Free Will in Psychological Research: Considerations on Methodic Procedure and Reproducibility of Results

Von der Fakultät Wirtschaftswissenschaften
der Leuphana Universität Lüneburg zur Erlangung des Grades

Doktorin der Philosophie
- Dr. phil. -

genehmigte Dissertation von
Wiebke Melcher (geb. Mahlfeld)

geboren am 03.02.1986 in Peine

Eingereicht am: 18. März 2019

Tag der Disputation: 05. August 2019

Betreuer & Erstgutachter: Prof. Dr. Hans-Rüdiger Pfister

Zweitgutachter: Prof. Dr. Rainer Höger

Drittgutachter: Prof. Dr. Oliver Genschow

Zusammenfassung

Die Publikation der berühmten Studie von Libet und Kollegen (1983) zur zeitlichen Reihenfolge von der neuronalen Vorbereitung einer Handlung, die dem Zeitpunkt der bewussten Handlungsentscheidung vorausgeht, säte erste Zweifel an der Existenz des freien Willens und markiert den Beginn einer kontroversen und bis heute offenen Debatte, die nicht nur über die Existenz des freien Willens an sich geführt wird, sondern auch über die Angemessenheit der eingesetzten Untersuchungsmethoden und Verlässlichkeit der Ergebnisse der Forschung zum freien Willen. In den letzten Jahrzehnten gab es nur wenige Beiträge empirischer Art zu dieser Debatte, sodass bisweilen der Anschein erweckt wird, dass Beiträge mehr auf der persönlichen Überzeugung der Beteiligten basieren als auf empirischen Belegen. Passenderweise ist auch der Glaube an den freien Willen zum Thema psychologischer Forschung geworden. In der Literatur finden sich Hinweise darauf, dass Laien verschiedener kultureller Hintergründe zumeist glauben, einen freien Willen zu haben und dass der persönliche Glaube an den freien Willen einen Einfluss auf das Erleben und Verhalten hat, mit der Tendenz zu positiverem Verhalten, wenn der Glaube an den freien Willen stärker ist. Empirische Befunde aus dem deutschsprachigen Raum sind selten, möglicherweise da es bisher kein geeignetes und validiertes Fragenbogeninstrument zur Erfassung des Glaubens an den freien Willen gibt. Die letzten Jahre sind geprägt von der Replikationskrise, die in weiten Teilen zu Zweifeln an der Belastbarkeit und Replizierbarkeit von psychologischer Forschung führte und auch zeigte, wie wichtig die Replikation von Forschungsergebnissen für den Erkenntnisgewinn ist. Um diese Krise zu bewältigen, ist es wichtig, die in einem Forschungsfeld genutzten Methoden kritisch zu hinterfragen, bereits publizierte Befunde zu replizieren und die Generalisierbarkeit der Ergebnisse auf einen größeren Kontext zu untersuchen. Ziel dieser Dissertation ist es, einige Aspekte der psychologischen Forschung zum freien Willen und zum Glauben an den freien Willen einer kritischen Prüfung zu unterziehen. Zwei Studien werden berichtet, deren Ziel es ist, das Libet-Paradigma auf eine

freiwillige Entscheidung zu übertragen, die reelle Konsequenzen für die handelende Person hat, da eine solche Studie bisher nicht veröffentlicht wurde. Außerdem wurde getestet, ob die Zeitschätzung der bewussten Handlungsintention einen direkten Einfluss auf das im Libet-Paradigma gemessene Bereitschaftspotential hat. Des Weiteren wird die Entwicklung eines ersten deutschsprachigen Fragebogeninventares zur Erfassung des Glaubens an den freien Willen beschrieben, welches auch einige methodische Schwierigkeiten der bestehenden englischsprachigen Instrumente überwindet. Zusätzlich werden Studien zur experimentellen Manipulierbarkeit des Glaubens an den freien Willen berichtet. Die Ergebnisse der hier berichteten Studien geben Einblick in den aktuellen Stand und die Belastbarkeit der Forschung zum freien Willen und Glauben an den freien Willen.

Abstract

When Libet and colleagues published their results on the temporal order of movement preparation and the reported time of conscious will to move in 1983, they shed some doubt on the existence of free will. This marked the beginning of a controversial and still ongoing debate, not only about the existence of free will, but also about the appropriateness of methods and validity of results from research on free will. Only a few empirical contributions were added to this debate in the last decades, so the discussion about the existence of free will sometimes seems to rely more on personal views than on empirical evidence. Opportunely, belief in free will was also discovered as psychological research topic. Literature on belief in free will shows some evidence that most laypersons across different cultural backgrounds believe that they have free will and that a person's belief in free will might have an impact on cognition and behavior, tending to positive outcomes with a greater belief in free will. Empirical findings from the German-speaking area are sparse, probably due to a lack of validated measurements assessing belief in free will available in the German language. For many psychological research fields, recent years have been characterized by the publication crisis. To overcome the crisis, it is important to critically scrutinize the methodological procedures used in a specific research field, to replicate published results, and to examine the ability to generalize these results to a broader context. The aim of this dissertation is to critically examine some aspects in psychological research on free will and the belief in free will. Two studies are reported that aim to generalize the Libet paradigm for a free and voluntary decision with consequences for the acting person, as this was never reported to have been researched in literature before, and to test the critical objection that the measurement of reporting the conscious intention to move has a direct effect on the result in the Libet paradigm. Furthermore, the construction of the first inventory measuring belief in free will in the German language is described. This inventory was also created with the aim of overcoming some methodological problems in the existing instruments in English language. Furthermore, studies on the

experimental manipulability of the belief in free will are reported. These findings provide implications in view of the current state of research on free will and belief in free will and its reliability.

Contents

List of Tables

List of figures

1. Introduction

The question of whether or not we have free will might be a question as old as human civilization, at least going back to ancient Greek philosophers like Socrates, Plato, and Aristotle (Doyle, 2011). More than three decades ago, neuroscientists and psychologists started to examine this question with the specific methods of their professions, resulting in an ongoing debate, not only about the existence of free will, but also about the appropriateness of methods and validity of results. The focus of this dissertation lies in a critical examination of the methodological processes in psychological research on free will and the reproducibility of its results.

No generally accepted definition of free will exists. For laypersons, free will often means that people are in conscious control of their actions and decisions, are morally responsible for it, and could do otherwise in the exact same situation. Philosophers developed an astonishing number of different positions on the existence and appearance of free will. Some philosophical positions on free will completely oppose each other, such as determinism and indeterminism, whereas in some positions, like compatibilism, aspects of other positions (e.g. of libertarianism and determinism) are entirely compatible. The debate surrounding free will is related to a lot of philosophical issues, e.g. moral agency and responsibility, limits of human freedom and self-control (Kane, 2011). Due to its close connection to terms like responsibility, insight, and guilt, the debate is not only highly relevant on a theoretical level, but also in more practical applications such as legislation, especially in German civil law, where the terms "free will" and "free determination of will" are explicitly used (Kawohl & Habermeyer, 2007). It is also crucial for criminal law and the concept of lawful penalty, because one must be responsible for an act in order to be found guilty and punished for an act. Not only could the existence of free will be important for legislation, but altering the belief in free will can change one's view on retributive punishment (Shariff et al., 2014) and therefore scientific findings on free will could have an impact on legislation by helping to shape public opinion. Despite its continued relevance and although the free will debate has had a long history in philosophy, the question itself seems to remain unsolved and this dissertation does not aim to settle that debate. However, it should shed some light on validity of results on free will and the belief in free will in psychological research.

Psychologists and neuroscientists have expanded the methodological diversity in the debate on free will. In general, psychological research on free will can be divided into two general fields, one field concentrating on the question of whether humans have a conscious free will and a second field focusing on how the abstract belief in free will affects human experience and behavior. A popular example from the first field is the well-known study by Libet, Gleason, Wright, and Pearl (1983), who developed an innovative approach to the question with a completely unprecedented method and challenged the impression of the human being as a consciously acting entity. Although Libet (1999) did not interpret these results as evidence against free will, their study took on a leading role among supporters of the position that people do not have free will. They found that a movement-related brain potential is already measurable before a person becomes aware of the intention to perform that movement. This result lead to the conclusion that free will could be just an illusion of the brain because the cerebral initiation of a voluntary movement begins unconsciously (Libet, Gleason et al., 1983), thus the whole volitional process is initiated unconsciously. But the study has often been criticized for its methodological procedure as well as for its artificial experimental setting, which might not elicit a free and voluntary decision. More recently, critics have also taken aim at the presumptions that were made concerning the movement-related cortical potential involved, which might not be what it was thought to be.

Not only were neuroscientific approaches to the free will debate developed in the last few decades, but sociopsychological research approaches as well. Sarkissian et al. (2010) found that the majority of people across different nations believe in an indeterministic universe with the existence of free will. This adds another aspect to the debate about free will and poses another question altogether, regardless of whether or not free will indeed exists: Do people believe that they have a free will and does this belief have an impact on cognitive functions and behavior? Many studies have addressed these topics: Less belief in free will has been shown to be associated with more aggressive and less helping behavior (Baumeister, Masicampo, & Dewall, 2009), more cheating (Vohs & Schooler, 2008), less learning from emotional experiences (Stillman & Baumeister, 2010), reduced self-control (Rigoni, Kuhn, Gaudino, Sartori, & Brass, 2012) and impaired action-monitoring (Rigoni, Wilquin, Brass, & Burle, 2013). People with lower levels of belief in free will were more likely to conform with other opinions (Alquist, Ainsworth, & Baumeister, 2013) and were less grateful (MacKenzie, Vohs, & Baumeister, 2014). However, there are not only negative consequences. Telling people that they do not have free will makes them more willing to forgive, mediated by a lower level of perceived moral responsibility (Brewer, 2011), and, as mentioned before, makes them less supportive of retributive punishment (Shariff et al., 2014). Therefore, belief in free will seems to have an impact on people's daily lives and on society as a whole.

The work of Rigoni, Kuhn, Sartori, and Brass (2011) connects neuroscientific research on free will with sociopsychological research on the belief in free will. They experimentally manipulated belief in free will and subsequently measured the preconscious motor preparation in the task developed by Libet, Gleason et al. (1983). They found that the preconscious motor preparation measured by readiness potential is altered in individuals who are induced to disbelieve in free will. To induce disbelief, it was sufficient to read a passage of the book *The Astonishing Hypothesis* (Crick, 1995), claiming that free will is just an illusion. This manipulation method was adapted from Vohs and Schooler (2008) and is widely used throughout the literature.

In contrast to the experimental manipulation studies, researchers have also measured belief in free will as a presumably stable trait (Baumeister & Brewer, 2012). Some inventories measuring belief in free will were developed in the English language (e. g. Nadelhoffer, Shepard, Nahmias, Sripada, & Ross, 2014; Paulhus & Carey, 2011; Rakos, Laurene, Skala, & Slane, 2008; Stroessner & Green, 1990), but no inventory in the German language has been published yet. Additionally, the available inventories often suffer from methodological deficits.

Recently psychological research has been struggling with doubts concerning the reproducibility of published findings, and this has grown to the point where it has become a full-scale crisis, especially since the publication of the results of the Open Science Collaboration (2015). The collaborative, crowdsourced project with more than 300 authors and volunteers aimed to replicate 100 findings published in three high-impact journals of psychology and provided empirical evidence on the topic of reproducibility, which is often neglected. The results were astonishing. Only 39% of the replication trials reproduced a significant effect comparable to the effect that was originally published (Open Science Collaboration, 2015) and the study by Vohs and Schooler (2008) was one of the 61% of studies that was not replicated successfully. The reasons for the low replication rate of 39% are certainly manifold, yet two of the more notable reasons are the publication bias and the definition of academic success. For the individual success of a scientist, it is far more important to make a novel contribution than to recheck effects in order to test how stable an effect is under different conditions or if it depends on a very specific setting and is not generalizable. This situation leads to less efficient knowledge accumulation because published results often remain unquestioned and therefore false results remain undiscovered (Nosek, Spies, & Motyl, 2012). Despite all the doubts that it raised, the replication crisis led to a positive outcome. It reminded the scientific community that replication is necessary and also led to a renewed effort to improve research practice. In light of recent events and the importance of the debate on free will, it might be useful to take a closer look at the validity and robustness of results.

Questions arising from literature on free will and the belief in free will are specifically the following: Is the movement used in the Libet paradigm comparable to a free and voluntary decision? Can a movement based on a voluntary decision be involved in the paradigm? How can belief in free will be measured reliably? With regard to the diverse philosophical positions, it is worth looking at the underlying structure of belief in free will and the correlations to other traits and constructs. Can belief in free will be altered by experimental interventions, or in other words, how stable is it? These questions will be addressed in this dissertation.

Objectives and structure of the dissertation

In chapter 2, the focus is placed on research conducted in the mindset of the Libet paradigm. Two studies are reported that aim to generalize the Libet paradigm for a free and voluntary decision with consequences for the acting person, as this was never reported to have been researched in literature before. These studies also aim to test the critical objection that the measurement of reporting the conscious intention to move has a direct effect on the result in the Libet paradigm. Chapter 3 describes the construction and validation of the first German-language inventory measuring belief in free will, which also aims to overcome some methodological problems in the existing instruments in the English language. Chapter 4 describes two attempts to manipulate belief in free will with two different methods that were reported to have been successful in literature. Chapter 5 summarizes the findings and discusses the implications in view of the current state of research on free will and belief in free will.

2. Research on Empirical Evidence Supporting or Refuting the Existence of Free Will Using the Libet Paradigm

Theoretical Background

In 1983, Libet, Gleason et al. took an experimental approach to answering the question of whether or not we have free will (Libet, 1999), thereby restarting the debate concerning the existence and nature of free will, albeit with a completely new method. Libet and colleagues developed a method to mark the onset time of a movement-related cortical potential, the so-called readiness potential, in relation to the self-reported time of the corresponding conscious intention to perform a voluntary motor act. Their results suggest that the onset of readiness potential precedes the conscious intention to move by approximately 350 ms, leading to the conclusion that the cerebral initiation of a free and voluntary movement begins unconsciously. This surprising result restarted the debate, often leading to the interpretation that the impression of free will in the sense of conscious agency is an illusion (Banks & Pockett, 2007).

Readiness potential

For their research paradigm, Libet and colleagues used a cortical potential known from neuroscientific research: Voluntary movements are preceded by a slowly increasing surface-negative cortical potential measurable at the cortex, which is called readiness potential, or *Bereitschaftspotential*, first observed and reported by Kornhuber and Deecke (1965). The readiness potential is reported to be associated with the preparation and initiation of voluntary muscle contractions such as flexion or extension of a finger (Shibasaki, Barrett, Halliday, & Halliday, 1980) but also quite similar in volitional motor inhibition and muscle relaxation (Shibasaki & Hallett, 2006). Therefore, it is used for detecting the participation of the voluntary motor system. The readiness potential starts up to 1.5 s (Jahanshahi & Hallett, 2003) or 2 s (Shibasaki & Hallett,

2006) before movement. It can be divided into two phases: The early, slowly-increasing segment of readiness potential, which shows no site-specificity and reaches its maximum at vertex, and a late phase, which starts about 400 ms before the onset of movement and shows a steeper negative increase until it reaches its maximum at the contralateral site of the motor cortex (Shibasaki & Hallett, 2006).

In the first publication about the readiness potential by Kornhuber and Deecke (1965), it was already mentioned that the amplitude of the readiness potential increases with higher attention and a higher level of intentional involvement in the task and decreases with indifference of the participants to the task. The amplitude of the readiness potential is also higher in more complex movements (Benecke, Dick, Rothwell, Day, & Marsden, 1985) and depends on the demands of planning to perform a movement, as the amplitude that was found tended to be higher for freely selected movements than for fixed repetitive movements (Dirnberger, Fickel, Lindinger, Lang, & Jahanshahi, 1998). Lang (2003) gives an overview on further factors influencing the readiness potential such as force and speed of movement and, as a consequence of the representation in the motor cortex, the body part performing the voluntary movement.

The Libet paradigm

In their experiment, Libet and colleagues (Libet, Gleason et al., 1983; Libet, Wright, & Gleason, 1982) recorded the readiness potential of six participants while letting them work on three different tasks. In all three tasks, the participants observed a clocklike display presented by a cathode ray oscilloscope: In the first condition, participants were asked to decide freely when to move their hand and to report afterwards at which time on the clock their conscious intention to move (will judgment) appeared. They were instructed to act spontaneously and to avoid pre-planning the movement. In the second condition, participants were asked to produce a hand movement at a specific time relative to the clock provided. In the third condition, participants were asked to judge at which position of the rotating clock hand they received a near-threshold external stimulus pulse on the back of their hand. Libet and colleagues reported some interesting findings: The onset of the readiness potential starts earlier when the task required preplanning or participants reported awareness of some sort of preplanning for the movement (Banks & Pockett, 2007; Libet et al., 1982). Furthermore, Libet, Gleason et al. (1983) compared the onset times of the readiness potential in the first condition with the reported time of conscious intention to move and the onset of the movement measured by EMG onset. They found that the conscious will preceded the beginning of the hand movement by about 200 ms, but the onset of the readiness potential already starts 550 ms before movement onset. Therefore there is already a measurable preparation for the movement at least 350 ms before a conscious will to move occurs, leading to the conclusion that the movement preparation starts unconsciously (Libet, Gleason et al., 1983). In the third condition, which served as a baseline for the time-judgment task, they found that participants reported the onset of the external stimulus about 50 ms earlier than it actually appeared. Taking this timing bias into account, the time difference between the conscious will and the actual movement would decrease to 150 ms, which means the onset of cortical movement preparation would precede the conscious will by 400 ms. As the movement preparation starts earlier than the conscious intention to move occurs and is therefore initiated unconsciously, the conscious will cannot be the cause of the motor preparation reflected by the readiness potential.

The conscious veto

Libet himself did not interpret the result as a threat to the existence of free will and instead proposed the conscious veto function (Libet, 1985, 1999). These findings only indicate that free will could not cause the initiation of the volitional motor action, but free will could be in causal relation to the actual execution of the motor action. Since the conscious intention to move is reported at least 150 ms earlier than the beginning of the movement, there would be enough time left for a conscious interruption of the movement. Libet drew this conclusion by comparing the readiness potential of preplanned movements with that of preplanned but "vetoed" movements, i.e. those that were not actually followed by movements (Libet, Wright Jr., & Gleason, 1983). He found a readiness potential in both, but the readiness potential in the vetoed movements tended to reverse direction to smaller amplitudes 150 to 250 ms before the actual movement. Nevertheless, the vetoed movement was studied in the context of preplanned movements and not in a self-initiated "free" movement, since it is not possible to compare the readiness potential of a stimulus-independent self-initiated movement, because it would lack a time point zero (i.e. the time of an actual movement) for averaging.

Recently, empirical evidence was found for the veto function proposed by Libet. Schultze-Kraft et al. (2016) trained a brain-computer interface to detect readiness potential in real time and presented a stop sign when an onset was detected. They found "a point of no return": A movement can be canceled even after the onset of the readiness potential if the stop signal occurred at least 200 ms before movement (Schultze-Kraft et al., 2016). Maybe free will is therefore preserved in the sense of a free won't. Nevertheless, the Libet paradigm remains controversial in terms of its methodological procedure.

Experimental replications, extensions, and empirical criticism of the Libet paradigm

There is a far greater amount of commentary and reinterpretations of Libet's original data than there is empirical evidence addressing critical points in the method of the Libet paradigm. Here the focus will be placed on the literature providing empirical insights to the debate. The results of Libet and colleagues have been reproduced and extended several times by independent research groups (Haggard & Eimer, 1999; Keller & Heckhausen, 1990; Miller, Shepherdson, & Trevena, 2011).

Keller and Heckhausen (1990) compared the readiness potential of conscious voluntary movements in a Libet-style task with the readiness potential of unconscious involuntary movements. For the Libet task, they obtained nearly the same result as reported by Libet, Gleason et al. (1983), with minor differences in the mean times for the onset of readiness potential and for reported intention to move. Furthermore, they found the same onset of readiness potential for both conditions, but a higher amplitude for voluntary movements. They mentioned that the difference in amplitude could alternatively be explained due to the different levels of applied force required for the movement (which was not measured during the experiment). Additionally, they found differences in scalp distribution, with the readiness potential reaching its maximum at FCz electrode (according to 10-20 system) for conscious movements and at Cc for unconscious movements. They conclude that the differences they obtained are the result of the task to introspectively monitor internal processes and therefore, the readiness potential obtained in the Libet-style tasks results from the instruction to perceive an "urge to move."

Several years later, Haggard and Eimer (1999) also reproduced the result of the Libet study but additionally compared a fixed movement condition with a free movement condition. In contrast to the fixed movement condition, which involved a predefined response finger, participants were allowed to decide at their own pace whether they use the left or the right index finger for their movement in the free movement condition. No difference was found in readiness potentials between the fixed and free movement conditions, but the readiness potential showed a tendency to occur later in trials with early reported awareness of movement initiation, raising doubts that the readiness potential is the cause for conscious awareness of movement initiation. In contrast, the lateralized readiness potential was found to covary with the time of conscious awareness of intention, thus indicating that earlier awareness occurs with earlier onsets of lateralized readiness potentials. Results suggest that the conscious awareness of intention corresponds to the choice of a specific action (here the decision to move the right or left index finger) characterized by the lateralized readiness potential, and corresponds not to the earliest initiation of movement process characterized by the readiness potential. However, the onset of the lateralized readiness potential also precedes the reported time of conscious awareness.

Miller et al. (2011) compared the readiness potential measured while working on a classic Libet task with the readiness potential in a task where a clock was not presented. They found that the amplitude was reliably more negative in the condition with a clock display than in the comparable condition without a clock. They conclude that the readiness potential might be merely an artifact of the clock, or more specifically that the amplitude of readiness potential is primarily caused by the requirements of tasks such as clock monitoring and time reporting, rather than caused by unconscious movement preparation. It remains unclear whether it might be just the visual presence or absence of a clock (and therefore the clock monitoring) that causes this effect or if it is the task of actively introspecting and reporting the time of conscious intention, because in the study conducted by Miller et al. (2011), the clock was present in one condition and absent in the other condition. A short note on the problem concerning the task of clock monitoring and time reporting as a confounding variable can be found in Libet et al. (1982), who present data on the readiness potential from one subject in an unchanged experimental situation that included a clock yet the subject "was told to ignore this usual requirement and not to report anything. This difference in reporting requirements appeared to make no obvious difference" (p. 327). Nevertheless, this was never tested systematically. Results from fMRI on activation of the pre-supplementary motor area (an area probably involved in generating the early readiness potential; Praamstra, Stegeman, Horstink, & Cools, 1996; Shibasaki & Hallett, 2006) show higher activation in the pre-supplementary motor area for a Libet task with report of intention to move (W judgment according to Libet, Gleason et al., 1983) than for a Libet task with a report of the time when the actual movement happened (Lau, Rogers, Haggard, & Passingham, 2004). So it might be possible that the effect reported by Miller et al. (2011) depends on the demands of introspection required to report the time of conscious intention or the demands of monitoring the clock, or both combined. If the task of reporting the time of intention alone had caused the effect, then the readiness potentials measured in a Libet task with a clock present and involving reporting the time of conscious intention to move should also differ from the readiness potentials in a task where the clock was present yet the time of conscious intention to move did not have to be reported.

A simple movement as a representation of a willful decision?

The kind of motor act used by Libet et al. was often criticized for being too artificial and not representing a willful act or free decision (Banks & Pockett, 2007; Brücher & Gonther, 2006; Kawohl & Habermeyer, 2007; Klemm, 2010; van de Grind, 2002). In most of the studies, the decision is restricted in time in order to meet the requirements of recording an event-related potential like the readiness potential: Participants were instructed to wait a few seconds before deciding to move (Haggard & Eimer, 1999; Libet et al., 1982; Libet, Wright Jr. et al., 1983) to prevent contaminating the pre-movement time span with artifacts from the previous trial. In order to obtain an event-related potential, several repetitions are necessary to reach a robust signal-to-noise ratio after averaging the individual movements. Therefore, it is necessary to perform the same movement repeatedly during the experiment. Usually, participants only have to decide when to execute a predefined motor act (except in the case of Haggard & Eimer, 1999, who allowed participants to decide whether to move the right or left index finger in one experimental condition). For these kinds of restricted movements, Keller and Heckhausen (1990) stated that the "general intention to move, therefore, has already been formed at the beginning of the experiment" (p. 352). A further common factor in all studies described using the classical Libet paradigm is that is the decision to move was not followed by any consequences in the experimental procedure for the acting person. Therefore, it is worth questioning whether or not the self-initiated movement in a Libet paradigm truly represents a willful act of deciding. It remains unclear whether the readiness potential also precedes a decision that is more personal than a simple finger movement since there are no studies addressing this question (Banks & Pockett, 2007).

Hypotheses

The aforementioned considerations led to the following research questions and hypotheses:

1. According to the chronological order published by Libet et al. (1983), it is expected that the onset of readiness potential precedes the time of the conscious intention to move.
2. Does the task of reporting time have an impact on the amplitude of the readiness potential? If the task of reporting the conscious intention to move, rather than clock-monitoring task, is the cause of the decreased amplitude reported when the clock was absent (Miller et al., 2011), then the amplitude in a classic Libet situation with a clock presented but without the intention-reporting task should be smaller than in the same situation with the intention-reporting task.
3. Is there a measurable readiness potential preceding a personal, voluntary decision with a real and immediate consequence for the individual performing an action? Furthermore, if so, does this readiness potential differ in terms of the onset time and amplitude from a readiness potential measured in a classical Libet task?

Study 1: The effect of a consequence on the readiness potential preceding a self-initiated motor act

Method

The methods were mainly based on those developed by Libet et al. (1982) with some modifications in order to test the assumptions made above. The full duration of the experiment was 2.5 to 3 h depending on the time needed to place the electrode cap on participants for the EEG measurement.

Participants

Data were recorded from 26 participants; 2 participants were removed from the sample due to technical problems, resulting in a sample of 24 participants (11 female, 13 male; ranging in age from 18-32 years, M=24.00, SD=3.52, one missing value for age). All participants had normal or corrected-to-normal vision and participated voluntarily. Most participants were students majoring in Psychology and received course credit for their participation. Two participants reported being left-handed, but were used to handling a regular gamepad.

Procedure

Participants were tested individually with an observer present. Participants sat in a comfortable chair with a monitor in front of them at a distance of 1.95 m. They performed a modified Libet-task with four conditions (presented in a fully balanced order) with 80 trials each. In the Libet task, participants fixated the center of a clock presented on the monitor. The radius of the clock was 6.1 cm, resulting in a visual angle of 1.8° subtended between the center and the moving cursor (see figure 1). The beginning of each trial was marked by a preparatory tone. 1500 ms later a cursor appeared and the clock rotation started. The duration of one clock revolution was 2400 ms. Clock hand rotation ended either after four revolutions or after the press of a button, with a random delay between 500 to 800 ms so as not to provide a visual cue for the time of movement. After the clock hand disappeared, the clear clock display was presented for the rest of the duration of four revolutions and afterwards a display was presented with a fixation mark in the center. The minimum duration of a trial from its start to the start of the next trial was 14 s. Participants were asked not to move or blink during the experimental trials. They were instructed to wait for one complete revolution and to perform a movement, i.e. pressing a button on a gamepad, in one of the following revolutions corresponding to the instructions of the respective experimental condition (specified below). During some conditions of the experiment, participants had the opportunity to take a break at the time of their own choosing by pressing a second button. If the break button was operated, the trial remained unfinished and the next trial was started following the end of the break. All movements (button presses) were performed with the thumb of the right hand on a gamepad. The software used for stimulus delivery and experiment control was Presentation® (Version 16.5; Neurobehavioral Systems, Inc., Berkeley, USA). Documentation of the self-written scripts used in the experiment can be found in electronic appendix.

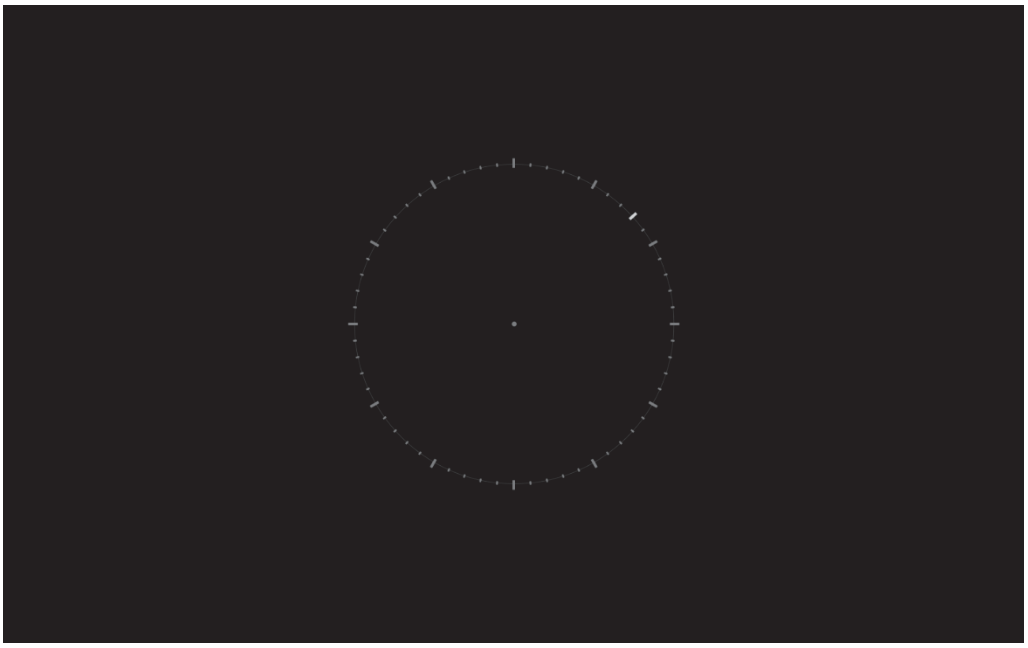


Figure 1. Clock display. Clock rotation started with a delay of 1500 ms. Each trial contained four clock revolutions, with a duration of 2400 ms per revolution. The visual angle subtended between the center and the moving curser was 1.8°.

Experimental conditions

Each participant worked on four different conditions with a duration of approximately 30 min each. The order of the experimental conditions was fully balanced. In condition A, participants were asked to make a voluntary movement spontaneously at a time of their own choosing. In condition B, participants made a voluntary movement at a time of their own choosing and were asked to report, based the clock display presented, the time that they felt the conscious intention to move (W judgment according to Libet, Gleason et al., 1983); in condition C, participants performed a movement at a pre-defined clock-time. In condition D, participants were asked to report the onset time of a tone delivered randomly. An overview of the experimental conditions is given in table 1.

Table 1: Experimental conditions of study 1.

<table>
<tr><td>A</td><td>Self-initiated movement:
Participants voluntarily pressed a key at a time of their own choosing.</td><td rowspan="3">Self-determined decision:
In conditions A, B & C, the participants had the possibility to request a break by pressing a button whenever they wanted to. The break was introduced to the participants as having a recreational function.</td></tr>
<tr><td>B</td><td>Self-initiated movement with time-report:
Participants voluntarily pressed a key at a time of their own choosing and were asked to report the time that they felt the conscious intention to move.</td></tr>
<tr><td>C</td><td>Preplanned movement:
Participants were asked to press a key at a pre-set time.</td></tr>
<tr><td>D</td><td colspan="2">Baseline measurement:
Participants were asked to report the onset time of a tone delivered randomly.</td></tr>
</table>

Conditions B and D involved having the participant report the clock time, which was done verbally and noted by the observer. Conditions B and C were analogous to the “self-initiated voluntary act” and “pre-set motor act” that were developed and described by Libet et al. (1982). Their procedure was supplemented by condition A in order to check for a possible influence of the time-reporting task. Additionally, a voluntary movement reflecting a free decision with realistic consequences was established within the experimental procedure: Participants had the opportunity to request a break at any time in condition A, B, and C by pressing a second button next to the movement button on the gamepad. Participants were instructed to use it as a respite from the monotonous experimental procedure, for instance to change seating position, to drink some tea, or to have a short conversation with the observer. They were encouraged to take a break often because relief from the task would improve the quality of data recording. This approach makes it possible to obtain data concerning a self-determined decision. Condition D functions as a baseline measurement to check the accuracy of the time reporting. In condition D, participants had the task of reporting the onset time of a tone that was delivered randomly, using the rotating clock hand as a basis.

Data acquisition and preprocessing

EEG was recorded with an Ag/AgCl electrode cap (waveguard™ cap) from positions Fp1, Fpz, Fp2, F3, Fz, F4, FCz, C3, Cz, C4, P3 and Pz, mounted according to the 10/20-system with AFz as ground electrode. Following the manufacturer's specifications, electrode impedance was kept below 20 kOhm (ANT Neuro B.V./eemagine Medical Imaging Solutions GmbH, 2013). Continuous EEG data were analyzed off-line using asalab™ software (Version 4.9.1; ANT Neuro, Enschede, Netherlands). The signal was digitized at 256 Hz, re-referenced to the average signal of the mastoids and filtered (notch filter of 50 Hz and band-pass filter from 0.016 to 70 Hz, 24 dB/octave attenuation). Epochs were time-locked to participants' button press using an interval from 2500 ms before and 300 ms after the button press. The interval between -2500 to -2300 ms was used for baseline correction. Epochs containing artifacts were excluded based on automatic artifact detection (peak to peak < 100 µV) and visual inspection. Onset of the readiness potential in Cz waveform was determined similarly to the definition provided by Verleger, Haake, Baur, and Śmigasiewicz (2016). When the press of the button, rather than the muscle activation measured by EMG, is used to mark the time-point zero, readiness potentials tend to reach their maximum negative amplitude earlier than time-point zero and the peak is more likely to appear around -300 ms (Verleger et al., 2016). The maximum amplitude of readiness potential is calculated as the mean amplitude within the time interval -400 to -200 ms. The onset of readiness potential is determined as the latest time-point prior to this interval when the amplitude in Cz finally rose above 20% of the maximum negative amplitude. In contrast to Verleger et al. (2016), the baseline was defined as a static interval (-2500 to -2300 ms). This procedure, using a relative criterion, is likely to underestimate the real onset of the readiness potential, however it gives an estimate of the latest plausible onset. Additionally, the R package *segmented* (Muggeo, 2008) was used to estimate a segmented regression model for each condition using the data of all individual averages for Cz amplitude (in long data format) as the criterion and the time relative to pressing the button as the predictor. This method requires an initial estimate of breakpoints to be specified. Theoretically the readiness potential is expected to have two components, the early rise from shortly after the baseline definition and the steeper late rise around 400 ms before the peak of the readiness potential. After the peak, the amplitude increases in a positive direction. As mentioned above, the peak of readiness potential is expected around -300 ms, therefore -2300 ms, -900 ms and -400 ms were entered as initial breakpoints since they were the earliest expected breakpoints. Data from after the button was pressed were excluded.

Analysis and Results

Analysis of EEG data / Differences in readiness potential according to the tasks of experimental condition

After rejecting artifacts, on average 39.1 epochs free of artifacts remained per condition (condition A: M=38.88, SD=13.46; condition B: 38.17, SD=14.22; condition C: M=40.21, SD=15.04; in condition A, one participant was removed from the sample because no artifact-free trials remained after artifact rejection). Figure 2 shows the grand average of 24 participants (23 participants in condition A) in the three experimental conditions: Self-initiated movement without reporting time (A, yellow) and with reporting time (B, red), as well as the preplanned movement (C, grey). For the statistical analysis of EEG data, an approach similar to the approach reported by Rigoni et al. (2011) was used: The epoch between 2500 ms and 100 ms prior to pressing the button was subdivided into 12 pre-movement time windows of 200 ms with t01 as the baseline (see table 2). The mean amplitude of every electrode location and condition was computed for each of these time windows. Means, standard deviations, and upper limits for 95% CI of amplitude at vertex (electrode position Cz) for time windows t01 to t12 are reported in table 3.

Table 2. Definition of time windows in the epoch event.

Time window	Time interval relative to button press	Note
t01	[-2,500 ms; -2,300 ms[	Baseline
t02	[-2,300 ms; -2,100 ms[	
t03	[-2,100 ms; -1,900 ms[	
t04	[-1,900 ms; -1,700 ms[	
t05	[-1,700 ms; -1,500 ms[	
t06	[-1,500 ms; -1,300 ms[	
t07	[-1,300 ms; -1,100 ms[	
t08	[-1,100 ms; -900 ms[	
t09	[-900 ms; -700 ms[	
t10	[-700 ms; -500 ms[	
t11	[-500 ms; -300 ms[	
t12	[-300 ms; -100 ms[	
t13	[-100 ms; +100 ms[	movement (0 ≙ button press)
t14	[+100 ms; +300 ms[	post-movement

Repeated-measures ANOVA with a 3 (experimental conditions: A, B, C) × 11 (time windows: t02 to t12) × 3 (laterality: C3, Cz, C4) design was conducted for mean amplitude as a dependent variable. In the case of a significant result for Mauchly's test of sphericity, Greenhouse-Geisser correction was used. Unsurprisingly, as the readiness potential increases in the time period prior to the movement, a significant main effect of time was found, $F(1.24, 27.18)=22.44$, $p=.000$, ${\eta_p}^2=.51$. No other main effect or interaction was found to be significant, therefore the experimental condition was not found to have an effect on the amplitude of the readiness potential.

Table 3. Mean amplitude at electrode position Cz for time windows of 200 ms between 2500 and 100 ms before button press (study 1).

	A			B			C		
	M	**SD**	**UL**	**M**	**SD**	**UL**	**M**	**SD**	**UL**
t01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t02	-0.50	0.58	-0.29*	-0.12	0.52	0.06	-0.18	0.77	0.09
t03	-0.63	1.08	-0.24*	-0.50	1.09	-0.12*	-0.40	1.21	0.02
t04	-1.08	1.54	-0.53*	-0.64	1.62	-0.07*	-0.99	1.40	-0.50*
t05	-1.33	1.86	-0.67*	-0.59	2.36	0.24	-1.01	1.90	-0.34*
t06	-1.62	2.42	-0.76*	-0.81	3.03	0.25	-1.37	2.02	-0.66*
t07	-1.76	2.64	-0.82*	-1.00	3.59	0.25	-1.95	2.54	-1.06*
t08	-2.00	2.89	-0.96*	-1.36	4.09	0.07	-2.26	2.97	-1.22*
t09	-2.17	3.29	-1.00*	-1.93	4.35	-0.41*	-2.75	2.99	-1.71*
t10	-2.42	3.63	-1.12*	-2.21	5.12	-0.42*	-3.30	3.53	-2.06*
t11	-2.90	4.11	-1.43*	-2.88	5.82	-0.85*	-4.07	3.71	-2.77*
t12	-2.96	4.50	-1.35*	-2.82	6.35	-0.54*	-4.06	4.04	-2.65*

Note. The upper limits (UL) for one-tailed 95% CI of the mean amplitude at Cz are presented. As the readiness potential is known to be negative, the lower limit is -∞. *One-tailed 95% CI does not contain 0. t01 reflects the baseline of EEG measurement.

Accuracy of the time report

In the baseline measurement (condition D), the mean deviation for the task of reporting the time of a tone delivered randomly was -41.8 ms (SD=49.19 ms). The time estimate differs significantly from the onset time of the tone ($t(23)=-4.161$, $p=.000$; $d=0.85$; 95% CI [-62.54, -21.00]) indicating that participants reported the onset of the tone earlier than it actually occurred.

Reported time of intentions and timing of mental events

In condition B, the averaged reported time of conscious intention to move (W judgment according to Libet, Gleason et al., 1983) was -144.9 ms before pressing the button (SD=96.35, Mdn=-132.3 ms), and therefore significantly earlier than pressing the button ($t(23)= -7.367$, $p=.000$, $d=1.5$, 95% CI [-185.57, -104.20]). In this experimental condition, the amplitude at vertex (Cz) started to reach significant negative values during the pre-movement time window t09 (the time window between -900 ms and -700 ms) and these remain permanently negative until the movement, see lower limit of the 95% CI for the time of conscious intention to move in table 3. The CI does not estimate the exact time point of the onset of readiness potential, but is an efficient method to show that the amplitude at Cz already differed significantly from zero in negative direction. The descriptive statistics of readiness potential onsets estimated with the relative criterion method are shown in table 4. The average onset of readiness potential in condition B is -1179.7. Mean onset of readiness potential occurred significantly earlier than the mean of reported conscious intention ($t(15)= -8.043$, $p=.000$, $d=2.0$) but did not covary with it ($r=.03$, $p=.918$). The onset times for the different conditions A, B, and C did not differ significantly ($F(2,26)=0.857$, $p=.436$). A noticeable finding is that in 26.8% of cases, no clear negativity preceding the button press was found.

Table 4. Onset of readiness potential in ms based on relative criterion method in study 1.

	Condition		
	A	B	C
n	17	16	19
n (no clear negativity)	6	8	5
M	-1317.10	-1179.69	-1347.45
SD	446.02	513.74	534.39
Mdn	-1324.22	-1152.34	-1457.03
Min	-1976.56	-1984.38	-2085.94
Max	-527.34	-464.84	-445.31

The segmented regression method was used for further investigation of the readiness potential that was measured in the different conditions. Estimated breakpoints for conditions A, B, and C are listed in table 5. For conditions B and C, the first breakpoint was estimated to occur at approximately -2300 ms, close to the end of baseline interval. For condition A, the first breakpoint of regression was estimated at a quite later point in time (-1667.9 ms), but in contrast to the other conditions, the initial slope for the first segment was already significantly negative (see table 6). The second breakpoint that was estimated might reflect the onset of the lateralized readiness potential. The third breakpoint marked the peak of the readiness potential and was found, as expected, approximately 250 to 300 ms before pressing the button. The regression segments are depicted in figure 4.

Table 5. Breakpoint estimations in ms for study 1 from segmented regression.

	Condition					
	A		B		C	
Breakpoint	*Estimate*	*SE*	*Estimate*	*SE*	*Estimate*	*SE*
1	-1667.92	150.72	-2373.70	155.86	-2303.73	69.92
2	-639.75	89.14	-1140.10	85.42	-888.22	69.61
3	-285.75	26.14	-298.69	36.14	-251.36	13.28

Table 6. Results for slopes in segmented regression in study 1.

	Condition A			
	Estimate	*SE*	*t*	*p*
Intercept	-4.1925	0.3576	-11.72	<.000
Initial slope for time relative to 0	-0.0017	0.0002	-10.14	<.000
Δ slope after breakpoint 1	0.0007	0.0002	3.47	
Δ slope after breakpoint 2	-0.0017	0.0006	-2.59	
Δ slope after breakpoint 3	0.0072	0.0010	6.90	
	Condition B			
	Estimate	*SE*	*t*	*p*
Intercept	-6.4068	9.2612	-0.69	0.489
Initial slope for time relative to 0	-0.0027	0.0038	-0.70	0.484
Δ slope after breakpoint 1	0.0019	0.0038	0.50	
Δ slope after breakpoint 2	-0.0017	0.0003	-6.42	
Δ slope after breakpoint 3	0.0060	0.0011	5.47	
	Condition C			
	Estimate	*SE*	*t*	*p*
Intercept	1.6109	3.3528	0.48	0.631
Initial slope for time relative to 0	0.0007	0.0014	0.48	0.631
Δ slope after breakpoint 1	-0.0024	0.0014	-1.74	
Δ slope after breakpoint 2	-0.0016	0.0003	-6.08	
Δ slope after breakpoint 3	0.0126	0.0010	12.63	

Self-determined decision

As intended, the break-request function apparently was not perceived as an experimental condition. Although participants were instructed not to move until the break-related screen saver was visible, most of the epochs around break requests contained artifacts. Overall, usage of the break-request button was sufficient, with participants pressing the break button on average 47.8 times during the three conditions in which the break-request function was enabled (SD=57.94). After artifact rejection, on average only 1.9 artifact-free instances of pressing the break button remained per participant (SD=3.30) and consequently not enough trials for averaging.

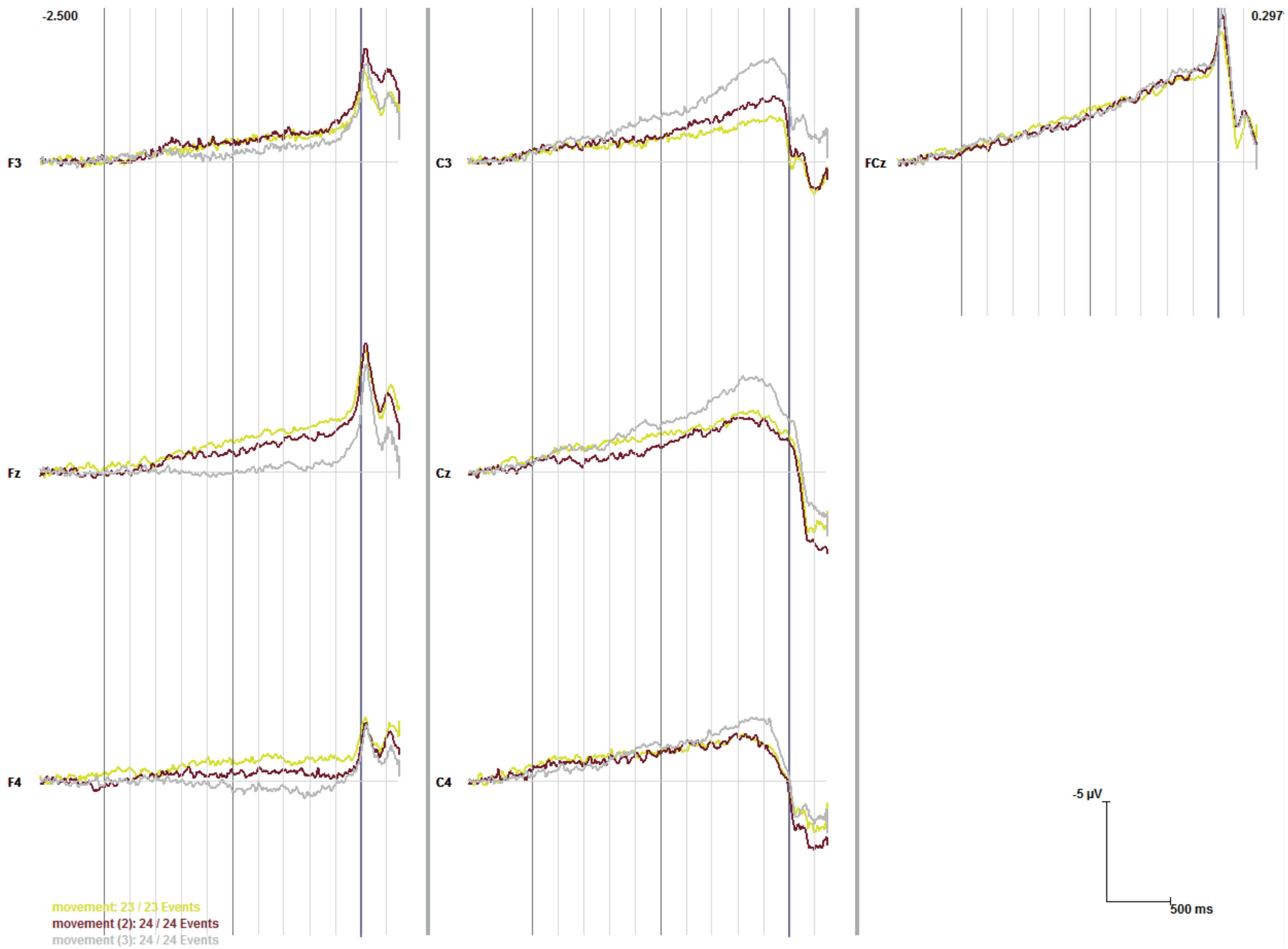


Figure 2. Grand average of 23/24 participants in 3 conditions: Self-initiated movement without time-reporting (A, yellow, 23 participants) and with time reporting (B, red) and preplanned movement (C, gray).

Preliminary Discussion

Readiness potentials did not differ significantly in amplitude or onset between the different experimental conditions, therefore, no significant effect was found for the time reporting task. Nevertheless, the amplitude in Cz in figure 4 appears slightly higher for the preplanned condition than for the other conditions. Two potential reasons for this are, firstly, that a possible effect of the condition is too small to be detected among a small sample, like the one reported here, or secondly, descriptive difference in amplitude could occur due to random effects. If the difference occurs systematically, it should be observable in other studies using the original conditions of Libet as well and should also occur in a repetition of this study.

Results for the baseline measurement for time reporting showed a mean difference of -42 ms between the onset of the tone and the reported time of stimulus onset. These results were comparable and surprisingly close to the results reported by Libet et al. (1983), who reported a mean of -47 ms for accuracy of the time report although they measured with another stimulus. They used a near-threshold electric stimulus on the muscle instead of an auditory stimulus, which was used in the study reported here.

As expected, the readiness potential starts earlier than the reported time of conscious intention to move. Different methods were used to estimate the onset. Firstly, a method based on the relative criterion was used for the estimation of individual readiness potential onsets in order to compare these to the individual means of reported intention. Secondly, a regression-based method was used for further investigation of the waveform. In contrast to the findings of Libet and colleagues, who reported onsets of -1050 ms for the preplanned movement and -550 ms for the self-initiated movement (Banks & Pockett, 2007; Libet et al., 1982; Libet, Gleason et al., 1983), the readiness potential onsets clearly occurred earlier in this study and did not differ across conditions. The earlier onset is consistent with other findings (Haggard & Eimer, 1999; Rigoni et al., 2011) and might occur due to pre-planning, which was reported by some participants in a qualitative post-inquiry. Because the onset of readiness potential fundamentally depends on the definition of the baseline period, another possibility could be that the readiness potential already starts prior to the baseline period. In the results of the segmented regression, the onset of readiness potential estimated by the first breakpoint either starts directly at end of the baseline period or, like in condition A, already showed a significant negative slope in the first segment. Therefore, it might be reasonable that the time interval between -2500 ms and zero (defined as when the button is pressed) only captures a late part of the readiness potential and that the real onset might be earlier.

Consistent with the findings of Haggard and Eimer (1999), the onset of readiness potential, as measured by a criterion-based method, was not found to covary with the time of conscious intention. Accordingly, the early readiness potential cannot be the cause of or a process directly related to the conscious awareness of intention to move. It is remarkable that, in the case of some participants, no clear negativity was found preceding the movement. Similar observations are reported in literature (Schurger, Sitt, & Dehaene, 2012), but rarely. If the readiness potential primarily reflects conscious motor preparation, negativity should be observable any time a movement is executed. A possible reason for the lack of readiness potential in some participants could also be errors in measurement. Therefore, it is questionable whether or not this phenomenon would be found again in a repetition of this study.

Unfortunately, no averaging for the decision with consequence could be performed due to the insufficient number of trials remaining after artifact rejection. There is no rule of thumb for the minimum number of trials necessary for an averaging. Kornhuber and Deecke (1965) mentioned that readiness potential is sometimes already observable with 20 trials, but often only after 100 trials. Measurement procedures and equipment have fundamentally changed over the course of the last half-century, but nevertheless, unfortunately, there is no literature reporting results on the lowest number of trials needed to obtain a stable readiness potential. However, there is evidence regarding other event-related potentials: Cohen and Polich (1997) report that 20 trials are sufficient for a reliable measurement of P300 amplitude and latency; 20 trials were also enough for a reliable measurement of the feedback-related negativity component in a non-clinical sample of students (Marco-Pallares, Cucurell, Münte, Strien, & Rodriguez-Fornells, 2011); for the error-related negativity component, a minimum of six to eight trials was shown to be sufficient (Olvet & Hajcak, 2009). Taking into consideration that the laboratory situation available for measurement is not optimal due to a lack of insulation from acoustic and electro-magnetic noise and other potential sources of error, the target for the remaining number of trials after artifact rejection for the self-determined decision in an upcoming study was set to a minimum of 20 trials. In order to reach this aim of obtaining more artifact-free decisions, the experimental procedure was modified in terms of the instruction, programming, and duration.

Study 2: Prolonged experimental procedure to provoke more self-paced decisions

Method

The duration of the whole experiment was extended to approximately 3.5 h, depending on the time needed to place the electrode cap for the EEG measurement.

Participants

Readiness potentials of 28 participants were measured. Data from 10 participants were excluded from analysis for various reasons: Data from five participants had to be excluded due to a loose contact in EEG cap adapter and therefore temporary data loss, data from 4 participants were excluded due to bad impedances while recording, and data from one participant was incomplete because of problems that occurred during data storage, resulting in a sample of n=18 participants (13 female, 5 male; average age = 21.28 years, SD=2.67). All participants had normal or corrected to normal eye-sight; three of them reported tendencies to left-handedness according to the Edinburgh Inventory (Oldfield, 1971), but were used to handling a regular gamepad. All participants were students majoring in Business Psychology and received course credit for their participation.

Procedure

Participants worked on the same modified Libet task used in study 1. The distance to the monitor was adjusted to one meter in this set-up, the radius of the clock was approximately 3.1 cm, resulting in a visual angle of 1.8°, as was the case in study 1. The sequence of every trial remained the same as in study 1, but the number of trials was extended to 120 trials for each of the four conditions, which led to a longer duration of approximately 40 minutes per condition. Therefore, the duration of the whole experiment was approximately 3.5 hours. All movements (button presses) were again performed with the thumb of the right hand on a gamepad. Some modifications were made for study 2 concerning programming the break request. In order to avoid artifacts, a delay of 800 ms was included after a break request. Participants were instructed that there would be a short delay after initiating a break and that they should refrain from blinking or moving until the break display appears. During this delay period, participants were presented a display that said, "please do not move" in order to remind participants of this instruction.

Experimental conditions

Each participant worked on five different conditions, the four conditions from study 1 and an additional 2-back task. An overview of the experimental conditions is provided in table 4. The order of the experimental conditions was changed. In contrast to study 1, the first condition was always the baseline condition in order to check the accuracy of time reporting (condition D), in which no EEG data was recorded. This change was intended to tighten the experiment duration, which had already been extended by increasing the trial number, resulting in a duration of 3.5 to 4 hours for the whole experiment, but the goal was also to reach passable impedances without long waiting times. After the baseline condition, participants worked on conditions A, B, and C in a fully balanced order. Subsequently, they completed an additional 2-back task with a duration of 30 min. The task was to compare the alphabetic letter currently presented to the letter presented penultimately (exposure duration was 1500 ms and latency between letters was 500 ms). If the letters matched, participants had to press the same button that served as movement button in the other conditions. In this condition, the requested break was restricted to 30 s and afterwards it was automatically switched back to the task. The breaks did not prolong the duration of the task, which incentivized their usage, as they provided a possibility to escape the demanding but monotonous task. The only purpose of the 2-back task was to prolong the duration of the experiment and to provide a demanding task, thus enabling the observation of a higher amount of break requests.

Table 7. Overview of the experimental conditions in study 2.

<table>
<tr><td>D</td><td colspan="2">Baseline measurement: Participants were asked to report the onset time of a tone delivered randomly.</td></tr>
<tr><td>A*</td><td>Self-initiated movement: Participants pressed a button voluntarily at the time of their own choosing.</td><td rowspan="4">Self-determined decision: In conditions A, B, C, and N, the participants had the possibility to request a break by pressing a button whenever they wanted to. The break was explained to the participants as having a recreational function (not as an experimental condition).</td></tr>
<tr><td>B*</td><td>Self-initiated movement with time report: Participants pressed a button voluntarily at the time of their own choosing and were asked to report the time when they felt the conscious intention to move.</td></tr>
<tr><td>C*</td><td>Preplanned movement: Participants were asked to press a button at a pre-set clock time</td></tr>
<tr><td>N</td><td>2-back task: Additional task to extend the experimental duration with the aim of increasing the number of break request observations</td></tr>
</table>

Note. *Conditions presented in fully balanced order.

Data Acquisition

EEG data was measured and processed in a similar way as in study 1, with a few modifications to the montage and sampling rate. Ag/AgCl electrode caps were used to record EEG at electrode positions Fp1, Fpz, Fp2, F1, F3, Fz, F2, F4, FC1, FC3, FCz, FC2, FC4, C1, C3, Cz, C2, C4, P3, Pz, and P4 according to the 10/20-system with AFz as ground electrode. Impedances were kept below 20 kOhm (ANT Neuro B.V./eemagine Medical Imaging Solutions GmbH, 2013). Data were analyzed offline using asalab™ software (Version 4.9.1; ANT Neuro, Enschede, Netherlands). As opposed to study 1, the sampling rate was 1024 Hz, but further processing remained the same as in study 1: Signal was re-referenced to linked ear mastoids, bandpass-filtered from 0.016 to70 Hz and notch-filtered, epochs were time-locked to button press using an interval between -2500 ms and +300 ms with the interval between -2500 ms to -2300 ms as a baseline. Epochs containing artifacts were rejected on the basis of visual inspection after automatic artifact detection (peak to peak <100 μV). Onset of the readiness potential in Cz waveform was determined as in study 1.

Analysis and Results

Differences in readiness potential according to the tasks of experimental condition

The number of artifact-free epochs is slightly higher than in study 1, on average resulting in 52.5 trials per participant and condition (condition A: M=53.22, SD=20.12; condition B: 49.61, SD=19.20; condition C: M=54.67, SD=15.81). Figure 3 shows the grand average of the 18 participants for self-initiated movements with (condition B, red line) and without reporting the time of conscious intention (condition A, yellow), and for preplanned movements (condition C, grey). Like in study 1, the epoch between 2500 ms and 100 ms prior to pressing the button was subdivided into the same pre-movement time windows defined in table 2. Means, standard deviations, and upper limits for 95% CI of the amplitude at vertex (electrode position Cz) for time windows t01 to t12 are reported in table 8.

Repeated-measures ANOVA with a 3 (experimental conditions: A, B, C) × 11 (time windows: t02 to t12) × 3 (laterality: C3, Cz, C4) design was conducted for mean amplitude as a dependent variable. As the Mauchly's test of sphericity was significant in all relevant cases, Greenhouse-Geisser correction was used. Again, as in study 1, a significant main effect was found for time, $F(1.41, 23.98)=8.11$, $p=.005$, $\eta_p{}^2=.32$, reflecting the negative increase of the amplitude prior to the movement. Additionally, a significant interaction was found for time × laterality, $F(2.11, 35.81)=5.47$, $p=.008$, $\eta_p{}^2=.24$; this effect indicates that the readiness potential evolved asymmetric across the scalp. No other main effect or interaction was found to be significant. Therefore, again, no effect of the experimental condition on the amplitude of the readiness potential was detected.

Table 8. Mean amplitude at electrode position Cz for time windows of 200 ms between 2,300 and 100 ms before pressing the button (study 2).

	A			B			C		
	M	**SD**	**UL**	**M**	**SD**	**UL**	**M**	**SD**	**UL**
t01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t02	-0.41	0.91	-0.04*	-0.33	0.98	0.07	-0.17	0.85	0.18
t03	-0.35	1.38	0.22	-0.81	1.86	-0.04*	-0.51	1.14	-0.04*
t04	-0.56	1.87	0.21	-1.11	2.60	-0.04*	-0.67	1.68	0.02
t05	-0.85	2.01	-0.02*	-1.59	3.08	-0.32*	-1.16	2.05	-0.32*
t06	-0.97	2.44	0.03	-1.98	3.51	-0.54*	-1.38	2.32	-0.43*
t07	-1.22	2.63	-0.14*	-2.37	4.13	-0.67*	-1.66	2.84	-0.50*
t08	-1.65	3.02	-0.41*	-2.55	4.38	-0.76*	-1.97	2.96	-0.75*
t09	-2.05	3.64	-0.56*	-2.93	4.95	-0.90*	-2.32	3.30	-0.96*
t10	-2.28	4.10	-0.60*	-3.34	5.16	-1.22*	-2.67	3.57	-1.20*
t11	-2.32	4.73	-0.38*	-3.49	5.39	-1.28*	-3.29	3.92	-1.68*
t12	-1.94	5.43	0.28	-3.19	5.43	-0.96*	-3.22	4.43	-1.40*

Notes. Presented are the upper limits (UL) for one-tailed 95% CI of the mean amplitude at Cz. As the readiness potential is known to be negative, the lower limit is $-\infty$. *One-tailed 95% CI does not contain 0. t1 reflects baseline of EEG measurement.

Accuracy of time reporting

The averaged deviation between the onset of the tone and the corresponding time report in the baseline condition was -35.58 ms (SD=43.43 ms), slightly smaller but comparable to the result from study 1. The time estimate again differs significantly from onset time of the tone ($t(17)=-3.475$, $p=.003$; $d=0.82$; 95% CI [-57.18, -13.98]), indicating that participants reported the onset earlier than it actually occurred consistently in both studies.

Reported time of intentions and chronological order of mental events

As expected from the results of study 1, the averaged time of conscious intention to move (W, -223.85 ms before button press, SD=153.91 ms) in condition B was significantly earlier than the button press (t(17)=-6.171, p=.000, d=1.45, 95% CI [-300.38, -147.31]). In this condition, the averaged amplitude at Cz had already reached significantly negative values during t03 (time interval ranging from -2100 ms to -1900 ms before button is pressed; see table 8). Onsets of readiness potentials estimated with the relative criterion method are reported in table 9 and also indicate an onset of readiness potential earlier than the reported time of intention judgment. As in study 1, mean onset of readiness potential occurred significantly earlier than the mean reported conscious intention ($t(12)= -8.736$, $p=.000$, $d=2.4$). No correlation between these time points was found ($r=.06$, $p=.851$). The onset times for conditions A, B, and C did not differ significantly ($F(2,18)=2.248$, $p=.134$). In 24.1% of cases, no clear negativity was found preceding the button press, which was comparable with the result of 26.8% in study 1.

Table 9. Onset times of readiness potential based on relative criterion method in study 2.

	Condition		
	A	B	C
n	13	13	15
n (no clear negativity)	5	5	3
M	-1377.10	-1501.95	-1192.64
SD	490.38	555.93	577.88
Mdn	-1399.41	-1563.48	-1254.88
Min	-2304.69	-2239.26	-2129.88
Max	-568.36	-601.56	-472.66

Estimated breakpoints from segmented regression for conditions A, B, and C are shown in table 10. Similar to results of study 1, the first breakpoints for conditions B and C were quite early, this time they even occur during the baseline period. For condition A, the first breakpoint of regression was estimated a bit later than before (-1204.55 ms instead of -1667.9 ms in study 1), but again with a significantly negative slope for the first segment (see table 11). Figure 4 displays the segmented regression lines for conditions A, B, and C from study 1 and 2. The courses of the lines in the corresponding conditions are generally similar, but have variability in the estimation of the breakpoints. For a better understanding of the graphs, the slopes for each segment are shown in table 12, converted to display the relative change in amplitude in μV per second. Slopes for the first segment of condition A were significantly negative in both studies. For the other conditions, they were not significantly different from zero. Slopes differ especially in condition A and in the second segment for condition B.

Table 10. Breakpoint estimations in ms for study 2 from segmented regression.

	Condition					
	A		B		C	
Breakpoint	*Estimate*	*SE*	*Estimate*	*SE*	*Estimate*	*SE*
1	-1204.55	63.55	-2418.16	45.12	-2341.06	54.09
2	-712.80	52.99	-1386.81	158.10	-698.02	61.10
3	-267.00	19.21	-338.50	22.06	-261.61	11.96

Table 11. Results for slopes in segmented regression in study 2.

	Condition A			
	Estimate	*SE*	*t*	*p*
Intercept	-2.2870	0.1052	-21.74	<.000
Initial slope for time relative to 0	-0.0009	0.0001	-16.63	<.000
Δ slope after breakpoint 1	-0.0012	0.0002	-5.11	
Δ slope after breakpoint 2	0.0019	0.0004	5.08	
Δ slope after breakpoint 3	0.0060	0.0007	9.22	
	Condition B			
	Estimate	*SE*	*t*	*p*
Intercept	6.3602	10.4382	0.61	0.542
Initial slope for time relative to 0	0.0026	0.0042	0.61	0.543
Δ slope after breakpoint 1	-0.0047	0.0042	-1.10	
Δ slope after breakpoint 2	0.0005	0.0001	3.80	
Δ slope after breakpoint 3	0.0052	0.0005	10.04	
	Condition C			
	Estimate	*SE*	*t*	*p*
Intercept	1.2511	2.7605	0.45	0.650
Initial slope for time relative to 0	0.0005	0.0011	0.45	0.652
Δ slope after breakpoint 1	-0.0020	0.0011	-1.78	
Δ slope after breakpoint 2	-0.0012	0.0003	-4.60	
Δ slope after breakpoint 3	0.0086	0.0006	14.50	

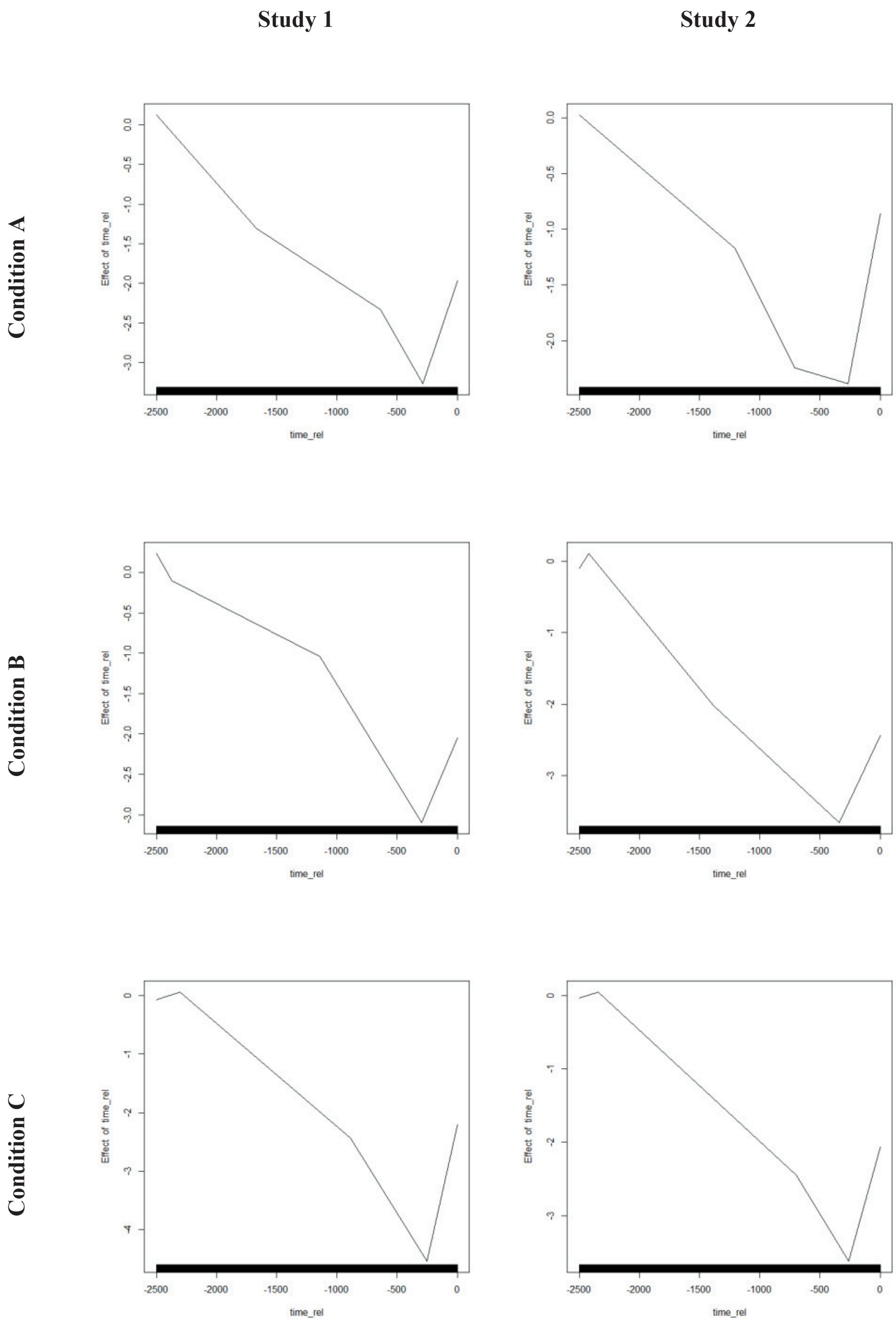


Figure 3. Segmented regressions lines for self-initiated movement (condition A), self-initiated movement with reporting time of conscious intention (condition B) and preplanned movement (condition C) from study 1 and 2.

Table 12. Slopes of segments in µV/s per study and condition.

	Study 1			Study 2		
Conditions	A	B	C	A	B	C
Segment 1	-1.7289	-2.6579	0.6707	-0.9259	2.5841	0.5137
Segment 2	-0.9942	-0.7646	-1.7651	-2.1749	-2.0673	-1.5192
Segment 3	-2.6481	-2.4463	-3.3168	-0.3239	-1.5624	-2.685
Segment 4	4.5675	3.5177	9.3133	5.7141	3.5947	5.9612

Self-determined decision

After artifact rejection, on average only 7.00 epochs for the break request per participant remained (*SD*=7.48, *Mdn*=4.5, *Min*=0, *Max*=25). Therefore, there were not enough epochs to calculate an average with a reliable signal-to-noise ratio. Although the participants were instructed to wait for a display announcing the beginning of the break, which was presented with a delay of 800 ms after pressing the button, most epochs that were time-locked to the break requests contained muscle or eye-movement artifacts.

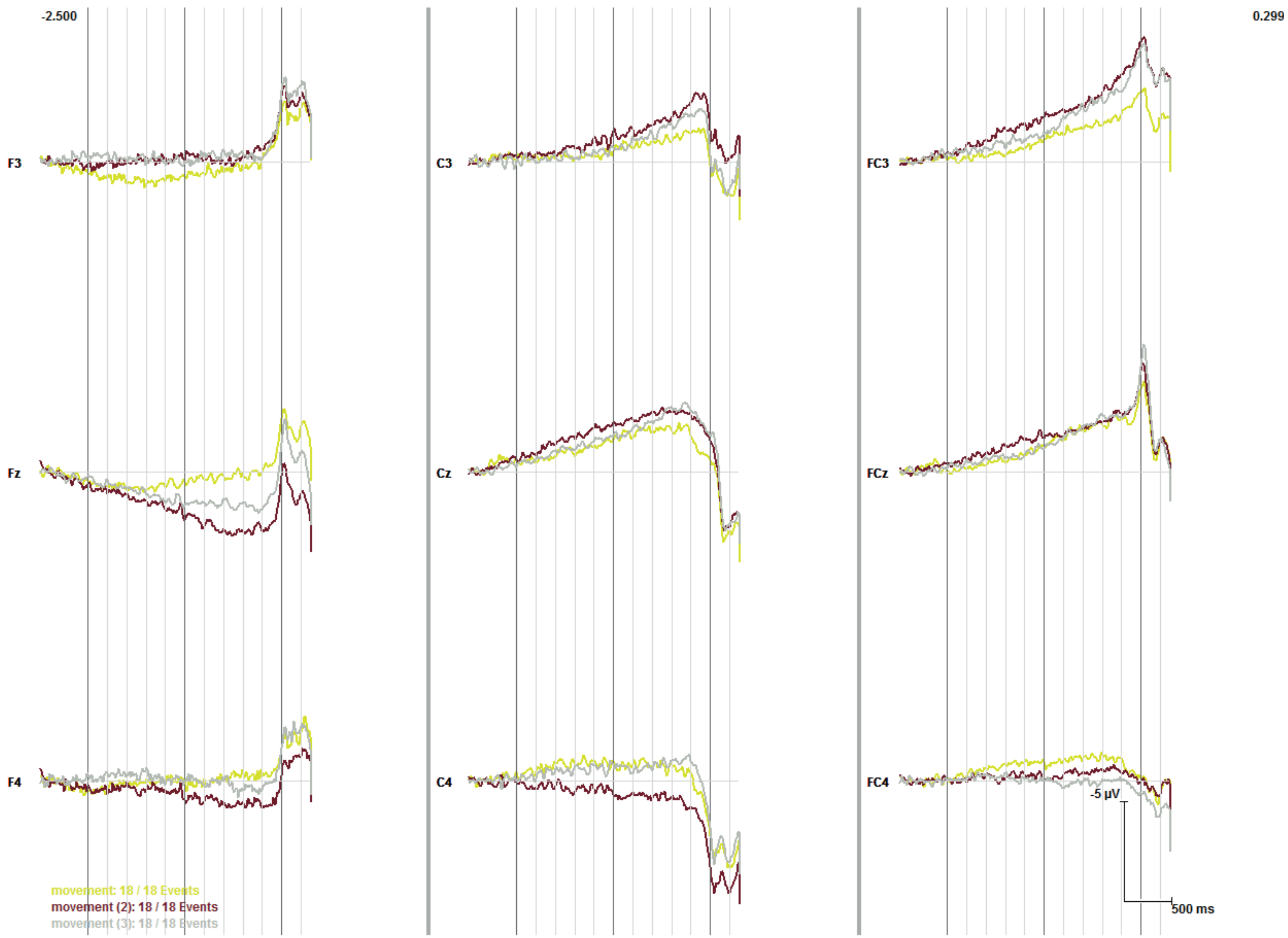


Figure 4. Grand average of 18 participants in 3 conditions: Self-initiated movement without time report (A, yellow) and with time report (B, red) and preplanned movement (C, gray).

Combined analysis of amplitude measured in study 1 and 2

Sample sizes in EEG studies are usually quite small due to the time-consuming measurements and high effort. The detection of a significant effect depends on the interplay of effect size, level of variance, and effect size of the estimated effect. Moreover, because of the limited possibility to influence the effect size or variance of an attribute, it is generally a good idea to increase the sample size. Therefore, an analysis was carried out with the combined data from studies 1 and 2. There are some differences in the experimental procedure in studies 1 and 2, leading to different trail numbers in the averages. Furthermore, the data of the two studies presented here were not collected for the purpose of a combined analysis, so any conclusions regarding this analysis are necessarily tentative. In order to compare the amplitudes at Cz between the two experiments, a mixed ANOVA was carried out with a 2 (between-factor experiment: study 1, study 2) × 3 (experimental conditions: A, B, C) × 11 (time windows: t02 to t12) × 3 (laterality: C3, Cz, C4) design. First of all, there was no significant between-subject effect regarding sample affiliation, $F(1, 39)=2.28$, $p=.139$, ${\eta_p}^2=.06$. Accordingly, measurements of study 1 and 2 did not differ substantially. Like in both single analyses, a main effect of time was found to be significant, $F(1.30, 50.52)=27.64$, $p=.000$, ${\eta_p}^2=.42$, indicating the slow negative increase of the readiness potential prior to the movement. Additionally, a significant main effect for laterality was found, $F(1.43, 55.92)=4.00$, $p=.036$, ${\eta_p}^2=.09$, indicating the asymmetrical progress of the readiness potential with a higher amplitude in the contralateral hemisphere. Unsurprisingly, the interaction between time × laterality was found to be significant, $F(2.11, 82.25)=8.30$, $p=.000$, ${\eta_p}^2=.18$, as the readiness potential is known to rise uniformly for both hemispheres in the early component of the potential and to rise more steeply and asymmetrically in the late component. Additionally, the interaction of condition × time was found to be significant, $F(3.01, 117.55)=4.20$, $p=.007$, ${\eta_p}^2=.10$. This effect indicates a different rise in readiness potential over time for the experimental conditions with different tasks plotted in figure 5. Mean amplitude at Cz did not differ significantly between condition A and B for all pre-movement time intervals (all p>.10).

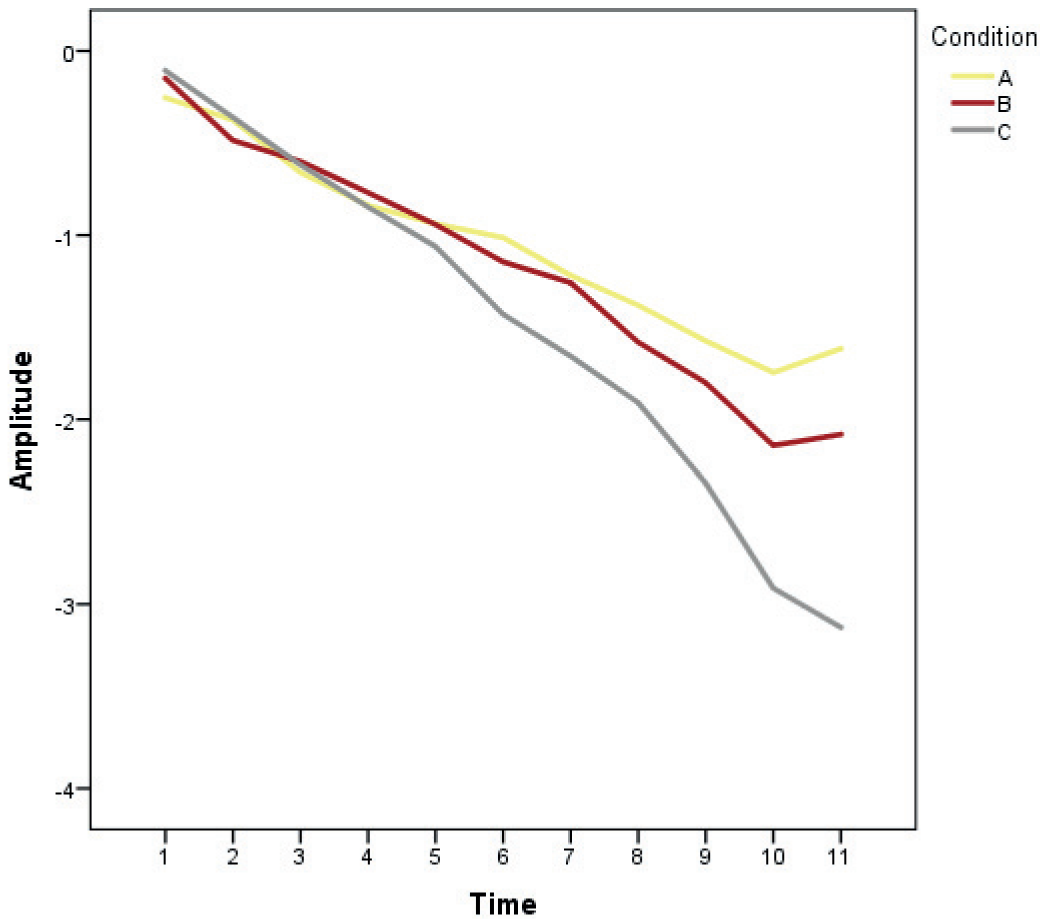


Figure 5. Plot of mean amplitudes from the combined analysis of both studies for the pre-movement time intervals.

Discussion

Summary of the results and limitations

The original experiment by Libet and colleagues did not truly contain a self-determined decision, but instead provided merely a choice of when to execute a pre-defined movement that had no consequences at all. Here an extension of the Libet experiment is reported that involves the possibility of making a free decision that has a realistic and immediate consequence for the acting subject. The experimental setting in an EEG laboratory is fairly restricted so as to prevent artifacts from unrelated movements, blinking, or other visual or auditory stimuli. In order to recognize systematic changes in neural activity derived from the cortex, which are small compared to the noise like the readiness potential, it is necessary to obtain several repetitions and to average these specific events. It is intrinsic to voluntary decisions that they cannot be enforced in an experimental setting; otherwise they would not be free and voluntary. The break-request function in this experimental procedure offers a simple possibility to include a realistic voluntary decision that is observable repetitively throughout the duration of several hours in the experiment, without a loss or shift in the meaning for the acting subject. No extension of the Libet paradigm to include a free decision has been reported in literature thus far, either because it has never been researched or perhaps due to the publication bias concealing unsuccessful attempts. Unfortunately, the attempt reported here was also not successful. No readiness potential preceding the decision could be investigated due to an insufficient number of epochs that were free of artifacts. Therefore, it remains unclear whether or not a movement that is operated as an expression of a conscious and voluntary decision is also preceded by a readiness potential.

The time report of an auditory stimulus delivered externally showed a small bias in reported time points, indicating that the time the stimulus occurred is dated back in time. Participants reported time points approximately 40 ms before the stimulus was actually given. The differences between reported times and actual times were comparable to the amount reported in the literature (Libet, Gleason et al., 1983) and stable throughout both studies. This suggests that participants were able to read the clock, or more precisely, were able to report reliable time points with just a small systematic deviation.

The Libet experiment has been replicated several times, as has its well-known finding, namely that the readiness potential as an indicator of motor preparation begins earlier than the intention to move is formed. Additionally, in both studies reported here, in the condition involving a self-initiated movement with time reporting, the onset time of readiness potential preceded the reported time of conscious intention to move, which means that the time order of the Libet experiment was replicated. However, it must be noted that there is some variation in the time of conscious intention throughout the literature (Banks & Pockett, 2007; Trevena & Miller, 2002), sometimes also ranging into post-movement times span like in Miller et al. (2011) who reported a mean of 30 ms (post-movement; range for will was from -358 ms to +421 ms) for time of will. In another study, the conscious decision to move was found temporally earlier than the onset of the lateralized readiness potential, suggesting that the preparation of the specific movement (e.g. movement execution with left or right hand) happens consciously (Trevena & Miller, 2002). Also in the two studies reported here, some variance in intention times was found from study 1 (95% CI [-185.57, -104.20] to study 2 (95% CI [-300.38, -147.31]). In general, the task of introspection and estimation of perceived will or urge to move is seen critically due to possible bias (Joordens, Spalek, Razmy, & van Duijn, 2004; Joordens, van Duijn, & Spalek, 2002). It has been shown that the judgment of will can be altered when a delay in visual or auditory feedback is established within the movement, indicating that the will judgment is more an inference of the perception of movement or of an available cue for the situation (Banks & Isham, 2009). Furthermore, it has been shown that not only efferent signals, but also sensory feedback (re-afferent signals) contribute to the awareness of movements (Obhi, Planetta, & Scantlebury, 2009) and it is suggested that both of these signals may also contribute to the experience of intentions of movements (Strother & Obhi, 2009).

Taken altogether, it is suggested that the judgment of the intention to move could contain systematic errors.

Another uncertainty is the definition of the onset of readiness potential. As a methods for onset determination, Libet, Gleason et al. (1983) used what was called the eyeball inspection (which was checked by a second investigator). As a statistical approach, they determined the time point according to 90% of the area under the readiness potential waveform preceding the movement onset. In the study reported here, the time point when the readiness potential finally rises over 20% of the maximum amplitude was chosen as the definition of onset. Additionally, a regression method was used because it was assumed that the first definition is less likely to produce smearing artifacts (Trevena & Miller, 2002) whereas the second method gives a better insight into the changes in slope over time. However, in all cases, the definition of onset depends on the choice of baseline, and results from segmented regression indicate that the readiness potential might start earlier than the baseline interval because the baseline artificially sets the waveform to zero, but the estimation for the first breakpoint was found to be near or even in the baseline, or rather the initial slope was found to be negative, as in condition A in both studies. Probably the readiness potential is just the summation of a slowly ongoing, long-lasting fluctuation of activation (Jo, Hinterberger, Wittmann, Borghardt, & Schmidt, 2013; Schmidt, Jo, Wittmann, & Hinterberger, 2016; Schurger et al., 2012).

The task of time reporting was not found to have an impact on the readiness potential, as there was no difference in readiness potentials measured in the self-initiated movement condition with and without active introspection and retrospectively reporting the time of conscious intention. Taken together with the findings of Miller et al. (2011) it seems reasonable that the readiness potential might be an artifact of the presence of the clock or the monitoring thereof, rather than of the time-reporting task or introspective searching for the intention to move, as the relevance of the clock was experimentally manipulated here and the presence was manipulated in Miller et al. (2011).

Recently the readiness potential has been put in connection with timing more than once. The readiness potential could be linked to the demands of internal timing (Verleger et al., 2016). They examined the readiness potential according to different time intervals between movements and found that as the minimum time interval between movements increased, the onset of readiness potential started earlier in relation to the movement. Following a movement, the readiness potential starts at approximately the same time regardless of the minimum time interval. They conclude that if the readiness potential reflected a process that is necessary and sufficient for movement but unrelated to timing, it should be the other way around: The interval between movement and the onset of readiness potential for the next movement should vary in accordance with the minimum interval between movements, and the length of readiness potentials should be even (Verleger et al., 2016).

There are increasing indications that the readiness potential might not be what it was thought to be. Readiness potentials have also been measured before a decision not to move and they do not differ in onset or amplitude (Trevena & Miller, 2010). However, in this experimental procedure, a visual clock representation was also involved. Freude, Ullsperger, Krüger, and Pietschmann (1988) found the readiness potential to be influenced by different amounts of mental load in an anticipated arithmetic task, tending to higher amplitudes in anticipation of a task with high mental workload and time pressure. This finding is also inconsistent with the view of readiness potential as motor preparation. An interesting finding of the two studies presented here is that in some participants no pre-movement negativity was found at all. If the readiness potential indeed reflects processes related to motor preparation, it should be measurable when a movement is executed. However, the observation that neutral or positive amplitudes precede a voluntary movement is comparable to the findings of Schmidt et al. (2016). They reported an alternative explanation of the readiness potential as an accompanying phenomenon resulting from averaging slow ongoing cortical potentials, not as an indicator of decision-making processes or motor preparation. They propose that the initiation of a movement is more likely to occur during and/or is facilitated by a negative increase or peak of these ongoing fluctuations, but that it is also possible during a positive phase. The assumption that it is more likely during negative phases is based on the data that show the averaged signal from many repetitions results in a negative waveform. Furthermore, it is reported that in 33% of trials, a voluntary movement was preceded by a positive readiness potential (Jo et al., 2013; Schmidt et al., 2016).

In a combined analysis of data from study 1 and 2, a significant interaction for time and condition was found. This effect could result from averaging different proportions of positive and negative readiness potentials as described above. An alternative explanation could simply be the level of boredom or "mental indifference" in the task, which was mentioned as having an effect on the amplitude of the readiness potential, with lower amplitudes corresponding with higher states of mental indifference (Kornhuber & Deecke, 1965). In the second study, the participants were asked to rate the experimental conditions according to how boring they are. 15 of 18 participants rated the preplanned condition C as the least boring (the other three rated the self-initiated movement condition with time reporting as the least boring). Therefore, the difference in amplitude could be due to motivational differences.

Conclusion

To summarize, it is possible to include a free, voluntary, and realistic decision in a laboratory EEG study, but it has not been possible to obtain a sufficient number of decisions that are free of artifacts to enable averaging. So it remains unclear whether or not the readiness potential also precedes free decisions. However, it became clear that there are more problems than just the generalization from an artificial movement to a realistic decision. Firstly, the onset is difficult to define and it could likely be that the readiness potential is much longer than expected or just one part of another ongoing activation. Furthermore, at this point it is unclear what exactly the readiness potential reflects or which processes it is associated with. The most likely theory currently seems to be that it is a process related to timing. More research is necessary in order to clarify the role of the readiness potential in general, not only in relation to voluntary movements.

The retrospective retrieval of the time of conscious intention is not uncontroversial. It has been criticized for being biased and the question has also been posed as to whether an intention to press a button necessarily has to be consciously accessible or if the general will to press buttons is already formed and expressed at the beginning of an experiment. Using data from Libet-style studies to draw a conclusion on the existence or non-existence of free will seems arbitrary. The issue will remain more of a question of one's belief in free will than an empirically driven conclusion due to the many critical issues in the methodological procedure and also because the nature and function of readiness potential has not yet been clarified unequivocally.

3. Measurement of Belief in Free Will: Development of a Belief in Free Will Inventory in the German language (WiF; Willensfreiheits-Fragebogen)

Theoretical Background

As of now, the question of whether or not humans have free will remains one that cannot be reliably answered using scientific methods that are uncontroversial. Nevertheless the majority of people across different cultural backgrounds tend to believe in an indeterministic universe and the existence of free will. In a cross-cultural study examining intuitions about free will and moral responsibility, using vignettes describing a deterministic and an indeterministic universe, the majority of the subjects from the United States, Hong Kong, India, and Colombia admitted that the indeterministic universe was most likely to be the universe we live in. Moreover, they reported that in a deterministic universe without free will, people could not be responsible for their actions. Therefore, moral responsibility was not seen as being compatible with determinism (Sarkissian et al., 2010). In contrast, Nahmias and colleagues report that the majority of their participants attributed moral responsibility and the possibility of acting according to one's own free will to agents in deterministic scenarios (Nahmias, Morris, Nadelhoffer, & Turner, 2005) which apparently would mean that they have a kind of compatibilistic view of free will, although perhaps more in an unsophisticated or layperson's view, not according to a strict philosophical definition of compatibilism. These vignette studies provide an indication that people "naturally" tend to believe in free will, and they also give an example of how the belief in free will or a sense of agency is considered to be in close connection with a concept of moral responsibility (Baumeister & Brewer, 2012). But the binary scale of measurement in vignette studies is not useful for providing deeper insight into the structure of belief in free will, determinism, and related positions. For a finely scaled picture of belief in free will, validated instruments are necessary. Alongside vignette studies, belief in free will was researched in experimental settings involving methods to manipulate the individual level of belief in free will. Vohs and Schooler (2008) gave the first report of a study like this: They found more cheating behavior in participants whose belief in free will was weakened. Adapting the manipulation methods of Vohs and Schooler, other researchers investigated further behavioral consequences of changed levels of belief in free will (e.g. Alquist et al., 2013; Baumeister et al., 2009; Brewer, 2011; MacKenzie et al., 2014; Rigoni et al., 2011; Rigoni et al., 2012; Rigoni et al., 2013; Shariff et al., 2014; Stillman & Baumeister, 2010). Most of these studies combine correlative and experimental parts to provide evidence that changes in belief in free will are the cause of behavioral effects found in correlative studies. There is a contrast between the ways free will is considered, on the one hand, as a presumably stable trait measurable by questionnaire procedures, and on the other hand, as being suggestible in the manipulation studies, which raises the question of the stability of the construct. Estimates for retest reliabilities have not been reported in the literature. Findings of Ent and Baumeister (2014) support the hypothesis that belief in free will is less stable than a trait: They found that people's bodily states such as perceived sexual desire, physical tiredness, and the urge to urinate affect their beliefs about free will.

In the past few decades, inventories with different conceptualizations of free will were developed (e.g. Nadelhoffer et al., 2014; Paulhus & Carey, 2011; Rakos et al., 2008; Stroessner & Green, 1990; Viney, Waldman, & Barchilon, 1982; for an overview see Nadelhoffer et al., 2014). Some of them conceptualized belief in free will as a one-dimensional construct ranging from a deterministic view to a libertarian view, such as in the Free Will and Determinism Scale by Rakos et al. (2008) and the Free Will-Determinism Scale (FWD) by Viney et al. (1982). The assumption that free will and determinism are mutually exclusive and represent opposite poles of one dimension seems arbitrary and is problematic as it excludes the possibility of perceiving free will and determinism as being compatible.

Stroessner and Green (1990) were the first to take the possible multidimensional nature of free will into account by identifying three factors using exploratory factor analysis

(EFA), which were termed religious-philosophical determinism, libertarianism and psychosocial determinism. The authors used a pool of 19 items, several of which were adapted from Viney et al. (1982). They used an orthogonal varimax rotation in their analysis. Therefore, intercorrelations between these factors remain unclear. An unsatisfying aspect of that study is that the conglomeration of determinism and religious beliefs into the religious-philosophical determinism scale (six items) only contains items relating to God or a higher power controlling actions in the sense of a deus ex machina, which means that the scale is not useful for surveying people without Judeo-Christian beliefs. Another problematic aspect is the shortness of the Libertarianism scale, which has a total of four items.

The Free Will And Determinism Scale FAD-Plus (Paulhus & Carey, 2011) is a widely used instrument (e.g. MacKenzie et al., 2014; Rigoni et al., 2013; Shariff et al., 2014) consisting of four relatively independent subscales (*free will*, *scientific determinism*, *fatalistic determinism,* and *unpredictability*). The factorial structure was identified by EFA and validated by confirmatory factor analysis (CFA). The subscales showed mostly acceptable internal consistencies and reasonable but modest correlations with the locus of control construct, indicating that these constructs are related but not congruent. Furthermore, the belief in *free will* subscale has been shown to be a predictor that is distinct from locus of control regarding life satisfaction, meaning in life, gratitude, and self-efficacy (Crescioni, Baumeister, Ainsworth, Ent, & Lambert, 2015). Positive but small correlations between FAD-Plus subscales and the Big Five personality traits were found: Paulhus and Carey (2011) reported positive correlations for *free will* subscale with *extraversion* and *agreeableness*; Stillman et al. (2010) reported positive correlations for *free will* with *openness*, *conscientiousness* and *emotional stability* (they used an unpublished former version of the FAD-Plus). Correlations were also reported for *free will* and *internal locus of control*, *self-efficacy*, and *satisfaction with life* (see table 1 for an overview). According to Nadelhoffer et al. (2014), the FAD-Plus also suffers from methodological problems, namely intercorrelations that are difficult to explain, (a small but positive correlation for the subscale *fatalistic determinism* with *free will* and with *unpredictability*), the intermixture of the free will scale with items on responsibility and blame, and the FAD-Plus does not provide a sound way to measure compatibilistic/incompatibilistic tendencies in participants (for an overview see Nadelhoffer et al., 2014). Since the compatibilism or incompatibilism of free will and determinism is one of the pivotal topics in contemporary free will debates (Kane, 2011), the identification of individual tendencies to (in)compatibilism should be a central function of an instrument measuring belief in free will.

Table 13. Correlations of belief in free will and personality traits reported in other studies.

	Paulhus & Carey (2011)	**Stillman et al. (2010)**	**Crescioni et al. (2015)**
N	177	143	44/47
Openness	.03 [a]	.17* [b]	
Conscientiousness	-.04 [a]	**.25**** [b]	
Extraversion	**.20**** [a]	.03 [b]	
Agreeableness	.17 [a]	.07 [b]	
Neuroticism	-.07 [a]		
Emotional Stability		**.21*** [b]	
Internal locus of control	**.35**** [c]	**.23**** [d]	**.28*** [e]
Self-efficacy			**.35*** [f]
Satisfaction with life		**.32*** [g]	**.59**** [g]

Notes. * p < 0.5; ** p < .01; Measurements: [a] Big Five Inventory (John & Srivastava, 1999), [b] Ten-Item Personality Inventory (Gosling, Rentfrow, & Swann, 2003), [c] Multidimensional Locus of Control Inventory (Levenson, 1973), [d] Internality subscale of Locus of Control Scale (Levenson, 1974, as cited in Stillman et al., 2010), [e] Internal Control Index (Duttweiler, 1984), [f] General Self-Efficacy Scale (Chen, Gully, & Eden, 2001), [g] Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985).

The Free Will Inventory (FWI, Nadelhoffer et al., 2014) is the most recently developed instrument and addressed to avoid problems of previous questionnaires. It consists of two parts. The first part measures belief in *free will* (e.g. “people always have the ability to do otherwise”)*, determinism* (e.g. “everything that has ever happened had to happen precisely as it did, given what happened before”), and *dualism* (e.g. “the fact that we have souls that are distinct from our bodies is what makes humans unique”). The second part measures beliefs about the nature of free will (e.g. *"to have free will means that a person's decisions and actions could not be perfectly predicted by someone else no matter how much information they had"*) and moral responsibility (e.g. *"people deserve to be blamed and punished for bad actions only if they acted of their own free will"*). This inventory is the first that aims to differentiate between a compatibilistic and incompatibilistic view of belief in free will. In the end, Nadelhoffer and colleagues (2014) found some items that might be able to differentiate between compatibilism and incompatibilism according to the correlation patterns with the subscales of their first part of the inventory, but they provide no applicable scale for usage in research. An interesting further result is that they found that most people largely agree both with explicit statements of compatibility and with explicit statements of incompatibility (Nadelhoffer et al., 2014), which could be the reason why it might be difficult to develop a scale measuring (in)compatibilism.

Summarizing the literature on belief in free will, some positively connoted personality traits were found to correlate with belief in free will, but no concept was found to be congruent to belief in free will. Greater belief in free will seems to have positive outcomes on behavior. The temporal stability of belief in free will has never been reported before and remains unclear. Empirical findings addressing belief in free will from the German-speaking area are sparse. One possible reason for this could be the lack of validated measurements assessing belief in free will in German language. The present work addresses this gap and aims to examine the structure of the belief in free will in a German sample, to develop an instrument measuring belief in free will in the German language, and to take the first steps towards validating this instrument. In the development of this instrument, some problematic aspects of other inventories will be avoided, such as the assumption of free will and determinism as opposite poles of one dimension or the assumption of the orthogonality of the resulting factors. Another aim is to avoid mixing belief in free will with religious aspects. Items measuring compatibilism or incompatibilism will also be part of the initial item pool in order to check if an independent factor of (in)compatibilism can be detected. Taking all of that into consideration, an initial item pool of 31 statements was generated. These statements derived from descriptions of philosophical perspectives on free will as described in Kane (2011) and Doyle (2010). An EFA was performed to explore the underlying structure of belief in free will (study 3). In study 4, this structure was validated using CFA and correlations of related concepts were tested. The temporal stability of belief in free will was estimated with a repeated measure design in study 5.

Study 3: Exploring the Underlying Structure of Belief in Free Will

Method

Sample

The sample consists of 172 participants whose age ranged from 18 to 70, with a mean age of 25.5 (SD=7.8 years). 74.4 % (n=128) reported being female, 22.1 % (n=38) reported being male, and 3.5 % (n=6) did not report their gender. Most of the participants were students (85.5 %, n=147), of which 36 % (n=53) were currently majoring in Psychology. Participation was voluntary and unpaid. As an incentive, three vouchers worth 50 € each were raffled among participants.

Materials

The initial item set consisted of 31 statements about free will which were generated from descriptions of philosophical points of view on free will (libertarianism, determinism, indeterminism, compatibilism, incompatibilism, and illusionism) and five items adapted from the FAD-Plus (Paulhus & Carey, 2011) for the German language: three items of the subscale *scientific determinism* and two items of the subscale *unpredictability*). Additionally, demographic characteristics (age, gender, educational background) were measured.

Table 14. Initial item pool.

Philosophical position	Statement
Libertarianism* (7 items)	I know the reasons for my own choices very well. The reasons that lead to a particular decision are always comprehensible for the person deciding. Belief in free will is essential for humanity. The fact that people can explain how they make decisions shows that they have free will. Decisions are not completely free but free among a range of options.

	Moral decisions are always made deliberately. Free will is the reason that human behavior and decisions are only partially predictable.
Determinism* (11 items)	People do not have free will. Today's actions change the world of tomorrow in a predetermined manner. There is a reason for everything - but it is possible that it cannot be explained yet. Even if an event is not explainable, it might follow certain logic. For every behavior, there is a specific combination of circumstances and reasons leading to this specific behavior. Under identical circumstances and facts in a decision-making process, the result will be the same. If all parameters in a situation involving a decision are exactly the same, the specific resulting decision will be made again. With complete awareness of the situation, every decision will be predictable. The future of a person is substantially predetermined by his or her genetic make-up and the environmental impact. The majority of our daily actions take place unconsciously – often subtly influenced by external factors. The behavior of a person is predefined by unconscious and uncontrollable forces, such as genes and the environment in which he or she lives.
Indeterminism* (4 items)	Evolution is a concatenation of chance events. Random events determine the history of mankind. The behavior of other people is not predicable. There is nothing you can do to change the future of humanity.
Compatibilism* (2 items)	With every action, people have the freedom to decide not to take action. Decisions are also influenced by unconscious motives.
Incompatibilism* (4 items)	People cannot be blamed for their actions if there is nothing like free will. If human actions are predetermined, humans cannot be responsible for their actions. If human actions are predictable by external circumstances, they cannot be free. Free will is the prerequisite for taking responsibility for one's actions.
Illusionism* (3 items)	Free will is just an illusion of the brain. The sense of personal responsibility for a decision just develops from weighing the alternatives in decision-making situation. People benefit from perceiving themselves as free humans.
Biological Determinism** (3 items)	Psychologists will eventually figure out all human behavior. People's biological makeup determines their talents and personality. As with other animals, human behavior always follows the laws of nature.
Unpredictability** (2 items)	Chance events seem to be the major cause of human history. Life is hard to predict because it is almost totally random.

Notes. *Items were derived from descriptions of the philosophical positions on free will as described in Kane (2011) and Doyle (2010); **Items from subscales of FAD which were translated to German.

Procedure

Data were collected by means of an online survey. Items were presented in randomized order. Participants were asked to rate their level of agreement with each statement concerning free will using a five-point Likert scale ranging from “strongly disagree (1)” to “strongly agree (5).”

Analysis and Results

Structural analysis

Eight items of the initial set had extreme means (<2.0 or >4.0; Likert-type scale ranging from 1 to 5 per item) and/or small variance (SD<0.9), indicating a high degree of consensus among participants, and therefore were excluded from further analysis (see table 15 for German wording, means, and standard deviations). Table 16 shows the items that were excluded from further analysis due to extreme means and small standard deviations, as these characteristics identify items with low discriminative power because there is great conformity in the answers of the participants.

Table 15. Means, standard deviations, and German wording of the WiF items.

No.	German wording of the item	English translation	Mean	SD
SD01	Sind alle Parameter einer Entscheidungssituation exakt gleich, wird genau dieselbe Entscheidung wieder getroffen.	If all parameters in a situation involving a decision are exactly the same, the specific resulting decision will be made again.	2.70	1.27
SD02	Sind alle Umstände und Gegebenheiten einer Entscheidungssituation identisch, führt es immer zur selben Entscheidung.	Under identical circumstances and facts in a decision-making process, the result will be the same.	2.50	1.22
SD03	Bei vollständiger Kenntnis der Situation lässt sich irgendwann jede Entscheidung vorhersagen.	With complete awareness of the situation, every decision will be predictable.	2.47	1.24
SD04	Für jedes Verhalten gibt es eine spezifische Zusammensetzung von Gegebenheiten und Gründen, die genau zu diesem speziellen Verhalten führen.	For every behavior, there is a specific combination of circumstances and reasons leading to this specific behavior.	3.55	1.13
SD05**	Psychologen werden irgendwann das menschliche Verhalten vollständig verstehen können.	Psychologists will eventually figure out all human behavior.	1.97	1.08
FW06	Die Tatsache, dass Menschen erklären können, warum sie Entscheidungen treffen, zeigt, dass es einen freien Willen gibt.	The fact that people can explain how they make decisions shows that they have free will.	2.84	1.13
FW07	Dass man menschliches Verhalten und Entscheidungen nur teilweise vorhersagen kann, liegt daran, dass es einen freien Willen gibt.	Free will is the reason that human behavior and decisions are only partially predictable.	3.29	1.14
FW08r	Der freie Wille ist nur eine Illusion des Gehirns.	Free will is just an illusion of the brain.	2.50	1.14
FW09r	Der Mensch hat keine Willensfreiheit.	People do not have free will.	2.01	0.97
FW10	Bei jeder Handlung bleibt als Entscheidungsfreiheit die Möglichkeit, diese Handlung nicht auszuführen.	With every action, people have the freedom to decide not to take action.	3.80	1.09
11*	Das Verhalten anderer Menschen ist nicht vorhersehbar.	The behavior of other people is not predicable.	2.85	0.96
12*	Entscheidungen werden nicht völlig frei getroffen, aber frei aus einer bestehenden Auswahl von Handlungsoptionen.	Decisions are not completely free, but free among a range of options.	3.92	0.92
IND13**	Zufälle scheinen die Hauptursache für die Entwicklung der Menschheit zu sein.	Chance events seem to be the major cause of human history.	2.86	1.03
IND14	Die Evolution ist eine Verkettung von Zufällen.	Evolution is a concatenation of chance events.	3.16	1.13
IND15	Zufällige Ereignisse bestimmen die Menschheitsgeschichte.	Random events determine the history of mankind.	3.29	1.01

IND16**	Das Leben ist schwer vorauszusagen, weil es fast komplett dem Zufall unterliegt.	Life is hard to predict because it is almost totally random.	2.59	0.99
17*	Es gibt für alles einen Grund - es kann allerdings sein, dass dieser noch nicht erklärbar ist.	There is a reason for everything - but it is possible that it cannot be explained yet.	3.73	1.11
BD18**	Größtenteils bestimmt die biologische Veranlagung eines Menschen seine Persönlichkeit und seine Begabungen.	People's biological makeup determines their talents and personality.	2.70	0.98
BD19	Die Zukunft eines Menschen ist maßgeblich durch seine genetischen Anlagen und seine Umwelt vorbestimmt.	The future of a person is substantially predetermined by his or her genetic make-up and the environmental impact.	3.13	0.99
BD20	Das Verhalten einer Person ist vorgegeben durch unbewusste, unkontrollierbare Kräfte wie Gene und die Umgebung, in der sie lebt.	The behavior of a person is predefined by unconscious and uncontrollable forces, such as genes and the environment in which he or she lives.	3.03	0.92
BD21**	Ebenso wie bei anderen Tieren, folgt das menschliche Verhalten den Gesetzen der Natur.	As with other animals, human behavior always follows the laws of nature.	3.24	0.98
22*	Der Großteil unseres täglichen Handelns läuft unbewusst ab - oft subtil von außen beeinflusst.	The majority of our daily actions take place unconsciously – often subtly influenced by external factors.	3.75	0.94
23*	Heutige Handlungen verändern die Welt von Morgen in vorbestimmter Weise.	Today's actions change the world of tomorrow in a predetermined manner.	3.01	1.10
INC24	Wenn es keinen freien Willen gibt, kann man Menschen für ihre Handlungen nicht verantwortlich machen.	People cannot be blamed for their actions if there is nothing like free will.	2.71	1.31
INC25	Wenn menschliche Handlungen vorbestimmt sind, kann man dafür nicht verantwortlich sein.	If human actions are predetermined, humans cannot be responsible for their actions.	2.50	1.20
INC26	Willensfreiheit ist die Voraussetzung um Verantwortung für Handlungen zu tragen.	Free will is the prerequisite for taking responsibility for one's actions.	3.60	1.05
27*	Wenn menschliche Handlungen aus äußeren Umständen vorhersagbar sind, sind sie nicht frei.	If human actions are predictable by external circumstances, they cannot be free.	2.81	1.15
28*	Moralische Entscheidungen werden immer willentlich getroffen.	Moral decisions are always made deliberately.	2.97	1.07

Notes. Subscales: SD=Situational Determinism, FW=Free Will, IND=Indeterminism/Chance, BD=Biological Determinism, INC=Incompatibilism; *Item was part of the first version of the inventory but is not included in the final version; **Item adapted from FAD-Plus (Paulhus & Carey, 2011); r: reversed item.

Table 16. Items excluded due to extreme means (<2 or >4) and/or small variance (SD<.9)

No.	German wording	English translation	MW	SD
29	Auch wenn ein Ereignis nicht erklärbar ist, kann es sein, dass es einer bestimmten Logik folgt.	Even if an event is not explainable, it might follow certain logic.	4.15	0.86
30	Es gibt nichts, was man tun kann, um die Zukunft der Menschheit zu verändern.	There is nothing you can do to change the future of humanity.	1.67	0.80
31	Ich kenne die Gründe für meine eigenen Entscheidungen sehr gut.	I know the reasons for my own choices very well.	3.47	0.85
32	Die Gründe, die zu einer bestimmten Entscheidung führen, sind für die Person, die sie trifft, immer nachvollziehbar.	The reasons that lead to a particular decision are always comprehensible for the deciding person.	2.12	0.85
33	Der Glaube an den freien Willen ist notwendig für die Menschheit.	Belief in free will is essential for humanity.	4.10	0.88
34	Aus dem Abwägen von Alternativen in Entscheidungssituationen wächst erst das Gefühl der persönlichen Verantwortung für die Entscheidung.	The sense of personal responsibility for a decision just develops from weighing the alternatives in decision-making situation.	4.04	0.76
35	Menschen profitieren davon, sich selbst als freie Wesen zu erleben.	People benefit from perceiving themselves as free humans.	4.24	0.75
36	Entscheidungen werden auch von unbewussten Motiven beeinflusst.	Decisions are also influenced by unconscious motives.	4.39	0.78

An EFA was conducted on the responses to the remaining 28 items using the ordinary least square extraction method with an oblique rotation (oblimin) allowing the factors to correlate. Ordinary least square method (OLS; 'minres' from R package *psych*; Revelle, 2016) was chosen because a cross-validation with a CFA was planned for a following study and the OLS method is similar to maximum likelihood method but more robust, even for badly behaved matrices (Revelle, 2016). An oblique rotation was applied because the assumption of uncorrelated factors is arbitrary.

Data fulfilled all requirements for using an EFA: Referring to Kaiser-Meyer-Olkin measure, sampling adequacy is appropriate (MSA=.74; (Kaiser, 1974)); MSA values for the individual items were acceptable (>.5, except for item 14, "evolution is a concatenation of chance events," KMO=.48). Bartlett's test of sphericity indicates that the correlation matrix differs significantly from the identity matrix ($\chi^2(378)=1472.98$; $p<.000$). A parallel analysis was run to determine the adequate number of factors (Horn, 1965). This method compares the eigenvalues of the obtained data with the eigenvalues of simulated data and is more reliable than the Kaiser criterion, which often overestimates the number of appropriate factors (Zwick & Velicer, 1986). After the fifth factor, the obtained eigenvalues dropped below the simulated eigenvalues, suggesting a five-factor solution (see figure 6). This solution can be interpreted well and accounted for 38% of the total variance. The factors were labeled *situational determinism*, *free will*, *indeterminism/chance*, *biological determinism* and *incompatibilism/moral responsibility*. Small intercorrelations were shown between r=-.26 for the factors *free will* and *biological determinism* and r=.26 for the two deterministic factors (*situational* and *biological determinism*; see table 17 for intercorrelations between factors). Table 18 contains the factor loadings of this five-factor solution.

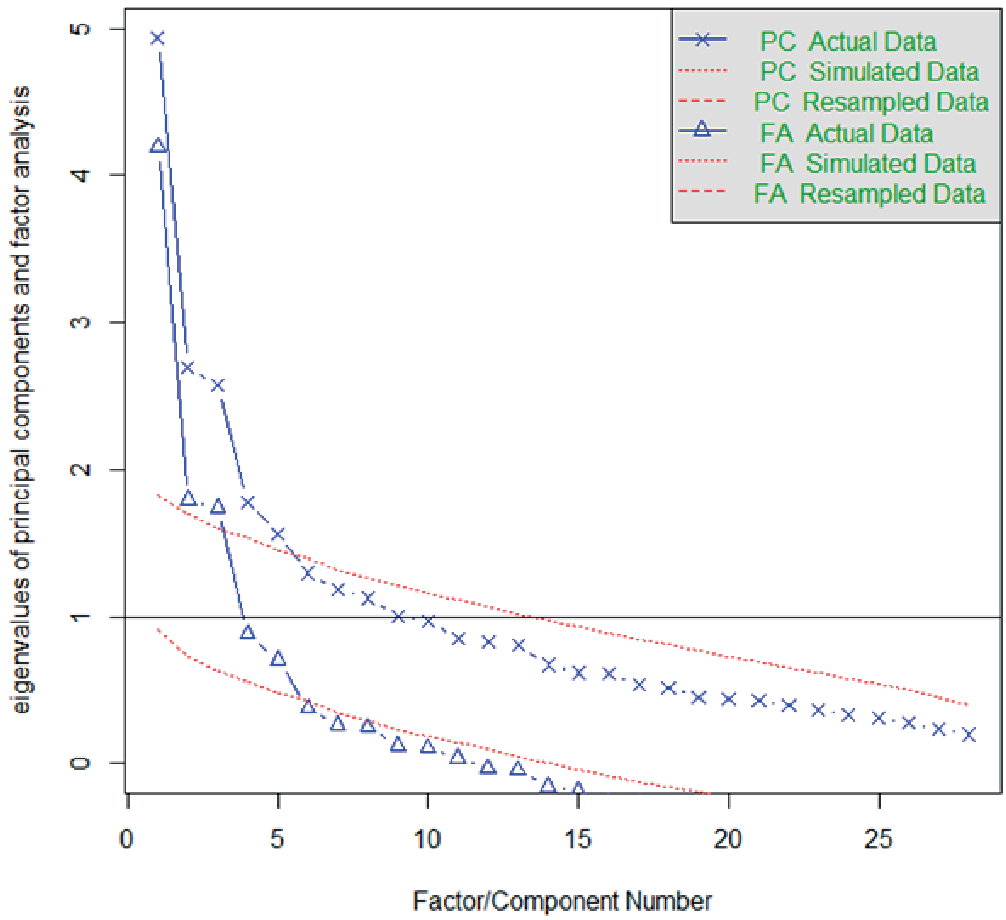


Figure 6. Scree plot of parallel analysis.

Table 17. Intercorrelations between factors.

Factors	1	2	3	4	5
1 Situational Determinism	1.00				
2 Free Will/Libertarianism	-0.24	1.00			
3 Indeterminism/Chance	0.04	-0.07	1.00		
4 Biological Determinism	0.26	-0.26	0.22	1.00	
5 Incompatibilism/Moral responsibility	0.11	0.23	0.00	-0.01	1.00

Table 18. Pattern matrix of EFA.

Items	SitD	FW	Indet	BioD	Inc
1. If all parameters in a situation involving a decision are exactly the same, the specific resulting decision will be made again.	**0.88**	0.12	0.05	-0.05	0.01
2. Under identical circumstances and facts in a decision-making process, the result will be the same.	**0.81**	-0.11	0.03	0.02	0.02
3. With complete awareness of the situation, every decision will be predictable.	**0.63**	-0.16	-0.11	0.07	0.10
4. For every behavior, there is a specific combination of circumstances and reasons leading to this specific behavior.	**0.47**	0.04	0.08	0.04	-0.14
5. Psychologists will eventually figure out all human behavior.	**0.41**	0.01	-0.15	0.25	0.01
6. The fact that people can explain how they make decisions shows that they have free will.	0.04	**0.81**	0.00	0.11	0.06
7. Free will is the reason that human behavior and decisions are just partially predictable.	-0.17	**0.66**	0.09	-0.04	0.09
8. Free will is just an illusion of the brain.	0.04	**-0.50**	0.13	0.32	0.12
9. People do not have free will.	0.02	**-0.48**	0.07	0.33	0.09
10. With every action, people have the freedom to decide not to take action.	0.00	**0.46**	0.00	-0.15	0.04
11. The behavior of other people is not predicable.	*-0.25*	***0.32***	*0.13*	*0.10*	*0.09*
12. Decisions are not completely free, but free among a range of options.	*-0.02*	***0.24***	*0.08*	*0.10*	*0.04*
13. Chance events seem to be the major cause of human history.	0.09	0.08	**0.75**	-0.01	-0.03
14. Evolution is a concatenation of chance events.	-0.04	-0.16	**0.74**	-0.12	0.07
15. Random events determine the history of mankind	0.04	0.08	**0.65**	0.19	-0.02
16. Life is hard to predict because it is almost totally random.	-0.15	0.11	**0.46**	0.17	0.02
17. There is a reason for everything – but it is possible that it cannot be explained yet.	*0.10*	*-0.16*	***-0.18***	*0.14*	*-0.08*
18. People's biological makeup determines their talents and personality.	0.00	0.12	-0.06	**0.64**	-0.01
19. The future of a person is substantially predetermined by his or her genetic make-up and the environmental impact.	0.01	-0.05	0.05	**0.61**	-0.09
20. The behavior of a person is predefined by unconscious and uncontrollable forces such as genes and the environment in which he or she lives.	0.13	-0.08	0.10	**0.52**	-0.07
21. As with other animals, human behavior always follows the laws of nature.	0.10	-0.09	0.14	**0.41**	0.06
22. The majority of our daily actions take place unconsciously – often subtly influenced by external influences.	*-0.10*	*-0.05*	*0.01*	***0.34***	*0.17*
23. Today's actions change the world of tomorrow in a predetermined manner.	*0.04*	*0.00*	*-0.13*	***0.30***	*0.07*
24. People cannot be blamed for their actions if there is nothing like free will.	0.05	0.10	0.05	-0.02	**0.76**
25. If human actions are predetermined, humans cannot be responsible for their actions.	0.03	-0.08	-0.01	-0.02	**0.68**
26. Free will is the prerequisite for taking responsibility for one's actions	-0.02	**0.42**	-0.17	-0.11	**0.42**

27. If human actions are predictable by external circumstances, they cannot be free.	*-0.08*	*-0.18*	*-0.02*	*0.16*	***0.37***
28. Moral decisions are always taken deliberately.	*0.09*	***0.13***	*0.03*	*-0.10*	***0.13***

Notes. Factor labeling: SitD=Situational determinism; FW=Free will; Indet=Indeterminism/chance; BioD=Biological determinism; Inc=Incompatibilism; N=172; Ordinary least square extraction with oblimin rotation. Bold loadings indicate the highest loading(s) of an item. Items with loading <.4 on all factors are printed in italic.

Loadings smaller than .4 are considered non-substantial (Stevens, 2002). Seven items showed loadings smaller than .4 on all five factors. These items were not interpreted as prototypical for any of the factors and were not included in any subscale. Item 26, ("free will is the prerequisite to take responsibility for actions,") shows a double loading on two factors (*free will* and *incompatibilism*) to the same extent. The item was originally intended for the subscale *incompatibilism* and is used exclusively as part of this subscale to avoid double usage in two subscales. The internal consistencies of the resulting subscales were sufficient (see table 4 for α coefficients, means, and SDs of subscales).

Intercorrelations of subscales were slightly higher but comparable to the intercorrelations of the factors (see table 5) and well explainable: As expected, the two deterministic subscales correlated positively (r=.31) with each other but negatively with the *free will* subscale (r= -.32 for *situational* and r= -.34 for *biological determinism*). As intended, they are not contrary parts of one dimension. Therefore, it is possible to obtain a high score on both the *free will* and deterministic subscales, as was expected for a highly compatibilistic view of free will and determinism. The subscale *incompatibilism/moral responsibility*, which uses thoughts about moral responsibility to measure the extent of the belief that free will and determinism are contradictory, correlated positively with free will (r=.34). Participants with a high level of belief in free will are more likely to believe that free will is required to be responsible for one's actions. What seems more surprising is the positive correlation between *indeterminism/chance* and *biological determinism*. In regard to the content of the subscales, it is coherent: People with a stronger belief in biological determinism are more likely to agree that evolution is based on chance events, but the evolutionary principal of the survival of the fittest forms the present genetic make-up of creatures, and the resulting genetic make-up of a single human determines that person's behavior and intellect. There is no contradiction between indeterminism or chance and biological determinism because both aspects are part of a functional evolutionary process.

Table 19. Internal consistencies of the resulting subscales (measured by Cronbach's α; Cronbach, 1951)

Subscale	Items	α	Mean (SD)
Situational Determinism	5	.80	2.6 (0.88)
Free Will / Libertarianism	5	.78	3.5 (0.80)
Indeterminism / Chance	4	.74	3.0 (0.78)
Biological Determinism	4	.69	3.0 (0.70)
Incompatibilism / Moral Responsibility	3	.69	2.9 (0.94)

Table 20. Intercorrelations of subscales.

Subscales	1	2	3	4
1 Situational Determinism				
2 Free Will/Libertarianism	-.32*			
3 Indeterminism/Chance	.03	-.11		
4 Biological Determinism	.31*	-.34*	.25*	
5 Incompatibilism/Moral Responsibility	.06	.34*	-.05	-.10

Note. * p<.001

Preliminary Discussion

Results of study 3 indicate that belief in free will consists of five distinct factors. Based on their content, they can be labeled *situational determinism*, *free will*, *indeterminism/chance*, *biological determinism,* and *incompatibilism/moral responsibility*. The factorial structure is not unexpected and can be interpreted well. The resulting subscales are consistent with previous research. They show proximity to subscales of other inventories on belief in free will, while avoiding some of their methodological problems. Similarities in content can be found between the *free will* subscale and the corresponding subscales from FWI and FAD-Plus, *biological determinism* and *scientific determinism* from FAD-Plus, *situational determinism* and *determinism* from FWI, *indeterminism/chance* and *unpredictability* from FAD-Plus, and the subscale *incompatibilism/moral responsibility* has content that is slightly similar to the moral responsibility items in the second part of FWI. Similarities to the FAD-Plus are not surprising, as five items were adapted from the FAD-Plus. The observed intercorrelations between subscales are able to be explained well. Although the positive correlation between *indeterminism/chance* and *biological determinism* may be surprising at first glance, it makes sense based on scientific thoughts on the evolutionary process. One beneficial difference to other inventories is the additional factor of incompatibilism/moral responsibility. Despite its brevity it provides a realistic possibility to identify tendencies towards compatibilism or incompatibilism. In order to confirm the factorial structure of the EFA, study 4 was conducted using a CFA.

Study 4: Confirmatory Factor Analysis and Correlates to Related Constructs

The study was planned and conducted in collaboration with Meyke (2015) and additionally contained a translation of FAD+ (Paulhus & Carey, 2011). Results concerning correlations of constructs possibly related to belief in free will with the translated scale are reported in the bachelor's thesis of Meyke (2015).

Method

Sample

The sample consisted of 358 participants whose age ranged from 16 to 68 years, with a mean age of 25.9 (SD=8.6 years). 70.4 % (n=252) reported being female, 28.8 % (n=103) reported being male, and 0.8 % (n=3) did not report their gender. As in study 3, most of the participants were students (79.1 %, n=283) and took part voluntarily, incentivized by the opportunity to win one of three vouchers worth 50 € for an ecological online bookstore.

Materials

The following instruments were part of the data acquisition:

WiF: Belief in free will was measured with the 21 items resulting from study 1 using a five-point Likert scale ranging from *strongly disagree* (1) to *strongly agree* (5).

BFI-K: The Big Five personality traits were measured by a short Big Five Inventory (BFI-K; Rammstedt & John, 2005) which consists of 21 items (four items each for conscientiousness, extraversion, neuroticism, and agreeableness, as well as five items for openness) with a five-point Likert scale ranging from *strongly disagree* (1) to *strongly agree* (5) as an answer format. Internal consistencies reported by the authors ranged from α=.59 (agreeableness) to α=.81 (extraversion), the mean retest-reliability was r=.84.

Self-efficacy: General self-efficacy was measured using a short form of a questionnaire by Collani and Schyns (2014) consisting of 10 items (e.g. "if I want to reach an aim, I will usually accomplish it"). Participants respond using a 6-point Likert scale ranging from *totally true* (1) to *not true at all* (6). The authors reported an internal consistency of α=.89 for this scale.

Satisfaction with life: Life satisfaction was measured using the German version of the Satisfaction With Life Scale (SWLS; Janke & Glöckner-Rist, 2012). It is a five-item scale (e.g. "in most ways, my life is close to ideal") that has a seven-point Likert answer format anchored by *strongly disagree* (1) and *strongly agree* (5). Internal consistency for this version is high (α=.92; Glaesmer, Grande, Braehler, & Roth, 2011).

Internal locus of control: Perceptions of an internal locus of control were measured by the Internal Control subscale from the German adaption of Internality, Powerful Others, and Chance Scales (IPC) by Hanna Levenson (α=.91-.96; Krampen, 1981). The scale consists of eight items with a 6-point Likert-type format ranging from *very wrong* (1) to *very right* (6).

Self-esteem: Self-esteem was measured by the revised version of the German Rosenberg scale (Collani & Herzberg, 2003a) consisting of 10 items with a four-point Likert-type scale anchored by *strongly disagree* (1) and *strongly agree* (4). Collani and Herzberg (2003a) reported α=.84 for internal consistency.

Procedure

As in study 3, data was conducted by means of an online survey. Items were presented in randomized order within each questionnaire, and the order of questionnaires was randomized among participants.

Analysis and Results

Structural analysis

22 participants had randomly missing data and were excluded from analysis, therefore n=336 cases were used in the CFA. Like in the EFA, the model allowed for covariation between all factors. The model requires estimations of loading and error term for each of the 21 observed variables and 10 covariances between the latent factors. Item 26 is allowed to load on two latent variables (*free will* and *incompatibilism*) as it already showed double loading in study 3. A standard maximum likelihood procedure was used to estimate the path weights (*cfa* function of *lavaan* package for R; Rosseel, 2012).

Different measurements are available to evaluate the fit of a model. X^2 statistic compares the observed covariance matrix with the matrix predicted by the model and should be non-significant. In this case, X^2 statistic is highly significant ($X^2(178)=305.66$; $p=.000$). But this statistic tends to be too stringent in personality research (Hopwood & Donnellan, 2010; Nadelhoffer et al., 2014; Paulhus & Carey, 2011; Raykov, 1998). X^2/df ratio is 1.71, indicating good model fit. An inferential fit measurement that is often recommended when considering model complexity is the Root Mean Square Error of Approximation (RMSEA) in conjunction with a descriptive (absolute) fit index like Standardized Root Mean Square Residual (SRMR). Hu and Bentler (1999) recommend cutoff values of 0.06 for RMSEA and 0.08 for SRMR, with smaller or equal values indicating a good model fit. These fit indices indicate a good model fit for the present data: RMSEA=.046 (90% CI=[.037;.055]); SRMR=.062. For the sake of completeness, the Comparative Fit Index (CFI=.911), Tucker-Lewis Index (TLI=.9), Goodness of Fit Index (GFI=.919), and Adjusted Goodness of Fit Index (AGFI=.895) were calculated and fulfill commonly used criteria ($\geq$.9) but not the strict criteria ($\geq$.95) demanded by Hu and Bentler (1999). The standardized regression weights of the model are shown in figure 7.

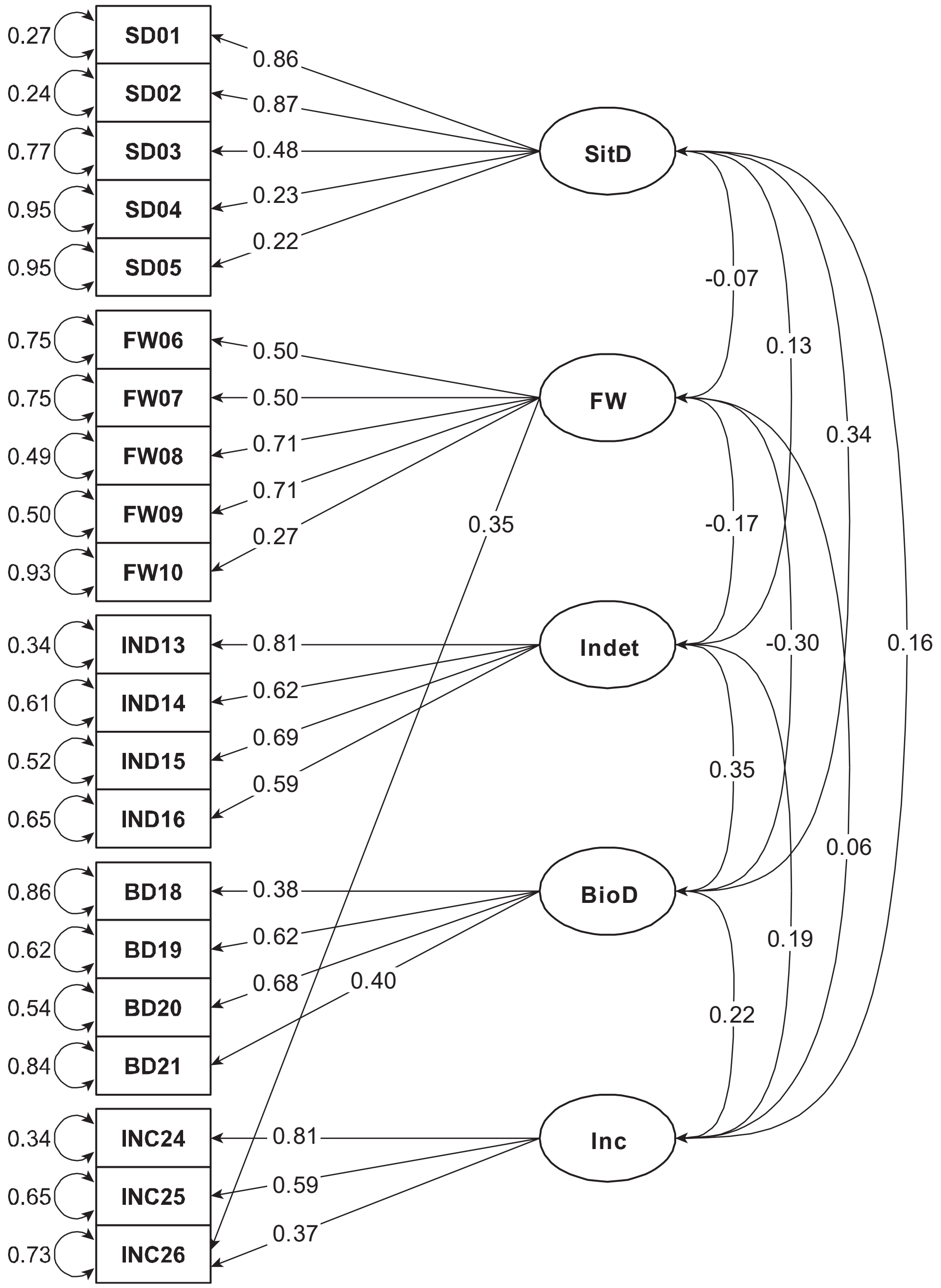


Figure 7. CFA model with standardized regression weights. Reversed items FW08 and FW09 were recoded. Factor labeling: SitD=Situational determinism; FW=Free will; Indet=Indeterminism/chance; BioD=Biological determinism; Inc=Incompatibilism/Moral Responsibility.

Table 21. Descriptive statistics of personality scales.

	N	missing	*M (SD)*
Situational determinism	350	8	2.61 (0.72)
Free will	354	4	3.65 (0.66)
Indeterminism / Chance	351	7	3.04 (0.82)
Biological determinism	351	7	2.98 (0.67)
Incompatibility / Moral responsibility	355	3	2.84 (0.88)
Openness	355	3	3.92 (0.65)
Conscientiousness	354	4	3.68 (0.68)
Extraversion	354	4	3.55 (0.90)
Agreeableness	354	4	3.05 (0.78)
Neuroticism	357	1	3.10 (0.87)
Self-efficacy	351	7	4.32 (0.72)
Internal control	349	9	4.32 (0.51)
Self-esteem	355	3	3.20 (0.55)
Satisfaction with life	355	3	5.13 (1.10)

Table 22. Correlations with Big Five traits and other possibly related constructs.

	Situational determinism	Free will	Indeterminism/ Chance	Biological determinism	Incompatibility/ Moral responsibility
Openness					
Conscientiousness		.15**	-.12*		
Extraversion					
Agreeableness			-.12*		
Neuroticism				.14**	
Self-efficacy		.14**			
Internal control	.28***	.24***		.12*	.15**
Self-esteem		.19***		-.13*	
Satisfaction with life		.13*			

Notes. Only significant correlations are shown; * $p<.05$, ** $p<.01$, *** $p<.001$.

Correlations with related constructs

Descriptive statistics for measured personality scales can be found in table 21. Correlations of WiF scales with measured and possibly related constructs and personality traits are shown in table 22. For the Big Five, only small correlations ($|r|\leq.15$) to WiF subscales are observable. For the subscale *free will,* only a correlation with *conscientiousness* (r=.15) was found. No other correlations to Big Five measurements were found to be significant for the subscale *free will*. Additionally, for the subscale *free will*, small positive correlations were found with self-efficacy, internal control, self-esteem, and satisfaction with life. The small and all positive correlations of internal control with *situational determinism*, *biological determinism* and *indeterminism/moral responsibility* are unexpected and contradictory.

Preliminary Discussion

The results of study 4 support the factorial structure and construct validity of the subscales of the WiF. Small correlations of the subscale *free will* to other related constructs like self-efficacy, internal locus of control and self-esteem were found. As expected from previous research, the *free will* subscale correlates positively with *self-efficacy* and *satisfaction with life,* however the correlations were much smaller than expected. A higher reported *self-esteem* correlates with a stronger belief in *free will,* but the correlation is smaller than reported by Rakos et al. (2008; they reported a correlation between belief in free will and self-esteem with r=.29, but measured belief in free will as a one-dimensional construct with determinism and free will as contrary poles). Regarding the Big Five, a correlation between *free will* and *conscientiousness* (r=.15) was found, which is also consistent with previous research, but lower than reported by Stillman et al. (2010, r=.25). All of the constructs measured each explain less than 8% of variance in belief in free will, which indicates that belief in free will is a distinct construct and not just a new measurement for an already well-known construct. The relationship between belief in *free will* and *internal control* is comparable to those reported in literature, but unexpected and contradictive is that also positive correlations between *internal control* and *situational determinism, biological determinism and incompatibility/moral responsibility* were found. One possible reason might be found in the psychometric characteristics of the subscale *internal control*. Some of the eight items only show small variances within this sample and therefore low discriminatory power. In this sample, the whole scale only reaches an internal consistency of α=.65, therefore the reliability is doubtful, perhaps due to the advanced age of the scale.

In summary, the factorial structure found in study 3 could be affirmed and belief in free will was found to be a discrete construct. Belief in free will was found to be associated with positive self-evaluations, which is consistent with the literature on belief in free will, although the correlations found here were smaller than reported elsewhere. An open question remains the temporal stability of belief in free will, as it is handled in the literature on the one hand as a presumably stable trait that is obviously measurable with questionnaire instruments, and on the other hand as a variable state that is manipulable using simple experimental interventions and/or changes depending on people's bodily states like sexual desire, physical tiredness, or hunger (Ent & Baumeister, 2014). Surprisingly, no trial to measure the stability of belief in free will over time can be found in literature. Therefore, it shall be the subject of study 5.

Study 5: Temporal Stability of Belief in Free Will

Method

Sample

N=64 undergraduate students at first assessment (t1) and n=46 undergraduate students at second assessment (t2) participated as part of a university lecture. Data was matched with a reproducible participation code to ensure anonymity. Unfortunately, only n=25 students who completed the questionnaire at both the first and second time of assessment were left after matching the data. 84% of them (n=21) reported being female, 16% (n=4) male. Mean age was approximately 20.8 years with a standard deviation of approximately 4.9 years (calculated from age intervals).

Materials and Procedure

A repeated measure design was used with a temporal lag of 13 weeks between t1 and t2. Participants completed a paper-and-pencil version of the WiF anonymously during a lecture at the beginning and the end of a semester. A reproducible participant code was used to match data from the measurement times t1 and t2.

Analysis and Results

Inspection of skewness and kurtosis as well as a Kolmogorov-Smirnov test led to doubts concerning the assumption of normal distribution for the subscales *free will* and *incompatibilism/moral responsibility*. Therefore, Spearman's rank correlation coefficients will additionally be reported to Person correlation coefficients. All subscales showed medium to high positive correlations between both measurement times (t1 and t2). See table 23 for estimates of retest reliabilities and descriptive statistics of both measurement times. Some subscales seem to be more stable than others, such as *indeterminism/chance*, *situational* and *biological determinism.* The scores of the first assessment of these subscales explain 40-56% of variance in the second assessment. For the subscales *free will* and *incompatibilism/moral responsibility*, retest stability is lower.

Table 23. Estimations for retest reliabilities.

Subscale	**t1** Mean (SD)	**t2** Mean (SD)	**r (t1, t2)**	**r_s (t1, t2)**
Situational determinism	2.66 (0.66)	2.60 (0.67)	*r*=.635, *p*=.001	*r*=.665, *p*=.000
Free will	3.68 (0.73)	3.59 (0.68)^{nn}	*r*=.524, *p*=.007	*r*=.467, *p*=.018
Indeterminism/chance	2.88 (0.72)	2.95 (0.82)	*r*=.748, *p*=.000	*r*=.755, *p*=.000
Biological determinism	2.92 (0.68)	2.80 (0.60)	*r*=.738, *p*=.000	*r*=.701, *p*=.000
Incompatibilism/moral responsibility	2.73 (0.89)	2.67 (0.80)^{nn}	*r*=.376, *p*=.064	*r*=.401, *p*=.047

Notes. ^{nn}: non-normal, data differ from normal distribution; r: Pearson correlation coefficient; $r_{s:}$ Spearman's rank correlation coefficient.

Discussion

Temporal stability of belief in free will

The correlations found for the subscales of WiF are lower than the classic retest reliabilities that are aspired to for stable personality traits. Due to the small sample size, the generalizability is fairly limited. In addition, systematic effects due to educational content during the semester cannot be excluded because all participants are enrolled in a Business Psychology major program. Nevertheless, the results give initial insight into the stability of belief in free will. Deterministic subscales and *indeterminism/chance* were found to be more stable than *free will* and *incompatibilism/moral responsibility*. The subscales *free will* and *incompatibilism/moral responsibility* appear less stable and possibly underlie more variability due to either internal causes like bodily states (as proposed by Ent & Baumeister, 2014) or external causes like biographical experiences, moral considerations, or experimental manipulations.

Conclusion

The development of the WiF, a questionnaire assessing belief in free will in the German language, was described. The final version consists of 21 items in five subscales: *situational determinism*, *free will*, *indeterminism/chance*, *biological determinism*, and *incompatibility/moral responsibility*. The subscales showed acceptable internal consistencies and the factorial structure observed by EFA was supported by CFA results, indicating a reliable factor structure. The first steps of construct validation were made. Correlations with related constructs were small, which indicates that belief in free will as measured here is a construct that is related to, but adequately distinct from, concepts like self-efficacy or internal locus of control. The WiF has a content proximity to instruments in the English language, especially the FAD-Plus, which is not surprising because the conceptual development followed the approach of FAD-Plus (Paulhus & Carey, 2011) and some of its items were adapted in the WiF. Like in the

FAD-Plus, FWI (Nadelhoffer et al., 2014), and the Free Will-Determinism Questionnaire by Stroessner and Green (1990), belief in free will and belief in determinism are not the extreme poles of a unidimensional scale, but instead are measured as different dimensions that are not necessarily independent. This approach also allows compatibilistic and incompatibilistic views on free will to be accessed, as it includes a scale measuring incompatibilism based on questions concerning moral responsibility (although it is very short scale). Belief in determinism is measured by two subscales, *situational* and *biological determinism*, which address either determinism though biological circumstances, like genes, or the reproducibility and predictability of a decision or behavior in a parallel situation. The subscale *indeterminism/chance* is comparable to the *unpredictability* subscale of the FAD-Plus. Overall, the intercorrelations were small ($r \leq .35$), but further research is necessary to clarify the relationship between these subscales. As all samples reported here consisted mostly of students, the factorial structure needs to be validated with a representative population sample. Also, cultural differences in the factorial structure could be of interest for further research, as the concept and understanding of free will might differ depending on religious affiliation and cultural background, which may include shared perspectives on free will. Whether belief in free will is a fluctuating state or a stable trait remains an unanswered question. Results of study 5 give a first indication, but generalizability is limited due to the small sample size in study 5. In spite of or due to these unanswered questions, we are hopefully providing a useful instrument for research in German speaking areas.

4. Experimental Manipulation of Belief in Free Will

Theoretical Background

Most people across different cultures intuitively believe that they possess free will (Sarkissian et al., 2010) and most people define free will as being capable of controlling their actions and decisions (Monroe & Malle, 2010). Despite that, changes in belief in free will might have an effect on experience and behavior, as literature on the belief in free will reports that this belief can be influenced by experimental manipulations and have an impact on different aspects of social interaction and behavior, emotions, and even cognitive functions (Alquist et al., 2013; Baumeister et al., 2009; Brewer, 2011; Clark et al., 2014; Lynn, Muhle-Karbe, Aarts, & Brass, 2014; Lynn, van Dessel, & Brass, 2013; MacKenzie et al., 2014; Rigoni et al., 2011; Rigoni et al., 2012; Rigoni et al., 2013; Shariff et al., 2014; Stillman & Baumeister, 2010; Vohs & Schooler, 2008). Belief in free will has been shown to have impacts on attributions of moral responsibility, e.g. weaker belief in free will leads to more cheating behavior (Vohs & Schooler, 2008), and it was shown that induced disbelief in free will provoked people to be more willing to forgive, as they perceived other persons to be less morally responsible for their actions (Brewer, 2011). Furthermore, it was shown that induced disbelief in free will reduced people's support for retributive punishment (Shariff et al., 2014) as they perceive others as being less blameworthy for their actions. A similar mechanism leads to less gratitude when people's belief in free will is undermined. The effect was mediated by attributions of free will in the benefactor: With lower belief in free will, the benefactor was perceived to be less sincerely or altruistically motivated, since the benefactor had no other choice but to help (MacKenzie, Vohs, & Baumeister, 2014). Belief in free was also shown to be associated with prosocial behavior, as lower belief in free will reduces the self-reported willingness to help others and leads to more aggressive behavior (Baumeister et al., 2009). Results of Alquist et al. (2013) suggest that higher levels of belief in free will make people act more autonomously and more resistant to pressures to conform with other opinions and suggestions. Furthermore, a lower level of belief in free will was shown to impair learning from emotional experiences (Stillman & Baumeister, 2010). Effects of an altered belief in free will are not only restricted to complex social behavior, but it also affects basic and unconscious cognitive functions. Induced disbelief in free will seems to alter the preconscious motor preparation measured by readiness potential due to an impaired feeling of being in control of one's own actions (Rigoni, Kuhn, Sartori, & Brass, 2011). Furthermore, it impairs the action-monitoring process, measured as reduced post-error slowing in a Simon task (Rigoni, Wilquin, Brass, & Burle, 2013). Lastly, it inhibits the effortful cancellation of prepotent behavior (Lynn et al., 2013; Rigoni et al., 2012).

The first study on the effects of belief in free will on social behavior was the study by Vohs and Schooler (2008), who induced disbelief in free will in two different ways: Firstly by reading of passages of *The Astonishing Hypothesis* (Crick, 1995) either claiming that free will is an illusion or a neutral passage not mentioning free will at all, and secondly by using an adaption of the classic Velten-style mood induction procedure (Velten, 1968), in which participants are confronted with statements on free will, determinism, or neutral statements. With the development of these manipulation methods, they started a new research field – the search for more behavioral effects of belief in free will. Most studies mentioned above used the original manipulation material or variations of it from the primal study of Vohs and Schooler (2008). For an overview of the manipulation methods applied in the studies mentioned above, as well as experimental conditions, see table 24. Except for the study of Shariff et al. (2014), variations of the original material were used in all studies. Most variance is in the number of statements involved in the Velten-style method.

Table 24. Applied methods for manipulating the level of belief in free will and experimental conditions in different studies.

Study	Crick method	Velten-style method
Alquist et al., 2013		Study 2 & 3 pro-free will, control & anti-free will group (10 statements)
Baumeister et al., 2009		Study 1 with pro-free will, control & anti-free will group, study 3 with pro-free will & anti-free will group (15 statements each)
Brewer, 2011	Study 2 & 3, anti-free will vs. control group	
Lynn et al., 2014	Main study, anti-free will vs. control group	
Lynn et al., 2013		Main study, anti-free will group vs. control (60 statements)
MacKenzie et al., 2014	Study 3 & 4, anti-free will vs. pro-free will	Study 2, anti-free will vs. control group (10 statements)
Rigoni et al., 2012		Study 1, anti-free will vs. control group (15 statements)
Rigoni et al., 2011	Main study, anti-free will vs. control	
Rigoni et al., 2013	Main study, anti-free will vs. control	
Shariff et al., 2014*	Study 2, anti-free will vs. control group	
Stillman & Baumeister, 2010		Study 2, 3 & 4, pro-free will, control & anti-free will group (15 statements)
Vohs & Schooler, 2008	Study 1, anti-free will vs. control group	Study 2, pro-free will, control & anti-free will group (15 statements)

Notes. *Unique manipulation methods: In study 3 through reading neuroscience passages vs. neutral science passages, in study 4 the between-group factor in the longitudinal design is the participation in an introductory cognitive-neuroscience vs. a geography class.

Published replication studies on these findings and empirical evidence from German-speaking samples are sparse, maybe due to the smaller community working on topics related to belief in free will or maybe due to a long-term lack of validated measurements of belief in free will available in the German language and manipulation material validated for German samples. It might also be explained by publication bias: Non-significant results or non-innovative results have an extremely lower chance of getting published. One key concept of the scientific process is replication, although unfortunately it is often neglected. Some effects decline when studies are repeated, and this may be due to the unavailability of unpublished results (Schooler, 2011). Replication studies might not be an innovative contribution, but every replication study contributes a unique impact, as it serves as proof of a published result in another or just slightly changed setting. Against the background of the replication crisis, it is highly important that new findings are subjected to critical examination before they achieve the status of textbook knowledge.

This work serves as preparation for replication studies on belief in free will with German-speaking samples. The aim of this work was to test whether methods to manipulate belief in free will reported in the literature are still successful when adapted to the German language. Therefore, two studies were conducted in order to test German versions of the two most common methods: manipulation by means of text excerpts from Crick's (1995) book *The Astonishing Hypothesis* and manipulation using a Velten-style procedure.

Study 6: Manipulation Study Using the Text Passages by Crick's *The Astonishing Hypothesis*

The study was planned and conducted in collaboration with Wett (2015) and contained a replication trial of study 2 from the article by Shariff et al. (2014) using translations of original material and of the FAD+ (Paulhus & Carey, 2011). Results concerning the translated belief in free will scale and the replication trial are reported in the bachelor's thesis of Wett (2015).

Method

Sample

In order to determine the required sample size, a power analysis was performed using G*power program (Faul, Erdfelder, Lang, & Buchner, 2007). Means and standard deviations were estimated based on results of Vohs and Schooler (2008), who reported descriptive statistics in each group for the free will subscale of the Free Will and Determinism Scale (FAD-4; Paulhus & Margesson, 1994, an unpublished preliminary version of FAD-Plus, cited in Paulhus & Carey, 2011, p. 96). Vohs and Schooler (2008) reported that participants showed weaker free-will beliefs (M=13.6, SD=2.66, n=15) than participants in the control condition (M=16.8, SD=2.67, n=15), indicating a large effect size of d=1.2. Assuming error probability for α=.01 and power=.99 using a one-tailed *t*-test for independent groups with equal sample sizes, the minimal required sample size was n=64 (n=32 for each group). Accordingly, 71 undergraduate students took part in this study to receive course credit for participation in empirical research. 73.2 % (n=52) reported being female, 26.8 % (n=19) male. Mean age was 21.9 years (SD=4.2 years).

Materials and Procedure

The experiment was conducted on three dates as group testing sessions. Participants were randomly assigned to experimental conditions, which resulted in group sizes of n=35 in the experimental group and n=36 in the neutral control group. Participants read one of two passages from the book *The Astonishing Hypothesis* (Crick, 1995), either claiming that free will is just an illusion (experimental group) or a neutral passage unrelated to free will. This procedure was developed and validated by Vohs and Schooler (2008) and has been successfully used in other studies (MacKenzie et al., 2014; Rigoni et al., 2011; Rigoni et al., 2013; Shariff et al., 2014). To ensure that participants read the passages carefully, they were informed that they would take part in a test of reading comprehension. After reading the passages, participants answered the 21 items of the WiF, two manipulation check items regarding determinism ("to what extent are your actions and decisions predetermined through biological, social or situational circumstances?") and free will ("how free do you perceive your actions and decisions?") measured on a visual analogue scale (with higher values indicating higher expression of the characteristic), and demographic questions. Additionally, participants answered German translations of the FAD-Plus (Paulhus & Carey, 2011) and the FWD (Rakos et al., 2008). They also read a fictional vignette involving a violent crime and were asked to recommend a prison sentence as hypothetical jurors. Results concerning these measurements were part of a study aiming to replicate the results of Shariff et al. (2014, study 2) and are reported elsewhere (Wett, 2015). In the final task, participants were asked to write a short summary of the passage they had read in the beginning of the experiment in order to ensure they had read it carefully.

Analysis and Results

One case has been removed from the sample due to a large amount of missing data and resulting doubts concerning the particpant's willingness to answer frankly. Only randomly missing data occurred in other cases, which were excluded from analysis concerning these data.

Control variables

No difference was found for age between groups: T(43.9)=1.8, p=.074 (df corrected for unequal variances; experimental group: M=21.0 years, SD=1.92; control group: M=22.78 years, SD=5.48). Gender and group assignment are independent ($X^2(1)=0.901$; *p*=.42). Subscale scores for WiF did not differ from normally distributed data except for the subscale *free will* in the control group. Therefore, parametrical and non-parametrical group differences will be reported. In some measurements, randomly missing data occurred, these cases were excluded from analysis pairwise.

Group differences

Contrary to expectations, no difference was found between the groups concerning the subscales of WiF and the visual analogue scale ratings. Table 24 shows the results of *t*-tests as well as means and standard deviations of the experimental groups regarding WiF subscales. Table 25 shows the results for the slider ratings.

Table 25. Group differences in WiF subscales (control vs. experimental group).

Subscale	**Group**	***n***	***M* (*SD*)**	***t*-Test**	**Mann-Whitney-*U*-Test**
Situational determinism	Control	35	2.59 (0.71)	*T*(67)=.17, *p*=.87	*U*=592.0, *p*=.97
	No free will	34	2.56 (0.73)		
Free will	Control	36	3.69 (0.63)	*T*(68)=-.59, *p*=.56	*U*=589.5, *p*=.79
	No free will	34	3.77 (0.51)		
Indeterminism / Chance	Control	36	2.98 (0.89)	*T*(68)=-.10, *p*=.92	*U*=584.5, *p*=.75
	No free will	34	3.00 (0.80)		
Biological determinism	Control	36	2.99 (0.61)	*T*(68)=-.75, *p*=.45	*U*=548.5, *p*=.46
	No free will	34	3.10 (0.61)		
Incompatibilism / Moral responsibility	Control	36	3.17 (0.84)	*T*(68)=-.50, *p*=.62	*U*=556.0, *p*=.51
	No free will	34	3.27 (0.96)		

Table 26. Group differences in determinism and free will slider ratings (control vs. experimental group).

Slider rating	Group	*n*	*M* (*SD*)	*t*-Test
Determinism	Control	35	4.76 (1.93)	*T*(67)=-.44, *p*=.66
	No free will	34	4.96 (1.87)	
Free will	Control	35	6.69 (1.86)	*T*(60.98*)=-1.75, *p*=.09
	No free will	34	7.37 (1.40)	

Notes. * Welsh-corrected *df*.

Preliminary Discussion

Although it was often reported to be successful, the manipulation of belief in free will by passages of Crick's book *The Astonishing Hypothesis* was not successful in this study. No differences in WiF subscales or manipulation check slider ratings were found between the experimental groups. This was unexpected but nevertheless interesting with regard to the current discussion about the reproducibility of psychological studies, initiated by the alarming findings of the Open Science Collaboration (2015). The generalizability of the results reported here is limited due to the sample size as well as educational background of participants, although it was chosen to be similar to the sample used by Vohs and Schooler (2008). Possible reasons for the failure of manipulation are manifold: Perhaps the participating students are less credulous or not easily convinced by a written text, or the "Crick method" is not suitable in the German language or not effective on German-speakers. Further studies should be conducted involving other manipulation methods like the Velten-style method also developed by Vohs and Schooler (2008). Furthermore, it could be possible that belief in free will as measured by WiF is more stable than it is an influenceable state. This may also be the reason for the failure of the attempt to manipulate belief in free will in this study. The excerpt of Crick's *The Astonishing Hypothesis* triggers a deterministic view on the free will debate. These subscales of WiF were found to be relatively stable in study 5. A method encouraging a stronger belief in free will, like in study 2 of Vohs and Schooler (2008), could be a more promising approach for experimental manipulation.

Study 7: Manipulation Study Using a Velten-Style Procedure

The study was planned and conducted in collaboration with Stöber (2016) and contained a direct replication trial of the study 3 reported in Crescioni et al. (2015) using German versions of the original material and a translation of the FAD+ (Paulhus & Carey, 2011). Results concerning the translated belief in free will scale and the replication trial are reported in the bachelor's thesis of Stöber (2016).

Method

Sample

As in study 6, a power analysis was conducted using the G*Power program (Faul et al., 2007). Estimations for means and standard deviations were taken from study 2 of Vohs and Schooler's report (2008) who conducted a between-design with three experimental groups (free-will condition: M=23.09, SD=6.42; control condition: M=20.04, SD=3.76; determinism condition: M=15.56, SD=2.79). Assuming equal sample sizes and equal standard deviations (which was estimated using the highest standard deviation reported), the effect size in a one-way ANOVA would be f=0.48. Minimum required sample size (with error probability α=.05 and power=.95) is n=72 (n=24 for each group). Accordingly, data of 75 undergraduate students were collected. 57.3 % (n=43) reported being female and the mean age was 21.7 years (SD=3.03 years). Participants who are enrolled in psychological fields of study (72.0 % of participants) received course credit for their participation, all others had the possibility to take part in a raffle with the chance to win a canteen voucher of 30 € as an incentive. Participants were randomly assigned to experimental conditions, resulting in subsample sizes of n=24 in the free will condition, n=25 in the determinism condition, and n=26 in the control group.

Materials and procedure

Data were collected as a group testing session with a duration of approximately 30 minutes. Participants were told that they were taking part in a study measuring memory performance. According to the experimental condition, they were shown a set of ten statements. Each statement was presented for 45 seconds, participants were asked to memorize it and then to rephrase the content in their own words in the 45 seconds that immediately followed. The sentences for all conditions were translated from those used by MacKenzie et al. (2014), who reported that they used a subset of the original statements from the study of Vohs and Schooler (2008). Example statements for the experimental conditions are: “I demonstrate my free will every day when I make decisions” for the pro-free will condition and “science has demonstrated that free will is an illusion” for the determinism condition. In the control condition, statements with facts that were irrelevant to free will were used, such as: “The Nile River in Africa is the world's longest river.” After memorizing and rewriting the statements, participants answered the 21 items of the WiF, an additional one-item manipulation check, (“I feel I am free to choose whatever I want to do in my life,” adapted from Baumeister et al., 2009), measured with a 5-point Likert scale ranging from strong disagreement to strong agreement, as well as demographic questions. Additionally, participants answered German translations of the FAD-Plus (Paulhus & Carey, 2011) and a questionnaire measuring meaning in life (Steger, Frazier, Oishi, & Kaler, 2006) in association with a related research question by Stöber (2016), which aimed to replicate a finding by Crescioni et al. (2015, study 3). Results concerning these questionnaires are reported in Stöber (2016).

Analysis and Results

No difference was found for age between groups: $F(2, 72)=1.03$, $p=.36$ (anti-free will group: $M=22.04$ years, $SD=2.76$; control group: $M=20.96$ years, $SD=1.86$; pro-free will group: $M=22.00$ years, $SD=4.16$). Gender and group assignment are independent

($\chi^2(2)$=0.21; *p*=.92). Subscale scores for WiF did not differ from normally distributed data. No significant differences between experimental groups were found, neither regarding the WiF subscales nor for the one-item manipulation check ($\chi^2(2)$=3.26, *p*=.20; *mdn*=4 for all groups). Descriptive statistics and results of ANOVA regarding WiF subscales are reported in table 26.

Table 27. Descriptive statistics and results of mean comparison.

					95% CI		
		n	*M*	*SD*	Lower	Upper	ANOVA result
Situational Determinism	Anti-free will	25	2.69	0.57	2.45	2.93	
	Control	25	2.45	0.77	2.13	2.77	
	Pro-free will	24	2.73	0.75	2.41	3.04	
	Overall	74	2.62	0.70	2.46	2.78	F(2,71)=1.13, *p*=.33
Free Will	Anti-free will	25	3.72	0.57	3.49	3.95	
	Control	26	3.51	0.61	3.26	3.75	
	Pro-free will	24	3.83	0.79	3.50	4.17	
	Overall	75	3.68	0.67	3.53	3.84	F(2,72)=1.58, *p*=.21
Indeterminism/ Chance	Anti-free will	25	3.22	0.83	2.88	3.56	
	Control	26	3.11	0.71	2.82	3.39	
	Pro-free will	24	3.14	0.91	2.75	3.52	
	Overall	75	3.15	0.81	2.97	3.34	F(2,72)=0.13, *p*=.88
Biological Determinism	Anti-free will	25	3.27	0.53	3.05	3.49	
	Control	26	3.13	0.67	2.85	3.40	
	Pro-free will	24	3.16	0.55	2.92	3.39	
	Overall	75	3.18	0.58	3.05	3.32	F(2,72)=0.42, *p*=.66
Incompatibilism/ Moral responsibility	Anti-free will	25	3.07	0.79	2.74	3.39	
	Control	26	3.26	0.98	2.86	3.65	
	Pro-free will	24	3.36	0.93	2.97	3.75	
	Overall	75	3,23	0,90	3,02	3,43	F(2,72)=0.67, *p*=.51

Discussion

Velten-style manipulation

As with the Crick manipulation method, no differences in WiF subscales or in the one-item manipulation check between the experimental groups were found for the Velten-style manipulation method either. Therefore the experimental manipulation using Velten-like statements was not found to be an effective way to manipulate belief in free will in this study. Like in study 6, sample size and educational background were chosen similar to the original study of Vohs and Schooler (2008) for a direct replication trial, but this sample composition also leads to limitations in the generalizability of results. Perhaps it would be easier to manipulate belief in free will when dealing with a more diverse sample in terms of age and educational background, or simply a sample without psychology undergraduates, as they could be suspicious.

Summary and conclusion

Successful experimental manipulations among samples of similar size and educational background to those reported in literature could not be replicated in the two trials reported here, neither with text excerpts from Crick's *The Astonishing Hypothesis* nor with the Velten-style statements. Furthermore, results regarding the unvalidated German translation of the FAD+ were not found to be significant in both studies either (Stöber, 2016; Wett, 2015). The reasons for this might be based on the translation of the material, small sample sizes, or the characteristics of the samples, such as age, educational background, and/or the proportion of students in the sample, although these easily accessible student samples were often used in literature that reported successful implementation of these manipulation methods. The paper-and-pencil method, which requires the participants to attend in person, was deliberately chosen as the method of data acquisition although it limits the reachable sample size. The sample sizes reported here were limited due to the size of the university where the studies were conducted.

The available student population is small, and the time span for conducting the procedure is short. (The longer the time span is, the higher the probability that the participants are biased, because they may have already heard of the study.) Besides these disadvantages, experimental effects of online studies are far less controllable. In the setting here, the risk of an online design would be a potentially higher drop-out rate in the anti-free will condition because participants are confronted with arguments that may be perceived as uncomfortable or unpleasant. This would lead to a sample with lower belief in free will, however, not as a consequence of the experimental manipulation, but instead because it would only consist of people who already had a low level of belief in free will prior to the experiment. This could be an additional explanation (besides the economic reasons and reachable sample size) for why most of the recently published studies involving belief in free will manipulations are online studies.

Although non-significant results are often not reported, the replication trials reported here are not the only unsuccessful attempts. Indications of the difficulties of replicating studies manipulating belief in free will arose in the replication project of the Open Science Collaboration (2015). Study 1 of the article by Vohs and Schooler (2016) was one of the 100 studies that the authors aimed to replicate, and was also one of the studies that could not be successfully replicated (Embley, Johnson, & Giner-Sorolla, 2015). A second unsuccessful attempt to manipulate belief in free will was reported by Macrae and Pitts (2015), who aimed to replicate the study of Rigoni et al. (2011), but failed to manipulate the belief in free will in their participants successfully. Due to publication policies, the number of failed and unpublished replication trials is suspected to be much higher. Furthermore, the effect size could be much smaller than reported. For example, Vohs and Schooler (2008) reported a high effect size of $d=1.2$, but newer studies like that of Genschow, Rigoni, and Brass (2017) reported much smaller effect sizes of $d=0.43$ (study 3a) and $d=0.22$ (study 3b). Therefore, attempting to replicate the study of, for instance, Rigoni et al. (2011) with sufficient power would be difficult to realize, because of the high effort involved in making EEG recordings for every single participant. Hence, in order to estimate a realistic effect size for manipulations of belief in free will, a meta-analysis containing published as well as unpublished datasets would be necessary and very helpful, and furthermore, it would be useful to plan manipulation studies that are not underpowered.

Another, less likely reason for the failed manipulation attempts might be that belief in free will as measured by WiF in German-speaking samples could be more stable than it is an influenceable state. If this is indeed the case, it would be reasonable that the manipulation attempts in this study failed. However, this is less likely because the results on the temporal stability of the WiF subscales presented in chapter 3, study 5, showed the WiF subscales to be less stable than personality traits. In particular, the subscales *free will* and *incompatibilism/moral responsibility* showed lower correlations between the two measurements, indicating that in these subscales, the belief in free will could be more situationally influenced. Furthermore, results of Ent and Baumeister (2014) indicate that even bodily states like tiredness or sexual desire can influence the belief in free will among participants.

Nevertheless, most studies reported in literature focusing on the manipulation of free will beliefs and its consequences used English-speaking samples from the United States of America. Generalizability on samples with other cultural backgrounds, other nationalities, or other languages is therefore limited and should be studied further. For replications studies with more diverse samples, effective manipulation methods in different languages are needed. Intercultural studies could also be useful to test whether there are structural differences in belief in free will based on different cultural backgrounds and the related thoughts and convictions. Last but not least, the general effective direction should be studied further, as the results of Clark et al. (2014) challenge the hypothesis of the belief in free will (or a lack of belief in free will) as a cause of behavioral effects (e.g. cheating, Vohs & Schooler, 2008) and judgments concerning moral responsibility. They tested the hypothesis that free will beliefs emerge at least in part from a desire to hold people morally responsible for their wrongful behaviors (Clark et al., 2014). In consequence, it could be worthwhile to further investigate the effect mechanism underling behavioral changes that are associated with belief in free will and moral responsibility.

5. Conclusion

Questions concerning the existence and manifestation of free will have been the focus of researchers from various disciplines for many centuries. Throughout the course of history, it has been a controversial topic, and presumably all scientists working in this field has their own beliefs or opinions on the existence of free will. Benjamin Libet, for example, questioned the existence of free will with his experiments, yet he stated and maintained the presumptive existence of free will as an important aspect in his veto hypothesis (Libet, 1999) even though he did not provide any empirical evidence for it. Theories and debates on free will might often seem to be driven by an individual's belief in free will and therefore it is not surprising that belief in free will itself became an important research topic.

Research on free will and belief in free will is highly relevant for different reasons. Jurisprudence is built upon the assumption that humans have free will and consequently are responsible for their actions and decisions. Reliable scientific evidence of the non-existence of free will in humans would not only have an impact on theoretical or applied jurisprudence, but also on the law enforcement system. Furthermore, it would have implications for all humans in terms of our emotions, experience, and behavior, and consequently have an impact on our whole civilization. Especially when dealing with such a highly controversial topic, with outcomes that have a potential to impact every human, it is imperative that empirical findings are analyzed critically with regard to their reliability, replicability, and generalizability. The aim of this work is to critically examine some of the findings in psychological research on free will and belief in free will, as well as to address the following questions:

1. Can a movement based on a voluntary decision be involved in the paradigm?
2. How can belief in free will be measured reliably in the German language?
3. Can belief in free will be altered by experimental interventions in German-speaking samples?

In chapter 2, two studies in the mindset of the Libet paradigm are reported. These aimed to generalize the Libet paradigm for a factual, free, and voluntary decision with consequences for the acting person, as well as to test the critical objection that the experimental condition of reporting the conscious intention to move has a direct effect on the readiness potential in the Libet paradigm. The implementation of the voluntary break request in the classical paradigm was successful and increased the scope of action for the participants. Unfortunately, this enhancement of the Libet paradigm was not used often enough, or rather, the action of requesting a break was accompanied by artefacts too often to calculate a readiness potential. The key finding of Libet et al. (1983) was replicated in both studies: The onset of the readiness potential occurred earlier than the reported conscious will to act. Furthermore, the task of reporting the time of the conscious will to act was not found to have a significant impact on the readiness potential. Yet it became clear that there are more methodological problems in the Libet paradigm, such as the definition of the onset time of the readiness potential and the general interpretation of the phenomenon readiness potential. Far more research is necessary in order to clarify whether the readiness potential really exists, or if it is just an artifact of the data processing and averaging process utilized for event-related potentials. If readiness potential exists, its role needs to be clarified precisely, not only in relation to voluntary movements. As it cannot be measured in every condition or every participant, it is neither a necessary nor sufficient condition for a voluntary movement. Due to these unclear or controversial presumptions in the experimental design, it is hardly possible to draw a conclusion on the existence or non-existence of free will based on Libet-style experiments. Furthermore, other experimental evidence based on the previously common interpretation of the readiness potential has to be reinterpreted.

In chapter 3, the construction and first steps toward validating the WiF (from German *Willensfreiheitsfragebogen*) inventory measuring belief in free will in the German language are described. The new instrument was developed with the aim of overcoming the methodological problems of some of the existing instruments in the English language and to provide a suitable instrument for German-speaking samples. The final instrument includes 21 items on five subscales: *situational determinism*, *free will*, *indeterminism/chance*, *biological determinism*, and *incompatibility/moral responsibility*. All scales show acceptable internal consistencies. Higher values for Cronbach's Alpha were not expected due to the shortness of the scales. The factorial structure was derived by an exploratory factor analysis and was supported by a confirmatory factor analysis in an independent sample. As expected, the subscales showed small and reasonable correlations to related constructs like *self-efficacy* and *locus of control*, indicating that belief in free will is a related but discrete construct. A multi-factorial, non-orthogonal solution was used in order to solve the problem of the unrealistic assumptions of belief in free will as either a unidimensional construct ranging from free will to determinism or as a multidimensional construct with independent factors. Measuring belief in free will as a unidimensional construct, like in the instrument of Rakos et al. (2008), does not allow the measurement of compatibilistic and/or incompatibilistic tendencies in individuals, which is currently an important topic in the modern philosophical discourse on free will. Including the short scale *incompatibility/moral responsibility* allows these tendencies to be measured not only indirectly (through high scorings on deterministic and free will scales), but also directly with three items. Results are promising so far. Moreover, the next useful step to validate the instrument further should be a validation using a sample that is representative for the German-speaking population, since it has only been used on student samples thus far. Furthermore, transcultural studies regarding the factorial structure could be valuable in providing a deeper understanding of belief in free will and the possible influence of different cultural backgrounds.

Two attempts to manipulate belief in free will using methods that were reported to have been successful in literature are reported in chapter 4, one using text excerpts from Crick (1995) and one using a Velten-style manipulation. Unfortunately, both attempts were unsuccessful, although the observed samples were similar in size, age, and proportion of students to the samples reported in literature. Reasons for the failure of these attempts could be manifold; it might be based on the translation of the materials, characteristics of the samples, such as educational background or gender proportions, the sample sizes, and/or an overestimated effect size in the literature. The manipulation attempts reported here are not the only unsuccessful attempts to manipulate belief in free will. For example, study 1 of Vohs and Schooler (2008) using a Crick manipulation could not be replicated as part of the Reproducibility Project of the Open Science Collaboration (Embley et al., 2015). Unsuccessful replication attempts are often inaccessible or not recognized due to the publication bias, therefore it is reasonable to assume that there are many more studies that failed to replicate a belief in free will manipulation. Therefore, the actual effect sizes of belief in free will manipulations might be highly overestimated in the current literature on this topic. A meta-analysis including published as well as unpublished datasets could contribute a realistic estimation of the effect sizes. This is necessary to design replication studies that are not underpowered and it would help test whether effects from English-speaking populations are generalizable for other populations as well. It would also test whether effects reported in individual studies are robust or if they merely represent single false-positive findings. Furthermore, the effect mechanisms as well as the effective directions underling behavioral changes that are associated with belief in free will should be studied in depth.

In the current state of research, many findings on free will and belief in free will appear less stable and less reliable than at a first glance, and some findings leave more questions than they answer, like findings from the Libet paradigm. If the readiness potential does not necessarily reflect a process crucial for voluntary motor preparation, then results that are based on the Libet paradigm have to be reinterpreted. For example the most recent results using the paradigm mentioned in the introduction, Rigoni et al. (2011), reported that belief in free will seems to alter early stages of motor preparation. They measured a slightly altered readiness potential in participants induced to disbelieve in free will, but to this point it is unclear what the altered amplitude of the readiness potential actually reflects. As lay persons generally tend to believe in free will (Nahmias et al., 2005), an alternative explanation of their results could be that reading the Crick texts not only alters the belief in free will but also threatens the world view of the participants, which could have an impact on motivation. Subsequently, these motivational issues could alter the involvement of the participants in the task, which might have an effect on the readiness potential and its amplitude (Kornhuber & Deecke, 1965). Taken together with the results of this work with respect to the study of Rigoni et al. (2011), it leads to the conclusion that the results of the study published in *Psychological Science* (impact factor for 2017: 6.128) are less reliable: A replication of the results is hardly possible. One reason for that is the difficulty of manipulating belief in free will in small samples, another is that alternative explanations for differences in the amplitude of readiness potential are possible as well as and the fact that the interpretation of the readiness potential itself remain unclear. Despite methodological problems, the findings were published by a high impact journal. John Oliver, an English satirist, writer, and political commentator, amusingly illustrated some problems of the current process of generating scientific knowledge in an episode on scientific studies of his late-night talk show *Last Week Tonight*, and mentioned that there is no noble prize for fact checking (Oliver & Pennolino, 2016). Although there is no reason to change that, consideration should be given to how to set incentives for good research practices like profound and solid designs and reporting stable and reliable results, instead of focusing on catchy headlines. Nevertheless, free will and belief in free will are important topics that even laypeople are curious about, but more research is necessary in order to provide deeper insight. Furthermore, it is also necessary to build a reliable theoretical framework and to check the assumptions of research designs. Otherwise free will remains just a matter of belief.

References

Alquist, J. L., Ainsworth, S. E., & Baumeister, R. F. (2013). Determined to conform: Disbelief in free will increases conformity. *Journal of Experimental Social Psychology*, *49*(1), 80–86. https://doi.org/10.1016/j.jesp.2012.08.015

ANT Neuro B.V./eemagine Medical Imaging Solutions GmbH. (2013). *asalab User Guide (Version 4.7.12)*.

Banks, W. P., & Isham, E. A. (2009). We infer rather than perceive the moment we decided to act. *Psychological Science*, *20*(1), 17–21. https://doi.org/10.1111/j.1467-9280.2008.02254.x

Banks, W. P., & Pockett, S. (2007). Benjamin Libet's Work on the Neuroscience of Free Will. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 657–670). Malden, MA: Blackwell. https://doi.org/10.1002/9780470751466.ch52

Baumeister, R. F., & Brewer, L. E. (2012). Believing versus Disbelieving in Free Will: Correlates and Consequences. *Social and Personality Psychology Compass*, *6*(10), 736–745. https://doi.org/10.1111/j.1751-9004.2012.00458.x

Baumeister, R. F., Masicampo, E. J., & Dewall, C. N. (2009). Prosocial benefits of feeling free: disbelief in free will increases aggression and reduces helpfulness. *Personality & Social Psychology Bulletin*, *35*(2), 260–268. https://doi.org/10.1177/0146167208327217

Benecke, R., Dick, J.P.R., Rothwell, J. C., Day, B. L., & Marsden, C. D. (1985). Increase of the Bereitschaftspotential in simultaneous and sequential movements. *Neuroscience Letters*, *62*(3), 347–352. https://doi.org/10.1016/0304-3940(85)90573-7

Brewer, L. E. (2011). Forging Freely: Perceptions of Moral Responsibility Mediate the Relationship Between Belief in Free Will and Willingness to Forgive (Master's Thesis). Florida State University, Tallahassee, Florida.

Brücher, K., & Gonther, U. (2006). Zum Verhältnis von Willensfreiheit und Neurobiologie. *Fortschritte Der Neurologie · Psychiatrie*, *74*(4), 194–202. https://doi.org/10.1055/s-2005-915592

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a New General Self-Efficacy Scale. *Organizational Research Methods*, *4*(1), 62–83. https://doi.org/10.1177/109442810141004

Clark, C. J., Luguri, J. B., Ditto, P. H., KNOBE, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: a motivated account of free will belief. *Journal of Personality and Social Psychology*, *106*(4), 501–513. https://doi.org/10.1037/a0035880

Cohen, J., & Polich, J. (1997). On the number of trials needed for P300. *International Journal of Psychophysiology*, *25*(3), 249–255. https://doi.org/10.1016/S0167-8760(96)00743-X

Collani, G. von, & Schyns, B. (2014). *Generalisierte Selbstwirksamkeitserwartung.*

Collani, G. von, & Herzberg, P. Y. (2003a). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *24*(1), 3–7. https://doi.org/10.1024//0170-1789.24.1.3

Collani, G. von, & Herzberg, P. Y. (2003b). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *24*(1), 3–7. https://doi.org/10.1024//0170-1789.24.1.3

Crescioni, A. W., Baumeister, R. F., Ainsworth, S. E., Ent, M., & Lambert, N. M. (2015). Subjective correlates and consequences of belief in free will. *Philosophical Psychology*, *29*(1), 41–63. https://doi.org/10.1080/09515089.2014.996285

Crick, F. (1995). *The astonishing hypothesis: The scientific search for the soul* (1st Touchstone ed.). *A Touchstone book*. New York, NY: Simon & Schuster.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Dirnberger, G., Fickel, U., Lindinger, G., Lang, W., & Jahanshahi, M. (1998). The mode of movement selection. Movement-related cortical potentials prior to freely selected and repetitive movements. *Experimental Brain Research*, *120*(2), 263–272.

Doyle, B. (2011). *Free will: The scandal in philosophy* (First edition). Cambridge, Mass.: I-Phi Press.

Duttweiler, P. C. (1984). The Internal Control Index: A Newly Developed Measure of Locus of Control. *Educational and Psychological Measurement*, *44*(2), 209–221. https://doi.org/10.1177/0013164484442004

Embley, J., Johnson, L. G., & Giner-Sorolla, R. (2015). *Reproducibility Project - Replication Report: Replication of Study 1 by Vohs & Schooler (2008, Psychological Science)*. Open Science Collaboration. Retrieved from https://osf.io/2nf3u/

Ent, M. R., & Baumeister, R. F. (2014). Embodied free will beliefs: some effects of physical states on metaphysical opinions. *Consciousness and Cognition*, *27*, 147–154. https://doi.org/10.1016/j.concog.2014.05.001

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Freude, G., Ullsperger, P., Krüger, H., & Pietschmann, M. (1988). The Bereitschaftspotential in preparation to mental activities. *International Journal of Psychophysiology*, *6*(4), 291–297. https://doi.org/10.1016/0167-8760(88)90016-5

Genschow, O., Rigoni, D., & Brass, M. (2017). Belief in free will affects causal attributions when judging others' behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(38), 10071–10076. https://doi.org/10.1073/pnas.1701916114

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German Version of the Satisfaction With Life Scale (SWLS): Psychometric Properties, Validity, and Population-Based Norms. *European Journal of Psychological Assessment*, *27*(2), 127–132. https://doi.org/10.1027/1015-5759/a000058

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126*(1), 128–133. https://doi.org/10.1007/s002210050722

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review : an Official Journal of the Society for Personality and Social Psychology, Inc*, *14*(3), 332–346. https://doi.org/10.1177/1088868310361240

Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrica*, *30*(2), 179–185.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jahanshahi, M., & Hallett, M. (2003). The Bereitschaftspotential: What does it measure and where does it come from? In M. Jahanshahi & M. Hallett (Eds.), *The Bereitschaftspotential: Movement-related Cortical Potentials.* New York: Kluwer Academic.

Janke, S., & Glöckner-Rist, A. (2012). *Deutsche Version der Satisfaction with Life Scale (SWLS).*

Jo, H.-G., Hinterberger, T., Wittmann, M., Borghardt, T. L., & Schmidt, S. (2013). Spontaneous EEG fluctuations determine the readiness potential: is preconscious brain activation a preparation process to move? *Experimental Brain Research*, *231*(4), 495–500. https://doi.org/10.1007/s00221-013-3713-z

John, O. P., & Srivastava, S. (1999). The Big Five Taxonomy: History, Measurement, and Theoretical Perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed.). New York: Guilford Press.

Joordens, S., Spalek, T. M., Razmy, S., & van Duijn, M. (2004). A Clockwork Orange: Compensation opposing momentum in memory for location. *Memory & Cognition*, *32*(1), 39–50. https://doi.org/10.3758/BF03195819

Joordens, S., van Duijn, M., & Spalek, T. M. (2002). When Timing the Mind One Should Also Mind the Timing: Biases in the Measurement of Voluntary Actions. *Consciousness and Cognition*, *11*(2), 231–240. https://doi.org/10.1006/ccog.2002.0559

Kaiser, H. F. (1974). An Index of Factorial Simplicity. *Psychometrika*, *39*(1), 31–36.

Kane, R. (2011). Introduction: The Contours of Contemporary Free-Will Debates. In R. Kane (Ed.), *Oxford Handbooks. Oxford Handbook of Free Will.* s.l.: Oxford University Press USA.

Kawohl, W., & Habermeyer, E. (2007). Free will: reconciling German civil law with Libet's neurophysiological studies on the readiness potential. *Behavioral Sciences & the Law*, *25*(2), 309–320. https://doi.org/10.1002/bsl.752

Keller, I., & Heckhausen, H. (1990). Readiness potentials preceding spontaneous motor acts: Voluntary vs. involuntary control. *Electroencephalography and Clinical Neurophysiology*, *76*(4), 351–361. https://doi.org/10.1016/0013-4694(90)90036-J

Klemm, W. R. (2010). Free will debates: Simple experiments are not so simple. *Advances in Cognitive Psychology*, *6*, 47. https://doi.org/10.2478/v10053-008-0076-2

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv Für Die*

Gesamte Physiologie Des Menschen Und Der Tiere, *284*(1), 1–17. https://doi.org/10.1007/BF00412364

Krampen, G. (1981). *IPC- Fragebogen zu Kontrollüberzeugungen ("Locus of control"): Deutsche Bearbeitung der IPC-Scales von Hanna Levenson*. Göttingen: Hogrefe.

Lang, W. (2003). Surface Recordings of the Bereitschaftspotential in Normals. In M. Jahanshahi & M. Hallett (Eds.), *The Bereitschaftspotential: Movement-related Cortical Potentials.* New York: Kluwer Academic.

Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science (New York, N.Y.)*, *303*(5661), 1208–1210. https://doi.org/10.1126/science.1090973

Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, *41*(3), 397–404. https://doi.org/10.1037/h0035357

Libet, B., Wright, E. W., JR, & Gleason, C. A. (1982). Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology*, *54*(3), 322–335.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*(04), 529. https://doi.org/10.1017/S0140525X00044903

Libet, B. (1999). Do we have free will? *Journal of Consciousness Studies*, *6*(8-9), 47–57.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of Conscious Intention to Act in Relation to Onset of Cerebral activity (Readiness-Potential). *Brain*, *106*(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Libet, B., Wright Jr., E. W., & Gleason, C. A. (1983). Preparation- or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalography and Clinical Neurophysiology*, *56*(4), 367–372. https://doi.org/10.1016/0013-4694(83)90262-6

Lynn, M. T., Muhle-Karbe, P. S., Aarts, H., & Brass, M. (2014). Priming determinist beliefs diminishes implicit (but not explicit) components of self-agency. *Frontiers in Psychology*, *5*, 1483. https://doi.org/10.3389/fpsyg.2014.01483

Lynn, M. T., van Dessel, P., & Brass, M. (2013). The influence of high-level beliefs on self-regulatory engagement: evidence from thermal pain stimulation. *Frontiers in Psychology*, *4*, 614. https://doi.org/10.3389/fpsyg.2013.00614

MacKenzie, M. J., Vohs, K. D., & Baumeister, R. F. (2014). You didn't have to do that: belief in free will promotes gratitude. *Personality & Social Psychology Bulletin*, *40*(11), 1423–1434. https://doi.org/10.1177/0146167214549322

Macrae, F., & Pitts, M. (2015, March). *Does Inducing Disbelief in Free WIll Alter Brain Correlates of Preconscious Motor Preparation? A Replication Study*. Poster presented at the 22nd Annual Meeting of the Cognitive Neuroscience Meeting. California, USA: San Francisco. Retrieved from http://www.reed.edu/psychology/scalp/assets/conference%20files/Macrae_Pitts_CNS_poster_2015.pdf [05.09.2017]

Marco-Pallares, J., Cucurell, D., Münte, T. F., Strien, N., & Rodriguez-Fornells, A. (2011). On the number of trials needed for a stable feedback-related negativity. *Psychophysiology*, *48*(6), 852–860. https://doi.org/10.1111/j.1469-8986.2010.01152.x

Meyke, N. (2015). Der Zusammenhang zwischen dem Glauben an den freien Willen und Persönlichkeitseigenschaften (Bachelor's Thesis). Leuphana Universität, Lüneburg.

Miller, J., Shepherdson, P., & Trevena, J. (2011). Effects of clock monitoring on electroencephalographic activity: is unconscious movement initiation an artifact of the clock? *Psychological Science*, *22*(1), 103–109. https://doi.org/10.1177/0956797610391100

Monroe, A. E., & Malle, B. F. (2010). From Uncaused Will to Conscious Choice: The Need to Study, Not Speculate About People’s Folk Concept of Free Will.

Review of Philosophy and Psychology, *1*(2), 211–224. https://doi.org/10.1007/s13164-009-0010-7

Muggeo, V. M.R. (2008). Segmented: An R package to fit regression models with broken-line relationships. *R News*, *8*(1), 20–25.

Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., & Ross, L. T. (2014). The free will inventory: measuring beliefs about agency and responsibility. *Consciousness and Cognition*, *25*, 27–41. https://doi.org/10.1016/j.concog.2014.01.006

Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying Freedom: Folk Intuitions about free will and moral responsibility. *Philosophical Psychology*, *18*(5), 561–584. https://doi.org/10.1080/09515080500264180

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science : a Journal of the Association for Psychological Science*, *7*(6), 615–631. https://doi.org/10.1177/1745691612459058

Obhi, S. S., Planetta, P. J., & Scantlebury, J. (2009). On the signals underlying conscious awareness of action. *Cognition*, *110*(1), 65–73. https://doi.org/10.1016/j.cognition.2008.10.009

Oldfield, R. C. (1971). The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.

Oliver, J. (W.), & Pennolino, P. (D.) (Author) (2016, May 8). *Last Week Tonight with John Oliver: Scientific Studies* [Television broadcast]. HBO. Retrieved from https://www.youtube.com/watch?v=0Rnq1NpHdmw [28.08.2017]

Olvet, D. M., & Hajcak, G. (2009). The stability of error-related brain activity with increasing trials. *Psychophysiology*, *46*(5), 957–961. https://doi.org/10.1111/j.1469-8986.2009.00848.x

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science (New York, N.Y.)*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Paulhus, D. L., & Carey, J. M. (2011). The FAD-Plus: measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment*, *93*(1), 96–104. https://doi.org/10.1080/00223891.2010.528483

Praamstra, P., Stegeman, D. F., Horstink, M.W.I.M., & Cools, A. R. (1996). Dipole source analysis suggests selective modulation of the supplementary motor area contribution to the readiness potential. *Electroencephalography and Clinical Neurophysiology*, *98*(6), 468–477. https://doi.org/10.1016/0013-4694(96)95643-6

Rakos, R. F., Laurene, K. R., Skala, S., & Slane, S. (2008). Belief in Free Will: Measurement and Conceptualization Innovations. *Behavior and Social Issues*, *17*, 20–39.

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, *51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Raykov, T. (1998). On the use of confirmatory factor analysis in personality research. *Personality and Individual Differences*, *24*(2), 291–293. https://doi.org/10.1016/S0191-8869(97)00159-1

Revelle, W. (2016, June). Package ‘psych’. Retrieved from http://personality-project.org/r/psych-manual.pdf [November 24, 16]

Rigoni, D., Kuhn, S., Gaudino, G., Sartori, G., & Brass, M. (2012). Reducing self-control by weakening belief in free will. *Consciousness and Cognition*, *21*(3), 1482–1490. https://doi.org/10.1016/j.concog.2012.04.004

Rigoni, D., Kuhn, S., Sartori, G., & Brass, M. (2011). Inducing disbelief in free will alters brain correlates of preconscious motor preparation: the brain minds whether we believe in free will or not. *Psychological Science*, *22*(5), 613–618. https://doi.org/10.1177/0956797611405680

Rigoni, D., Wilquin, H., Brass, M., & Burle, B. (2013). When errors do not matter: weakening belief in intentional control impairs cognitive reaction to errors. *Cognition*, *127*(2), 264–269. https://doi.org/10.1016/j.cognition.2013.01.009

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Sarkissian, H., Chatterjee, A., Brigard, F. d., Knobe, J., Nichols, S., & Sirker, S. (2010). Is Belief in Free Will a Cultural Universal? *Mind & Language*, *25*(3), 346–358. https://doi.org/10.1111/j.1468-0017.2010.01393.x

Schmidt, S., Jo, H.-G., Wittmann, M., & Hinterberger, T. (2016). 'Catching the waves' - slow cortical potentials as moderator of voluntary action. *Neuroscience and Biobehavioral Reviews*, *68*, 639–650. https://doi.org/10.1016/j.neubiorev.2016.06.023

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature*, *470*(7335), 437. https://doi.org/10.1038/470437a

Schultze-Kraft, M., Birman, D., Rusconi, M., Allefeld, C., Gorgen, K., Dahne, S., . . . Haynes, J.-D. (2016). The point of no return in vetoing self-initiated movements. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(4), 1080–1085. https://doi.org/10.1073/pnas.1513569112

Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(42), E2904-13. https://doi.org/10.1073/pnas.1210467109

Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. J., Schooler, J. W., . . . Vohs, K. D. (2014). Free will and punishment: a mechanistic view of human nature reduces retribution. *Psychological Science*, *25*(8), 1563–1570. https://doi.org/10.1177/0956797614534693

Shibasaki, H., Barrett, G., Halliday, E., & Halliday, A.M. (1980). Components of the movement-related cortical potential and their scalp topography. *Electroencephalography and Clinical Neurophysiology*, *49*(3–4), 213–226. https://doi.org/10.1016/0013-4694(80)90216-3

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *117*(11), 2341–2356. https://doi.org/10.1016/j.clinph.2006.04.025

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, *53*(1), 80–93. https://doi.org/10.1037/0022-0167.53.1.80

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4. ed.). Mahwah NJ u.a.: Erlbaum.

Stillman, T. F., Baumeister, R. F., Vohs, K. D., Lambert, N. M., Fincham, F. D., & Brewer, L. E. (2010). Personal Philosophy and Personnel Achievement: Belief in Free Will Predicts Better Job Performance. *Social Psychological and Personality Science*, *1*(1), 43–50. https://doi.org/10.1177/1948550609351600

Stillman, T. F., & Baumeister, R. F. (2010). Guilty, free, and wise: Determinism and psychopathy diminish learning from negative emotions. *Journal of Experimental Social Psychology*, *46*(6), 951–960. https://doi.org/10.1016/j.jesp.2010.05.012

Stöber, K. (2016). Der Glaube an den freien Willen und wahrgenommener Sinn im Leben: Eine Replikationsstudie im deutschsprachigen Raum (Bachelor's Thesis). Leuphana Universität, Lüneburg.

Stroessner, S. J., & Green, C. W. (1990). Effects of Belief in Free Will or Determinism on Attitudes Toward Punishment and Locus of Control. *The Journal of Social Psychology*, *130*(6), 789–799.

Strother, L., & Obhi, S. S. (2009). The conscious experience of action and intention. *Experimental Brain Research*, *198*(4), 535–539. https://doi.org/10.1007/s00221-009-1946-7

Trevena, J., & Miller, J. (2010). Brain preparation before a voluntary action: evidence against unconscious movement initiation. *Consciousness and Cognition*, *19*(1), 447–456. https://doi.org/10.1016/j.concog.2009.08.006

Trevena, J. A., & Miller, J. (2002). Cortical Movement Preparation before and after a Conscious Decision to Move. *Consciousness and Cognition*, *11*(2), 162–190. https://doi.org/10.1006/ccog.2002.0548

Van de Grind, W. (2002). Physical, Neural, and Mental Timing. *Consciousness and Cognition*, *11*(2), 241–264. https://doi.org/10.1006/ccog.2002.0560

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, *6*(4), 473–482. https://doi.org/10.1016/0005-7967(68)90028-4

Verleger, R., Haake, M., Baur, A., & Śmigasiewicz, K. (2016). Time to Move Again: Does the Bereitschaftspotential Covary with Demands on Internal Timing? *Frontiers in Human Neuroscience*, *10*, 642. https://doi.org/10.3389/fnhum.2016.00642

Viney, W., Waldman, D. A., & Barchilon, J. (1982). Attitudes Toward Punishment in Relation to Beliefs in Free Will and Determinism. *Human Relations*, *35*(11), 939–950.

Vohs, K. D., & Schooler, J. W. (2008). The value of believing in free will: encouraging a belief in determinism increases cheating. *Psychological Science*, *19*(1), 49–54. https://doi.org/10.1111/j.1467-9280.2008.02045.x

Wett, M. K. (2015). Die Auswirkung einer Schwächung des Glaubens an den freien Willen auf die Entscheidung in einem Bestrafungsszenario (Bachelor's Thesis). Leuphana Universität, Lüneburg.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*(3), 432–442. https://doi.org/10.1037/0033-2909.99.3.432

Appendix

Study 1: Instructions

Liebe(r) Versuchsteilnehmer(in),

vielen Dank, dass Sie sich bereit erklärt haben, an dieser **EEG-Studie** teilzunehmen. In diesem Experiment wird das **Bereitschaftspotential** untersucht. Das Bereitschaftspotential ist ein elektrophysiologisch messbares Phänomen, welches vor willkürlichen Bewegungen in bestimmten Arealen der Großhirnrinde auftritt und als Vorbereitung der Bewegung interpretiert wird.

Sie werden gleich eine Elektrodenkappe aufgesetzt bekommen. Mithilfe dieser Kappe zeichnen wir die **Spannungsschwankungen an Ihrer Kopfoberfläche** auf. Ursache dieser Spannungsschwankungen sind physiologische Vorgänge einzelner Gehirnzellen, die durch ihre elektrischen Zustandsänderungen zur Informationsverarbeitung des Gehirns beitragen. Zur besseren Leitfähigkeit der Elektroden wird eine **Elektrolyt-Paste** zwischen Elektrode und Kopfhaut aufgetragen. Diese Paste ist völlig ungefährlich und lässt sich mit Wasser wieder ausspülen. Im Experiment werden die Spannungen (im Mikrovoltbereich), die Ihr Gehirn selbst produziert, aufgezeichnet. Zusätzlich wird die Muskelaktivität am Daumen mit zwei Klebeelektroden abgeleitet.

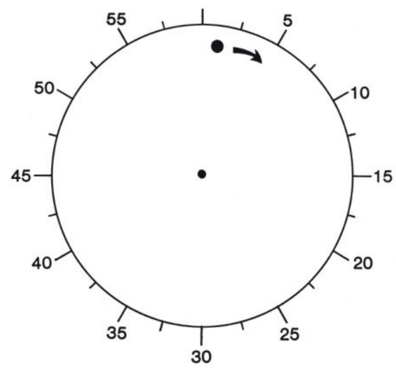
Alle Daten werden anonym gespeichert und vertraulich behandelt. Teilen Sie dazu dem Versuchsleiter bitte folgende Versuchspersonen-Kennung mit:

- dritter Buchstabe des Vornamens der Mutter
- letzte zwei Ziffern des Geburtsjahrs
- dritter Buchstabe des eigenen Nachnamens
- erster Buchstabe des eigenen Vornamens

Insgesamt wird das Experiment **etwa zwei (maximal drei) Stunden dauern**. **Bitte schalten Sie ihr Telefon aus** (und alle anderen elektrischen Geräte), da sie sie Messung stören können.

Sie werden im Laufe des Versuchs **vier unterschiedliche Aufgaben** bekommen, die jeweils etwa 80 Mal wiederholt werden. Bei allen Durchgängen sehen Sie eine Art **Uhr, auf der sich ein Strich im Uhrzeigersinn** bewegt. Zu Beginn eines jeden Durchlaufs hören Sie einen Ton. Dieser **„Get ready"-Ton** signalisiert, dass der Uhrumlauf gleich startet. Wenn er ertönt, entspannen Sie sich bitte (besonders die Muskeln von Kopf, Nacken und Armen) und **bitte blinzeln Sie nicht** während des Uhrumlaufs. Wichtig: Bitte verfolgen Sie nicht den sich bewegenden Punkt, sondern **fixieren Sie die gesamte Zeit**

den zentralen Punkt der Uhr. Nach einem Tastendruck bzw. nach einem Ton verschwindet der bewegte Punkt nach Ablauf eines zufälligen Zeitintervalls.



Schematische Darstellung der Uhr

Es ist für den Versuch besonders wichtig, dass Sie sich auf die Aufgabenstellung konzentrieren. Da Sie sich während des Versuchs nicht bewegen und nicht blinzeln dürfen und der Ablauf recht monoton ist, ist es unvermeidbar, dass es zu **Ermüdungserscheinungen** kommt (auch wenn man es selbst noch nicht bemerkt). Daher bitten ich Sie **häufig Pausen zu machen**. Drücken Sie dazu einfach jederzeit im Verlauf des Versuchs den Pause-Knopf. Zögern Sie nicht, den Versuch zu unterbrechen, da es für die Erhebung der Daten vor allem wichtig ist, dass Sie **sehr konzentriert die Aufgaben bearbeiten**. Machen Sie lieber einmal mehr Pause als einmal zu wenig.

Haben Sie Fragen?

A

Ihre Aufgabe ist es, **zu einem selbstgewählten Zeitpunkt die mit einem X markierte Taste des Gamepads zu betätigen**. Bitte warten Sie dazu die erste Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Falls Sie zwischendurch doch geblinzelt haben, warten Sie einfach noch eine Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Bitte versuchen Sie möglichst spontan und schnell zu klicken. **Bitte versuchen Sie, dem Drang zu Handeln einfach freien Lauf zu lassen. Planen Sie nicht auf einen Zeitpunkt voraus, sondern handeln Sie spontan sobald Sie den Willen dazu verspüren.** Der Zeitpunkt des Tastendrucks beeinflusst nicht die Gesamtlänge des Experiments. Wir starten mit drei Übungsdurchläufen. Danach folgen 80 Wiederholungen. **Vergessen Sie bitte nicht, Pausen zu machen.**

B

Ihre Aufgabe ist es, **zu einem selbstgewählten Zeitpunkt die mit einem X markierte Taste des Gamepads zu betätigen**. Bitte warten Sie dazu die erste Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Falls Sie zwischendurch doch geblinzelt haben, warten Sie einfach noch eine Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Bitte versuchen Sie möglichst spontan und schnell zu klicken. **Bitte versuchen Sie, dem Drang zu Handeln einfach freien Lauf zu lassen. Planen Sie nicht auf einen Zeitpunkt voraus, sondern handeln Sie spontan sobald Sie den Willen dazu verspüren.** Der Zeitpunkt des Tastendrucks beeinflusst nicht die Gesamtlänge des Experiments. Hinterher bitte ich Sie, **den Zeitpunkt** (entsprechend der Uhrposition) **zu berichten, an dem Sie den Entschluss gefasst haben, die Taste zu betätigen**. Wir starten mit drei Übungsdurchläufen. Danach folgen 80 Wiederholungen. **Vergessen Sie bitte nicht, Pausen zu machen.**

C

Ihre Aufgabe ist es, **zu einem bestimmten Zeitpunkt die mit einem X markierte Taste des Gamepads zu betätigen**. Bitte warten Sie dazu die erste Umdrehung der Uhr ab und drücken Sie dann die Taste zu dem Zeitpunkt, der Ihnen vorab mitgeteilt wurde. Drücken Sie bitte schnell und **zeitlich so genau wie möglich**. Wir starten mit drei Übungsdurchläufen. Danach folgen 80 Wiederholungen. **Vergessen Sie bitte nicht, Pausen zu machen.**

D

Ihre Aufgabe ist es, **den Zeitpunkt eines zufälligen Tons anhand des Uhrumlaufs zu berichten**. Sie brauchen keine Taste zu betätigen. Nach dem Ton verschwindet der bewegte Punkt nach Ablauf eines zufälligen Zeitintervalls. Hinterher bitte ich Sie, **den Zeitpunkt** (entsprechend der Uhrposition) **zu berichten, an dem Sie den Ton gehört haben.** Wir starten mit drei Übungsdurchläufen. Danach folgen 80 Wiederholungen. **Jeweils nach 20 Durchgängen gibt es eine Pause.**

Study 1: Qualitative Observations

Bedingungsabfolge (A: Selbst-initiierte Handlung ohne Zeitrückmeldung; B: Selbst-initiierte Handlung mit Zeitrückmeldung; C: vorgeplante Handlung; D: Externe Stimulierung)

Nachbefragung (A):

- Haben Sie vorgeplant? Waren Sie sich einer Vorplanung bewusst?
- Fiel es schwer, sich nicht auf die Handlung vorzubereiten?
- Waren Sie selbst überrascht vom Zeitpunkt Ihrer Handlung?
- Ist Ihnen irgendetwas aufgefallen?

Nachbefragung (B):

- Haben Sie vorgeplant? Waren Sie sich einer Vorplanung bewusst?
- Fiel es schwer, sich nicht auf die Handlung vorzubereiten?
- Waren Sie selbst überrascht vom Zeitpunkt Ihrer Handlung?
- Fiel es Ihnen schwer, durchgängig den Mittelpunkt zu fixieren?
- Fiel es Ihnen schwer, den Zeitpunkt Ihrer Handlung zu schätzen?
- Ist Ihnen irgendetwas aufgefallen?

Nachbefragung (C):

- Fiel es Ihnen schwer, sich auf die Handlung vorzubereiten?
- Ist Ihnen irgendetwas aufgefallen?

Nachbefragung (D):

- Fiel es Ihnen schwer, durchgängig den Mittelpunkt zu fixieren?
- Fiel es Ihnen schwer, den Zeitpunkt Ihrer Handlung zu schätzen?
- Ist Ihnen irgendetwas aufgefallen?

Study 2: Instructions

Liebe(r) Versuchsteilnehmer(in),

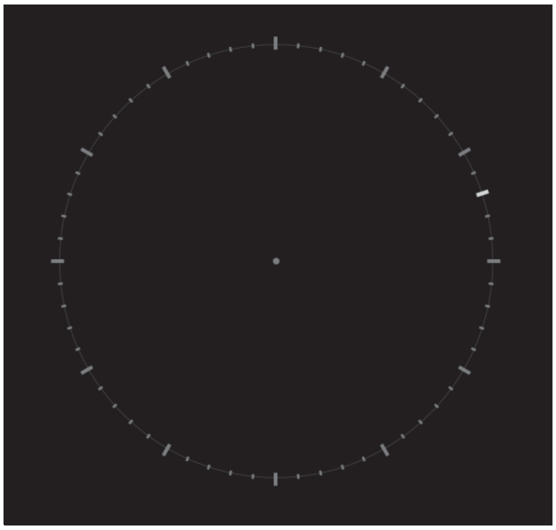
vielen Dank, dass Sie sich bereit erklärt haben, an dieser **EEG-Studie** teilzunehmen. In diesem Experiment wird das **Bereitschaftspotential** untersucht. Das Bereitschaftspotential ist ein elektrophysiologisch messbares Phänomen, welches vor willkürlichen Bewegungen in bestimmten Arealen der Großhirnrinde auftritt und als Vorbereitung der Bewegung interpretiert wird.

Sie werden gleich eine Elektrodenkappe aufgesetzt bekommen. Mithilfe dieser Kappe zeichnen wir die **Spannungsschwankungen an Ihrer Kopfoberfläche** auf. Ursache dieser Spannungsschwankungen sind physiologische Vorgänge einzelner Gehirnzellen, die durch ihre elektrischen Zustandsänderungen zur Informationsverarbeitung des Gehirns beitragen. Zur besseren Leitfähigkeit der Elektroden wird eine **Elektrolyt-Paste** zwischen Elektrode und Kopfhaut aufgetragen. Diese Paste ist völlig ungefährlich und lässt sich mit Wasser wieder ausspülen. Im Experiment werden die Spannungen (im Mikrovoltbereich), die Ihr Gehirn selbst produziert, aufgezeichnet. Zusätzlich wird die Muskelaktivität am Daumen mit zwei Klebeelektroden abgeleitet.

Alle Daten werden anonym gespeichert und vertraulich behandelt.

Insgesamt wird das Experiment **etwa vier Stunden dauern**. **Bitte schalten Sie ihr Telefon aus** (und alle anderen elektrischen Geräte), da sie sie Messung stören können.

Sie werden im Laufe des Versuchs **fünf unterschiedliche Aufgaben** bekommen, die jeweils etwa 80 Mal wiederholt werden. Bei allen Durchgängen sehen Sie eine Art **Uhr, auf der sich ein Punkt im Uhrzeigersinn** bewegt. Zu Beginn eines jeden Durchlaufs hören Sie einen Ton. Dieser **„Get ready"-Ton** signalisiert, dass der Uhrumlauf gleich startet. Wenn er ertönt, entspannen Sie sich bitte (besonders die Muskeln von Kopf, Nacken und Armen) und **bitte blinzeln Sie nicht** während des Uhrumlaufs. Wichtig: Bitte verfolgen Sie nicht den sich bewegenden Punkt. Nach einem Tastendruck bzw. nach einem Ton verschwindet der bewegte Punkt nach Ablauf eines zufälligen Zeitintervalls.



Schematische Darstellung der Uhr

Es ist für den Versuch besonders wichtig, dass Sie sich auf die Aufgabenstellung konzentrieren. Da Sie sich während des Versuchs nicht bewegen und nicht blinzeln dürfen und der Ablauf recht monoton ist, ist es unvermeidbar, dass es zu **Ermüdungserscheinungen** kommt (auch wenn man es selbst noch nicht unbedingt bemerkt). Daher bitte ich Sie **häufig Pausen zu machen**. Drücken Sie dazu einfach jederzeit im Verlauf des Versuchs den Pause-Knopf. Zögern Sie nicht, den Versuch zu unterbrechen, da es für die Erhebung der Daten vor allem wichtig ist, dass Sie **sehr konzentriert die Aufgaben bearbeiten**. Machen Sie lieber einmal mehr Pause als einmal zu wenig. Die Pause-Funktion verlängert die Versuchsdauer nicht. Die Pause startet mit einer Verzögerung von 800 ms. Bitte beachten Sie, dass Sie sich erst bewegen und blinzeln, wenn auf dem Bildschirm „PAUSE" angezeigt wird.

Haben Sie Fragen?

A

Bitte halten Sie Ihren Blick die gesamte Versuchsdauer über auf den Mittelpunkt der Uhr gerichtet. Bitte verfolgen Sie nicht den umlaufenden Punkt.

Ihre Aufgabe ist es, **zu einem selbstgewählten Zeitpunkt die mit einem X markierte Taste des Gamepads zu betätigen**. Bitte warten Sie dazu die erste Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Falls Sie zwischendurch doch geblinzelt haben, warten Sie einfach noch eine Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Bitte versuchen Sie möglichst spontan und schnell zu klicken. **Bitte versuchen Sie, dem Drang zu Handeln einfach freien**

Lauf zu lassen. Planen Sie nicht auf einen Zeitpunkt voraus, sondern handeln Sie spontan sobald Sie den Willen dazu verspüren. Der Zeitpunkt des Tastendrucks beeinflusst nicht die Gesamtlänge des Experiments.

Wir starten mit drei Übungsdurchläufen. Danach folgen 120 Wiederholungen.

Denken Sie bitte daran, sich nicht zu bewegen und während des Uhrumlaufs **nicht zu blinzeln .Vergessen Sie bitte nicht, Pausen zu machen** (idealerweise mindestens alle 10-15 Trials)**.** Bitte achten Sie darauf, dass Sie sich erst bewegen und blinzeln, wenn auf dem Bildschirm „PAUSE“ angezeigt wird.

B

Bitte halten Sie Ihren Blick die gesamte Versuchsdauer über auf den Mittelpunkt der Uhr gerichtet. Bitte verfolgen Sie nicht den umlaufenden Punkt.

Ihre Aufgabe ist es, **zu einem selbstgewählten Zeitpunkt die mit einem X markierte Taste des Gamepads zu betätigen**. Bitte warten Sie dazu die erste Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Falls Sie zwischendurch doch geblinzelt haben, warten Sie einfach noch eine Umdrehung der Uhr ab und drücken dann zu einem frei gewählten Zeitpunkt. Bitte versuchen Sie möglichst spontan und schnell zu klicken. **Bitte versuchen Sie, dem Drang zu Handeln einfach freien Lauf zu lassen. Planen Sie nicht auf einen Zeitpunkt voraus, sondern handeln Sie spontan sobald Sie den Willen dazu verspüren.** Der Zeitpunkt des Tastendrucks beeinflusst nicht die Gesamtlänge des Experiments. Hinterher bitte ich Sie, **den Zeitpunkt** (entsprechend der Uhrposition) **zu berichten, an dem Sie den Entschluss gefasst haben,** die Taste zu betätigen (nicht den Zeitpunkt des Tastendrucks selbst).

Wir starten mit drei Übungsdurchläufen. Danach folgen 120 Wiederholungen.

Denken Sie bitte daran, sich nicht zu bewegen und während des Uhrumlaufs **nicht zu blinzeln .Vergessen Sie bitte nicht, Pausen zu machen** (idealerweise mindestens alle 10-15 Trials)**.** Bitte achten Sie darauf, dass Sie sich erst bewegen und blinzeln, wenn auf dem Bildschirm „PAUSE“ angezeigt wird.

C

Bitte halten Sie Ihren Blick die gesamte Versuchsdauer über auf den Mittelpunkt der Uhr gerichtet. Bitte verfolgen Sie nicht den umlaufenden Punkt.

Ihre Aufgabe ist es, **zu einem bestimmten Zeitpunkt die mit einem X markierte Taste des Gamepads zu betätigen**. Bitte warten Sie dazu die erste Umdrehung der Uhr ab und drücken Sie dann die Taste zu dem Zeitpunkt, der Ihnen vorab mitgeteilt wurde. Jeweils nach 10 Durchgängen bekommen Sie einen neuen Zeitpunkt genannt. Drücken Sie bitte schnell und **zeitlich so genau wie möglich**.

Wir starten mit drei Übungsdurchläufen. Danach folgen 120 Wiederholungen.

Denken Sie bitte daran, sich nicht zu bewegen und während des Uhrumlaufs **nicht zu blinzeln**. **Vergessen Sie bitte nicht, Pausen zu machen** (idealerweise mindestens alle 10-15 Trials)**.** Bitte achten Sie darauf, dass Sie sich erst bewegen und blinzeln, wenn auf dem Bildschirm „PAUSE“ angezeigt wird.

D

Bitte halten Sie Ihren Blick die gesamte Versuchsdauer über auf den Mittelpunkt der Uhr gerichtet. Bitte verfolgen Sie nicht den umlaufenden Punkt.

Ihre Aufgabe ist es, **den Zeitpunkt eines zufälligen Tons anhand des Uhrumlaufs zu berichten**. Sie brauchen keine Taste zu betätigen. Nach dem Ton verschwindet der bewegte Punkt nach Ablauf eines zufälligen Zeitintervalls. Hinterher bitte ich Sie, **den Zeitpunkt** (entsprechend der Uhrposition) **zu berichten, an dem Sie den Ton gehört haben.**

Wir starten mit drei Übungsdurchläufen. Danach folgen 120 Wiederholungen. **Jeweils nach 20 Durchgängen gibt es eine Pause.**

n-back

Ihnen wird im Experiment eine **Abfolge von Buchstaben** präsentiert. Ihre Aufgabe besteht darin, die Reaktionstaste des Gamepads zu betätigen, **wenn der angezeigte Buchstabe mit dem vorletzten Buchstaben übereinstimmt**. Beispiel:

A B C D <u>C</u> F G H <u>G</u>

Die unterstrichenen Buchstaben stimmen jeweils mit dem vorletzten Buchstaben überein, das heißt hier sollte jeweils mit die Reaktions-Taste betätigt werden.

Sie haben auch hier die Möglichkeit, **selbstständig Pausen zu machen**. Die Pause-Funktion verlängert die Versuchsdauer nicht. **Sie bearbeiten die Aufgabe genau 30 Minuten, unabhängig von der**

Pausenanzahl. Die Pause startet mit einer Verzögerung von 800 ms. Bitte achten Sie darauf, dass Sie sich erst bewegen und blinzeln, wenn auf dem Bildschirm „PAUSE“ angezeigt wird. Sie können die Pause entweder selbst über erneutes Betätigen des Pause-Buttons beenden, andernfalls wird die Pause nach 30 Sekunden automatisch beendet.

Der Proband, der die wenigsten Fehler macht, gewinnt einen Gutschein. Das heißt, Sie können sich durch häufige Pausen einen Vorteil verschaffen.

Viel Erfolg!

Study 2: Qualitative Observations

Bedingungsabfolge (A: Selbst-initiierte Handlung ohne Zeitrückmeldung; B: Selbst-initiierte Handlung mit Zeitrückmeldung; C: vorgeplante Handlung; D: Externe Stimulierung)

Nachbefragung (A):

- Haben Sie vorgeplant? Waren Sie sich einer Vorplanung bewusst?
- Fiel es schwer, sich nicht auf die Handlung vorzubereiten?
- Waren Sie selbst überrascht vom Zeitpunkt Ihrer Handlung?
- Ist Ihnen irgendetwas aufgefallen?

Nachbefragung (B):

- Haben Sie vorgeplant? Waren Sie sich einer Vorplanung bewusst?
- Fiel es schwer, sich nicht auf die Handlung vorzubereiten?
- Waren Sie selbst überrascht vom Zeitpunkt Ihres Willens/ Ihrer Handlung?
- Fiel es Ihnen schwer, durchgängig den Mittelpunkt zu fixieren?
- Fiel es Ihnen schwer, den Zeitpunkt Ihres Handlungswillens zu schätzen?
- Ist Ihnen irgendetwas aufgefallen?

Nachbefragung (C):

- Fiel es Ihnen schwer, sich auf die Handlung vorzubereiten?
- Ist Ihnen irgendetwas aufgefallen?

Nachbefragung (D):

- Fiel es Ihnen schwer, durchgängig den Mittelpunkt zu fixieren?
- Fiel es Ihnen schwer, den Zeitpunkt des Tons zu schätzen?
- Ist Ihnen irgendetwas aufgefallen?
- Welche der Bedingungen fanden Sie am anstrengendsten (Reihenfolge aufschreiben)? Warum?
- Welche der Bedingungen fanden Sie am langweiligsten (Reihenfolge aufschreiben)?
- Warum?

N-back

- Hatten Sie eine Strategie, wie Sie sich für die Pausen entschieden haben?

Study 6: Manipulation Materials

Anti-free will condition

Lesen Sie den Text bitte aufmerksam durch.

Francis Crick war der britische Physiker und Biochemiker, der zusammen mit James D. Watson an der Entdeckung der DNA-Molekularstruktur arbeitete, wofür sie im Jahr 1962 den Nobelpreis erhielten. Er ist der Autor von What Mad Pursuit, Life Itself und Of Molecules and Men. Dr. Crick dozierte weltweit vor Fach- und Laienpublikum und war ein anerkannter Professor in der Forschung am Salk Institut in La Jolla, CA. Dr. Cricks (untenstehendes) Essay stammt aus seinem Buch Was die Seele wirklich ist (The Astonishing Hypothesis).

Ein Postskriptum über den Freien Willen

„Sie", Ihre Freuden und Leiden, Ihre Erinnerungen, Ihre Ziele, Ihr Sinn für Ihre eigene Identität und Willensfreiheit – bei alle dem handelt es sich in Wirklichkeit nur um das Verhalten einer riesigen Ansammlung von Nervenzellen und dazugehörigen Molekülen. Sie sind nichts weiter als ein Haufen Neuronen.

Den meisten Religionen zufolge existiert irgendeine Art von Freiem Willen, der in einem gewissen Maße das Wesen des einzelnen Menschen ausmacht. Die Religionen stimmen in den Einzelheiten häufig nicht überein, trotzdem gibt es eine breite Übereinstimmung, dass Menschen einen Freien Willen haben.

Die heute gängige Meinung zeichnet eine völlig andere Sicht und hält die Vorstellung von einer Seele, die nichts Körperliches ist und den bekannten Gesetzen der Wissenschaft nicht unterworfen ist, für ein Märchen. Es ist leicht zu verstehen, wie solche Märchen ohne detailliertes Wissen über die Natur von Materie und Strahlung, sowie über die biologische Evolution entstehen konnten. Zum Beispiel glaubte vor ungefähr 4000 Jahren fast jedermann, dass die Erde eine Scheibe sei. Nur durch den Fortschritt der modernen Wissenschaft wurde uns verdeutlicht, dass die Erde in Wirklichkeit rund ist.

Von der modernen Wissenschaft wissen wir, dass alle Lebewesen, von den Bakterien bis zu uns selbst, auf der biochemischen Ebene eng verwandt sind. Wir wissen, dass sich viele Pflanzen und Tierarten über die Zeit entwickelt haben. Wir können die grundlegenden Vorgänge der Evolution heute sowohl in der freien Wildbahn, als auch in unseren Reagenzgläsern beobachten, deshalb braucht es die religiöse Vorstellung eines freien Willens nicht, um das Verhalten von Menschen und anderen Lebewesen zu erklären. Neben den Wissenschaftlern sind viele gebildete Menschen der Überzeugung, dass der Freie Wille eine Metapher ist.

Die meisten Menschen setzen den Freien Willen als selbstverständlich gegeben voraus, denn sie haben den Eindruck, dass es ihnen gewöhnlich freisteht, so zu handeln, wie es ihnen gefällt.

Drei Annahmen können über den freien Willen getroffen werden. Die erste Annahme ist: Ein Teil des Gehirns ist damit beschäftigt, Pläne für das Handeln in der Zukunft zu machen, die natürlich nicht unbedingt ausgeführt werden. Die zweite Annahme ist: Man ist sich der „Berechnungen“ nicht bewusst, die der betreffende Teil des Gehirns anstellt, sondern nur der „Entscheidungen“, die sich daraus ergeben – d.h. der resultierenden Pläne, die natürlich auch von den Inputs abhängen, die dieser Teil des Gehirns zu diesem Zeitpunkt von anderen Teilen des Gehirns empfängt. Die dritte Annahme ist, dass für die Entscheidung, den einen oder den anderen Plan auszuführen, das gleiche gilt, also, dass der Inhalt der Entscheidung, aber nicht die Berechnungen, die in sie eingegangen sind, unmittelbar abgerufen werden können.

Obwohl es uns so erscheint als hätten wir einen Freien Willen, ist es in Wirklichkeit so, dass unsere Entscheidungen schon vorab für uns getroffen wurden und wir dieses nicht ändern können. Die Ursache der Entscheidung kann scharf umrissen sein, sie kann aber auch deterministisch und zugleich chaotisch sein, d.h., eine sehr kleine Schwankung kann für das Endresultat sehr viel ausmachen. Daraus würde der Anschein entstehen, der Wille sei „frei“, denn das Ergebnis wäre dann ja eigentlich unvorhersagbar. Natürlich können auch bewusste Aktivitäten Einfluss auf den Entscheidungsmechanismus haben.

Unser Selbst kann versuchen, sich zu erklären, warum es eine bestimmte Entscheidung getroffen hat. Gelegentlich gelangen wir vielleicht zur richtigen Schlussfolgerung. In anderen Fällen werden wir es nicht wissen oder, was wahrscheinlicher ist, einfach konfabulieren (erfundene Erlebnisse als selbst erlebt darstellen), weil es ja von den „Gründen“ für die Entscheidung keinerlei bewusste Kenntnis hat. Daraus ergibt sich, dass es einen Mechanismus des Konfabulierens geben muss, und das bedeutet, dass ein Teil des Hirns angesichts einer gewissen Menge von Anhaltspunkten (die vielleicht irreführend sind, vielleicht aber auch nicht), blindlings den einfachsten Schluss daraus zieht.

Control condition

Lesen Sie den Text bitte aufmerksam durch.

Francis Crick war der britische Physiker und Biochemiker, der zusammen mit James D. Watson an der Entdeckung der DNA-Molekularstruktur arbeitete, wofür sie im Jahr 1962 den Nobelpreis erhielten. Er ist der Autor von What Mad Pursuit, Life Itself und Of Molecules and Men. Dr. Crick dozierte weltweit vor Fach- und Laienpublikum und war ein anerkannter Professor in der Forschung am Salk Institut in

La Jolla, CA. Dr. Cricks (untenstehendes) Essay stammt aus seinem Buch Was die Seele wirklich ist (The Astonishing Hypothesis).

Die allgemeine Natur des Bewusstseins

Psychologen haben gezeigt, dass uns der gesunde Menschenverstand im Hinblick auf die Funktionsweisen des Geistes in die Irre führen kann. Als die Psychologie begann, eine experimentelle Wissenschaft zu werden, vornehmlich in der zweiten Hälfte des 19. Jahrhunderts, bestand ein reges Interesse am Bewusstsein. Man hoffte die Psychologie könnte dadurch wissenschaftlicher werden, dass man die Introspektion (Selbstbeobachtung) zu einer zuverlässigen Technik verfeinerte.

Da das Problem des Bewusstseins derart zentral ist und das Bewusstsein so geheimnisvoll erscheint, könnte man vielleicht erwarten, dass Psychologen und Neurowissenschaftler sich heutzutage im großen Rahmen um ein Verständnis des Bewusstseins bemühen. Dies ist allerdings bei weitem nicht der Fall. Die Mehrzahl der modernen Psychologen vermeidet jederlei Erwähnung des Problems, obwohl vieles von dem, wozu sie Untersuchungen anstellen, für das Bewusstsein eine Rolle spielt. Die meisten Neurowissenschaftler ignorieren das Problem.

Der amerikanische Psychologe William James hat sich in seinem Werk „The Principles of Psychology" (1890) ausführlich mit dem Bewusstsein auseinandergesetzt und beschreibt fünf Eigenschaften dessen, was er „Denken" nennt. Jeder Gedanke, so schrieb er, hat die Tendenz, Teil des personalen Bewusstseins zu sein. Das Denken befindet sich in ständiger Veränderung, hat eine spürbare Kontinuität und scheint sich mit Objekten zu befassen, die von ihm unabhängig sind. Zudem konzentriert sich das Denken auf gewisse Objekte und blendet dabei andere aus. Anders gesagt: zum Denken gehört Aufmerksamkeit. Über die Aufmerksamkeit machte James folgende Feststellung: „Der Geist nimmt einen Gegenstand klar und lebhaft in Besitz, obwohl zur gleichen Zeit verschiedene Gegenstände (oder Gedankensequenzen) präsent sind. Zur Aufmerksamkeit gehört es, sich von gewissen Dingen zurückzuziehen, um sich mit andern wirkungsvoller auseinanderzusetzen."

Viele Psychologen glaubten, dass gewisse Prozesse unterschwellig oder unbewusst seien. Ein Beispiel ist, dass die Wahrnehmung im Hinblick auf ihre logische Struktur dem gleicht, was wir normalerweise als Schlussfolgerung bezeichnen, auch wenn dies weitgehend unbewusst geschieht. Drei Grundideen des Bewusstseins waren geläufig. Erstens, nicht alle Tätigkeiten des Hirns haben eine Entsprechung im Bewusstsein. Zweitens, das Bewusstsein umfasst irgendeine Art von Gedächtnis, wahrscheinlich ein Ultrakurzzeitgedächtnis. Drittens, Bewusstsein hängt eng mit Aufmerksamkeit zusammen.

Leider entstand in der akademischen Psychologie eine Bewegung, die dem Begriff des Bewusstseins jede Nützlichkeit für die Psychologie absprach. Das lag zum Teil daran, dass es so schien, als würden

Experimente, in denen Introspektion eine Rolle spielt, zu gar nichts führen, und zum Teil auch daran, dass man hoffte, die Psychologie werde dadurch wissenschaftlicher werden, dass man unzweideutig experimentell beobachtbares Verhalten untersucht. Diese Bewegung wurde Behaviorismus genannt. Über geistige Ereignisse zu sprechen, war tabu. Verhalten musste ganz und gar durch Reiz und Reaktion erklärt werden.

Wie können wir das Bewusstsein wissenschaftlich angehen? Das Bewusstsein nimmt viele Formen an, doch wie ich bereits erläutert habe, lohnt es sich normalerweise, sich auf diejenige Form zu konzentrieren, die allem Anschein nach am leichtesten zu untersuchen ist. Christof Koch und ich haben das visuelle Bewusstsein allen anderen Bewusstseinsformen (wie z.B. Schmerz oder Selbst-Bewusstsein) vorgezogen, weil Menschen ausgesprochen visuelle Lebewesen sind und weil unser visuelles Bewusstsein besonders lebhaft und reich an Information ist. Außerdem ist der Input hier oft hochgradig strukturiert und dennoch leicht zu kontrollieren. Deshalb gibt es auch schon viele experimentelle Untersuchungen zum visuellen Bewusstsein.

Study 7: Manipulation Materials

Free will condition

1. Ich beweise meinen freien Willen jeden Tag, wenn ich Entscheidungen treffe.
2. Ich bin in der Lage, die genetischen Faktoren und Umweltfaktoren, die manchmal mein Verhalten beeinflussen, zu überwinden.
3. Ich bereue es, wenn ich schlechte Entscheidungen treffe, weil ich weiß, dass ich letztendlich für meine Handlungen verantwortlich bin.
4. Ich bin stolz auf gute Entscheidungen, die ich in der Vergangenheit getroffen habe, weil ich weiß, dass ich zu der Zeit die Freiheit hatte, auch schlechte Entscheidung treffen zu können.
5. Um Versuchungen zu widerstehen, ist es erforderlich, dass ich meinen freien Willen einsetzte/anwende.
6. Letztendlich können Menschen ihre eigenen Handlungen niemand anderem vorwerfen als sich selbst.
7. Ich habe den freien Willen, meine Handlungen und letztendlich mein Schicksal zu kontrollieren.
8. Ich bin mehr als ein Roboter, der von Genen und Umwelt programmiert wurde, egal was einige Wissenschaftler behaupten.
9. Menschen sind verantwortlich für ihr Verhalten, weil sie den freien Willen haben, um ihre Handlungen zu kontrollieren.
10. Unsere Handlungen und Gedanken sind nicht einfach das Ergebnis vorheriger Erfahrungen.

Anti-free will condition

1. Letztendlich sind wir biologische Computer – von der Evolution gestaltet, durch Genetik gebaut und vom Umfeld programmiert.
2. Das Gehirn ist eine komplexe Maschine, die im Stande ist, extrem hochentwickelte Verhaltensweisen auszuführen.
3. Die Wissenschaft hat bewiesen, dass der freie Wille eine Illusion ist.
4. Es ist wahrscheinlich, dass Wissenschaftler irgendwann verstehen werden, wie das Gefühl persönlicher Erfahrung aus dem Feuern von Neuronen im Gehirn entsteht.
5. Alles was eine Person tut ist eine direkte Folge ihres Umfelds und ihres Erbguts.
6. Sobald Wissenschaftler die physischen Gesetzmäßigkeiten, die dem Verhalten zugrunde liegen, genug verstehen, sollten sie in der Lage sein, das zukünftige Verhalten einer Person einzig und

allein basierend auf den Erbinformationen der Person und früheren Erfahrungen genau vorhersagen zu können.

7. Unsere Handlungen sind davon bestimmt, was wir in der Vergangenheit erlebt haben, kombiniert mit dem spezifischen Erbgut, das wir haben.
8. Wie alles andere im Universum, ergeben sich alle menschlichen Handlungen aus früheren Ereignissen und können letztendlich in Form von Bewegung der Moleküle verstanden werden.
9. Glaube an den freien Willen widerspricht der bekannten Tatsache, dass das Universum von gesetzmäßigen Prinzipien der Wissenschaft geregelt wird.
10. Unsere geistigen Aktivitäten sind ausschließlich das Produkt der physikalischen Prozesse.

Control condition

1. Die Ozeane bedecken 71% der Erdoberfläche.
2. Alkaline Batterien halten generell länger als normale Batterien.
3. Monarchschmetterlinge fliegen langsam, wurden aber hunderte Meilen auf hoher See gesichtet.
4. Die Olympischen Spiele werden alle vier Jahre veranstaltet.
5. Einen halben Tag Bootsfahrt von Athen entfernt liegt die kleine Insel Mykonos.
6. Zuckerrohr und Zuckerrüben werden in 112 Ländern angebaut.
7. Viele der Berggipfel in den Rocky Mountains sind höher als 4270m.
8. Die Appalachen sind abgenutzte Berge und Plateaus, die sich vom nördlichen Alabama bis zum St.-Lawrence-Fluss in Kanada erstrecken.
9. Die größte Entfernung zwischen Erde und Sonne beträgt 152.005.760 km.
10. Der Nil in Afrika ist der längste Fluss der Welt.

Danksagung

In den letzten Jahren habe ich von vielen Seiten Unterstützung bei der Anfertigung dieser Arbeit erfahren. Bedanken möchte ich mich besonders bei meinem Doktorvater Prof. Dr. Hans-Rüdiger Pfister für die Betreuung meines Dissertationsprojekts, für die konstruktive Kritik und dafür, dass er es mir ermöglichte, frei und unabhängig zu forschen, sowie bei Prof. Dr. Rainer Höger für die Unterstützung und die Begutachtung meiner Dissertation. Dass ich in meiner Zeit am Institut für Experimentelle Wirtschaftspsychologie zu dieser Fragestellung gekommen bin, verdanke ich beiden. Ich danke Prof. Dr. Oliver Genschow für die Übernahme des Drittgutachtens und Prof. Friedrich Müller für die vielen inspirierenden Gespräche. Besonderer Dank gilt Anne Lange, die diese Arbeit in ihren Anfängen maßgeblich mit vorangebracht hat und mich in allen Phasen des Projekts begleitet und gestärkt hat. Ich danke Natascha Meyke, Marie-Kristin Wett und Katharina Stöber dafür, dass auch sie dieses Thema begeistert hat, sie durch ihre Bachelorarbeiten einen Teil des Weges mit mir gegangen sind und diese Arbeit tatkräftig unterstützt haben. Ich danke Anja Huemer und Daniela Gundert für die zahllosen Diskussionen, die das Profil dieser Arbeit geschärft haben. Ich danke Andreas Melcher für seine Nachsicht in schwierigen Zeiten und dafür, dass er mir geistreich und ermutigend auch auf diesem Weg zur Seite stand. Danken möchte ich auch meiner Familie und meinen Freunden, und mich auf ganz unterschiedliche Weise in allen Phasen meines akademischen wie nicht-akademischen Lebens unterstützt haben.