FatigueSet: A Multi-modal Dataset for Modeling Mental Fatigue and Fatigability

Manasa Kalanadhabhatta[1], Chulhong Min[2], Alessandro Montanari[2], and Fahim Kawsar[2]

[1] College of Information & Computer Sciences, University of Massachusetts Amherst, USA
manasak@cs.umass.edu
[2] Nokia Bell Labs, Cambridge, UK
{chulhong.min,alessandro.montanari,fahim.kawsar}@nokia-bell-labs.com

Abstract. A comprehensive understanding of fatigue and its impact on performance is a prerequisite for fatigue management systems in the real world. However, fatigue is a multidimensional construct that is often poorly defined, and most prior work does not take into consideration how different types of fatigue collectively influence performance. The physiological markers associated with different types of fatigue are also underexplored, hindering the development of fatigue management technologies that can leverage mobile and wearable sensors to predict fatigue. In this work, we present FatigueSet, a multi-modal dataset including sensor data from four wearable devices that are collected while participants are engaged in physically and mentally demanding tasks. We describe the study design that enables us to investigate the effect of physical activity on mental fatigue under various situations. FatigueSet facilitates further research towards a deeper understanding of fatigue and the development of diverse fatigue-aware applications.

Keywords: fatigue · multi-modal sensing · cognitive performance.

1 Introduction

Fatigue is a complex psychophysiological condition that is characterized by experiential feelings of tiredness or sleepiness, suboptimal performance, and a broad range of physiological changes [33]. Fatigue has a detrimental effect on physical and mental performance, leading to reduced decision making and planning abilities, reduced alertness and vigilance, loss of memory, increased risk-taking and errors in judgment, and increased sick time, incident rates, and medical costs [5]. Therefore, fatigue management systems have received much attention for managing potential risks from fatigue in organizations and for promoting individual wellbeing [38, 6, 15]. In common, they aim to detect individuals' fatigue in time and intervene to mitigate any resulting lapses in performance.

There have been extensive efforts to study the causes and temporal dynamics of fatigue in the domain of physiology, medicine, and neuroscience [43, 39, 42]. However, since these attempts have been mostly made with medical devices in a clinical setting, it is difficult to adopt the findings from these studies into ubiquitous computing for healthcare and wellbeing. While there have been attempts to computationally model fatigue using physiological signals from wearable sensors in the mobile computing domain [15, 26], they have mostly been limited to the consideration of a single type of fatigue. However, different types of fatigue have been shown to influence each other in significant ways [41, 27, 20].

Motivated by these observations, we present FatigueSet[3], a multi-modal dataset for modeling the interplay between physical and mental fatigue and its impact on cognitive performance. As a first step for computationally modeling this interplay, in this paper, we collect and introduce a dataset for exploring the impact of physical activity on mental fatigue and associated cognitive performance; we leave the impact of mental fatigue on physical fatigue to be explored in future work. We recruit 12 participants and collect multi-modal sensor data while inducing different levels of neuromuscular burden as well as cognitive load, and observing their physiological responses and performance on cognitive tasks. To enable a comprehensive study, we include a variety of physiological sensors – electroencephalography (EEG), photoplethysmography (PPG), electrocardiography (ECG), electrodermal activity (EDA), skin temperature sensor, accelerometer, and gyroscope – on four different wearable devices (an earable prototype based on Nokia eSense [2], Empatica E4 wrist band [1], Muse S EEG headband [3], and Zephyr BioHarness 3.0 ECG chestband [4]). We hope this dataset will expedite studies towards a deeper understanding of fatigue in the research community as well as facilitate the development of diverse fatigue-relevant applications.

2 Background and Related Work

2.1 Defining Fatigue

Fatigue is a multifaceted construct that lacks a single clear definition. Prior work has attempted to define fatigue in terms of at least three sets of characteristics – experiential, behavioral, and physiological [33].

Experiential definitions of fatigue emphasize feelings of tiredness, exhaustion, and lack of energy, along with low levels of motivation and a disinclination to continue a task. Fatigue as an experiential construct is measured in terms of individuals' self-reported feelings, often on one of many standardized fatigue measurement scales (see [11] for a review). Targeting an experiential measure of fatigue can be a valid treatment goal for clinically fatigued individuals and is also a desirable outcome for fatigue management technologies for the general population. However, there may be a gap between an individual's perception of tiredness and exhaustion and the external consequences resulting from it.

Behavioral definitions of fatigue focus on these consequences, emphasizing decline in performance as a fundamental indication of fatigue. Prior work has conceptualized fatigue as decrements measured either on a *primary* task or a *probe* task. Performance measures on primary tasks are those that are recorded

[3] https://www.esense.io/datasets/fatigueset

as participants engage in the fatiguing task of interest. For instance, fatigue-related effects of time on task have been measured in terms of reaction times or lapses on the psychomotor vigilance test (PVT) [42]. On the other hand, *probe* tasks are interspersed with the primary task and used to obtain momentary performance levels at different points in time. A probe task such as the PVT or the Mackworth Clock Test can be administered several times between trials of a different fatiguing task, or at regular intervals throughout a workday [23].

Prior work has also operationalized fatigue in terms of the neurophysiological changes that occur either to cause it, or as a response to a fatigued condition. This provides an opportunity to objectively measure fatigue in terms of its physiological markers, which is necessary for fatigue management systems based on wearable or environmental sensors. The exact physiological responses depend on the type of fatigue under consideration, and will be discussed Section 2.3.

2.2 Operationalizing Fatigue for Fatigue Management Technologies

From the above discussion, it is clear that attempts to define fatigue solely in terms of either experiential, behavioral, or physiological variables present an incomplete view. Prior research has also shown dissociation between fatigue measurements across two or more dimensions (no change in physiological responses even as individuals report higher levels of fatigue, or dissociation between experiential and performance measures, for example) [33]. This has led to increased interest in a dynamic, multidimensional definition of the concept of fatigue.

We adopt the taxonomy proposed by Kluger et al. [19] to define this concept in terms of two complementary constructs – *fatigue* and *fatigability.* For the remainder of this paper, we use the term "fatigue" to refer to the subjective sensations and perceptions of tiredness and exhaustion. We use the term "fatigability" to refer to objective changes in performance resulting from fatigue and the underlying mechanisms driving it.

Based on the adopted taxonomy, fatigue and fatigability are often co-occurring and are accompanied by associated neurophysiological responses. An ideal fatigue management system should target both these constructs separately as well as consider how they influence each other. Also, physiological measures should be closely monitored and their relationship with both fatigue and fatigability must be individually assessed. FatigueSet is an attempt in this direction, measuring both fatigue and fatigability along with physiological sensor data.

2.3 Types of Fatigue

Prior work has identified two primary types of fatigue based on its causes, physiological markers, and symptomatology: physical and mental fatigue. *Neuromuscular* or *physical fatigue* is fatigue induced by physical exercise, which leads to a decline in muscle power or exerted force [43]. Physical fatigue is associated with changes in EMG activity in the muscles [14]. Other potential indicators of physical fatigue include biomarkers related to the metabolism of adenosine triphosphate, oxidation, or inflammation in the body [43].

Mental fatigue is in turn experienced during and after prolonged periods of demanding cognitive activity. Mental fatigue is characterized by feelings of tiredness, lack of concentration, and performance decline on cognitive tasks [40]. It is associated with increased EEG alpha and theta wave activity in all regions of the cortex, and an increase in beta activity in frontal sites as individuals attempt to maintain vigilance under fatigue [9].

While a majority of prior studies have focused exclusively on either physical or mental fatigue, there is evidence that the two processes influence each other. Studies investigating physical fatigability following mental fatigue have found that mental fatigue or time on task led to less adequate preparation for new tasks and more errors [24]. Mental fatigue significantly reduced time to physical fatigue during a cycling task, though physiological responses to exercise remained unchanged. It was also associated with higher subjective perception of effort [27]. On the other hand, studies have observed both a decline and improvement in different cognitive functions after different physical exercises. The type of physical activity and the level of physical fatigue (low-to-medium activity vs maximal exertion), duration of activity, type of mental fatigability investigated, and initial levels of physical fitness are all thought to be deciding factors [20]. A variety of theories have been proposed to explain this relationship, but prior works lack consensus on how this effect is manifested on various cognitive tasks. In this work, we study this relationship with a focus on two tasks that require different levels of attentional and processing resources.

2.4 Datasets for Fatigue Detection

While a few datasets for fatigue modeling are currently available, most of these are inadequate for deeply understanding the interplay between physical and mental fatigue and between fatigue and fatigability. Luo et al. presented a dataset for the assessment of fatigue using wearable sensors [25]. While they collected longitudinal sensor data from 27 subjects with various sensors, they lack fatigability measures which are essential to understand changes in performance resulting from fatigue and its underlying mechanisms. Other fatigue-related datasets are extremely domain-specific, e.g., Gjoreski et al. presented datasets [13] to infer cognitive loads on mobile games and physiological tasks on a PC using wearable sensors. Elshafei et al. presented a dataset [12] for modeling bicep fatigue during gym activities. Our dataset focuses on task-independent, lower-level cognitive performance and how it is influenced by physical *and* cognitive activity.

3 Methodology

3.1 Study Design

Twelve participants (9 male, 3 female) between the ages of 21 and 40 (mean age: 30.75 years, SD: 5.78 years) completed the present study. One participant had mild asthma and another had seasonal asthma, while none of the others had any current or past health conditions. All participants completed an informed consent before the study and were compensated with a £30 gift card upon completion.

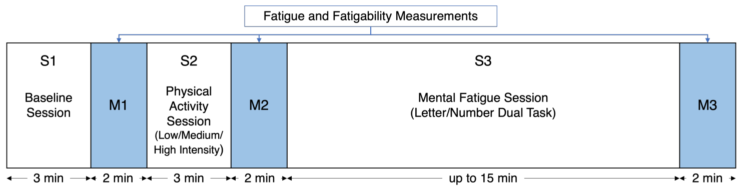


Fig. 1: Study protocol

The study consisted of three sessions conducted on three different days, with a gap of up to 19 days between sessions. Figure 1 shows the protocol for each session. All sessions for a participant were conducted at roughly the same time of the day whenever possible in order to control for circadian effects. Before the first session, participants were asked to fill a preliminary demographic questionnaire. We assessed participants' personality on the short Big Five Inventory (BFI-10) scale [34] and their chronotype (early bird or late owl-ness) using the Munich Chronotype Questionnaire (MCTQ; [35]). Participants also reported the impact of fatigue on their daily functioning using the Fatigue Severity Scale (FSS; [21]) and their general fitness levels on the International Fitness Scale (IFIS; [30]).

Participants began each session by reporting their current sleepiness levels on the Stanford Sleepiness Scale (SSS; [37]) and their baseline vigor and affect on the Global Vigour and Affect Scale (GVAS; [28]). The SSS gives a score between 1 to 7, with 1 corresponding to minimal sleepiness and 7 to highest sleepiness. The GVAS requires participants to rate various aspects of vigor, mood, and affect on visual analogue scales (VAS), which are then converted to separate scores for vigor and affect. Our implementation of the GVAS used a 10-point rating instead of a VAS for easier administration and scoring.

Participants were then fitted with four wearable devices to monitor physiological signals, each of which are described in Section 3.5. They were seated at rest in a comfortable position for three minutes while baseline physiological data was recorded ($S1$). Following the baseline recording period $S1$, participants completed a survey to measure physical and mental fatigue and completed two cognitive tasks to measure baseline cognitive performance for mental fatigability at later stages in the experiment (henceforth referred to as $M1$). This was followed by a 3-minute physical activity session ($S2$), where participants were assigned to one of three conditions (low, medium, or high intensity activity) on a given day. Our study followed a within-subjects design, with all participants completing one session corresponding to each condition. The order in which these conditions are performed was counterbalanced across participants using a balanced Latin Square design.

The period of physical activity in each session was followed by a second measurement of mental fatigue and cognitive performance, $M2$, as shown in Figure 1. Subsequently, participants completed a mental fatigue-inducing task that lasted approximately 15 minutes ($S3$), followed by a third fatigue and fatigability measurement ($M3$). In total, each session lasted up to 30 minutes, and physiological data was recorded for this entire duration. The following subsections provide more details about each part of the study session.

3.2 Physical Activity Protocol

Based on prior research that theorizes an inverted-U relationship between physical activity and cognitive performance [8], we were interested in investigating the effect of low, medium and high intensity physical activity on the development of mental fatigue. We use the metabolic equivalent of task (MET) as an objective indicator of the intensity of physical activity. A MET is defined as the resting metabolic rate, or the amount of oxygen consumed while sitting at rest [16]. The amount of energy required to perform a given physical activity can be quantified in terms of METs, e.g., work requiring twice the resting metabolism is said to be 2 METs. Activities demanding 1-4 METs, 5-8 METs, and > 8 METS, are considered light, medium, and high intensity activities, respectively.

We therefore selected walking at 5 km/hr (3.2 METs), jogging at 7 km/hr (5.3 METs), and jogging at 9 km/hr (8.8 METs) as our low, medium, and high intensity physical activities respectively [16]. During $S2$, participants were asked to walk or run at the given speed on a treadmill without incline for three minutes. Activity sessions were ended early if participants reported a rating equal to or above 10, 14, and 16 on Borg's Rating of Perceived Exertion (RPE) scale [7] during low, medium, and high intensity conditions respectively [29] to avoid overexertion and ensure the safety of the participants.

3.3 Inducing Mental Fatigue

Following the physical activity session $S2$, we used the well-validated dual letter/number task switching paradigm [36] to induce cognitive fatigue in $S3$. Switching between dual tasks has been shown to require additional cognitive overhead and induce fatigue faster than a single cognitive task [31]. The letter/number task was implemented using the PsychoPy framework [32] for experiment design and was administered on a laptop while participants were seated. During $S3$, participants were presented a 2×2 square grid on a grey background. On each trial, a combination of two characters – a letter followed by a number – appeared in one of the squares (see Figure 2). If the characters appear in one of the *top* two squares, participants had to respond to the *letter* and ignore the number. In this case, they were asked to press the 'c' key on their keyboard if the letter was a consonant and the 'm' key if it was a vowel. On the other hand, if the characters appeared in one of the *bottom* two squares, participants were required to respond to the *number* and ignore the letter. Based on whether the number was even or odd, participants had to respond by pressing 'c' or 'm' respectively. The task consisted of 200 trials and lasted approximately 15 minutes, which has been shown to be enough to induce mental fatigue [31]. Performance on the dual task was not analyzed, since the objective was only to induce mental fatigue by virtue of time on task.

3.4 Fatigue and Fatigability Measurements

To measure the impact of physical activity on mental fatigue and fatigability, we obtained self-reported fatigue scores and performance on two distinct cognitive

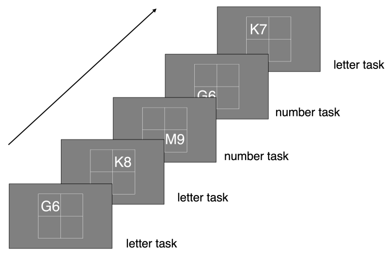


Fig. 2: Dual task to induce mental fatigue.

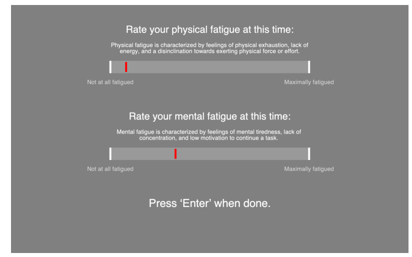


Fig. 3: Fatigue visual analog scales.

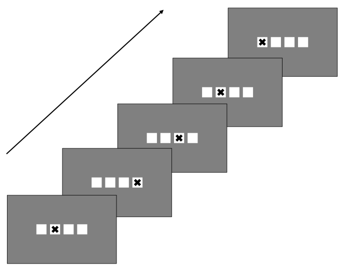


Fig. 4: Choice reaction time task.

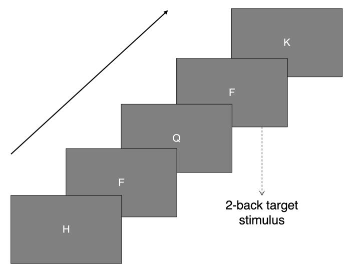


Fig. 5: N-back task.

tasks three times during each session - before physical activity ($M1$), between physical activity and mental fatigue induction ($M2$), and after the induction of mental fatigue ($M3$).

Measurement of fatigue: Participants were asked to report their physical and mental fatigue on two computerized VAS scales ranging from "Not at all fatigued" to "Maximally fatigued" (see Figure 3). Prior work has validated the use of simple VAS scales to measure fatigue, suggesting their utility over more complex multi-dimensional scales [11]. Participants were instructed to rate their levels of fatigue *at this time* by clicking or dragging the mouse along the scale, and were provided the following definitions of physical and mental fatigue in an attempt to ensure a similar understanding of the terms across participants:

Physical fatigue is characterized by feelings of physical exhaustion, lack of energy, and a disinclination towards exerting physical force or effort. Mental fatigue is characterized by feelings of mental tiredness, lack of concentration, and low motivation to continue a task.

Physical and mental fatigue ratings were converted to integers between 0-100 based on the distance from the "Not at all fatigued" end of the scale.

Measurement of mental fatigability: Participants were asked to perform two short cognitive tasks, and their performance was measured in terms of reaction times and errors committed. The difference in performance as compared to the baseline measurement $M1$ was used as an indicator of fatigability.

The first task was the Deary-Liewald Choice Reaction Time (CRT) task [10], which requires participants to select and make the appropriate response to each of several stimuli. The CRT task has been used as an indicator of processing speed, and reaction times have been shown to be affected by physical exertion [22]. In our study, participants were presented with four white squares stacked horizontally on a grey background (see Figure 4). The squares were each mapped to a different key on the keyboard – 'z', 'x', ',' (comma), and '.' (period) respectively from left to right. During each trial, a black cross appeared in one of the squares and participants had to press the corresponding key as soon as possible after the appearance of the cross. The stimulus stayed on the screen until a key was pressed. Once responded, the stimulus disappeared and the next one appeared after a random inter-stimulus interval of 1 to 3 seconds. Each performance measurement consisted of 36 trials of the CRT task, with the stimulus appearing in each of the 4 boxes an equal number of times.

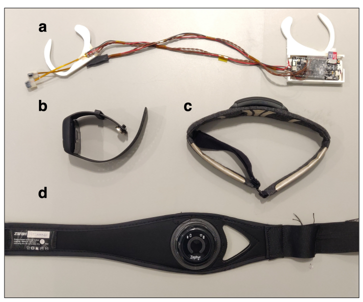


Fig. 6: Wearables for data collection: (a) our earable prototype, (b) E4 wristband, (c) Muse S headband, (d) BioHarness ECG chest band.

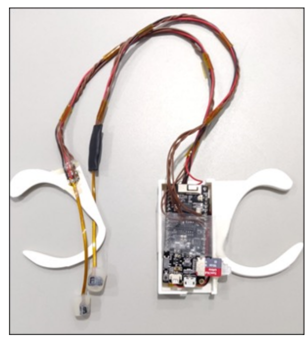


Fig. 7: Earable prototype with IMU and PPG in each earbud.

In addition to the CRT task, participants were also asked to complete a 2-back task to assess their working memory [17]. In this task, a sequence of letters appeared briefly at the center of the screen for 0.5 seconds, with a 2-second gap between letters (see Figure 5). Participants were asked to respond with the 'm' key on their keyboard if the current letter was the same as the one that appeared two letters before it. If not, they were asked to withhold their response and not press any key. Each round of fatigability measurement consisted of 20 trials of the 2-back task, lasting approximately 50 seconds with target trials (which required a response) occurring four times.

3.5 Physiological Measurements

A range of physiological signals were recorded throughout the experiment sessions using four different wearable devices (see Figure 6): (i) our earable prototype with inertial measurement units and photoplethysmographic sensors in each earbud, developed based on Nokia eSense [2, 18], (ii) a Muse S EEG headband [3], (iii) a Zephyr BioHarness 3.0 chestband [4], and (iv) an Empatica E4 wristband [1]. Table 1 has a detailed description of the sensors on each device.

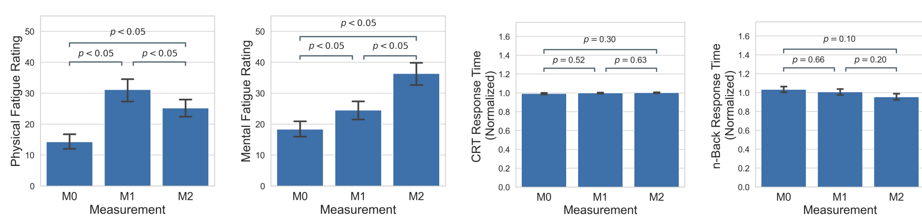
Table 1: Sensor data collected from each wearable device.

Sensor	Units/Range	Sampling Rate
Earable prototype		
Accelerometer	g {-2:+2}	100 Hz
Gyroscope	°/s {-500:+500}	100 Hz
PPG - green, infrared, and red channels	-	100 Hz
Muse S EEG headband		
Accelerometer	g {-2:+2}	52 Hz
Gyroscope	°/s {-245:+245}	52 Hz
EEG raw waveform	uV {0.0:1682.815}	256 Hz
EEG absolute band power (alpha, beta, delta, gamma, theta bands)	Bels	10 Hz
Zephyr BioHarness 3.0 chest band		
Accelerometer	bits {0-4094}	100 Hz
Breathing sensor raw output	bits {1:16777215}	25 Hz
Breathing rate	breaths per minute {4:70}	1 Hz
Breath-to-breath interval	ms	-
ECG raw waveform	bits {0:4095}	250 Hz
Heart rate	beats per minute {25:240}	1 Hz
Heart rate variability	ms {0:65534}	1 Hz
RR interval	ms {0:32767}	-
Posture	degrees from vertical {-180:180}	1 Hz
Empatica E4 wristband		
Accelerometer	g {-2:+2}	32 Hz
Blood volume pulse	-	64Hz
Average heart rate	1 Hz	
Inter-beat interval	ms	-
Electrodermal activity	microsiemens	4 Hz
Skin temperature	C	4 Hz

4 Preliminary Results

The collected dataset consisted of 36 sessions – twelve sessions each of low, medium, and high physical activity – with a total duration of almost 13 hours of physiological and behavioral recordings. The average duration of each recording was 21.24 minutes (SD: 3.23 minutes). No significant difference in session length was observed across the physical activity conditions ($F = 1.57$, $p = 0.24$).

4.1 Fatigue and Fatigability Measurements



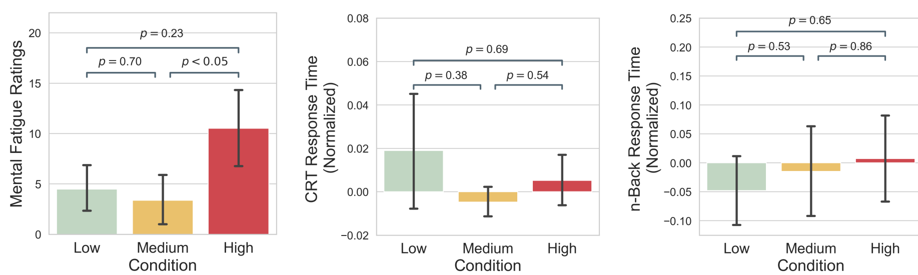
(a) Physical fatigue rating (b) Mental fatigue rating (c) Response time on CRT task (d) Response time on n-Back task

Fig. 8: Fatigue and fatigability measures at baseline ($M1$), after physical activity ($M2$), and after cognitive task ($M3$).

We first investigated the overall difference between fatigue and fatigability measurements at different stages – at baseline ($M0$), after physical activity ($M1$), and after cognitive task ($M2$) – pooling all experimental conditions together. As shown in Figure 8a, physical fatigue ratings increased significantly following the treadmill activity and decreased following the cognitive task. This is expected since the cognitive task was completed while participants were seated, allowing them to use this extended period of seating to recover from the physical activity session. Mental fatigue ratings showed a small increase following physical activity, and a larger increase after the cognitive task (see Figure 8b). We failed to observe a significant difference in response times measured at different points during the experimental session for either the CRT or the n-back task (see Figures 8c and 8d). The above analysis shows that the study design was able to successfully induce physical and mental fatigue, but significant mental fatigability was not observed when not accounting for physical activity conditions.

We also found no significant correlation ($p > 0.05$) between fatigue scores and response times on either the CRT or n-back tasks, indicating that participants' perception of fatigue did not correspond to their objective cognitive performance.

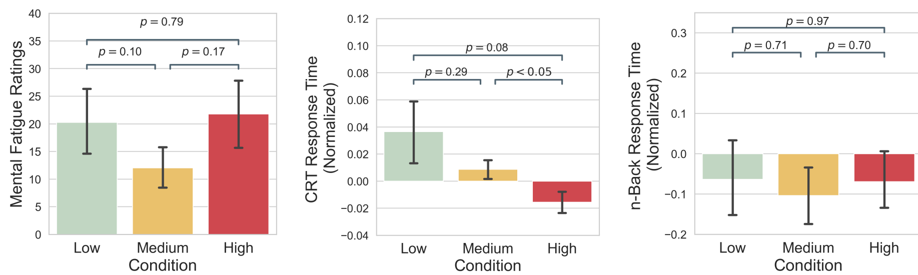
4.2 Effect of Physical Activity on Fatigue and Fatigability



(a) Mental fatigue rating (b) CRT response time (c) n-Back response time

Fig. 9: Mental fatigue and fatigability after physical activity ($M2 - M1$).

Next, we explore the effect of the level of physical activity on mental fatigue and fatigability. To this end, we first look into mental fatigue and fatigability after participants completed physical activity on the treadmill by calculating the difference between $M2$ and the baseline measurement $M1$ for each activity condition. As shown in Figure 9a, there was a positive trend in mental fatigue ratings across all conditions, but the increase in self-reported fatigue was significantly higher during the "High" physical activity condition compared to the "Medium" condition. In terms of fatigability, average response times on the CRT task increased during "Low" and "High" intensity activity and decreased during the "Medium" condition, but differences across conditions were not significant (see Fig. 9b). For the n-back task, response times decreased during "Low" and "Medium" conditions and increased slightly during "High" activity, though no significant differences were observed across conditions (see Fig. 9c). “Medium”-level activity was associated with both the least increase in subjective fatigue and slight improvements in performance on both cognitive tasks following physical activity.



(a) Mental fatigue rating (b) CRT response time (c) n-Back response time

Fig. 10: Mental fatigue and fatigability after cognitive activity ($M3 - M1$).

We also investigated mental fatigue and fatigability following the subsequent cognitive task (difference between $M3$ and $M1$). We found that all physical activity conditions were associated with an increase in fatigue ratings after the dual cognitive task, though the difference between conditions was not found to be significant (Figure 10a). In terms of fatigability, “High” physical activity exhibited a significant decline in CRT response times compared to the other two conditions (see Figure 10b). No significant differences between conditions were found on the more cognitively-demanding n-back task (Figure 10c), where participants may have overcome performance declines by expending more effort.

5 Conclusion

In this work, we present FatigueSet, a multi-modal dataset for understanding the impact of physical and cognitive activity on the development of mental fatigue and fatigability. Based on a preliminary analysis of experimental data recorded from twelve participants over 36 sessions, we show that cognitive performance and fatigability are poorly associated with individuals’ perception of fatigue. This demonstrates the need to independently consider experiential and behavioral dimensions of fatigue while developing fatigue-aware applications. Our analysis also reveals a difference in mental fatigue and fatigability across different physical activity conditions, illustrating the importance of accounting for the interplay between physical and mental fatigue. We hypothesize that these goals can be achieved by taking into account physiological correlates of fatigue and fatigability. Our publicly available dataset is an effort in this direction, and contains EEG, ECG, PPG, EDA, skin temperature, accelerometer, and gyroscope data from four devices at different on-body locations to facilitate a deeper understanding of mental fatigue and fatigability in daily life.

References

1. e4 wristband. https://www.empatica.com/research/e4/
2. esense overview. https://www.esense.io/
3. Introducing muse s. https://choosemuse.com/muse-s/
4. Zephyr^{TM} performance systems. https://www.zephyranywhere.com/
5. Fatigue (Jul 2021), https://www.ccohs.ca//oshanswers/psychosocial/fatigue.html
6. Ahlstrom, C., Nyström, M., Holmqvist, K., Fors, C., Sandberg, D., Anund, A., Kecklund, G., Åkerstedt, T.: Fit-for-duty test for estimation of drivers' sleepiness level: Eye movements improve the sleep/wake predictor. Transportation research part C: emerging technologies **26**, 20–32 (2013)
7. Borg, G.: Perceived exertion as an indicator of somatic stress. Scandinavian journal of rehabilitation medicine (1970)
8. Chmura, J., Nazar, K., Kaciuba-Uścilko, H.: Choice reaction time during graded exercise in relation to blood lactate and plasma catecholamine thresholds. International journal of sports medicine **15**(04), 172–176 (1994)
9. Craig, A., Tran, Y., Wijesuriya, N., Nguyen, H.: Regional brain wave activity changes associated with fatigue. Psychophysiology **49**(4), 574–582 (2012)
10. Deary, I.J., Liewald, D., Nissan, J.: A free, easy-to-use, computer-based simple and four-choice reaction time programme: the deary-liewald reaction time task. Behavior research methods **43**(1), 258–268 (2011)
11. Dittner, A.J., Wessely, S.C., Brown, R.G.: The assessment of fatigue: a practical guide for clinicians and researchers. Journal of psychosomatic research **56**(2), 157–170 (2004)
12. Elshafei, M., Shihab, E.: Towards detecting biceps muscle fatigue in gym activity using wearables. Sensors **21**(3), 759 (2021)
13. Gjoreski, M., Kolenik, T., Knez, T., Luštrek, M., Gams, M., Gjoreski, H., Pejović, V.: Datasets for cognitive load inference using wearable sensors and psychological traits. Applied Sciences **10**(11), 3843 (2020)
14. Häkkinen, K.: Neuromuscular fatigue and recovery in male and female athletes during heavy resistance exercise. International journal of sports medicine **14**(02), 53–59 (1993)
15. Janveja, I., Nambi, A., Bannur, S., Gupta, S., Padmanabhan, V.: Insight: monitoring the state of the driver in low-light using smartphones. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **4**(3), 1–29 (2020)
16. Jetté, M., Sidney, K., Blümchen, G.: Metabolic equivalents (mets) in exercise testing, exercise prescription, and evaluation of functional capacity. Clinical cardiology **13**(8), 555–565 (1990)
17. Kane, M.J., Conway, A.R., Miura, T.K., Colflesh, G.J.: Working memory, attention control, and the n-back task: a question of construct validity. Journal of Experimental Psychology: Learning, Memory, and Cognition **33**(3), 615 (2007)
18. Kawsar, F., Min, C., Mathur, A., Montanari, A.: Earables for personal-scale behavior analytics. IEEE Pervasive Computing **17**(3), 83–89 (2018)
19. Kluger, B.M., Krupp, L.B., Enoka, R.M.: Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. Neurology **80**(4), 409–416 (2013)
20. Krausman, A.S., Crowell III, H.P., Wilson, R.M.: The effects of physical exertion on cognitive performance. Tech. rep., ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD (2002)
21. Krupp, L.B., LaRocca, N.G., Muir-Nash, J., Steinberg, A.D.: The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. Archives of neurology **46**(10), 1121–1123 (1989)

22. Levitt, S., Gutin, B.: Multiple choice reaction time and movement time during physical exertion. Research Quarterly. American Association for Health, Physical Education and Recreation **42**(4), 405–410 (1971)
23. Li, F., Chen, C.H., Xu, G., Khoo, L.P., Liu, Y.: Proactive mental fatigue detection of traffic control operators using bagged trees and gaze-bin analysis. Advanced Engineering Informatics **42**, 100987 (2019)
24. Lorist, M.M., Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G., Meijman, T.F.: Mental fatigue and task control: planning and preparation. Psychophysiology **37**(5), 614–625 (2000)
25. Luo, H., Lee, P.A., Clay, I., Jaggi, M., De Luca, V.: Assessment of fatigue using wearable sensors: A pilot study. Digital biomarkers **4**(1), 59–72 (2020)
26. Maman, Z.S., Yazdi, M.A.A., Cavuoto, L.A., Megahed, F.M.: A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. Applied ergonomics **65**, 515–529 (2017)
27. Marcora, S.M., Staiano, W., Manning, V.: Mental fatigue impairs physical performance in humans. Journal of applied physiology **106**(3), 857–864 (2009)
28. Monk, T.H.: A visual analogue scale technique to measure global vigor and affect. Psychiatry research **27**(1), 89–99 (1989)
29. Norton, K., Norton, L., Sadgrove, D.: Position statement on physical activity and exercise intensity terminology. Journal of science and medicine in sport **13**(5), 496–502 (2010)
30. Ortega, F.B., Ruiz, J.R., Espana-Romero, V., Vicente-Rodriguez, G., Martínez-Gómez, D., Manios, Y., Béghin, L., Molnar, D., Widhalm, K., Moreno, L.A., et al.: The international fitness scale (ifis): usefulness of self-reported fitness in youth. International journal of epidemiology **40**(3), 701–711 (2011)
31. O'Keeffe, K., Hodder, S., Lloyd, A.: A comparison of methods used for inducing mental fatigue in performance research: Individualised, dual-task and short duration cognitive tests are most effective. Ergonomics **63**(1), 1–12 (2020)
32. Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J.K.: Psychopy2: Experiments in behavior made easy. Behavior research methods **51**(1), 195–203 (2019)
33. Phillips, R.O.: A review of definitions of fatigue–and a step towards a whole definition. Transportation research part F: traffic psychology and behaviour **29**, 48–56 (2015)
34. Rammstedt, B., John, O.P.: Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. Journal of research in Personality **41**(1), 203–212 (2007)
35. Roenneberg, T., Wirz-Justice, A., Merrow, M.: Life between clocks: daily temporal patterns of human chronotypes. Journal of biological rhythms **18**(1), 80–90 (2003)
36. Rogers, R.D., Monsell, S.: Costs of a predictible switch between simple cognitive tasks. Journal of experimental psychology: General **124**(2), 207 (1995)
37. Shahid, A., Wilkinson, K., Marcu, S., Shapiro, C.M.: Stanford sleepiness scale (sss). In: STOP, THAT and one hundred other sleep scales, pp. 369–370. Springer (2011)
38. Shen, K.Q., Li, X.P., Ong, C.J., Shao, S.Y., Wilder-Smith, E.P.: Eeg-based mental fatigue measurement using multi-class support vector machines with confidence estimate. Clinical neurophysiology **119**(7), 1524–1533 (2008)
39. Silverman, M.N., Heim, C.M., Nater, U.M., Marques, A.H., Sternberg, E.M.: Neuroendocrine and immune contributors to fatigue. PM&R **2**(5), 338–346 (2010)
40. Smith, M.R., Chai, R., Nguyen, H.T., Marcora, S.M., Coutts, A.J.: Comparing the effects of three cognitive tasks on indicators of mental fatigue. The Journal of psychology **153**(8), 759–783 (2019)

41. Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., Roelands, B.: The effects of mental fatigue on physical performance: a systematic review. Sports medicine **47**(8), 1569–1588 (2017)
42. Van Dongen, H., Belenky, G., Krueger, J.M.: Investigating the temporal dynamics and underlying mechanisms of cognitive fatigue. (2011)
43. Wan, J.j., Qin, Z., Wang, P.y., Sun, Y., Liu, X.: Muscle fatigue: general understanding and treatment. Experimental & molecular medicine **49**(10), e384–e384 (2017)