

Change in intraindividual variability over time as a key metric for defining performance-based cognitive fatigability



Chao Wang^a, Mingzhou Ding^a, Benzi M. Kluger^{b,*}

^a J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL, United States

^b Departments of Neurology and Psychiatry, University of Colorado Denver, Aurora, CO, United States

ARTICLE INFO

Article history:

Accepted 7 January 2014

Available online 30 January 2014

Keywords:

Fatigability

Stroop

Fatigue

Intraindividual variability

Cognition

ABSTRACT

Cognitive fatigability is conventionally quantified as the increase over time in either mean reaction time (RT) or error rate from two or more time periods during sustained performance of a prolonged cognitive task. There is evidence indicating that these mean performance measures may not sufficiently reflect the response characteristics of cognitive fatigue. We hypothesized that changes in intraindividual variability over time would be a more sensitive and ecologically meaningful metric for investigations of fatigability of cognitive performance. To test the hypothesis fifteen young adults were recruited. Trait fatigue perceptions in various domains were assessed with the Multidimensional Fatigue Index (MFI). Behavioral data were then recorded during performance of a three-hour continuous cued Stroop task. Results showed that intraindividual variability, as quantified by the coefficient of variation of RT, increased linearly over the course of three hours and demonstrated a significantly greater effect size than mean RT or accuracy. Change in intraindividual RT variability over time was significantly correlated with relevant subscores of the MFI including reduced activity, reduced motivation and mental fatigue. While change in mean RT over time was also correlated with reduced motivation and mental fatigue, these correlations were significantly smaller than those associated with intraindividual RT variability. RT distribution analysis using an ex-Gaussian model further revealed that change in intraindividual variability over time reflects an increase in the exponential component of variance and may reflect attentional lapses or other breakdowns in cognitive control. These results suggest that intraindividual variability and its change over time provide important metrics for measuring cognitive fatigability and may prove useful for inferring the underlying neuronal mechanisms of both perceptions of fatigue and objective changes in performance.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Although fatigue is a normal reaction to prolonged performance, it limits human performance with safety implications for many professions, including transport, military and medicine (Caldwell et al., 2009; Gander, Purnell, Garden, & Woodward, 2007; Shen et al., 2006). Fatigue is also a significant cause of disability for many medical and neurological illnesses (Kluger, Krupp, & Enoka, 2013). Unfortunately, progress in understanding fatigue mechanisms and developing interventions has been slow and is partly attributable to issues of definition and measurement. Fatigability is used to refer to objective changes in performance or physiological capacity, whereas perceived fatigue refers to subjective sensations of exhaustion (Kluger et al., 2013). Similar to anxiety, perceived fatigue may be divided into immediate state

perceptions and longer-term trait tendencies (Kluger et al., 2013). Fatigability and perceived fatigue are dissociable and linking objective measures to perceptions has proven challenging (Bailey et al., 2007; Krupp and Elkins 2000).

Compared to the motor domain, relatively little is known about the phenomenology and physiology of cognitive fatigability. Previous studies in this area demonstrate three important findings. First, the specific task used to induce fatigue influences the pattern of performance change. Persson, Welsh, Jonides, and Reuter-Lorenz (2007) for example found that performing an intensive training task only interfered with a subsequent but distinct probe task if both tasks relied on similar neural resources. Second, measures of cognitive control, including top-down control of attention, error monitoring, and resistance to interference are particularly vulnerable to deterioration with prolonged task performance (Boksem & Tops, 2008). Finally, neuroimaging and neurophysiology studies support a key role for medial frontal structures in compensating for early effects of fatigue and dysfunction of these structures appear to mediate many of the behavioral effects associated with prolonged cognitive performance (Boksem & Tops, 2008).

* Corresponding author. Address: Department of Neurology, University of Colorado Denver, Mail Stop B-185, 12631 East 17th Avenue, Aurora, CO 80045, United States. Fax: +1 (303) 724 2212.

E-mail address: benzi.kluger@ucdenver.edu (B.M. Kluger).

Studies of fatigability of cognitive performance typically use a block design comparing mean reaction time (RT) and/or accuracy on a probe task performed before and after a fatigue-inducing task; or comparing these metrics between the first and last block of the fatigue-inducing task itself. These studies assume a general slowing or change in accuracy as the fundamental characteristic of cognitive fatigue. Although mean performance levels have shed light on aspects of phenomenology and physiology of cognitive fatigue, they have not correlated with perceived state or trait fatigue.

Intraindividual variability, which refers to within-person fluctuations in performance across trials, is a long-recognized behavioral feature in individuals with brain lesions (Burton, Hultsch, Strauss, & Hunter, 2002; de Frias, Dixon, Fisher, & Camicioli, 2007; Head, 1926). Heightened intraindividual variability is also associated with aging (Bielak, Hultsch, Strauss, MacDonald, & Hunter, 2010b; Martin & Hofer, 2004), and predicts a higher risk for progression to dementia in older adults with mild cognitive impairment (Bielak, Hultsch, Strauss, MacDonald, & Hunter, 2010a). In attention-deficit hyperactivity disorder, intraindividual variability was shown to be more sensitive in classifying patients versus controls than mean RT or accuracy (Klein, Wendling, Huettner, Ruder, & Peper, 2006). Neuroimaging studies suggest these effects may in part be mediated by frontal lobe dysfunction and impaired regulation of the default mode network (Kelly, Uddin, Biswal, Castellanos, & Milham, 2008; MacDonald, Li, & Backman, 2009).

There are many similarities between conditions associated with heightened intraindividual variability and fatigability of cognitive performance including: (1) Associations with aging and other neurological illnesses (Eldadah, 2010; Kluger et al., 2013); (2) Neuroimaging evidence for similar neuronal resources (Boksem & Tops, 2008; Chaudhuri & Behan, 2004; MacDonald, Li, et al., 2009); and (3) Reduction by similar interventions (Edgar, Pace-Schott, & Wesnes, 2009; Spencer et al., 2009). Review of reaction time variability from other fatigue studies also supports this contention in some (van der Linden & Eling, 2006) but not all (Lorist & Jolij, 2012) studies. Finally, a recent study looking at intraindividual response variability on traditional (as opposed to fatiguing) cognitive tasks found that intraindividual response variability was correlated with cognitive trait fatigue in multiple sclerosis (Bruce, Bruce, & Arnett, 2010). We thus hypothesized that changes in intraindividual variability over time may represent a useful metric for understanding the phenomenology, subjective correlates, and physiology of cognitive fatigability.

We had three specific objectives to address through distinct analyses of a single behavioral experiment. First was to determine whether intraindividual variability was more sensitive to changes over time than response accuracy or mean response time. Second was to determine whether change in intraindividual variability, accuracy or mean reaction time correlated with trait measures of perceived fatigue in healthy adults. Third was to model changes in response time distributions to better understand the sources of intraindividual variability change.

2. Materials and methods

2.1. Participants

Sixteen college students, free from neurological disorders and with normal or corrected-to-normal vision were recruited in exchange for course credits. All participants were right-handed and native English speakers. One participant was unable to finish the experiment. Data from the remaining fifteen participants (8 females, 7 males), aged 18–33 (mean = 19.4; SD = 3.8), were included. All participants provided written informed consent in accordance with the University of Florida Institutional Review

Board approved protocol and were informed that they would be asked to perform a cognitive task without breaks for 3 h. To reduce potential cognitive changes associated with diurnal factors or arousal, participants were instructed to get a full night's sleep prior to testing and all testing began at 9:00 AM. Participants were asked to refrain from consuming caffeine or nicotine on the day of testing.

2.2. Perceived trait fatigue questionnaires

Two standardized scales were used for the assessment of perceived trait fatigue. The Fatigue Severity Scale (FSS; Krupp, LaRocca, Muir-Nash, & Steinberg, 1989) is a seven item questionnaire which queries overall trait fatigue severity. The Multidimensional Fatigue Inventory (MFI; Smets, Garssen, Bonke, & De Haes, 1995) is a 20 item questionnaire, covering five dimensions of trait fatigue: general fatigue, physical fatigue, reduced activity, reduced motivation and mental fatigue. Participants completed these questionnaires before the cognitive task on the morning of testing. State perceptions were not assessed due to concerns of interrupting the continuous performance test.

2.3. Objective cognitive fatigability

Testing was performed in an acoustically and electrically shielded room. Instructions and stimuli were displayed on a PC using the Experimental Run Time System (Berisoft Corporation, Frankfurt, Germany). A computerized cued Stroop task (Cohen, Barch, Carter, & Servan-Schreiber, 1999) was performed continuously for 3 h (see Fig. 1). In this task an instructional cue is presented on each trial (“word” or “color”) and followed by a 1, 3 or 5 s cue-target interval and an imperative color word stimulus (“red”, “blue” or “green”) written in colored letters. For the “word” task participants were instructed to read the word, while in the “color” task participants were instructed to name the color of the letters. 60% of the trials were congruent (e.g. “red” written in red letters) and 40% were incongruent (e.g. “red” written in green letters). Reaction times were determined by voice activation software and verbal responses were manually recorded by a research assistant. After a brief practice session (30 trials), participants were fitted with EEG electrodes (EEG data not presented here) and asked to perform the task for a single 3-h session. Breaks for any purpose were taken only if requested and resulted in participant exclusion.

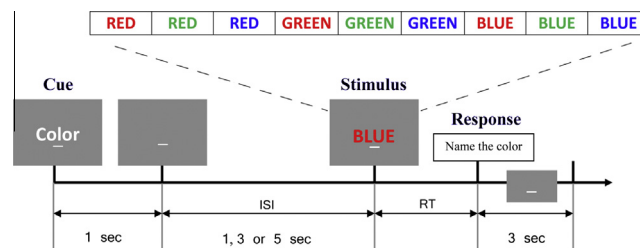


Fig. 1. The Stroop task used for testing fatigability of cognitive performance. In the beginning of each trial, a “word” or “color” instructional cue was presented. Followed by a 1, 3 or 5 s cue-target interval, an imperative color word stimulus (“red”, “blue” or “green”) written in colored letters would appear. For the “word” task participants were instructed to read the word, while in the “color” task participants are instructed to name the color of the letters. On 60% of the trials the color of letters and the word were congruent, but on the other 40% the letters and word were incongruent. Any response would trigger a 3 s inter trial interval prior to the next cue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. Intraindividual variability

Two parameters were used to measure intraindividual variability: the standard deviation of reaction time (SDRT) and the coefficient of variation of reaction time (CVRT). The coefficient of variation is a normalized measure of variation, which is defined as the ratio of the standard deviation to the mean. Compared with SDRT, CVRT has a low bias from mean RT which allows us to study variability changes relatively independently of changes in overall mean performance (Saville et al., 2011).

2.5. Ex-Gaussian distribution

As empirical RT distributions usually are not normally distributed but rather positively skewed, mean and variance in these cases do not fully describe the distribution. To better understand the influence of fatigue on RT distribution, we fitted the ex-Gaussian distribution to the data. The ex-Gaussian distribution results from the convolution of a Gaussian and exponential distribution. Its probability density function is given by

$$f(x|\mu, \sigma, \tau) = \frac{1}{\tau\sqrt{2\pi}} \exp[\sigma^2/2\tau^2 - (x - \mu)/\tau] \cdot \int_{-\infty}^{(x-\mu)/\sigma - \sigma/\tau} \exp(-y^2/2) dy \quad (1)$$

where μ is the mean of the Gaussian component, σ is the standard deviation of the Gaussian component, and τ is both the mean and the standard deviation of the exponential component. Changes in μ result in a left or rightward shift of the distribution, changes in σ result in a widening or narrowing of the overall distribution, and changes in τ result in stretching of the right tail of the distribution. An example of the ex-Gaussian distribution is shown in Fig. 6B.

The ex-Gaussian distribution has been demonstrated to provide excellent fits for empirical RT distributions (Heathcote, Popiel, & Mewhort, 1991; Luce, 1986; Ratcliff & Murdock, 1976). More importantly, the three ex-Gaussian parameters can be mapped onto the mean and the variance of the distribution. Specifically, the mean of the distribution equals the sum of μ and τ , while the variance of the distribution equals the sum of σ^2 and τ^2 . This property enables us to correlate changes in mean RT and intraindividual variability with changes in these independent distribution parameters, thereby providing insights into the source of the changes.

2.6. Statistical analyses

Trials with RTs abnormally slow (>5 s) or fast (<300 ms) or with incorrect responses were excluded from RT analysis. The five second cut-off was chosen as a compromise between outliers at 3 standard deviations for all trials (4.6 s) and trials from the last hour (5.4 s) and resulted in excluding 2.3% of trials. Temporal changes in performance were assessed by dividing the 3-h experiment into 6 non-overlapping 30-min blocks. Each 30-min block contains approximately 190 trials to allow reliable measurement of accuracy, mean RT, SDRT and CVRT (Saville et al., 2011). Cognitive fatigue was quantified as the slope of a linear regression for each performance outcome measures against time blocks. To test the reliability of this measure we calculated a regression slope using a moving window method with window size of 30 min and step size of 1 min resulting in 151 overlapping blocks and found very high correlation ($r = 0.99$, $p < 0.001$) between moving window and non-overlapping block results.

To determine which performance measures are most sensitive to changes over time, we estimated the effect size for the change of each performance measure from the first block (baseline) over time. The effect size, which quantifies the size of performance changes, was calculated by dividing the population mean differ-

ences between performance measures recorded at baseline and the block of interest by the standard deviation of the differences (Gibbons, Hedeker, & Davis, 1993). To test the statistical significance of the differences between effect sizes for different performance measures, a bootstrap method (Efron, 1982) with 1000 times resampling was applied to estimate the 95% confidence interval of the differences. Differences in which the 95% confidence interval did not include zero were considered statistically significant.

Correlations with trait fatigue scales was performed using Pearson's correlation coefficient. Statistical significance was tested by Fisher-transforming correlation coefficients to create a t statistic having $N - 2$ degree of freedom, where N is the number of participants. Differences between correlation coefficients between outcome measures was tested using a Z-test (Meng, Rosenthal, & Rubin, 1992).

For RT distribution analyses, the ex-Gaussian function was fitted to the RT data within each 30-min time block for each subject. A MATLAB toolbox named DISTRIB (Lacouture & Cousineau, 2008) was used for estimating the parameters of ex-Gaussian distribution. The differences of the estimated parameters from the first and the last time blocks were tested across subjects using paired t -tests.

3. Results

3.1. Overall task performance

The overall task performance error rate was 3.85% (SD = 2.00%). The mean number of trials per subject used for RT analyses was 1104 (SD = 103). Mean RT was 1266.87 ms (SD = 186.63 ms) and mean SDRT was 507.33 ms (SD = 184.77 ms). Mean RT and error rate from incongruent trials (mean RT = 1410.72 ms, error rate = 8.83%) were significantly larger ($p < 0.011$) than that from congruent trials (mean RT = 1180.57 ms, error rate = 0.52%). SDRT was also larger for incongruent trials ($p = 0.017$). CVRT did not show significant difference between congruent and incongruent trials ($p = 0.495$). By comparing performance data from color-naming trials and word-reading trials, we found more errors were made in color-naming condition ($p = 0.013$), while no significant difference was found for mean RT ($p = 0.577$). Regarding RT variability, both SDRT and CVRT were significantly smaller for color-naming trials ($p < 0.014$).

3.2. Objective 1: Changes in metrics of task performance over time

As shown in Fig. 2A–C, the CVRT, SDRT and mean RT continuously increased with time blocks. Response error rate did not consistently change over time blocks (Fig. 2D).

Fig. 2E–H summarized the changes over time blocks by task type for CVRT, SDRT, mean RT and error rate, respectively. Since temporal effects on different task types (e.g. word vs. color cue) were similar, a combined dataset including all task types was used for remaining analyses.

To further examine which performance measures are more sensitive to cognitive fatigue, we estimated the effect sizes for the changes of each performance measure from the first block over time. As shown in Fig. 3, the two measures of intraindividual variability, CVRT and SDRT, exhibit the largest effect size across time blocks. Statistical comparison based on bootstrapping found that effect sizes computed from CVRT were significantly larger than effect sizes from mean RT or error rate for block 3 and 4 ($p < 0.039$). Effect sizes computed from SDRT were significantly larger than effect sizes from mean RT or error rate for block 3, 4 and 5 ($p < 0.028$).

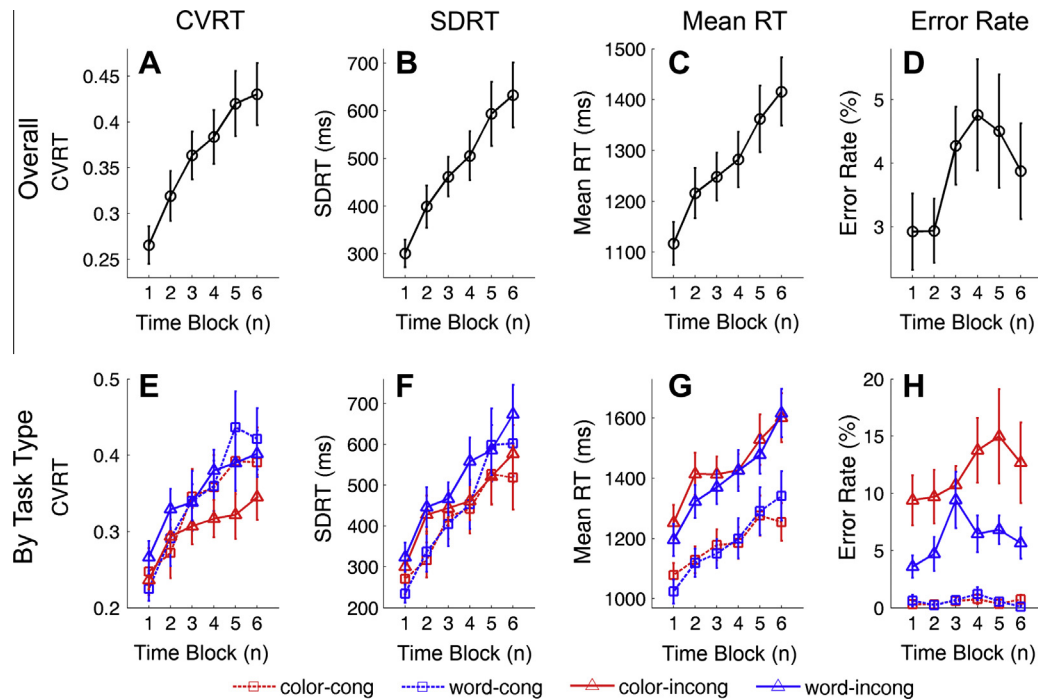


Fig. 2. Changes of behavioral performance over time blocks. The 3 h of performance data was divided into 6 time blocks with each time block lasting 30 min. (A) CVRT (B) SDRT (C) mean RT and (D) error rate from each time block across all task types. (E) CVRT (F) SDRT (G) mean RT and (H) error rate from each time block for different task types.

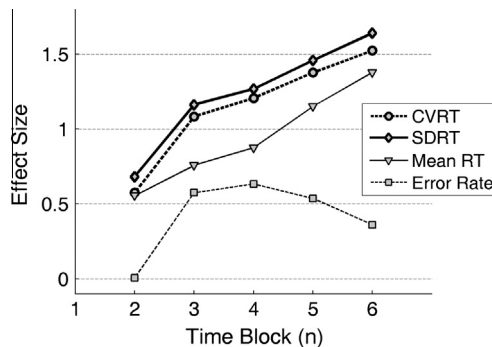


Fig. 3. The effect size for the change of each performance measure from the first block over time blocks. A larger effect size indicates that the statistical significance can be reached by a smaller sample size or shorter task. The effect sizes computed from CVRT were significantly larger than the effect sizes from mean RT and error rate for block 3 and 4 ($p < 0.039$). The effect sizes computed from SDRT were significantly larger than the effect sizes from mean RT and error rate for block 3, 4 and 5 ($p < 0.028$).

3.3. Objective 2: Correlation of cognitive fatigability and perceived trait fatigue

Fig. 4A shows the scatter plots displaying correlations between the cognitive fatigability using CVRT slope and trait fatigue measured by the MFI-20 before task. Quite strikingly, CVRT slope showed very strong correlations with the reduced activity scale ($r = 0.78$, $p < 0.001$), the reduced motivation scale ($r = 0.90$, $p < 0.001$) and the mental fatigue scale ($r = 0.76$, $p < 0.001$). Cognitive fatigability using SDRT slope showed similar strong correlations with these subjective fatigue scales (Fig. 4B).

Fig. 4C shows the correlations between the cognitive fatigability using mean RT slope and the subjective fatigue scales. No significant correlation was found between mean RT slope and the reduced activity scale across participants ($r = 0.38$, $p = 0.160$). For

the reduced motivation scale and the mental fatigue scale, correlations with mean RT were significant (reduced motivation scale: $r = 0.51$, $p < 0.050$; mental fatigue scale: $r = 0.52$, $p = 0.048$). However, CVRT slope and SDRT slope showed significantly stronger correlation with the reduced activity scale ($p < 0.016$), the reduced motivation scale ($p < 0.012$) and the mental fatigue scale ($p < 0.044$) than mean RT slope.

For the FSS and the MFI-20 subscales in the dimensions of general fatigue and physical fatigue, their correlations with the cognitive fatigability measures were not significant ($p > 0.353$). No significant correlation was found between error rate slope and the subjective fatigue scales $p > 0.262$.

3.4. Objective 3: Fatigue effects on RT distribution

Fig. 5A is the scatter plot displaying trial-by-trial RTs (converted into z-scores) from all 15 subjects during the 3 h Stroop task demonstrating an increase in RT variability with a relatively unchanged floor.

The changes of RT distribution over time were further explored by fitting the ex-Gaussian function to the RT data within each 30 min time block. Fig. 5B shows the probability density functions (PDFs) of the estimated distributions for the first 30 min block and the last 30 min block across participants showing most notably an increase in the rightward tail with fatigue. Fig. 5C is a bar plot comparing the fitted ex-Gaussian parameters between the first block and the last block. The parameter τ , which reflects the right-hand tail of the RT distribution, was significantly larger for the last block than the first block ($p < 0.001$). No significant difference was found in μ ($p = 0.200$) or σ ($p = 0.385$). Fig. 5D further showed the estimated parameters for each time block. We can see that τ was the only parameter to significantly increase with time blocks.

Fig. 6 shows the scatter plots for the correlations between the regression slopes of parameter τ against time blocks and the three MFI-20 subscales. Similar to what we have found for CVRT slope and SDRT slope, τ slope significantly correlated with the reduced activity scale ($r = 0.73$, $p = 0.002$), the reduced motivation scale

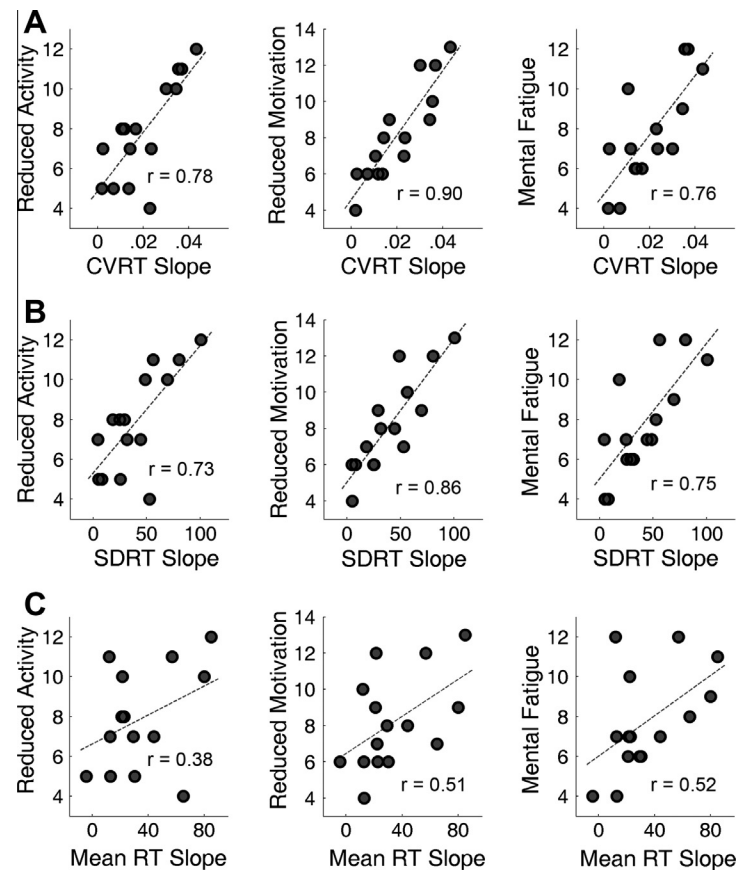


Fig. 4. Scatter plots displaying the correlation between cognitive fatigability, measured as the slope of performance versus time block regression line, and the subjective fatigue scales (MFI-20 subscales). Each dot represents one of the 15 participants. (A) Correlations between CVRT slope and the subjective fatigue scales. CVRT slope significantly correlated with the reduced activity scale ($r = 0.78$, $p < 0.001$), the reduced motivation scale ($r = 0.90$, $p < 0.001$) and the mental fatigue scale ($r = 0.76$, $p < 0.001$). (B) Correlations between SDRT slope and the subjective fatigue scales. Similar to CVRT slope, SDRT slope significantly correlated with the reduced activity scale ($r = 0.73$, $p = 0.002$), the reduced motivation scale ($r = 0.86$, $p < 0.001$) and the mental fatigue scale ($r = 0.75$, $p = 0.001$). (C) Correlations between mean RT slope and the subjective fatigue scales. Significant correlations were found for the reduced motivation scale ($r = 0.51$, $p = 0.050$) and the mental fatigue scale ($r = 0.52$, $p = 0.048$), but not for the reduced activity scales ($r = 0.38$, $p = 0.160$).

($r = 0.86$, $p < 0.001$) and the mental fatigue scale ($r = 0.75$, $p = 0.001$). For the FSS and the MFI-20 subscales in the dimensions of general fatigue and physical fatigue, their correlations with τ slope were not significant ($p > 0.459$). No significant correlation was found for μ or σ ($p > 0.132$).

4. Discussion

Changes in intraindividual variability over time provide information regarding cognitive fatigability above that provided by mean performance levels and may contribute to our understanding of the effects of fatigue on performance and underlying mechanisms. Specifically, we found that the intraindividual variability of RT increased linearly over time with continued task performance and demonstrated larger effect sizes in change from baseline than did the more traditional metrics of mean RT and accuracy. Moreover, these changes in RT variability significantly correlated with trait perceptions of many aspects of mental fatigue and were more strongly correlated with these measures than either mean RT or accuracy changes.

Regarding our first objective, the finding that intraindividual variability was more sensitive to changes over time than response accuracy or mean response time is in accord with studies in aging populations and patients with neuropsychiatric disorders which also find that intraindividual variability more sensitively identifies

populations or individuals with performance deficits than mean RT (Bielak et al., 2010b; Burton et al., 2002; de Frias et al., 2007; Head, 1926; Klein et al., 2006). There are several potential mechanisms for the association of intraindividual variability and cognitive fatigability. First, both phenomena are associated with changes in pre-frontal networks, particularly medial frontal structures (Boksem & Tops, 2008; Hahn, Ross, & Stein, 2007). Literature from both fields can be interpreted through a link to cognitive control mechanisms; regarding fatigue this challenge to cognitive control may be viewed as increases in mental effort to compensate for decrements in other networks while in intraindividual variability it may be seen as a response to unpredictability. Second, both phenomena are hypothesized to be associated with alterations in catecholamines, including both dopamine and norepinephrine levels and/or receptors (Curtin, Walker, & Schulz, 1996; Foley & Fleschner, 2008; Klass et al., 2012; MacDonald, Cervenka, Farde, Nyberg, & Backman, 2009; MacDonald, Karlsson, Rieckmann, Nyberg, & Backman, 2012; Meeusen, Watson, Hasegawa, Roelands, & Piacentini, 2006; Moeller, Tomasi, Honorio, Volkow, & Goldstein, 2012). Finally, we would consider the present study to offer evidence in support of a “neural noise hypothesis of fatigability” where neural noise builds over time of continuous task performance, similar to neural noise theories of aging which are also hypothesized to mediate intraindividual performance changes and linked to catecholamines through effects on neuronal signal to noise ratios (Li, Lindenberger, & Sikstrom, 2001; MacDonald, Li, et al., 2009). Notably, CVRT did

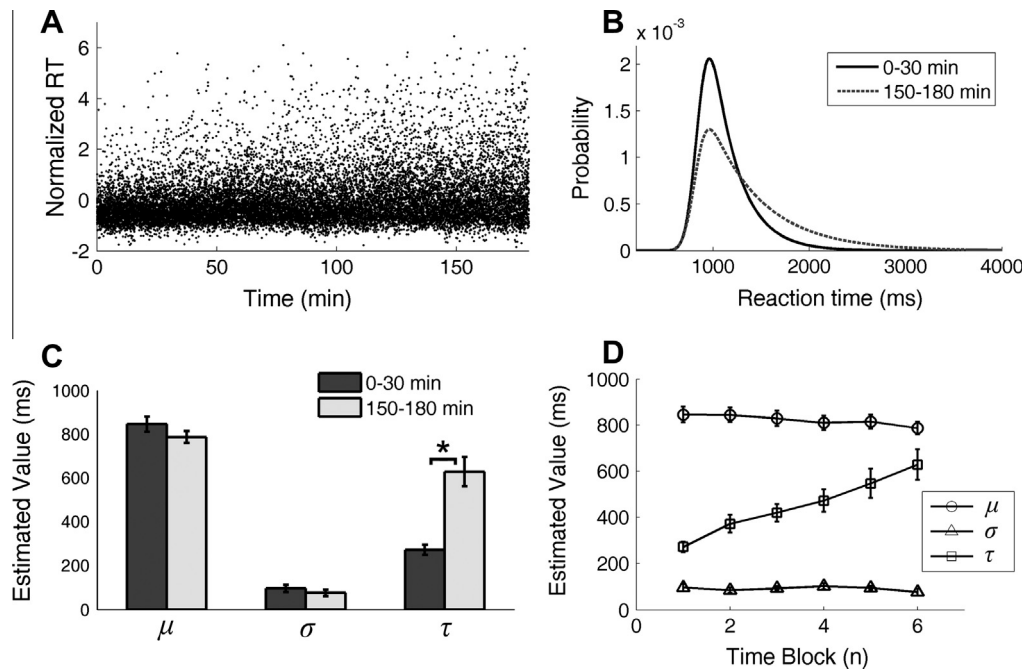


Fig. 5. RT distribution analyses using ex-Gaussian function. (A) Scatter plot of trial-by-trial RTs from all 15 subjects. The values of RT were normalized into z-scores within each subject. (B) The estimated ex-Gaussian probability density functions for the RT distributions within the first 30 min time block and the last 30 min time block. The parameters of ex-Gaussian distribution were first estimated within each subject and then averaged across participants. (C) The comparison between the estimated parameters of ex-Gaussian distribution for the first 30 min and for the last 30 min. No significant difference was found in μ and σ ($p > 0.05$). The value of τ estimated from the ending 30 min was significantly larger than that from the beginning 30 min ($p = 0.001$). (D) The estimated parameters of ex-Gaussian distribution for each 30 min time block. Only τ increased with time blocks.

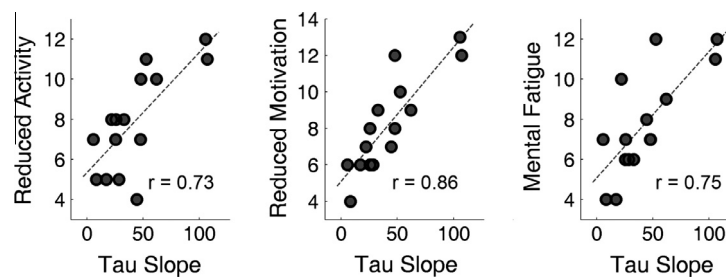


Fig. 6. Scatter plots displaying the correlation between the slope of parameter τ versus time block regression line and the subjective fatigue scales. Similar to SDRT slope and CVRT slope, τ slope significantly correlated with the reduced activity scale ($r = 0.73$, $p = 0.002$), the reduced motivation scale ($r = 0.86$, $p < 0.001$) and the mental fatigue scale ($r = 0.75$, $p = 0.001$).

not differ between congruent and incongruent trials despite significant differences in SDRT and mean reaction time further arguing that fatigue effects are not simply due to scaling effects and are mediated by distinct neural resources.

Regarding our second objective, we found that changes in intraindividual variability over time correlated with subjective measures of trait fatigue in healthy adults, and were a better predictor of these perceptions than changes in mean reaction time. Importantly, these correlations were specific to cognitive aspects of fatigue (mental fatigue, motivation and activity) but not general or physical fatigue. To date, studies in healthy populations generally do not show strong correlations between objective metrics of motor or cognitive performance and subjective state or trait measures, suggesting that psychological or motivational factors may play a larger role than performance or physiological changes in driving these perceptions (Boksem & Tops, 2008; Marcora, 2010). To our knowledge, intraindividual variability has not been previously studied in relation to trait fatigue in healthy populations and an individual's tendency to experience mental fatigue may be related to intraindividual variability in task performance. Moreover, our results are consistent with studies in multiple sclerosis where trait

fatigue complaints do not correlate with objective measures of cognitive or physical fatigability but do correlate with intraindividual variability measured at a single time point (Bruce et al., 2010; Kinsinger, Lattie, & Mohr, 2010; Surakka et al., 2004). Similarly, intraindividual variability showed greater differences than mean RT or accuracy across a range of tasks in chronic fatigue syndrome versus healthy controls (Fuentes, Hunter, Strauss, & Hultsch, 2001).

Regarding our third objective, changes in RT distributions quantified by the ex-Gaussian model suggest that changes in intraindividual variability were due to changes in τ parameter, which reflects the positive skewness of the distribution. The standard deviation of RT specifically reflects the parameter τ , as it equals to $\sqrt{\sigma^2 + \tau^2}$ and the value of σ was small and did not significantly change over time (Fig. 5D). Although the increase in positive skew also directly contributed to mean RT as the mean RT equals to $\mu + \tau$, its effect was attenuated by μ since the value of μ was relatively large (Fig. 5D). Ex-Gaussian models have been previously used to understand the source of reaction time changes in psychological tasks (Burnham, 2012; Heathcote et al., 1991). Given that top-down control processes are slower than stimulus driven processes, it has been hypothesized that changes in the exponential

component τ are more likely due to alterations in top-down influences (Burnham, 2012). This interpretation would be consistent with other evidence suggesting that fatigue specifically affects aspects of tasks and neurophysiological markers of cognitive control (Boksem, Meijman, & Lorist, 2006). It may also be consistent with theories of neural noise, in which increasing noise leads to an increased probability of breaks in network function and/or attentional lapses (Esterman, Noonan, Rosenberg, & Degutis, 2012). However, it remains possible that bottom-up and arousal processes may also contribute to these findings and should be examined including behavioral manipulations and neurophysiologic measures in future studies.

Limitations of this work include the fact that only trait perceptions of fatigue were measured rather than state fatigue. This suggests that our findings may relate to subjects' propensity for fatigue but may not reflect momentary change in perceptions of fatigue. Using multiple brief probes during task performance may prove effective to assess the temporal evolution of perceived fatigue and correlate it with objective measures. This strategy, however, should be handled carefully as our pilot studies suggested that even small breaks in task performance may affect the development of objective performance changes. Second, it should also be recognized that changes in task demands may alter how prolonged performance changes over time (Klass, Levenez, Enoka, & Duchateau, 2008) and that not all cognitive fatiguing paradigms report increases in variability (Lorist & Jolij, 2012). Moreover, we have found that continued performance of certain tasks (e.g. visual search) may be more likely to result in decrements in accuracy with no change in mean RT or RT variability (unpublished data). Given the potential link between executive functions and intraindividual variability, it may be that tasks which stress these functions are particularly susceptible to changes in intraindividual variability. Third, our small sample size should be considered as another limitation of this study. Despite the large effect sizes reported here, replication in larger cohorts is needed to increase confidence in these results and improve our understanding of their significance, particularly the relationship between fatigue perceptions and objective measures. Repeated testing would also be beneficial to understanding the reliability and potential test–retest effects of these measures. Finally, the long duration of testing is demanding for both participants and researchers and may preclude use of this testing for clinical purposes and in certain clinical populations. Although this duration of testing is consistent with prior studies of cognitive fatigue (Lorist & Jolij, 2012) and had a high completion rate (94% for this study), our results suggest that significant changes from baseline may be detected earlier, with a large effect size (>1.0) noted at 90 min. We would also predict that more difficult tasks may induce performance changes faster. Thus, depending on the task, sample size and goals of a study shorter durations of testing may be possible.

Future studies of cognitive fatigue should consider intraindividual variability and RT distributions as important potential outcomes. Variations in task demands, motivation and physiological interventions (e.g. administration of stimulants) will add further insight into the significance of these findings and there neuronal basis, as will neuroimaging studies. On a clinical note, intraindividual variability may be a sensitive objective marker for populations plagued by fatigue complaints and these complaints may respond to interventions shown to reduce intraindividual fatigue including cognitive training (Gajewski & Falkenstein, 2012).

Acknowledgments

This material is supported by the U.S. Army Research Laboratory and the U.S. Army Research Office under contract/grant number W911NF-10-1-0192, the Colorado Clinical and Translational

Sciences Institute KL2 program (NIH/NCATS 8 KL2 TR000156-05), and the National Institutes of Health (NIMH MH097320).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandc.2014.01.004>.

References

- Bailey, A., Channon, S., & Beaumont, J. G. (2007). The relationship between subjective fatigue and cognitive fatigue in advanced multiple sclerosis. *Multiple Sclerosis*, 13(1), 73–80.
- Bielak, A. A., Hultsch, D. F., Strauss, E., MacDonald, S. W., & Hunter, M. A. (2010a). Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. *Neuropsychology*, 24(6), 731–741.
- Bielak, A. A., Hultsch, D. F., Strauss, E., MacDonald, S. W., & Hunter, M. A. (2010b). Intraindividual variability is related to cognitive change in older adults: Evidence for within-person coupling. *Psychology and Aging*, 25(3), 575–586.
- Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology*, 72(2), 123–132.
- Boksem, M. A., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, 59(1), 125–139.
- Bruce, J. M., Bruce, A. S., & Arnett, P. A. (2010). Response variability is associated with self-reported cognitive fatigue in multiple sclerosis. *Neuropsychology*, 24(1), 77–83.
- Burnham, B. R. (2012). Using response time distributions to examine top-down influences on attentional capture. *Attention, Perception & Psychophysics*.
- Burton, C. L., Hultsch, D. F., Strauss, E., & Hunter, M. A. (2002). Intraindividual variability in physical and emotional functioning: Comparison of adults with traumatic brain injuries and healthy adults. *The Clinical neuropsychologist*, 16(3), 264–279.
- Caldwell, J. A., Mallis, M. M., Caldwell, J. L., Paul, M. A., Miller, J. C., & Neri, D. F. (2009). Fatigue countermeasures in aviation. *Aviation, Space and Environmental Medicine*, 80(1), 29–59.
- Chaudhuri, A., & Behan, P. O. (2004). Fatigue in neurological disorders. *Lancet*, 363(9413), 978–988.
- Cohen, J. D., Barch, D. M., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology*, 108(1), 120–133.
- Curtin, F., Walker, J. P., & Schulz, P. (1996). Day-to-day intraindividual reliability and interindividual differences in monoamines excretion. *Journal of Affective Disorders*, 38(2–3), 173–178.
- de Frias, C. M., Dixon, R. A., Fisher, N., & Camicioli, R. (2007). Intraindividual variability in neurocognitive speed: A comparison of Parkinson's disease and normal older adults. *Neuropsychologia*, 45(11), 2499–2507.
- Edgar, C. J., Pace-Schott, E. F., & Wesnes, K. A. (2009). Approaches to measuring the effects of wake-promoting drugs: A focus on cognitive function. *Human Psychopharmacology*, 24(5), 371–389.
- Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM.
- Eldadah, B. A. (2010). Fatigue and fatigability in older adults. *PM & R: The Journal of Injury, Function, and Rehabilitation*, 2(5), 406–413.
- Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2012). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex*.
- Foley, T. E., & Fleshner, M. (2008). Neuroplasticity of dopamine circuits after exercise: Implications for central fatigue. *NeuroMolecular Medicine*, 10(2), 67–80.
- Fuentes, K., Hunter, M. A., Strauss, E., & Hultsch, D. F. (2001). Intraindividual variability in cognitive performance in persons with chronic fatigue syndrome. *The Clinical Neuropsychologist*, 15(2), 210–227.
- Gajewski, P. D., & Falkenstein, M. (2012). Training-induced improvement of response selection and error detection in aging assessed by task switching: Effects of cognitive, physical, and relaxation training. *Frontiers in Human Neuroscience*, 6, 130.
- Gander, P., Purnell, H., Garden, A., & Woodward, A. (2007). Work patterns and fatigue-related risk among junior doctors. *Occupational and Environmental Medicine*, 64(11), 733–738.
- Gibbons, R. D., Hedeker, D. R., & Davis, J. M. (1993). Estimation of effect size from a series of experiments involving paired comparisons. *Journal of Educational Statistics*, 18(3), 271–279.
- Hahn, B., Ross, T. J., & Stein, E. A. (2007). Cingulate activation increases dynamically with response speed under stimulus unpredictability. *Cerebral Cortex*, 17(7), 1664–1671.
- Head, H. (1926). *Aphasia and kindred disorders of speech*. Cambridge, UK: Cambridge University Press.
- Heathcote, A., Popiel, S., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, 109, 340–347.

- Kelly, A. M., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2008). Competition between functional brain networks mediates behavioral variability. *NeuroImage*, 39(1), 527–537.
- Kinsinger, S. W., Lattie, E., & Mohr, D. C. (2010). Relationship between depression, fatigue, subjective cognitive impairment, and objective neuropsychological functioning in patients with multiple sclerosis. *Neuropsychology*, 24(5), 573–580.
- Klass, M., Levenez, M., Enoka, R. M., & Duchateau, J. (2008). Spinal mechanisms contribute to differences in the time to failure of submaximal fatiguing contractions performed with different loads. *Journal of Neurophysiology*, 99(3), 1096–1104.
- Klass, M., Roelands, B., Levenez, M., Fontenelle, V., Pattyn, N., Meeusen, R., et al. (2012). Effects of noradrenaline and dopamine on supraspinal fatigue in well-trained men. *Medicine and Science in Sports and Exercise*, 44(12), 2299–2308 (discussion 456–457).
- Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry*, 60(10), 1088–1097.
- Kluger, B. M., Krupp, L. B., & Enoka, R. M. (2013). Fatigue and fatigability in neurologic illnesses: Proposal for a unified taxonomy. *Neurology*, 80(4), 409–416.
- 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.
- Krupp, L. B., & Elkins, L. E. (2000). Fatigue and declines in cognitive functioning in multiple sclerosis. *Neurology*, 55(7), 934–939.
- Lacouture, Y., & Cousineau, D. (2008). How to use MATLAB to fit the ex-Gaussian and other probability functions to response times. *Tutorials in Quantitative Methods for Psychology*, 4(1), 35–45.
- Li, S. C., Lindenberger, U., & Sikstrom, S. (2001). Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences*, 5(11), 479–486.
- Lorist, M. M., & Jolij, J. (2012). Trial history effects in stroop task performance are independent of top-down control. *PLoS One*, 7(6), e39802.
- Luce, R. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- MacDonald, S. W., Cervinka, S., Farde, L., Nyberg, L., & Backman, L. (2009). Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. *Neuropsychologia*, 47(11), 2299–2304.
- MacDonald, S. W., Karlsson, S., Rieckmann, A., Nyberg, L., & Backman, L. (2012). Aging-related increases in behavioral variability: Relations to losses of dopamine D1 receptors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(24), 8186–8191.
- MacDonald, S. W., Li, S. C., & Backman, L. (2009). Neural underpinnings of within-person variability in cognitive functioning. *Psychology and Aging*, 24(4), 792–808.
- Marcora, S. (2010). Counterpoint: Afferent feedback from fatigued locomotor muscles is not an important determinant of endurance exercise performance. *Journal of Applied Physiology*, 108(2), 454–456.
- Martin, M., & Hofer, S. M. (2004). Intraindividual variability, change, and aging: Conceptual and analytical issues. *Gerontology*, 50(1), 7–11.
- Meeusen, R., Watson, P., Hasegawa, H., Roelands, B., & Piacentini, M. F. (2006). Central fatigue: The serotonin hypothesis and beyond. *Sports Medicine*, 36(10), 881–909.
- Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation-coefficients. *Psychological Bulletin*, 111(1), 172–175.
- Moeller, S. J., Tomasi, D., Honorio, J., Volkow, N. D., & Goldstein, R. Z. (2012). Dopaminergic involvement during mental fatigue in health and cocaine addiction. *Translational Psychiatry*, 2, e176.
- Persson, J., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2007). Cognitive fatigue of executive processes: Interaction between interference resolution tasks. *Neuropsychologia*, 45(7), 1571–1579.
- Ratcliff, R., & Murdock, B. (1976). Retrieval processes in recognition memory. *Psychological Review*, 83, 190–214.
- Saville, C. W. N., Pawling, R., Trullinger, M., Daley, D., Intriligator, J., & Klein, C. (2011). On the stability of instability: Optimising the reliability of intra-subject variability of reaction times. *Personality and Individual Differences*, 51(2), 148–153.
- Shen, J., Botly, L. C., Chung, S. A., Gibbs, A. L., Sabanadzovic, S., & Shapiro, C. M. (2006). Fatigue and shift work. *Journal of Sleep Research*, 15(1), 1–5.
- Smets, E. M., Garssen, B., Bonke, B., & De Haes, J. C. (1995). The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research*, 39(3), 315–325.
- Spencer, S. V., Hawk, L. W., Jr., Richards, J. B., Shiels, K., Pelham, W. E., Jr., & Waxmonsky, J. G. (2009). Stimulant treatment reduces lapses in attention among children with ADHD: The effects of methylphenidate on intra-individual response time distributions. *Journal of Abnormal Child Psychology*, 37(6), 805–816.
- Surakka, J., Romberg, A., Ruutinen, J., Virtanen, A., Aunola, S., & Maentaka, K. (2004). Assessment of muscle strength and motor fatigue with a knee dynamometer in subjects with multiple sclerosis: A new fatigue index. *Clinical Rehabilitation*, 18(6), 652–659.
- van der Linden, D., & Eling, P. (2006). Mental fatigue disturbs local processing more than global processing. *Psychological Research*, 70(5), 395–402.