flanker condition. Two-tailed, one-sample \(t\) tests revealed significant differences in RT between incongruent and congruent condition, \(M = 111, CI = 101–122, t(52) = 20, p < .001\), between incongruent and neutral condition, \(M = 124, CI = 112–136, t(52) = 20, p < .001\) and between congruent and neutral condition, \(M = 12, CI = 3.7, t(52) = 3.6, p < .001\).

There was no significant association between FSS and mean RT across \((r = .09, p = .48)\) or within conditions (incongruent flanker: \(r = .05, p = .67\), congruent flanker: \(r = .11, p = .47\), neutral flanker: \(r = .12, p = .37\)). There was no association between FSS and error rate \((r = −.09, p = .48)\).

3.2 Associations between FSS, time and RT

Table 4 shows the summary statistics from a linear mixed-effects model testing for associations between RT and FSS, time, sex, age and flanker condition for all conditions simultaneously. The model including FSS score performed significantly better than the model not including FSS score as indicated by ANOVA model comparison, supporting the predictive value of FSS \((L.ratio(1) = 19.09, p < .001\), see also Table S2\).

The model presented in Table 4 was also run with lesion volume and lesion location as independent variables to control for effects related to lesion characteristics. As both volume and location displayed low predictive value and did not improve model fit, they were not included in the final analyses. Results from the linear mixed model including lesion volume and lesion location are presented in Table S3.

| Table 3 | Error rates and mean RTs by flanker condition |
|---------|-----------------------------------------------|
| Flanker | Mean RT in ms |                          |                        |
|         | Total (SD)   | Error (SD) | Accurate (SD) | Accuracy (%) |
| Congruent | 668 (187)    | 841 (329) | 667 (185)    | 99.2          |
| Incongruent | 773 (199)    | 645 (300) | 776 (195)    | 97.7          |
| Neutral  | 655 (180)    | 713 (245) | 654 (179)    | 99.1          |
Table 5 presents summary statistics for models estimated for each flanker condition separately, estimating the effect of FSS score, time (trial number) and the interaction effect between time and FSS. Figure 3 shows the estimated RT (output from Table 5) plotted by group (PSF vs. non-PSF patients, based on mean FSS score ≥ 4). Briefly, after Bonferroni correction for multiple comparisons (corrected alpha 0.5/8 = .006), the interaction between time and FSS was significant in the neutral and incongruent condition, as was the association between age and RT, indicating that age was associated with increased RT across conditions.

There was no significant main effect of FSS on RT in any condition.

Corresponding linear mixed model with group (PSF status) instead of FSS score revealed similar associations with the various independent factors as presented in Table 4, except for identifying a negative main effect of time ($\beta = -0.05$, $SE = .01$, $t = -3.31$, $p < .001$). The interaction effect between PSF status and time was comparable to that of FSS score and time, albeit smaller ($\beta = 0.06$, $SE = .01$, $t = 2.27$, $p = .022$), and only nominally significant. All results from the model with PSF status as predictor are presented in Table S4.

### Table 5  Linear mixed-effects models by flanker stimulus, one model per condition

|                | Neutral | Incongruent | Congruent |
|----------------|---------|-------------|-----------|
| Beta           | $t$     | $p$         | $t$       | $p$       | $t$   | $p$    |
| FSS            | 8.31    | 0.50        | .614      |           |       |       |
| Time           | -0.03   | -1.68       | .091      | -0.05     | -2.27 | .023   | 13.67  | 0.78  | .472   |
| Time:FSS       | 0.09    | 4.22        | <.001*    | 0.10      | 4.16  | <.001* | 0.04   | 2.08  | .037   |
| Sex            | -4.78   | -0.13       | .896      | 0.14      | 0.00  | .997   | 2.06   | 0.05  | .958   |
| Age            | 8.22    | 3.71        | <.001*    | 8.42      | 3.26  | .002*  | 7.95   | 2.08  | .037   |

*p-Values that remained significant after Bonferroni correcting for multiple comparisons.

### Table 4  Linear mixed-effects models for whole sample, all flanker conditions

|               | $t$     | Beta CI               | $p$       |
|---------------|---------|-----------------------|-----------|
| Intercept     | 0.63    | 109.49 (−229.4, 448.5) | .526      |
| FSS_z         | 0.37    | 7.02 (−29.9, 43.6)    | .298      |
| Time          | -1.03   | -0.03 (−0.09, 0.03)   | .870      |
| Time:FSS      | 2.65    | 0.07 (0.02, 0.14)     | .008      |
| Sex           | 0.04    | 1.92 (−77.34, 81.2)   | .962      |
| Age           | 3.34    | 8.15 (3.37, 12.93)    | .001*     |
| Incongruent flanker | 40.5  | 111.52 (106.14, 116.93) | <.000*   |
| Neutral flanker | -4.31 | -11.78 (−17.15, −6.43) | <.000*   |

*p-Values that remained significant after Bonferroni correcting for multiple comparisons.

### Figure 3  Estimated RT from linear mixed-effects models plotted by PSF status [Colour figure can be viewed at wileyonlinelibrary.com]
Mixed-effects models with PHQ score included instead of FSS did not indicate significant interaction effects between depressive symptoms and time on RT in any flanker condition.

Table 6 shows summary statistics from linear mixed models estimating the main effect of sustained performance (time) on RT in the various flanker conditions, conducted separately for patients with and without PSF. In this model, that included only trial number (time) and not FSS score as predictor, the results suggested that patients without PSF demonstrated more speeded RTs in the incongruent condition during the course of the experiment, while patients with PSF did not show any significant changes in RT in any condition.

3.3 | Associations between FSS and other clinical measures

There was no correlation between FSS score and months since stroke \( (r = .00, p = .97) \), between FSS score and lesion volume, indicated by number of voxels affected \( (r = -.14, p = .30) \) or FSS score and stroke severity, indicated by NIHSS score \( (r = .10, p = .46) \). FSS score was positively correlated with PHQ score \( (r = .47, p < .001) \).

3.4 | Associations between ANT network scores and FSS

One-sample \( t \) tests revealed significant group-level network score effects for executive control network \( (M = 0.18, CI = 0.17–0.20, t = 21.57, p < .001) \), orienting network \( (M = 0.06, CI = 0.05–0.08, t = 11.27, p < .001) \) and alerting network \( (M = 0.04, CI = 0.03–0.06, t = 7.47, p < .001) \). Table 7 shows summary statistics from linear models estimating the associations between ANT network scores and FSS. Whereas the analyses revealed a nominally significant negative association between FSS and the executive network score \( (t = −2.23, p = .03) \) and a negative effect of age on the alerting network \( (t = −2.17, p = .03) \), no associations remained significant after correction for multiple comparisons.

Table 8 shows linear models testing associations between ANT network efficiency change over time (network slope) and FSS score, age and sex. Results suggested a (nominally) significant association between executive slope (network efficiency change over time, where positive score indicate efficiency) and FSS \( (t = 2.24, p = .029) \). No associations remained significant after correction for multiple comparisons. See Figure 4 for network slopes plotted against FSS scores. Follow-up linear models with PSF status as predictor instead of FSS did not support a significant main effect of PSF status \( (t = −0.73, p = .466) \) on executive network slope.

3.5 | hDDM regression models

The best fitting model that showed adequate convergence allowed drift rate \( (v) \) to vary across flanker conditions, non-decision time \( (t) \) to vary across warning cue conditions and time while boundary separation \( (a) \) was kept constant (“\( v \sim \) flanker,” “\( t \sim \) warningcue +time”). In this group-level model, no Gelman–Rubin statistics (R-hat values) were > 1.1, and chains and autocorrelations confirmed adequate convergence for all parameters.

A less restricted model where all parameters were allowed to vary by the FSS*time interaction term generated a slightly better model fit (DIC value −16,608 vs. −16,602), but worse convergence in terms of (R-hat values > 1.1), chains and autocorrelations. This model was therefore discarded as not sufficiently valid. Estimations of the best fitting model (“\( v \sim \) flanker,” “\( t \sim \) warningcue +time”) on the individual level produced the best fit in terms of DIC values, but posterior predictive checks indicated that the models did not sufficiently reproduce observed patterns in the data and the standard deviations for \( t:FSS \) and \( t:time:FSS \) showed suboptimal convergence.

| TABLE 6 | Linear mixed-effects models estimating RT by time for PSF/non-PSF patients separately |

|                | Neutral |          | Incongruent |          | Congruent |          |
|----------------|---------|----------|-------------|----------|-----------|----------|
|                | Beta    | \( t \)  | \( p \)     | Beta     | \( t \)  | \( p \)  | Beta     | \( t \)  | \( p \)  |
| Non-PSF patients |         |          |             |          |           |          |         |          |           |          |
| Time           | −0.06   | −2.17    | .029        | −0.10    | −3.22    | .001*    | −0.01    | −0.55    | .575      |
| Sex            | 34.1    | 0.66     | .509        | 10.9     | 0.18     | .854     | 26.5     | 0.48     | .632      |
| Age            | 10.1    | 2.90     | .007        | 8.96     | 2.21     | .035     | 8.90     | 2.39     | .024      |
| PSF patients   |         |          |             |          |           |          |         |          |           |          |
| Time           | −0.00   | −0.12    | .377        | 0.00     | 0.18     | .851     | 0.00     | 0.17     | .859      |
| Sex            | −47.4   | −0.76    | .454        | −9.77    | −0.13    | .897     | −25.7    | −0.38    | .707      |
| Age            | 7.86    | 2.26     | .036        | 8.36     | 1.99     | .062     | 8.17     | 2.15     | .045      |

*\( p \)-Values that remained significant after Bonferroni correcting for multiple comparisons.
3.6 Effect of time and FSS on $t$, non-decision time

Figure 5 shows the posterior distributions for non-decision time, $t$. hDDM provided support for a negative main effect of time ($P(t_{\text{time}} < 0) = 0.98$) on non-decision time, indicating that time needed for stimulus encoding and response execution decreased during the course of the test. hDDM did not identify a main effect of FSS on non-decision time ($P(t_{\text{FSS}} > 0) = 0.77$). In contrast, the model provided evidence for a positive interaction effect between time on task and FSS on non-decision time ($P(t_{\text{time:FSS}} > 0) = 1.00$), suggesting that the association between FSS and non-decision time increased during the course of the experiment, so that patients with high levels of fatigue were more negatively affected by time on task (resulting in higher non-decision times), than patients low on fatigue. The interaction effect is small, but robust (posterior distribution not overlapping the null, model displaying good convergence), and it is in the opposite direction of the main effect of time when FSS is not accounted for.

3.7 Effect of warning cue on $t$, non-decision time

Figure 6 (left) shows the posterior probability plot for non-decision time ($t$) as a function of warning cue (intercept: center cue). Non-decision time was lowest for cue conditions “up” and “down”. “No cue” resulted in the highest non-decision time out of all cue conditions. Thus, model evidence suggests that the presence of cues facilitated the process of stimulus encoding and response execution, and most efficiently so when the cues provided both temporal and spatial information (“up” and “down”).

3.8 Effect of flanker conditions on drift rate

Figure 6(right) shows the posterior probability plot for the drift rate ($v$) estimated by flanker condition (intercept: congruent condition). The model provided strong evidence supporting that drift rate was lower in the incongruent condition compared to both congruent and neutral condition ($P(v_{\text{Incongruent}} < v_{\text{Congruent}}) = 1.0$, and $P(v_{\text{Incongruent}} < v_{\text{Neutral}}) = 1.0$), suggesting lower rates of evidence accumulation in the cognitively most demanding condition (incongruent flanker with cognitive conflict). Drift rate was highest in the neutral condition ($P(v_{\text{Neutral}} > v_{\text{Congruent}}) = 1.0$, $P(v_{\text{Neutral}} > v_{\text{Incongruent}}) = 1.0$).

4 DISCUSSION

Post-stroke fatigue is a common and debilitating symptom in stroke patients, yet its mechanisms are poorly understood. Many patients suffering from PSF report increased fatigue

| Table 7 Linear regression models by ANT network |
|-----------------------------------------------|
|                      | Orienting | Alerting | Executive |
|----------|---------|---------|-----------|
| $t$      | $p$     | $t$     | $p$       | $t$     | $p$     |
| FSS      | 0.57    | 0.567   | −1.03     | 0.305   | −2.23   | 0.030*   |
| Age      | 1.10    | 0.276   | −2.17     | 0.034   | −2.03   | 0.243    |
| Sex      | 1.66    | 0.103   | −0.41     | 0.681   | −0.03   | 0.975    |

*Nominal significance of $p$-values.

| Table 8 Linear regression models by ANT network slope |
|-----------------------------------------------|
|                      | Executive | Alerting | Orienting |
|----------|---------|---------|-----------|
| $t$      | $p$     | $t$     | $p$       | $t$     | $p$     |
| FSS      | 2.24    | 0.029*  | 0.18      | 0.85    | 0.55    | 0.580    |
| Age      | 0.832   | 0.409   | −0.54     | 0.587   | 1.44    | 0.156    |
| Sex      | −0.98   | 0.332   | 0.697     | 0.489   | 0.27    | 0.787    |

Note: Higher network values indicate lower relative network efficiency.
*Nominal significance of $p$-values.

FIGURE 4 Estimated attention network slopes plotted against $z$-normalized FSS scores [Colour figure can be viewed at wileyonlinelibrary.com]
and cognitive difficulties when engaging in cognitive tasks, but previous studies have largely failed to establish robust associations between subjective fatigue and cognitive performance. The scarcity of evidence may be due to the use of instruments lacking cognitive sensitivity and specificity, and tests that do not account for the effect of time on task.

In the current study, we aimed to characterize the relationship between subjective fatigue and attentional function, taking duration of effort into account. To this end, we collected behavioral data using ANT and self-reported fatigue using FSS from 53 chronic stroke patients. First, we tested the assumption that FSS scores would interact negatively with time on task, manifesting in a performance decline for patients with high fatigue relative to patients with low fatigue. Results from linear mixed models provided support for this hypothesis, identifying significant interactions between FSS score and time on RT in the neutral and incongruent flanker conditions. In these whole sample models, no significant main effects of time or FSS were identified. Interestingly, when examining the main effect of time separately for patients with and without PSF, results revealed that non-PSF patients significantly improved
RTs over time in the most cognitively demanding condition, while the PSF group did not demonstrate significant improvement.

These findings underscore the relevance of taking time on task into account and measure sustained performance when addressing fatigue. Because the study design does not allow causal inference, the observed interaction between subjective fatigue and RT may be either a manifestation of fatigue, a cause of fatigue (i.e., that attentional difficulties give rise to fatigue), or both. Providing a speculative theoretical context, the coping hypothesis (Van Zomeren, Brouwer, & Deelman, 1984; Van Zomeren & Van den Burg, 1985) offers one explanatory framework for the observed interaction. Originally articulated in relation to traumatic brain injury patients, this view suggests that the chronic effort needed to compensate for subtle, cognitive deficiencies gives rise to secondary symptoms, hereunder fatigue. Hence, subtle cognitive deficits associated with stroke may be temporarily disguised by a compensating and temporary increase in cognitive effort. However, this compensation comes with the cost of increased feeling of fatigue, in particular during sustained effort. In line with this, the current interaction between time on task and fatigue may be understood as a result of increased cognitive effort, producing increased tiredness over time, resulting in suboptimal performance. The concept of “cognitive compensation” also ties well with evidence from the split sample analysis indicating that non-PSF patients’ performance benefitted from practice (sustained performance) in the most cognitively demanding condition, while the PSF group did not improve with practice. This may reflect a weakening of learning effects due to cognitive compensation costs as described above, or, alternatively, a failure to benefit from practice due to increasing fatigue.

The interaction between FSS and time can also be mediated by motivation, with high levels of fatigue leading to reduced motivation and suboptimal performance. Accordingly, the role of motivation is implied by the high scores on the FSS item reflecting reduced motivation when feeling fatigued. Regardless of the specific theoretical account, the results can be understood as lending support to Holtzer’s definition of cognitive fatigue as “an executive failure to monitor and optimize performance over acute but sustained cognitive effort resulting in performance that is lower and more variable than the individual’s optimal ability” (Holtzer et al., 2010, p. 123).

It should be noted that the interaction between time on task, self-reported fatigue and RT did not change when depressive symptoms were added to the model. Moreover, when testing the model with PHQ score on the interaction term instead of FSS, we did not find any interaction effects between depressive symptoms and time. These results suggest that although fatigue and depression are overlapping and correlated clinical phenomena, the specific characteristics of fatigue may be more strongly associated with sustained attentional performance during the course of a demanding cognitive task.

Results did not reveal any significant association between stroke location/laterality or lesion volume and outcome variables (RT or FSS), suggesting that, in this sample, lesion location and volume are not strong predictors of subjective fatigue or attentional function as measured by ANT. Whereas the lack of a robust relationship between lesion location/lesion volume and FSS score is in line with previous reports (Choi-Kwon, Han, Kwon, & Kim, 2005; Mead et al., 2011), the literature is not conclusive, and right hemispheric lesions are frequently associated with attentional dysfunction and neglect (Robertson, Ridgeway, Greenfield, & Parr, 1997; Spaccavento et al., 2019; Vallar & Perani, 1986). The current lack of predictive value of stroke location highlights the complex etiology of attentional function in chronic stroke patients. However, we cannot rule out that different operationalizations of attentional dysfunction or alternative categorizations of lesion location could reveal stronger associations.

Our hypothesis that FSS scores would be associated with overall reduced executive network efficiency was not supported, and no associations between fatigue and attentional networks remained after correcting for multiple comparisons. This finding does not support previous studies on fatigue in neurological conditions, linking fatigue to reduced efficiency of the ANT executive network (Holtzer et al., 2010; Togo et al., 2015). There was, however, a nominally significant association between change in executive network efficiency over time (network slope) and fatigue, indicating that patients with higher levels of fatigue exhibited a larger decline in executive network efficiency with sustained effort than patients reporting lower levels of fatigue. Although these findings did not remain after corrections for multiple comparisons, they may suggest that subjective fatigue is less associated with reduced executive attention per se, and more with an increased susceptibility to distractors when the attentive system is put under sustained pressure.

Suggestive FSS network effects were only observed in the executive network. It is unclear whether the lack of alerting and orienting network effects reflects that subjective fatigue is related to executive attention exclusively, or rather reflects psychometric properties of the ANT networks. A psychometric evaluation of the ANT networks based on 15 previous studies (MacLeod et al., 2010) reported that the power to identify significant effects varied across networks, while network reliability was consistently highest for executive network effects, and low to medium for alerting and orienting network effects.

Because traditional analyses based on observed data alone do not allow for any inference regarding the specific cognitive processes that may underpin differences in RT, we performed an exploratory analysis where we fitted a hierarchical drift diffusion model (hDDM) to the ANT behavioral data.
This computational dissection of the ANT data indicated that the interaction between fatigue and time on RT was best explained by non-decision time, and not the speed of evidence accumulation (drift rate) or response style (boundary separation). hDDM revealed no main effect of FSS on any of the model parameters, but provided evidence of an interaction between time and FSS on non-decision time, indicating increasing effects of FSS during the course of the experiment. In this respect, the results concurred with the linear mixed-effects models on RT data, suggesting stronger associations between fatigue and hDDM parameters with more sustained performance, and indicate that hDDM is sensitive to fatigue in a cognitive context when explicitly modeling the interactions with time.

Non-decision time (t) comprises both sensory encoding and motor response output (Ratcliff & Smith, 2010). The fact that model evidence was stronger for the models where the interaction between FSS and time was estimated on non-decision time, rather than on drift rate or boundary separation, indicates that fatigue may be specifically associated with non-decision aspects of the response process, such as stimulus encoding or response execution rather than with the speed or efficiency of the evidence accumulation or with the decision threshold (i.e., how much information is required before making a decision). Previous studies have reported higher non-decision times in older compared to younger individuals (Ratcliff et al., 2001), and in this respect, patients reporting high fatigue are responding more like elderly individuals, but only after sustained exertion.

It is also interesting to note that the negative main effect of time (in non-decision time) suggested by the current model is in line with previous drift diffusion research on practice effects, identifying a reduction in the non-decision component across trials (Dutilh, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2009). In this context, the current positive interaction between time and fatigue on non-decision time may be understood as fatigue counteracting the otherwise beneficial effects of practice.

This ties well with results from the linear mixed-effects models, suggesting that patients with PSF did not improve performance over time, in contrast to patients without PSF who got faster with during the course of the session. However, this effect was found in the incongruent condition, defined by flankers, while in the reported hDDM results, non-decision time primarily accounts for variance introduced by cues. One explanation could be that responses in the incongruent condition require an inhibition of the dominant motor response after the decision is made and that the identified interaction between fatigue and time on non-decision time is driven by a stronger slowing of these inhibitory responses in patients with higher levels of fatigue.

Research aiming to delineate the nervous system pathophysiology of PSF may further inform hypotheses about this apparent link between non-decision time and fatigue. Applying transcranial magnetic stimulation (TMS), a previous study (Kuppuswamy, Clark, Turner, Rothwell, & Ward, 2014) reported higher motor thresholds in stroke patients with high fatigue and suggested that patients with PSF experience diminished excitability of motor pathways, regarding both corticospinal outputs and facilitatory inputs. In this respect, the current observation of an interaction between time and fatigue on non-decision time might reflect altered neuronal excitability. However, how such neurophysiological mechanisms would translate into the subjective perception of fatigue remains unclear. Here, the perception of effort might be central, in the sense that subjective fatigue may manifest when volitional motor cortex input does not longer produce the expected output due to reduced excitability (Kuppuswamy et al., 2014).

These explorative results based on computational modeling provide a novel account of the specific cognitive underpinnings of PSF. When the task context is appropriate, DDM parameters can be interpreted directly (Froehlich et al., 2016) and thus provide insight into the modular and temporal evolution of the decision process. Decision boundary separation (a) adjusts the trade-off between speed and accuracy (Pedersen, Frank, & Biele, 2017). Large estimates of (a) are typically interpreted as indicative of a conservative decision style, associated with higher RTs but more accurate responses (Pedersen et al., 2017). Larger estimates of drift rate (v) are typically interpreted as more efficient information processing and are expected to vary by “the quality of the information extracted from the stimulus” (Ratcliff & McKoon, 2008, p. 3), implying that experimental conditions varying in difficulty should produce different drift rates (Ratcliff & McKoon, 2008). In line with this, and in agreement with previous studies estimating the effect of stimulus difficulty on drift rate (Voss et al., 2004), hDDM identified a credible effect of flanker type on drift rate, with the more cognitively demanding incongruent condition resulting in the lowest drift rate, while neutral flankers yielded the highest drift rate.

To sum, results from linear mixed-effects models suggested that subjective fatigue interacts with time on task, possibly counteracting practice effects, in particular in the most cognitively demanding incongruent flanker condition. Group analyses revealed that patients without PSF improved performance over time in the incongruent condition, while the PSF group did not. Additionally, higher FSS scores were associated with declining efficiency in the executive network over time. However, the effect was small and was not associated with PSF status. Lastly, hDDM modeling identified an interaction between fatigue scores and time on non-decision time.

Some limitations should be considered when interpreting the results of the current study. In line with most clinical studies, the study design does not allow for causal inference. Still,
the findings may pave the way for future clinical or experimental studies examining possible causal mechanisms and subsequent interventions. Moreover, as subjective fatigue can manifest as both a normal and a pathological phenomenon, and no universally accepted definition or criteria of PSF exists, we adopted an explorative approach, aiming to characterize the relationship between subjective fatigue and sustained attentional performance by a continuous measure symptom scale. While our main objective was not to identify case–control differences between stroke patients and healthy controls, but rather to characterize the cognitive correlates of post-stroke fatigue using computational modeling of response patterns, future studies adding a healthy control group may provide stronger interpretations regarding the clinical sensitivity of the computational behavioral parameters.

The distribution of NIHSS scores indicates that the current patients were sampled from a relatively healthy part of the full population of stroke patients. It is possible that a higher fatigue symptom burden on the group level could reveal associations that were not expressed in this relatively well-functioning patient sample. Reported fatigue levels are, however, comparable with what has been reported in other studies (Wang, Wang, Wang, & Chen, 2014) and higher than what is reported in healthy control samples (Valko, Bassetti, Bloch, Held, & Baumann, 2008). Further studies are needed to test the generalizability of the findings to different and more severely affected patient samples.

Although the classical version of ANT (Fan et al., 2002) appears to be a suitable paradigm to target cognitive aspects of PSF, as performance requires sustained attentional and executive resources (Holtzer et al., 2010), other versions of the test, like the ANT-I Vigilance task (Roca, Castro, López-Ramón, & Lupiáñez, 2011), could offer a more comprehensive account of relevant, associated processes like vigilance. It should also be noted that the error rate in the sample was low. This might have implications for the validity of the results from hDDM model, because the model estimates parameters based on distributions of both RT and accuracy and assumes different RT distributions for correct versus erroneous responses. Moreover, ANT is not frequently applied in hDDM modeling and may not be ideal for such due to existence of flankers and cues. However, our model displayed adequate convergence, and a recent hDDM study reported encouraging results for ANT data with no error responses (O’Callaghan et al., 2017).

In conclusion, the current study represents a novel approach to assess the cognitive phenotype of fatigue in stroke patients. The results indicate a relationship between the subjective experience of fatigue and response time distributions from a sustained attention task and demonstrate the significance of considering the sustained nature of the task when targeting fatigue in a neuropsychological context, intuitively indicating that the cognitive phenotype of fatigue entails an increased vulnerability to sustained effort. It is encouraging that the evidence suggests a link between self-reported fatigue and performance in a computerized, standardized paradigm, as it may contribute to bridging the gap between subjective experience and behavioral performance in this complex and prevalent stroke sequela. The explorative application of an advanced computational model on the temporal evolution of response times enabled the possibility to parse the observed response time patterns into specific cognitive processes. In general, the use of computational approaches in the neuropsychological workup may offer a dissection of the specific cognitive processes underlying observed behavioral differences, with clinical relevance.

ACKNOWLEDGEMENTS
This study was supported by the Norwegian ExtraFoundation for Health and Rehabilitation (2015/FO5146), the Research Council of Norway (249795, 262372), the South-Eastern Norway Regional Health Authority (2014097, 2015044, 2015073, 2018037), the European Research Council under the European Union’s Horizon 2020 research and Innovation program (ERC StG Grant 802998), and the Department of Psychology, University of Oslo.

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS
KMU conceived the study, collected the data, analyzed the data and wrote the manuscript. DA was involved in the design of the work, data analysis and interpretation, and critical revision of the paper. KK, GR and AMS were involved in data collection and quality control and contributed to drafting the paper. ESD was involved in data collection and contributed to drafting the paper. HIH was involved in interpreting of results and critical revision of the paper. MLP was involved in data analysis, interpreting of results and critical revision of the paper. ST was involved in interpreting of results and critical revision of the manuscript. JEN was involved in data collection and critical revision of manuscript. LTW was involved in conception of the study, interpretation of results and critical revision of manuscript.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.14861.

DATA AVAILABILITY STATEMENT
Data can be supplied from the corresponding authors on reasonable request.
REFERENCES

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. W. J. E. Birren (Ed.), *Behavior, aging and the nervous system* (pp. 114–149). Springfield, IL: Thomas.

Chang, Y.-K., Pesce, C., Chiang, Y.-T., Kuo, C.-Y., & Fong, D.-Y. (2015). Antecedent acute cycling exercise affects attention control: An ERP study using attention network test. *Frontiers in Human Neuroscience, 9*, 156. https://doi.org/10.3389/fnhum.2015.00156

Choi-Kwon, S., Han, S. W., Kwon, S. U., & Kim, J. S. (2005). Poststroke diseases. *Cerebrovascular Diseases*, 156. https://doi.org/10.3389/fnhum.2015.00156

Duncan, F., Wu, S., & Mead, G. E. (2012). Frequency and natural history of fatigue after stroke: A systematic review of longitudinal studies. *Journal of Psychosomatic Research, 73*(1), 18–27. https://doi.org/10.1016/j.jpsychores.2012.04.001

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review, 16*(6), 1026–1036. https://doi.org/10.3758/BF03203267

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

François, O., & Laval, G. (2011). Deviance information criteria for model selection in approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology, 10*(1), 1–25. https://doi.org/10.2202/1544-6115.1678

Froehlich, E., Liebig, J., Ziegler, J. C., Braun, M., Lindenberger, U., Heekeren, H. R., & Jacobs, A. M. (2016). Drifting through basic subprocesses of reading: A hierarchical diffusion model analysis of age effects on visual word recognition. *Frontiers in Psychology, 7*, 1863. https://doi.org/10.3389/fpsyg.2016.01863

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Glader, E.-L., Stegmayr, B., & Asplund, K. (2002). Poststroke fatigue. *Stroke, 33*(5), 1327–1333. https://doi.org/10.1161/01.STR.000014248.28711.D6

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

Johansson, B., & Rönnbäck, L. (2012). Cognitive impairment after an almost neurological recovered stroke. *ISRN Psychiatry, 2012*, 1–7. https://doi.org/10.5402/2012/686425

Johansson, B., & Ronnback, L. (2014). Evaluation of the mental fatigue scale and its relation to cognitive and emotional functioning after traumatic brain injury or stroke. *International Journal of Physical Medicine & Rehabilitation, 2011*, 7.

Kolskaa, K. K., Richard, G., Alnaes, D., Dorum, E. S., Sanders, A.-M., Ulrichsen, K. M., … Westlye, L. T. (2019). A randomized, double blind study of computerized cognitive training in combination with transcranial direct current stimulation (tDCS) in stroke patients – Utility of fMRI as marker for training outcome. bioRxiv, 603985.

Koopman, K., Uyttenboogaart, M., Vroomen, P. C., van der Meer, J., De Keyser, J., & Luijckx, G.-J. (2009). Long-term sequelae after cerebral venous thrombosis in functionally independent patients. *Journal of Stroke and Cerebrovascular Diseases, 18*(3), 198–202.

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology, 46*(10), 1121–1123. https://doi.org/10.1001/archner.1989.00520 460115022

Kruschke, J. (2014). Doing Bayesian data analysis: A tutorial with R, *JAGS, and Stan*. Burlington, MA: Academic Press.

Kuppusswamy, A., Clark, E. V., Turner, I. F., Rothwell, J. C., & Ward, N. S. (2014). Post-stroke fatigue: A deficit in corticromotor excitability? *Brain, 138*(1), 136–148. https://doi.org/10.1093/brain/awu306

Kulubaev, M. A., Shenkin, S. D., Farrall, A. J., Duncan, F. H., Lewis, S. J., Greig, C. A., … Mead, G. E. (2013). CT and clinical predictors of fatigue at one month after stroke. *Cerebrovascular Diseases Extra, 3*(1), 26–34. https://doi.org/10.1159/000347113

Lagogianni, C., Thomas, S., & Lincoln, N. (2018). Examining the relationship between fatigue and cognition after stroke: A systematic review. *Neuropsychological Rehabilitation, 28*(1), 57–116. https://doi.org/10.1080/09602011.2015.1127820

Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology, 7*, 1324. https://doi.org/10.3389/fpsyg.2016.01324

Lerche, V., & Voss, A. (2017). Experimental validation of the diffusion model based on a slow response time paradigm. *Psychological Research, 83*, 1–16.

Lerdal, A., Bakken, L. N., Kouwenhoven, S. E., Pedersen, G., Kirkevold, M., Finset, A., & Kim, H. S. (2009). Poststroke fatigue—A review. *Journal of Pain and Symptom Management, 38*(6), 928–949.

MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology, 24*(5), 637. https://doi.org/10.1037/a0019803
Mead, G. E., Graham, C., Dorman, P., Bruins, S. K., Lewis, S. C., Dennis, M. S. … IST, U. K. C. o. (2011). Fatigue after stroke: baseline predictors and influence on survival. Analysis of data from UK patients recruited in the International Stroke Trial. *PLoS One*, 6(3), e16988. https://doi.org/10.1371/journal.pone.0016988

Michael, K. (2002). Fatigue and stroke. *Rehabilitation Nursing*, 27, 89–94. https://doi.org/10.1002/jf.2048-7940.2002.tb01995.x

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review*, 20(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Naess, H., Lunde, L., Brogger, J., & Waje-Andreassen, U. (2012). Fatigue among stroke patients on long-term follow-up. The Bergen Stroke Study. *Journal of the Neurological Sciences*, 312(1–2), 138–141.

O’Callaghan, C., Hall, J. M., Tommassini, A., Muller, A. J., Walpola, I. C., Naess, H., Lunde, L., Brogger, J., & Waje-Andreassen, U. (2012). Fatigue among stroke patients recruited in the International Stroke Trial. *Journal of the Neurological Sciences*, 312(1–2), 138–141.

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, 24(4), 1234–1251. https://doi.org/10.3758/s13423-016-1199-y

Pilhaja, R., Uimonen, J., Mustanoja, S., Tatlisumak, T., & Pouitainen, E. (2014). Post-stroke fatigue is associated with impaired processing speed and memory functions in first-ever stroke patients. *Journal of Psychosomatic Research*, 77(5), 380–384. https://doi.org/10.1016/j.jpsychores.2014.08.011

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & the R Development Core Team. (2013). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-131.

Pollock, A., St George, B., Fenton, M., & Firkins, L. (2012). Top ten research priorities relating to life after stroke. *The Lancet Neurology*, 11(3), 209. https://doi.org/10.1016/S1474-4422(12)70029-7

Ponchel, A., Bombois, S., Bordet, R., & Hénon, H. (2015). Factors associated with poststroke fatigue: A systematic review. *Stroke Research and Treatment*, 2015, 1–11. https://doi.org/10.1155/2015/347920

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. https://doi.org/10.1080/00335558008248231

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Smith, P. L. (2010). Perceptual discrimination in static and dynamic noise: The temporal relation between perceptual encoding and decision making. *Journal of Experimental Psychology: General*, 139(1), 70. https://doi.org/10.1037/a0018128

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Ratcliff, R., Smith, P. L., & McKoon, G. (2015). Modeling regularities in response time and accuracy data with the diffusion model. *Current Directions in Psychological Science*, 24(6), 458–470. https://doi.org/10.1177/1069193115596228

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16(2), 323. https://doi.org/10.1037/0882-7974.16.2.323

Robertson, I. H., Ridgeway, V., Greenfield, E., & Parr, A. (1997). Motor recovery after stroke depends on intact sustained attention: A 2-year follow-up study. *Neuropsychology*, 11(2), 290. https://doi.org/10.1037/0894-4105.11.2.290

Roca, J., Castro, C., López-Ramón, M. F., & Lupiáñez, J. (2011). Measuring vigilance while assessing the functioning of the three attentional networks: The ANTI-Vigilance task. *Journal of Neuroscience Methods*, 198(2), 312–324. https://doi.org/10.1016/j.jneumeth.2011.04.014

Salthouse, T. A. (1985). A theory of cognitive aging. Amsterdam, The Netherlands: North-Holland.

Schepers, V. P., Visser-Meily, A. M., Ketelaar, M., & Lindeman, E. (2006). Poststroke fatigue: Course and its relation to personal and stroke-related factors. *Archives of Physical Medicine and Rehabilitation*, 87(2), 184–188. https://doi.org/10.1016/j.apmr.2005.10.005

Spaccavento, S., Marinelli, C. V., Nardulli, R., Macchitella, L., Bivona, U., Piccardi, L., … Angelelli, P. (2019). Attention deficits in stroke patients: The role of lesion characteristics, time from stroke, and concomitant neuropsychological deficits. *Behavioural Neurology*, 2019, 1–12. https://doi.org/10.1155/2019/7835710

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Patient Health Questionnaire Primary Care Study, G. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA*, 282(18), 1737–1744.

Strobel, C., & Engedal, K. (2008). MMSE-NR Manual. Retrieved from ab-web-prod.s3.amazonaws.com

Stulemeijer, M., Fosatti, L., & Bleijenberg, G. (2005). Fatigue after Stroke. In J. Deluca (Ed.), *Fatigue as a window to the brain* (pp. 73–87). Cambridge, MA: The MIT Press.

Tang, W. K., Chen, Y. K., Mok, V., Chu, W. C., Ungvari, G. S., Ahuja, A. T., & Wong, K. S. (2010). Acute basal ganglia infarcts in post-stroke fatigue: An MRI study. *Journal of Neurology*, 257(2), 178–182. https://doi.org/10.1007/s00415-009-5284-2

Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Stroke Research and Treatment*, 2013, 1–11. https://doi.org/10.1155/2013/789579

Togo, F., Lange, G., Natelson, B. H., & Quigley, K. S. (2015). Attention deficits in stroke patients: The role of lesion characteristics, time from stroke, and concomitant neuropsychological deficits. *Behavioural Neurology*, 2019, 1–12. https://doi.org/10.1155/2019/7835710

ULRICHSEN ET AL.
its relation to depressive symptoms and disease characteristics. *European Neurology, 45*(1), 28–33. https://doi.org/10.1159/000052085

van Eijsden, H. M., van de Port, I. G. L., Visser-Meily, J. M. A., & Kwakkel, G. (2012). Poststroke fatigue: Who is at risk for an increase in fatigue? *Stroke Research and Treatment, 2012*, 1–8. https://doi.org/10.1155/2012/863978

Van Zomeren, A. H., Brouwer, W. H., & Deelman, B. G. (1984). Attentional deficits: The riddles of selectivity, speed, and alertness. In N. Brooks (Ed.), *Closed head injury: Psychological, social, and family consequences* (pp. 74–107). New York, NY: Oxford University Press.

Van Zomeren, A. H., & Van den Burg, W. (1985). Residual complaints of patients two years after severe head injury. *Journal of Neurology, Neurosurgery & Psychiatry, 48*(1), 21–28. https://doi.org/10.1136/jnnp.48.1.21

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*(7), 1206–1220. https://doi.org/10.3758/BF03196893

Wang, S.-S., Wang, J.-J., Wang, P.-X., & Chen, R. (2014). Determinants of fatigue after first-ever ischemic stroke during acute phase. *PLoS One, 9*(10), e110037. https://doi.org/10.1371/journal.pone.0110037

Westlye, L. T., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2010). Associations between regional cortical thickness and attentional networks as measured by the attention network test. *Cerebral Cortex, 21*(2), 345–356. https://doi.org/10.1093/cercor/bhq101

Whitehead, L. (2009). The measurement of fatigue in chronic illness: A systematic review of unidimensional and multidimensional fatigue measures. *Journal of Pain and Symptom Management, 37*(1), 107–128.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis.

Wiecki, T. V. (2016). HDDM 0.6.0 documentation. Retrieved from http://ski.clps.brown.edu/hddm_docs/howto.html#assess-model-convergence

Wiecki, T. V., Sofer, L., & Frank, M. J. (2013). HDDM: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics, 7*, 14. https://doi.org/10.3389/fninf.2013.00014

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Ulrichsen KM, Alnæs D, Kolskår KK, et al. Dissecting the cognitive phenotype of post-stroke fatigue using computerized assessment and computational modeling of sustained attention. *Eur J Neurosci*. 2020;52:3828–3845. https://doi.org/10.1111/ejn.14861