

University of Groningen

## Mental Fatigue and Motivation

Herlambang, Mega Bagus

DOI:  
[10.33612/diss.135819347](https://doi.org/10.33612/diss.135819347)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Herlambang, M. B. (2020). *Mental Fatigue and Motivation: Effects and mechanisms*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen. <https://doi.org/10.33612/diss.135819347>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Chapter 3

---

## The Effects of Intrinsic Motivation on Mental Fatigue

M. B. Herlambang  
F. Cnossen  
N. A. Taatgen

*Manuscript submitted for publication*

## Abstract

There have been many studies attempting to disentangle the relation between motivation and mental fatigue. Mental fatigue occurs after performing a demanding task for a prolonged time, and many studies have suggested that motivation can counteract the negative effects of mental fatigue on task performance. To complicate matters, most mental fatigue studies looked exclusively at the effects of extrinsic motivation but not intrinsic motivation. Individuals are said to be extrinsically motivated when they perform a task to attain rewards and avoid punishments, while they are said to be intrinsically motivated when they do for the pleasure of doing the activity.

To assess whether intrinsic motivation has similar effects as extrinsic motivation, we conducted an experiment using subjective, performance, and physiological measures (heart rate variability and pupillometry). In this experiment, 28 participants solved Sudoku puzzles on a computer for three hours, with a cat video playing in the corner of the screen. The experiment consisted of 14 blocks with two alternating conditions: low intrinsic motivation and high intrinsic motivation.

The main results showed that irrespective of condition, participants reported becoming fatigued over time. They performed better, invested more mental effort physiologically, and were less distracted in high-level than in low-level motivation blocks. The results suggest that similarly to extrinsic motivation, time-on-task effects are modulated by the level of intrinsic motivation: With high intrinsic motivation, people can maintain their performance over time as they seem willing to invest more effort as time progresses than in low intrinsic motivation.

**Keywords:** mental fatigue, intrinsic motivation, time-on-task, effort, heart rate variability

### 3.1. Introduction

Mental fatigue is a subjective feeling and a psychobiological condition after doing a demanding cognitive task for a long time (Gergelyfi, Jacob, Olivier, & Zenon, 2015; Marcora, Staiano, & Manning, 2009; van der Linden et al., 2003). Typically, when an individual is mentally fatigued, performance levels, e.g., in sustained attention and working memory tasks, decline (Helton & Russell, 2015, 2017; Lorist et al., 2000; Warm, Parasuraman, & Matthews, 2008), and the inclination to look for more rewarding activities increases (Boksem & Tops, 2008; Herlambang, Taatgen, & Cnossen, 2019; Hockey, 2011; Hopstaken, van der Linden, Bakker, & Kompier, 2015). In addition, prolonged mental fatigue impairs physical performance (Brown et al., 2020; Marcora et al., 2009; van Cutsem et al., 2017) and leads to safety issues, such as traffic accidents and errors in the workplace (Mizuno et al., 2011; Ting, Hwang, Doong, & Jeng, 2008; van der Linden, 2011).

In general, there are two theories for explaining the effects of mental fatigue: declining resources and lack of motivation (Helton & Russell, 2017; Herlambang et al., 2019). The first theory suggests that finite resources and failure in allocating resources cause performance to decrease. Support for this theory is found in an experiment where doing a demanding task for a long time suppressed the brain activity (Helton et al., 2007). Moreover, the resource theory has been the prime explanation for performance decrement especially in vigilance tasks (Craig & Klein, 2019).

In contrast, the second theory suggests that impaired performance is caused by amotivation, which is the lack of desire to continue doing an activity (Boksem & Tops, 2008; Earle, Hockey, Earle, & Clough, 2015). In most cases, performing a cognitively demanding task for a long time increases the feeling of fatigue. As it increases, one will be less willing to stay engaged with the task, i.e., less motivated to continue performing the task (Müller and Apps, 2019). As a result, a lower level of motivation impairs performance (Hockey, 2011; Müller and Apps, 2019). However, prolonged tasks do not necessarily lower performance if the individual who performs the task is motivated: The individual is able to maintain or find another source of motivation. For instance, after doing a demanding task for a few hours, performance levels decreased but returned to the initial level when participants were offered external rewards at the end of the block (Hopstaken et al., 2015). A similar study by Boksem and colleagues (2006), where they offered participants a monetary reward 20 min before the experiment ended, showed that performance increased significantly during this last block. In our own study, where we manipulated rewards continuously for 2.5 hr and asked participants to do a demanding working memory task to count and calculate the total number of vowels, showed that in reward blocks, and in contrast to nonreward blocks, participants were able to maintain performance, invested more mental effort (subjectively and physiologically by using heart rate variability), and were less distracted (Herlambang et al., 2019). These results show a clear effect of motivation on mental fatigue, which are difficult to explain by the resource theory.

Motivation is different from, but related to, effort. Motivation refers to the “activating orientation of current life pursuits toward a positively evaluated goal state” (Rheinberg & Vollmeyer, 2018, p. 15). It drives an individual to perform a particular activity, and the individual may behave differently according to the level of motivation. Effort, on the other hand, refers to “the degree of engagement with demanding tasks” (Westbrook & Braver, 2015, p. 396) and reflects a feeling resulting from the cost/benefit calculation of doing a task that later determines performance (Kurzban, Duckworth, Kable, & Myers, 2013). For instance, when one feels fatigued, which corresponds to a high level of perceived effort (Brown et al., 2020; van Cutsem et al., 2017), one will choose to continue doing the task, maintaining performance, or let one’s performance level drop based on the result of the calculation (Boksem & Tops, 2008; Kurzban et al., 2013).

Furthermore, the effects of a highly motivated individual on his/her performance may be mediated by increased effort. A study by Gendolla and colleagues (2012) showed that the level of effort correlated with the subjective difficulty of the task according to its feasibility and reasonability. More specifically, if performing a task is viewed as beneficial and viable, the more difficult the task is, the higher the effort will be (Wright, Mlynski, & Carbajal, 2019). Their notion came from the motivation intensity theory (Brehm & Self, 1989), which posits that human beings attempt to avoid using unnecessary resources and have what is called potential motivation, which is the level of motivation that determines how much effort an individual is willing to invest based on the difficulty, feasibility, and benefits of performing the task, which resembles the cost/benefit calculation (Kurzban, et al., 2013).

With regard to the motivational theory of mental fatigue, broadly, there are two types of motivation: extrinsic and intrinsic motivation (Deci & Ryan, 2008; Di Domenico & Ryan, 2017). Ryan and Deci (2000) defined the former in their self-determination theory (SDT) as a type of motivation to attain distinct outcomes (i.e., to attain rewards or to avoid punishments). Two examples are an employee who works overnight to get overtime payments, and an engineer who works hard so that his or her family does not suffer. On the other hand, intrinsic motivation is defined as a type of motivation to do an activity because of the inherent enjoyment of the activity rather than to attain distinct outcomes. Moreover, adopted as part of the theory, DeCharms (1968) pointed out that intrinsic motivation reflects a psychological need for competence and autonomy. The first refers to the sense of acquiring skills in activities that are optimally challenging, whereas autonomy refers to the sense that the behavior is authentic rather than internally intimidated or externally compelled. When these two needs are supported, intrinsic motivation may last, but when these needs are not satisfied, the intrinsic motivation is undermined. Furthermore, with regards to fatigue, Hockey (2011), in his motivational control theory, states that self-initiated activities or tasks (i.e., with high intrinsic motivation) are unlikely to cause mental fatigue.

Intrinsic motivation, however, is often confounded with another type of motivation, namely, achievement motivation (Locke & Schattke, 2018). In their notion, they divide motivation into three types: intrinsic, achievement, and extrinsic.

They argue that intrinsic motivation should be viewed as a type of motivation that arises from the enjoyment of doing an activity per se, and that it should be separated from other motives, e.g., to attain skills, to make progress, or to meet some quality standards. They emphasize that achievement motivation does not focus on the pleasure of the activity itself but of achieving new standards and keeping progress; for example, a doctor who wants to become even better at diagnosing patients. Therefore, to avoid confusion, a study of intrinsic motivation should solely focus on the enjoyment of doing the activity and attempt not to mix it with achievement motives such as competition (Ryan & Moller, 2017).

Even though the effects of intrinsic motivation in prolonged tasks are evident in daily life, e.g., game players can play computer games for many hours regularly (Hainey, Connolly, Stansfield, & Boyle, 2011), we have not found any studies aimed at finding connections between mental fatigue and intrinsic motivation.

To investigate the links between intrinsic motivation and mental fatigue, we performed an experiment where we asked participants who liked Sudoku to solve Sudoku puzzles on a computer screen in two alternating conditions: low-level motivation (LL) and high-level motivation (HL). We hypothesized that if intrinsic motivation were an important component in mental fatigue, participants would be able to maintain performance and attention to the task in HL conditions over time. More specifically, individuals who like doing an activity because of the sense of satisfaction they receive from the activity itself would not show any effects of mental fatigue. On the other hand, in LL conditions, where the Sudoku involved less enjoyment, we predicted that performance would decline and be susceptible to distractions over time.

## **3.2. Method**

### **3.2.1. Participants**

Prior to the experiment, the sample size was calculated using G-power statistical software analysis (Faul, Erdfelder, Lang, & Buchner, 2007). The experiment was within-subject and designed to have a power of .90 (type II error = .10), a significance level of .05, and a large effect-size ( $d = .80$ ) (based on a similar experiment from Herlambang et al., 2019). Therefore, the required sample size using these parameters was 19. To avoid problems with very small sample sizes, we decided to aim for a slightly larger sample of 30.

Thirty-two healthy university students joined the experiment and received monetary reward for participating. Four participants gave up during the experiment. Heart rate data were lost due to equipment problems in two participants. Therefore, the final sample consisted of 28 participants (17 male; mean age = 24.57 years,  $SD = 4.21$ ), and 26 for heart rate variability (HRV) analysis. All participants included in the study gave written informed consent.

### 3.2.2. Procedure

University students who liked playing Sudoku were invited to join our experiment. After registration, we asked how often they played Sudoku. If they indicated having played Sudoku at least five times within the last month, we asked them to solve a difficult Sudoku puzzle within six hours, which they received by email. If they did, they were eligible to participate.

A few days before the experiment, all participants received another email informing them of the details of the study. The email did not mention mental fatigue. It asked participants not to drink coffee 24 hours, not to consume heavy meals or perform any exercise an hour before the experiment started. Also, participants were required to have enough sleep. The email stated that participants were not allowed to participate if they had heart abnormalities.

On the day of the experiment, participants were seated 60 cm in front of an LCD monitor. They were asked to attach the heart rate monitor on their chest. In the case of chest hair, we asked them to shave it in order to attach the heart rate monitor properly. Afterward, we asked them to rest for five minutes and checked their resting heart rate: if their resting heart rate was normal (Palatini, 1999), they were allowed to proceed to the next step.

Next, they were asked to hand over their wristwatches, turn off their phones, and sign an informed consent form. Afterward, they were requested to put their chin on a chin rest of the eye-tracker, and we performed calibration and drift correction before the experiment started. During the experiment, participants were not allowed to move, except when they felt tired, but they had to remain seated in the chair.

Before the main experiment, participants performed a practice session to familiarize them with the Sudoku until they were ready to proceed (max. 15 min). On the left side of the table, we put 14 sheets of subjective ratings to be filled in every time a block ended and asked them to put it on their right side when finished with a block.

After the experiment finished or when participants decided to give up, they had a debriefing session, in which we explained the purpose of the study.

### 3.2.3. Task

Participants were asked to play Sudoku puzzles on a computer screen continuously for three hours without rests. The experiment consisted of 14 blocks of two alternating conditions: low-level (LL) intrinsic motivation in the odd blocks and high-level (HL) intrinsic motivation in the even blocks. The duration for each block was 13 min. To indicate the two different conditions, the Sudoku puzzles in the LL blocks were drawn in green, whereas in the HL blocks, they were drawn in black. After a block ended, the computer screen proceeded to a subjective-ratings screen (**Figure 1**).

The Sudoku puzzles had 3 x 3 boxes, and each box consisted of 3 x 3 cells; therefore, the total number of cells was 81. At the start of the puzzle, a number of cells already had numbers in them between one and nine. The task of the participants was to fill in the remaining cells with numbers. Each box should include all the numbers from one to nine, and each number should only occur once in a row and once in a column.

To solve the Sudoku, participants were provided with a mouse. Participants had to click on an empty cells in the Sudoku and were then shown several buttons of numbers from one to nine (from left to right) positioned in the center of the screen, and then pressed one of the numbers. If the answer was incorrect, participants would hear a 'beep' sound, and the cell would remain empty. If the answer was correct, there was no sound, and the cell would be filled in with the chosen number.

In the LL condition (the odd blocks), a new Sudoku puzzle was generated every five trials, that is, after five clicks within the Sudoku box. After 13 min had elapsed, participants had to fill in three subjective rating scales on a sheet of paper for 20 s. Participants then continued with a new block. In HL blocks, if participants could not complete a Sudoku puzzle after 13 min, they continued solving the same Sudoku puzzle in the next HL block. In contrast, if participants were able to finish a Sudoku puzzle in a block, the algorithm would generate a new Sudoku puzzle immediately. This design assumed that being able to finish a complete puzzle was more motivating than just solving a few steps. It allowed for more long-term planning and, therefore, a more motivating mental investment in the task.

### 3.2.4. Materials

We used the Eyelink Duo from SR Research to obtain pupil diameter, eyeblinks, gaze positions, and saccades with a sample rate of 250 Hz by measuring participants' dominant eye during the experiment. To measure heart rate, we used the Cortrium C3 holter monitor. The device has been tested and is valid and reliable to be used as an instrument to measure heart rate (Kunkels, van Roon, Wichers, & Riese, 2019; Lindhart, 2018).

For subjective measures, we printed the Rating Scale Mental Effort (RSME), the National Aeronautics and Space Administration Task Load Index (NASA-TLX), and the Visual Analog Scale (VAS) on a two-sided page with RSME as the first measure on the first page followed by NASA-TLX and VAS as the second and third measure on the second page.

The display used in the experiment was a 19-inch square LCD monitor. We played Simon's cat video, a black-and-white animation of a cat, continuously with a resolution of 320 x 180 in the top right of the screen as a distractor. We were granted to use the video by Simon's Cat Ltd.

To present the Sudoku puzzles to participants, we used OpenSesame with a resolution of 1,280 x 960 (Mathôt, Schreij, & Theeuwes, 2012) and PyGaze (Dalmaijer, Mathôt, & van der Stigchel, 2014) to give commands to the eye-tracker.



The puzzles were designed to be equally difficult in both conditions. To generate a new puzzle, we used constraint propagation and searching algorithms (Norvig, n.d.) and filled in 26 cells randomly (see **Figure 1**). All puzzles were solvable, regardless of condition. The difficulty to solve the puzzles was chosen as moderate.

### **3.2.5. Measures**

#### **3.2.5.1. Subjective measures**

We used VAS to measure the subjective feeling of fatigue. This instrument has high validity and reliability to measure fatigue (Mizuno et al., 2011). In addition, we used RSME to measure mental effort (Zijlstra & van Doorn, 1985), and NASA-TLX to measure physical load, mental load, temporal load, frustration level, performance and effort (Galy, Paxion, & Berthelon, 2017; Hart & Staveland, 1988). Both RSME and NASA-TLX are reliable and valid to be used as subjective assessments of effort and mental workload respectively (Hill et al., 1992; Longo, 2018).

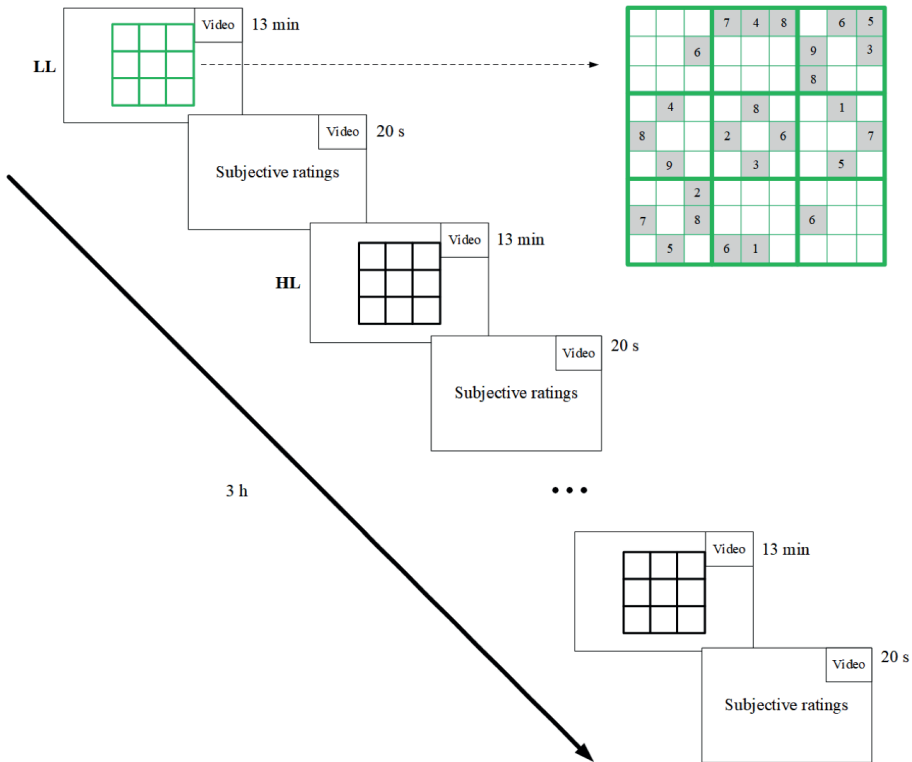
For the first seven participants, we used NASA-TLX as a measure of workload and effort. Later, we added RSME due to lack of sensitivity of the effort scale of the NASA-TLX. Therefore, RSME data were complete for only 21 participants.

#### **3.2.5.2. Performance measures**

We measured reaction time (RT) as the time between mouse clicks within the Sudoku puzzle. A click outside the Sudoku was not considered a response, and RT was not recorded. We excluded the response time of the first click of each Sudoku puzzle from the analysis because of its high variability. In addition, we measured the number of clicks for each block. To measure accuracy, we expressed it as the percentage of correct clicks for each block.

#### **3.2.5.3. Physiological measures**

Heart rate variability (HRV) is the variability between consecutive heartbeats and reflects how individuals react to environmental and internal changes (Acharya et al., 2006; Berntson et al., 1997). HRV provides information on individuals' autonomous nervous system over time (Evans et al., 2013; Kang et al., 2016). Moreover, the mid-frequency (MF) band of HRV (0.07–0.14 Hz) is commonly used as a measure of mental effort (Aasman, Mulder, & Mulder, 1987; Mulder & Mulder, 1981). In this experiment, participants were predicted to exert mental effort differently in LL and HL conditions, and an increase in effort is suggested to correlate with high motivation (Wright et al., 2019). Therefore, measuring participants' mental effort using the MF band of HRV could provide information on how participants performed and responded to the experimental manipulation over time.



**Figure 1.** The flow of the experiment. The LL Sudoku (LL) is indicated by green edges. Initial filled-in cells had a gray-colored background. After a block ends, it continues to a screen which asks participants to fill in the subjective ratings. A video distractor is played continuously in the top right of the screen.

Raw data of the heart rate signal from the Cortrium C3 were preprocessed using PreCAR to detect and correct R-peak artefacts. Afterward, we used CARSPAN (Mulder, Hofstetter, & van Roon, 2009) to determine heart rate variability in the MF band. Power data for each block were normalized by dividing the power of each block by the average power across the experiment.

Pupillometry used in this study consisted of several measures: pupil diameter, eyeblinks, gaze positions, and saccades. We obtained raw data of the pupillometry from Eyelink Duo. Next, we used EDF2ASC (a software package from SR Research) to convert the raw data to ASCII format and used Eyelinker (Barthelme, 2016) to convert ASCII format to a more structured format to be analyzed in R (Version 3.4.2). We filtered all pupillometry data from the start of each block to the end of the block.

Pupil diameter is commonly used to measure cognitive load and control (Karatekin, 2004). In addition, a recent study of mental fatigue used pupil dilation to measure task engagement in which the pupil dilates when participants re-engaged with a task motivated by extrinsic rewards (Hopstaken et al., 2015). In this experiment,

we measured the exploitation-exploration effect (the engagement-disengagement effect) on pupil dilation in LL and HL conditions. We normalized pupil diameter for each block by dividing the average of that block by the average of the entire experiment.

Eyeblinks have been used as a measure of fatigue and workload (Martins & Carvalho, 2015), and we used eyeblinks to measure mental fatigue in this experiment. We predicted that eyeblinks would increase over time regardless of conditions in the experiment. For data analysis, we calculated the mean of eyeblink frequency and eyeblink duration.

To measure distractibility by a video distractor, i.e., Simon's cat video, we used eye gaze positions. We hypothesized that a decrease in motivation leads to more distractions (Hockey, 2013; Hopstaken et al., 2015). We predicted that if the intrinsic motivation were essential to keep participants engaged (i.e., stay motivated) with the Sudoku puzzles, they would be less distracted in HL conditions than in LL conditions. Each time the coordinate of eye gaze was within the cat video for at least 200 ms, we used it as an instance of visual distraction. For each block, we calculated the mean of visual distraction frequency and visual distraction duration.

To measure attention to the task, we used in-task eye saccades, i.e., when the starting and ending point of saccades were within the Sudoku puzzle. We assumed that participants made saccades movement frequently to solve Sudoku puzzles, searching for the right number for the right cell. Afterward, we calculated the mean of saccades frequency and saccades amplitude.

### 3.2.6. Statistical analysis

We used linear mixed-effects models for all measures by using Lme4 package (Bates, Mächler, Bolker, & Walker, 2015) of R (Version 3.4.2). For visual distraction frequency analysis, we applied log-transformation because the dataset was not normally distributed. Moreover, we used the Car package in R to obtain *p* values (Fox & Weisberg, 2011).

To determine the best fitting model, we compared Akaike criteria from the simplest model to more complex models and used the function *anova* in R. First, we compared time-on-task with condition; both are fixed effects in the models. All models used participants as the random effect. Next, we compared the chosen model with a model of interaction between time-on-task and condition. In addition, we examined the residuals and fitted values to comply with the assumption of constant variance.

## 3.3. Results

Tables show the best-fitted model in each measure, regardless of significance.

### 3.3.1. Subjective measures

#### 3.3.1.1. Fatigue

To check our fatigue manipulation, we used the Visual Analog Scale (VAS). Time-on-task had a significant effect on the fatigue score (**Table 1**), which increased linearly from the first to the last block (see **Figure 2A**). Including condition as a fixed effect did not improve the model.

**Table 1.** The mixed-effect result of the subjective measure of fatigue and effort from the best fitted model

Measure	Variable	Mean	Standard error	p Value	95% CI	
					Lower limit	Upper limit
Fatigue	Intercept	3.11	0.41			
	Time	0.26	0.01	< .001	0.23	0.29
Effort	Intercept	72.27	4.55			
	Time	0.67	0.21	< .01	0.26	1.07
	Condition	-6.37	1.66	< .001	-9.63	-3.11

#### 3.3.1.2. Effort

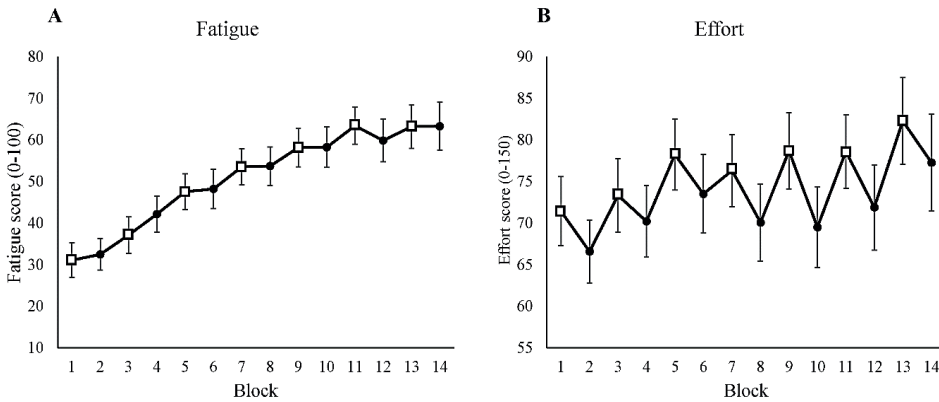
To measure subjective mental effort, we used the Rating Scale Mental Effort (RSME). **Table 1** shows that both time-on-task and condition had a significant effect on subjective mental effort. **Figure 2B** shows that participants reported higher ratings in low-level motivation (LL) blocks, and that subjective mental effort increased over time. Including the interaction as a fixed effect did not improve the model. In addition, the RSME score showed a significant correlation with the frustration scale of NASA-TLX  $r(12) = .86, p < .01$ .

#### 3.3.1.3. NASA-TLX

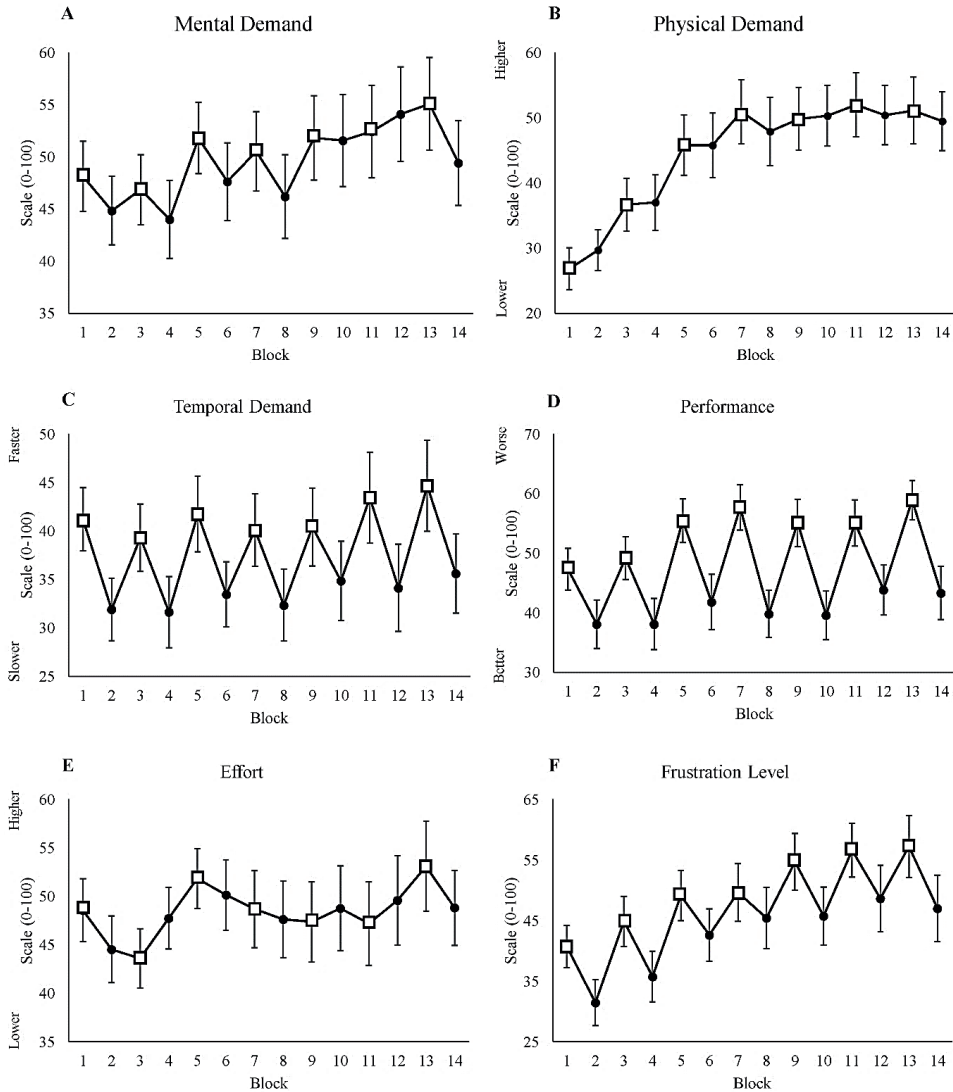
**Table 2** shows that time-on-task had a significant effect on mental demand, physical demand, performance, and frustration level. In addition, **Figures 3A, 3B, 3D, and 3F** show that all these dimensions increase over time. Moreover, the effect of condition on mental demand, temporal demand, performance, and frustration level was also significant. Participants reported higher frustration levels in LL blocks. However, we did not find any significant effect on the effort scale, which is not shown in the table.

**Table 2.** The mixed-effect results of NASA-TLX from the best fitted model

Scale	Variable	Mean	Standard error	p Value	95% CI	
					Lower limit	Upper limit
Mental Demand	Intercept	46.59	3.53			
	Time	0.65	0.13	< .001	-5.41	-1.23
	Condition	-3.32	1.06	< .01	0.36	0.88
Physical Demand	Intercept	31.53	4.15			
	Time	1.74	0.14	< .001	1.45	2.02
Temporal Demand	Intercept	41.55	3.15			
	Condition	-8.12	1.28	< .001	-10.64	-5.59
Performance	Intercept	49.65	3.01			
	Time	0.63	0.21	< .01	0.22	1.04
	Condition	-14.09	1.67	< .001	-17.38	-10.81
Frustration Level	Intercept	40.78	3.95			
	Time	1.37	0.18	< .001	1.01	1.73
	Condition	-9.41	1.48	< .001	-12.33	-6.48



**Figure 2.** The subjective measure of fatigue and effort. **(A)** Average fatigue scores for each block using the VAS. The y axis shows subjective fatigue scores from 0 to 100. **(B)** Average effort ratings for each block using the RSME. The y axis shows the subjective mental effort scores from 0 to 150 All x axes show blocks, where odd blocks represented by square markers are low motivation blocks. Error bars in each block represent standard errors.



**Figure 3.** NASA-TLX scales. **(A)** Average score of mental demand for each block. **(B)** Average score of physical demand for each block. **(C)** Average score of temporal demand for each block. **(D)** Average score of performance for each block. **(E)** Average score of subjective effort for each block. **(F)** Average score of frustration level for each block. All figures' y axes show the score for each scale from 0 to 100 and x axes show blocks, where odd blocks represented by square markers are the low motivation blocks. Error bars in each block represent standard errors.

3

### 3.3.2. Performance measures

#### 3.3.2.1. Response time

Response time decreased significantly over time (**Table 3**). In addition, RTs were significantly slower in LL blocks and faster in HL blocks (see **Figure 4A**). Including the interaction between time-on-task and condition did not improve the model.

**Table 3.** The mixed-effect result of performance measures from the best fitted model

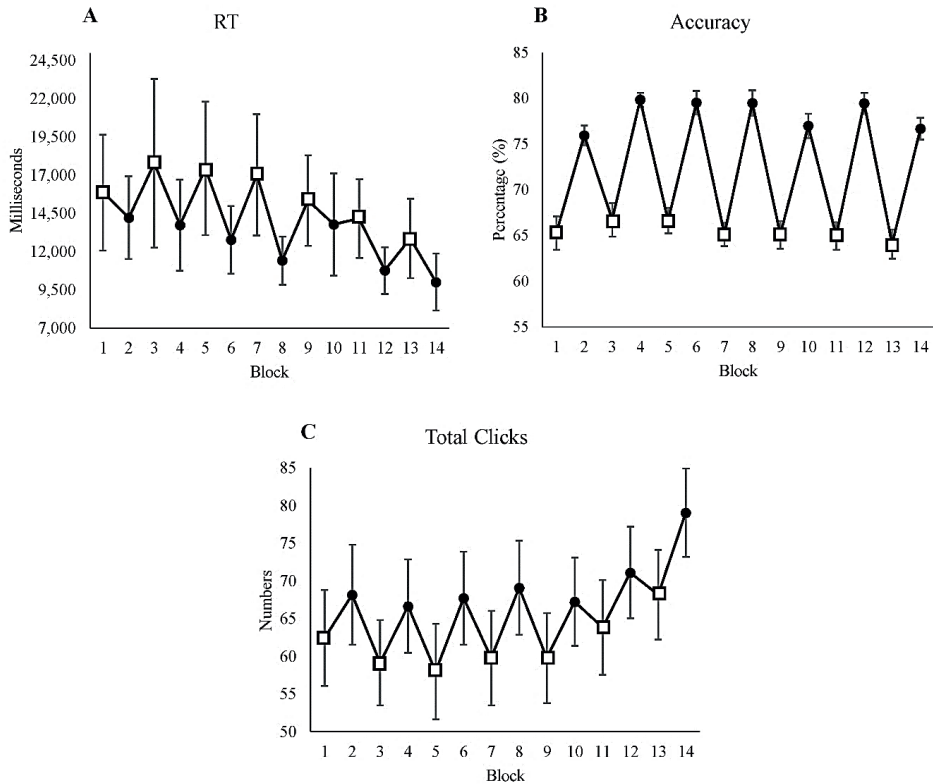
Measure	Variable	Mean	Standard error	<i>p</i> Value	95% CI	
					Lower limit	Upper limit
RT	Intercept	18,031.01	2,765.88			
	Time	-320.91	123.86	< .01	-564.31	-77.49
	Condition	-3,078.78	998.61	< .01	-5,041.19	-1,116.38
Accuracy	Intercept	65.41	0.87			
	Condition	12.88	0.62	< .001	11.64	14.11
Total clicks	Intercept	57.23	5.78			
	Time	0.62	0.14	< .001	0.32	0.91
	Condition	7.67	1.21	< .001	5.31	10.03

#### 3.3.2.2. Accuracy

We found a significant effect of condition on accuracy. Including time-on-task as a fixed effect did not improve the model (**Table 3**). Accuracy was lower in LL blocks and higher in HL blocks (see **Figure 4B**).

#### 3.3.2.3. Total clicks

This measure shows the number of clicks for each block regardless of accuracy. Time-on-task had a significant effect on total clicks, where participants made more clicks over time (**Table 3**). In addition, the effect of condition was significant, where participants made more clicks in the HL condition (see **Figure 4C**). Including the interaction did not improve the model.



**Figure 4.** Performance measures. **(A)** Average response time for each block. **(B)** Average accuracy for each block. **(C)** The number of clicks for each block. All figures' y axes show their value respectively, and x axes show blocks, where odd blocks represented by square markers are the low motivation blocks. Error bars in each block represent standard errors.

### 3.3.3. Physiological measures

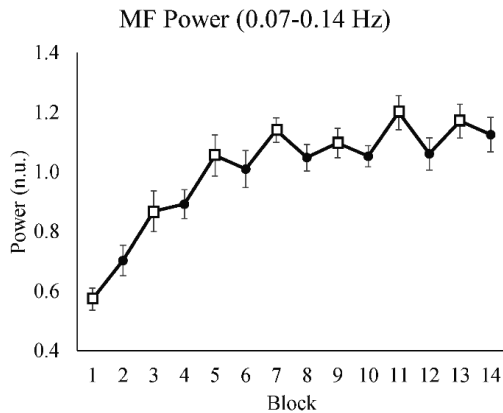
#### 3.3.3.1. HRV – MF band

To measure cognitive mental effort, we calculated the power in the MF band of HRV and expressed power in a block as a percentage of the average power of the whole experiment. Higher power in the MF band indicates that participants invested less mental effort (see Mulder & Mulder, 1981; Aasman et al., 1987). Conversely, lower power indicates the opposite. **Table 4** shows that time-on-task had a significant effect on MF power, which indicates that participants invested less mental effort over time. In addition, the interaction between time-on-task and condition was significant: The difference between the two conditions increased over time (see **Figure 5**).



**Table 4.** The mixed-effect result of MF power from the best fitted model

Variable	Mean	Standard error	p Value	95% CI	
				Lower limit	Upper limit
Intercept	0.71	0.04			
Time	0.04	0.01	< .001	0.03	0.05
Condition	0.04	0.06		-0.07	0.17
Time x Condition	-0.01	0.01	< .05	-0.03	-0.01

**Figure 5.** Average power of HRV in the MF band The y axis shows the normalized value of the MF power. The x axis shows blocks, where odd blocks represented by square markers are low motivation blocks. Error bars in each block represent standard errors.

### 3.3.3.2. Pupil diameter

To measure the involvement of cognitive control, we used pupil diameter. Time-on-task showed a significant effect on pupil diameter (see **Table 5**). **Figure 6A** shows that pupil diameter decreased over time, indicating lower cognitive control over time. Furthermore, the effect of condition on pupil diameter was significant, with the pupil diameter dilating more in HL blocks, indicating higher cognitive control. Including the interaction between time-on-task and condition did not improve the model.

### 3.3.3.3. Visual distraction frequency

We used visual distraction frequency to measure how often participants shifted their attention to the video distractor. Time-on-task had a significant effect on visual distraction frequency (see **Table 5**), where participants watched the cat video more often over time. In addition, time and condition had a significant interaction

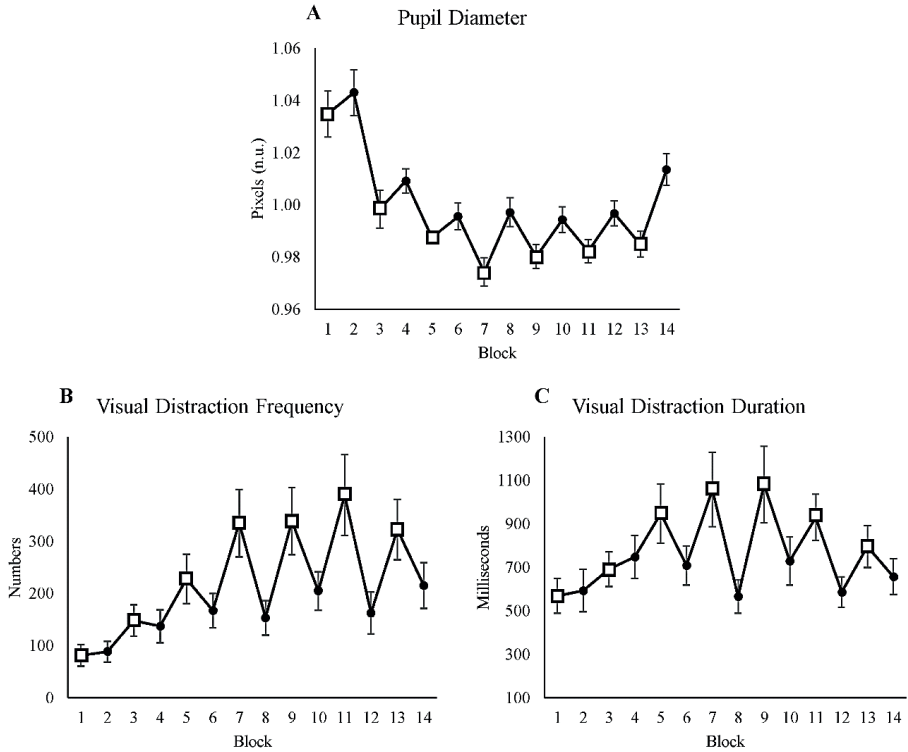
effect on visual distraction frequency: Over time, visual distraction frequency increased more in LL blocks than in HL blocks (see **Figure 6B**).

**Table 5.** The mixed-effect result of pupillometry measures from the best fitted model

Measure	Variables	Mean	Standard error	p Value	95% CI	
					Lower limit	Upper limit
Pupil diameter	Intercept	1.01	0.01			
	Time	-0.01	0.01	< .001	-0.02	-0.01
	Condition	0.02	0.01	< .001	0.01	0.03
Visual distraction frequency	Intercept	1.70	0.11			
	Time	0.05	0.01	< .001	0.04	0.07
	Condition	-0.01	0.09		-0.18	0.16
	Time x Condition	-0.03	0.01	< .01	-0.05	-0.01
Visual distraction duration	Intercept	706.16	92.02			
	Time	22.98	8.79	< .01	5.71	40.25
	Condition	-35.47	105.87		-243.52	172.57
	Time x Condition	-24.93	12.43	< .05	-49.35	-0.49
Eyeblink frequency	Intercept	181.78	27.92			
	Time	6.05	0.61	< .001	4.86	7.25
Eyeblink duration	Intercept	232.77	70.76			
	Time	4.69	2.37	< .05	0.03	9.35
Saccades frequency	Intercept	1,876.42	54.64			
	Time	-23.23	4.16	< .001	-31.41	-15.04
	Condition	-65.65	50.15		-164.22	32.91
	Time x Condition	19.18	5.89	< .01	7.61	30.76
Saccades amplitude	Intercept	3.19	0.07			
	Condition	-0.14	0.02	< .001	-0.18	-0.11

### 3.3.3.4. Visual distraction duration

Visual distraction duration measured how long participants watched the cat video. Time-on-task had a significant effect on visual distraction duration (see **Table 5**), where over time participants watched the cat video longer. Moreover, time-on-task and condition had a significant interaction effect on visual distraction duration. **Figure 6C** shows the interaction between time-on-task and condition. Visual distraction duration remained relatively stable in HL blocks, but it increased over time in LL blocks.



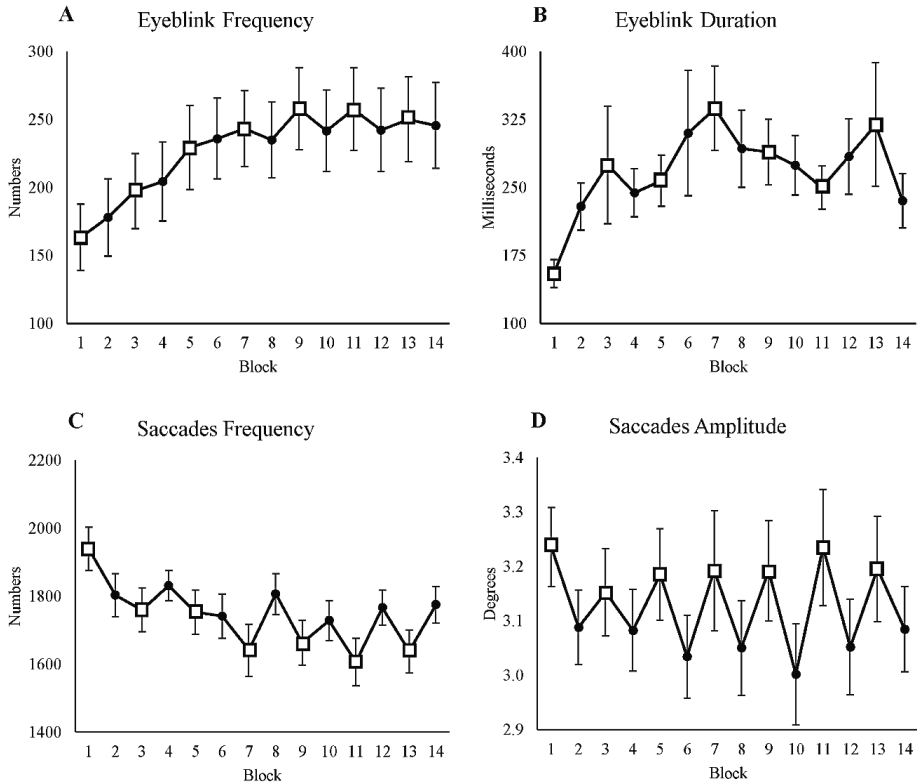
**Figure 6.** Pupil diameter and visual distraction. **(A)** Average pupil diameter for each block. **(B)** Average visual distraction frequency for each block. **(C)** Average visual distraction duration for each block. All figures' y axes show their value respectively, and x axes show blocks, where odd blocks represented by square markers are the low motivation blocks. Error bars in each block represent standard errors.

**3.3.3.5. Eyeblick frequency**

We used eyeblick frequency as an indicator of fatigue. **Table 5** shows that time-on-task had a significant effect on eyeblick frequency, with participants blinking more often over time (see **Figure 7A**). Including condition as a fixed effect did not improve the model.

**3.3.3.6. Eyeblick duration**

The purpose of eyeblick duration was similar to that of eyeblick frequency. Eyeblick duration significantly increased (see **Table 5** and **Figure 7B**); including condition as a fixed effect did not improve the model.



**Figure 7.** Eyeblinks and saccades. **(A)** Average eyeblink frequency for each block. **(B)** Average eyeblink duration for each block. **(C)** Average saccades frequency for each block. **(D)** Average saccades amplitude for each block. All figures' y axes show their value respectively, and x axes show blocks, where odd blocks represented by square markers are the low motivation blocks. Error bars in each block represent standard errors.

### 3.3.3.7. Saccades frequency

We used saccades frequency to measure participants' attention to solving Sudoku puzzles. Saccades frequency significantly decreased over time (see **Figure 7C**). In addition, time-on-task and condition had a significant interaction effect on saccades frequency: Over time, the difference between the two conditions increased (see **Table 5**).

### 3.3.3.8. Saccades amplitude

To measure participants' attention towards the task, we used saccades amplitude. Participants made significantly smaller saccade movements in HL blocks (see **Table 5** and **Figure 7D**), suggesting that participants were more careful in HL

blocks, searching for the right number for the right cell. Including time-on-task did not improve the model.

## **3.4. Discussion**

### **3.4.1. Hypothesis and results**

In this study, we investigated the effect of intrinsic motivation on mental fatigue by conducting a 3-hr experiment. We manipulated levels of motivation by including low-level motivation (LL) and high-level motivation (HL) blocks. We hypothesized that participants who liked playing Sudoku would not show effects of mental fatigue doing the task, particularly in HL blocks. To enable in-depth analyses, we used three types of measures: subjective, performance, and physiological.

The subjective measure of fatigue (i.e., VAS) showed that participants reported becoming fatigued over time regardless of condition, which suggested that our fatigue manipulation was successful. However, even though participants felt fatigued, they were able to maintain performance and attention in HL blocks.

In designing the task, we wanted to avoid that intrinsic motivation was confounded with achievement motivation (a sense of progress and motives to achieve particular standards (Locke & Schattke, 2018)). Therefore, we avoided instructions that would make participants perceive the game as a competition, or make them believe they needed to meet a particular standard: Participants performed the experiment freely, could give responses anytime without restriction, and were able to stop trying at any moment during the experiment. We reasoned that low-level motivation (LL) blocks lowered intrinsic motivation, even in people who liked playing Sudoku, since only being able to perform a few steps makes it impossible to do long-term planning. On the other hand, in high-level motivation (HL) blocks, the task was designed to be engaging and enjoyable, and participants should enjoy doing the task. This was confirmed by the lower frustration ratings on the NASA-TLX in HL blocks.

Several measures showed the effects of our manipulation of intrinsic motivation on performance. For instance, over time, accuracy remained at the same level in HL blocks (high motivation) but was lower in LL blocks (low motivation). The difference in the accuracy between the two conditions may lie in the fact that participants were more inclined to use a guessing strategy in the LL blocks. In contrast to paper-and-pencil Sudoku, where guessing is a suboptimal strategy, in our experiment, guessing could be beneficial to participants, as they received immediate feedback after choosing a number: Participants could still solve the puzzle with no increase in effort. Although accuracy was lower in LL blocks, it did not decrease over time. Therefore, the increase in subjectively reported mental fatigue (VAS) did not lead to decreased performance over time. However, we do see evidence of the effects of mental fatigue in other measures. In HL blocks, participants were less distracted (visual distraction frequency and visual distraction duration were lower), more attentive toward the task (saccades frequency was more frequent) and more

conscientious to solve the puzzles (saccades amplitude was smaller). In contrast, participants gradually became more susceptible to distractions and less attentive to the task in LL blocks. Also, the MF power of HRV suggests that participants invested more mental effort in HL blocks, whereas in LL blocks, they invested less mental effort. This result is in line with the motivation intensity theory (Brehm & Self, 1989). Because success was not possible in the LL blocks, the exerted effort in these blocks would be lower compared with HL blocks. On the other hand, the exerted effort was high in HL blocks, since success was possible and beneficial in these blocks (see Wright et al., 2019).

Regarding effort, RSME, the MF power of HRV, and the effort scale of NASA-TLX showed different effects. Participants reported higher RSME effort in LL blocks, but the power in the MF band physiologically suggests that participants exerted less mental effort in these conditions. A possible explanation for this difference may lie in participants' difficulty in rating subjective effort as a separate entity. In LL blocks, the Sudoku puzzles were less satisfying than in HL blocks because participants had to reorient themselves with a new puzzle every five trials in these blocks. Since they knew that the puzzles would be tedious, and they needed to stay alert until the experiment ended, they may have rated this as increased effort on the RSME scale. Moreover, Veltman and Gaillard (1996) already reported that RSME was more sensitive to measure mental effort than NASA-TLX, which may explain why the effort scale of NASA-TLX and RSME showed different effects in this experiment.

An interesting measure in this experiment was pupil diameter, which was larger in the first blocks, indicating a higher level of cognitive control in the first blocks. It has been suggested that a large pupil diameter indicates exploitation of the task, i.e., finding out how the task works (Hopstaken et al., 2015). We believe that although participants had experience with solving Sudoku puzzles, they still needed some time to fully understand how to do the task on the computer. In contrast to normal paper-and-pencil Sudoku, where people can write several candidate numbers in a cell, this was not possible here on the computer screen. This required finding a slightly different strategy, which was reflected in the larger pupil diameters in the first few blocks.

In addition, RTs decreased over time in both conditions. This can probably be attributed to a learning effect, which obscures any possible effects of fatigue on performance. As a result, participants had more chances to solve the Sudoku within a block, which was reflected in an increase in the total clicks over time.

It is possible to explain the results of this study using the resource theory. As part of the theory, rest can help individuals to recharge resources and thus maintain performance (Helton & Russell, 2017). Since the accuracy in LL blocks remained stable, participants might have used these blocks to rest and restore their resources. Another possibility was that participants used the moment when they filled in subjective measures (i.e., 20 s after each block ended) to recover. However, the explanation of the resource account assumes that the recovery takes place within a short amount of time and can occur any time regardless of task duration.

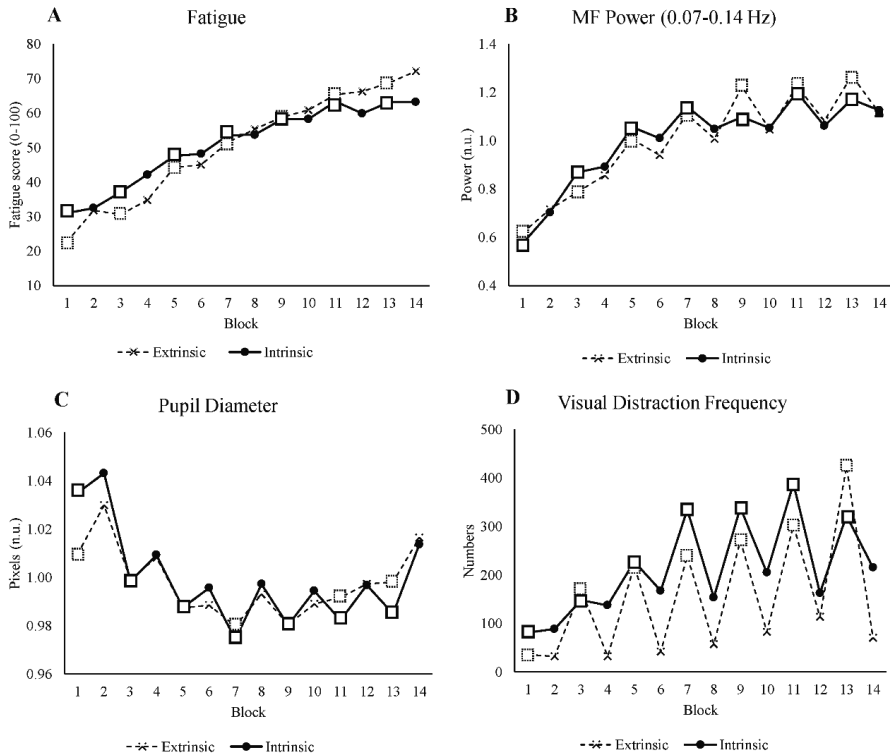
### 3.4.2. Extrinsic and intrinsic motivation

Many studies have tried to find links between mental fatigue and motivation (Kurzban et al., 2013). However, these studies were limited to a single type of motivation, namely extrinsic motivation. To investigate whether intrinsic motivation has the same effects as extrinsic motivation, we compared the results of this study with the results of our previous study in which we incorporated extrinsic rewards in the experiment (Herlambang et al., 2019).

Overall, the results of the current experiment are similar to the results of the extrinsic motivation experiment in that participants were more motivated in high motivation blocks (see **Figure 8**). In both studies, the feeling of fatigue increases over time, which shows that participants did feel fatigue over time regardless of condition. Moreover, participants invested more cognitive effort when the condition was more motivating, which was reflected by similarities in the MF power of HRV in both studies. Cognitive effort seems to be the mediator of the relationship between motivation and performance (see Goodman et al., 2011; Westbrook & Braver, 2015): Highly-motivated people invest more effort, and this maintains performance. In addition, we found the same effect of exploitation and exploration in the pupil size as other authors have (Hopstaken et al., 2015). In the first phase of both experiments, participants learned and tried to do the tasks well (exploitation) that manifested in smaller pupil size over time. Later, the pupil dilated more, searching for more rewarding activities (exploration). Furthermore, participants were more susceptible to distractions when they were not motivated, which is also in line with motivation theory (Hockey, 2011).

In summary, the results suggest that intrinsic motivation, as with extrinsic motivation, is essential for explaining the effects of mental fatigue on performance. We propose that intrinsic motivation and extrinsic motivation in both studies share the same process. As time progresses, people will invest more effort and exert more control in performing a task that they like (e.g., playing a game, solving a puzzle, working on a hobby—intrinsic motivation) and/or is favorable (it offers more rewards than costs—extrinsic motivation) compared to tasks for which they lack motivation.

Although this study shows important results, it was limited to a laboratory environment. Therefore, for future research, it is beneficial to conduct a study with real-life tasks that involves vigorous motivation, such as online gameplay (Hainey et al., 2011). In addition, it is favorable to have a control group to check a clear comparison between intrinsically motivational condition and nonmotivational one. Also, it would be interesting to investigate the effects of relief: In both experiments, we noticed that pupil diameter increased again in the last blocks. We assume that this is caused by participants' expectation that the experiment would end soon. By including these factors, a more robust theory of mental fatigue and motivation can be developed.



**Figure 8.** The results comparison between the study of the extrinsic motivation experiment represented by a dashed line and the intrinsic motivation experiment represented by a solid line. **(A)** Fatigue scale. **(B)** MF power of HRV. **(C)** Pupil diameter. **(D)** Visual distraction frequency. All x-axes show blocks, where odd blocks represented by square markers are low motivation blocks (nonreward condition in the extrinsic motivation experiment and LL condition in the intrinsic motivation experiment). The figures of the extrinsic motivation experiment were adapted from “The Role of Motivation as a Factor in Mental Fatigue”, by M. B. Herlambang, N. A. Taatgen, & F. Cnossen, 2019, *Human Factors*, 61(7), p. 1176, 1179, 1180. doi: 10.1177/0018720819828569. Sage Publishing. Creative Commons License (CC BY-NC 4.0).

### 3.5. Funding

This work was supported by Lembaga Pengelola Dana Pendidikan (LPDP) from the Indonesia government awarded to Mega B. Herlambang.









A gift from Almira on Father's Day

