A meta-analysis of Libet-style experiments

Moritz Nicolai Braun[a,*], Janet Wessler[a], Malte Friese[a]

[a] Department of Psychology, Saarland University, Campus, Geb. A2.4, 66123 Saarbrücken, Germany

[*] Corresponding author.

E-mail addresses: moritz.nicolai.braun@gmail.com (M. N. Braun), janet.wessler@uni-saarland.de (J. Wessler), malte.friese@uni-saarland.de (M. Friese).

Declarations of interest: none.

This article is in press at Neuroscience and Biobehavioral Reviews.

https://doi.org/10.1016/j.neubiorev.2021.06.018

© 2021. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

Abstract

In the seminal Libet experiment (Libet et al., 1983), unconscious brain activity preceded the self-reported, conscious intention to move. This was repeatedly interpreted as challenging the view that (conscious) mental states cause behavior and, prominently, as challenging the existence of free will. Extensive discussions in philosophy, psychology, neuroscience, and jurisprudence followed, but further empirical findings were heterogeneous. However, a quantitative review of the literature summarizing the evidence of Libet-style experiments is lacking. The present meta-analysis fills this gap. The results revealed a temporal pattern that is largely consistent with the one found by Libet and colleagues. Remarkably, there were only $k = 6$ studies for the time difference between unconscious brain activity and the conscious intention to move — the most crucial time difference regarding implications about conscious causation and free will. Additionally, there was a high degree of uncertainty associated with this meta-analytic effect. We conclude that some of Libet et al.'s findings appear more fragile than anticipated in light of the substantial scientific work that built on them.

Word count: 169

Keywords: Libet study, volition, intentional action, readiness potential, meta-analysis, free will, W-time, M-time

1. Introduction

In their seminal experiment, Libet and colleagues (1983) found that unconscious brain activity preceded the conscious intention to perform a simple motor action, which in turn preceded subjective awareness of this action and the action itself. This empirical pattern stimulated two prominent and intertwined lines of debate. First, it is discussed in the realm of conscious mental states and their ability to cause behavior (Baumeister, 2008; Baumeister et al., 2011; List, 2019). In these debates, the empirical pattern found by Libet et al. has repeatedly been interpreted to question conscious causation (e.g., Roediger et al., 2008; Pockett, 2006, cf. Mele, 2009). Second, the pattern has been prominently interpreted as questioning the existence of free will (e.g., Bargh, 2008; Wegner, 2002). What is more, the experiment not only spurred debates in several academic disciplines, including philosophy, jurisprudence, neuroscience, and psychology. It also continues to make its way into popular media (e.g., Gholipour, 2019; Overbye, 2007; Stafford, 2015).

Even though the Libet experiment has inspired scholars' thinking and empirical research for nearly 40 years, a comprehensive quantitative review of the literature is lacking. Some studies confirmed Libet et al.'s findings (e.g., Dominik et al., 2018; Haggard & Eimer, 1999), while others yielded divergent findings (e.g., Dominik et al., 2017; Sanford et al., 2020), leading to heterogeneity in the literature. To address these challenges, the present work provides the first meta-analysis of the temporal pattern found in the Libet experiment. We sought to provide answers to two pivotal questions: 1) Is the temporal pattern found by Libet and colleagues robust and, consequently, a suitable foundation for theoretical positions? 2) Are there methodological or other moderating conditions that systematically alter the temporal pattern, potentially calling into question the validity of Libet-style experiments?

1.1. The seminal Libet experiment

Debates on free will have a long philosophical tradition, preoccupying the likes of Plato and Aristotle in ancient Greece and continuing to keep thinkers busy in the 21st century (e.g., Mele, 2014; Pereboom, 2001; van Inwagen, 2008). Theorizing about the existence and nature of free will is directly connected to theorizing about moral responsibility (Caruso, 2013; Strawson, 1994; Vargas, 2013). Law, politics and morality all draw their legitimacy from the premise that humans possess free will. Together with our ability to speak, free will distinguishes us from other animals (Chomsky, 2006).

For a long time, free will was exclusively the object of non-empirical, philosophical debates. Things changed in 1983, when Libet and colleagues presented the first promising empirical neuropsychological approach to the concept of free will (Libet et al., 1983). The discovery of and research on the readiness potential prepared the ground for this breakthrough. The readiness potential is a slow negative potential shift that begins up to a second or more before a self-paced movement that is recorded from the scalp (Gilden et al., 1966; Kornhuber & Deecke, 1965). It has been interpreted as a specific causal precursor of voluntary actions (Deecke et al., 1976; Kornhuber & Deecke, 1965; Libet et al., 1982; Travers et al., 2020).

Libet and colleagues wondered about the temporal relation between the onset of the readiness potential and the conscious intention to move in self-paced, voluntary movements. In their seminal experiment, they asked participants to sit in front of a clock and perform a "quick, abrupt flexion of the fingers and/or the wrist of their right hand" without preplanning their action (Libet et al., 1983, p. 625). They detected the onset of this movement using an electromyograph (EMG), a device recording the electrical activity of skeletal muscles. To assess the readiness potential, they monitored participants' brain activity using an electroencephalograph (EEG).

In some series of trials, participants reported when they *wanted* to perform the movement (i.e., the time of their conscious intention to move; often called "W" or "W-time"). In another series of trials, participants reported when they first became aware that they had *actually moved* (i.e., the time of their awareness of the movement; often called "M" or "M-time").

This procedure yielded the following temporal pattern: On average, onset times of the readiness potential preceded the conscious intention to move by about 335 ms, leading Libet and colleagues to conclude that "the initiation of voluntary action begins before an individual makes a conscious decision to move and by that starts unconsciously" (Libet et al., 1983, p. 640). This was the key finding of this study. In addition, and as expected, the conscious intention to move preceded subjective awareness of the movement by on average 115 ms. Libet and colleagues further found, expectedly, that the conscious intention to move preceded the actual onset of the corresponding movement by about 200 ms. They interpreted this gap as allowing some conscious veto power to potentially abort the unconsciously initiated movement. The temporal pattern found by Libet and colleagues is summarized in Figure 1.

1.2. Impact of Libet et al.'s (1983) findings

The findings reported by Libet et al. (1983) were widely interpreted to question that conscious mental states cause behavior and thus an intuitive concept of free will, according to which agents themselves – not unconscious brain activity – are the causes of their actions (e.g., Roediger et al., 2008; Pockett, 2006, cf. Mele, 2009). This is especially true for the time difference between the onset of the readiness potential and the conscious intention to move. The results therefore sparked debates in various scientific disciplines (e.g., Baumeister et al., 2011; Beckermann, 2008; Fischborn, 2016; Nahmias, 2015).

In philosophy, the idea that unconscious brain activity precedes any conscious intention for movement continues to make its way into debates about the role of conscious intention and mental states in volitional movements (List, 2019; Nahmias, 2005; Schlosser, 2012) and about the existence and nature of free will (e.g., List, 2019; Mele, 2014; Pereboom, 2001). Relatedly, in the realm of jurisprudence, Libet et al.'s findings have led scholars to contemplate the question of (moral) responsibility for one's actions if free will were found to be only an illusion (Caruso, 2013; Focquaert et al., 2013; Vargas, 2013).

In psychology, influential works built on Libet et al.'s findings to argue that free will does not exist (e.g., Bargh, 2008; Wegner, 2002). Libet et al.'s findings also inspired discussions of conscious causation (e.g., Baumeister et al., 2011) and stimulated research examining the consequences of (measured and manipulated) beliefs in free will (for an overview, see Baumeister & Brewer, 2012).

Finally, neuroscientific studies of volition and agency have been heavily influenced by the Libet experiment (see Wolpe & Rowe, 2014, for an overview). The intentional binding paradigm (Christensen et al., 2019; Haggard et al., 2002; Suzuki et al., 2019), to mention just one example, builds upon essential components of the original Libet paradigm (see Moore & Obhi, 2012, for a review). Its core result, intentional binding, refers to the effect that participants report shorter intervals between the onset of an action and its sensory consequence when the action is voluntary compared to involuntary.

To what extend the Libet experiment (Libet et al., 1983) and its results actually speak to the question of free will is debated (e.g., Gomes, 1998; Mele, 2009; Moore & Haggard, 2008). Some theorists assume that free will evolved for adaptive reasons to enable humans to successfully function in culture (e.g., Baumeister, 2005, 2008). From these perspectives, free will is especially important for meaningful, reasoned actions that may have moral or social implications. To make such decisions, conscious deliberation and planning seem central. However, such properties are explicitly left out of Libet-style experiments. In these experiments, participants solely decide about the timing of simple, predefined (finger or wrist) movements (see also Brass et al., 2019). Thus, it is unclear what implications these experimental setups have for contexts in which free will arguably matters most.

A meta-analysis of studies relying on a particular experimental paradigm cannot establish what this paradigm is or is not able to reveal about the role of free will in decision-making. This is a property of the paradigm itself and not a question of the empirical data obtained across studies. However, what can be confidently said is that the Libet experiment has become a linchpin for hosts of scientific research, medial presence, and public debate with far-reaching implications for social life. Given this impact of the Libet experiment and its successors, it is essential to comprehensibly scrutinize the empirical evidence that is the basis of all discussion. Whatever the interpretation of the temporal pattern in Libet-style experiments might be, it will only be reputable if it is based on robust, generalizable empirical data.

Maybe surprisingly, the Libet experiment can lay no claim to such data. Data from only five participants were included in the original experimental analysis. This meager data base stands in sharp contrast to the far-reaching conclusions, extensive research programs and major implications the findings have inspired. Fortunately, on the one hand, a variety of detailed research has been conducted in the tradition of the Libet experiment. On the other hand, this research is considered far less often than the Libet experiment itself. This becomes especially clear when comparing the more than 6,800 citations of Libet's original articles on Google Scholar (Libet et al., 1983, 3542 citations; Libet, 1985, 3307 citations) to the 1,394 citations of the second most cited article included in the present meta-analysis (date of retrieval: June 03, 2021). Thus, the temporal pattern found in the original Libet experiment (Libet et al., 1983) still lays at the heart of the debate. Therefore, it is all the more important to investigate whether these results are robust.

Subsequent research on the Libetian temporal pattern (Libet et al., 1983) is characterized by a variety of methodological alterations. These alterations not only yield heterogeneous results but also point to potential moderating variables. Such methodological moderators might even substantially challenge the temporal pattern and turn the whole interpretation upside down. For example, one experiment examined the onset times of the lateralized readiness potential, a lateralized version of the readiness potential that is associated with hand-specific movement preparation (Trevena & Miller, 2002). Here, the conscious intention to move preceded the lateralized readiness potential. This indicates – in direct contrast to the Libetian temporal pattern – that specific movement preparation as indicated by the lateralized readiness potential might start *after* the conscious decision to move.

In another series of studies, one group of participants first completed a set of trials in which they reported the time of their movement awareness before completing a set of trials in which they reported the time of their intention to move. In another group, the sequence was reversed (Dominik et al., 2017; Sanford et al., 2020). Brain activity was not assessed in these studies. Only the group of participants who first reported their awareness and then their intention yielded a Libetian-like temporal pattern. In the group who had first reported their intention, intention and awareness preceded actual movement by a similar amount of time. These results question (part of) the Libetian temporal pattern as well as its interpretations.

Further methodological alterations yielded temporal patterns that partially differed from the original one (Libet et al., 1983). Some studies used a randomized stream of letters instead of a clock as the method of time measurement (e.g., Haynes, 2011; Parés-Pujolràs et al., 2019). Some studies used functional magnetic resonance imaging (fMRI) instead of an EEG to monitor participants' brain activity with a higher spatial resolution (e.g., Hirose et al., 2018; Soon et al., 2008), while others only collected behavioral and not neural data (e.g., Giovannelli et al., 2016; Lush et al., 2016). In other studies, participants performed actions other than the original flexion of the wrist and/or fingers of the right hand (Libet et al., 1983). These actions included a mouse click (e.g., Douglas et al., 2015; Lau et al., 2007), a button press (e.g., Doñamayor et al., 2018; Lafargue & Duffau, 2008) or more complex actions like typing strings consisting of one, three or five letters (Haggard et al., 1999). All these variations may moderate the original temporal pattern found by Libet and colleagues (1983).

1.3. The present meta-analysis

Although some narrative reviews of Libet-syle experiments (Saigle et al., 2018), neuronal correlates of intentional action (e.g., Krieghoff et al., 2011), and the validity of the neuroscientific challenge that Libet-style experiments pose to the existence of free will exist (Brass et al., 2019), a systematic quantitative review is lacking. We conducted a meta-analysis to fill this gap. Note again that this meta-analysis is neither capable of nor seeking to answer the question whether or not free will exists or even whether experiments in the Libet tradition can unequivocally speak to this question. These questions are beyond what can be achieved with a meta-analysis. Instead, what we sought to accomplish was to summarize all available evidence on the Libetian temporal pattern (Libet et al., 1983) that is the basis for much scientific and public debate. That is, we examined the temporal relation between the onset of unconscious brain activity, the time of the conscious intention to move and the time of subjective awareness of movement onset in participants who are asked to freely choose the time of one specific movement. Based on results from brain imaging studies, it has repeatedly been proposed that intentional action consists of a what, a when and a whether component (Brass & Haggard, 2008; Haggard, 2008). These components reflect the decisions of what action to perform, when to perform this action and whether to perform this action at all (so-called "veto"), respectively. According to this view, the Libet experiment (Libet et al., 1983) focused on the when component. In line with this view, we argue that experiments investigating the what or whether component of intentional action focus on different decisions involved in intentional actions. Hence, the present meta-analysis focuses on experiments that investigated the when component of intentional action. We thereby aimed at answering the following research questions:

1. What is the mean time difference between the onset of unconscious brain activity and the conscious intention to move? (Onset Brain Activity minus Intention)
2. What is the mean time difference between the onset of unconscious brain activity and the actual onset of movement? (Onset Brain Activity)
3. What is the mean time difference between the conscious intention to move and the actual onset of movement? (Intention)
4. What is the mean time difference between subjective awareness of the onset of movement and the actual onset of movement? (Awareness)
5. What is the mean time difference between the conscious intention to move and subjective awareness of the onset of movement? (Intention minus Awareness)

While the magnitude of these differences is unclear, their direction is uncontroversial for two of the research questions. One would assume unconscious brain activity to precede the actual onset of a movement (RQ 2) and the intention to move to precede the actual onset of a movement (RQ 3). Whether subjective awareness of the onset of a movement should precede its actual onset (RQ 4) is less clear. Intuitively, one would assume a delay between a movement and the subjective awareness of its onset. However, in the original Libet experiment the subjective awareness preceded the onset of movement (Libet et al., 1983). Further, it is unclear whether the intention to move should precede subjective awareness of the onset of movement (RQ 5). The most crucial research question by far for discussions about conscious causation and free will is RQ 1: Does unconscious brain activity precede the conscious intention to move, and if so, by how much? It is this particular finding by Libet et al. (1983) that together with the interpretation of the brain activity as a causal precursor of the movement (Libet et al., 1982) evoked doubts about conscious causation and free will, much more than the findings examined by the other research questions.

Note that in our preregistration of RQs 1 and 2 we used the expression “preparatory brain activity” instead of “unconscious brain activity”. “Preparatory” implies temporal precedence and a causal function of the brain activity. There is an ongoing debate about these issues (e.g., Maoz et al., 2019; Schurger et al., 2012; Travers et al., 2020) which is not the focus of the present article. We now use the expression “unconscious brain activity” throughout the manuscript because we sought to remain agnostic toward these temporal and functional aspects.

2. Methods

We followed reporting guidelines for meta-analyses outlined in the PRISMA statement (Moher et al., 2009) and recommendations for the reproducibility of meta-analyses (Lakens et al., 2016). The study was preregistered under the international prospective register of systematic reviews (PROSPERO; registration number: CRD42019141909, http://www.crd.york.ac.uk/prospero/). To facilitate future updates of this work, we made all data, code, a complete list of all coded characteristics, the complete coding sheet, and supplementary materials available on the Open Science Framework (OSF, https://osf.io/z962m/).

2.1. Eligibility Criteria

To be as comprehensive as possible, we included both published and unpublished data. Published studies were included if they met eligibility criteria on both the abstract and the full-text level.

We checked the abstract for the following criteria: Studies were eligible for inclusion if they were (a) written in the English or German language, (b) reported original empirical research (i.e., reviews and theoretical papers were excluded), (c) focused on free will or similar notions (e.g., voluntary [motor] actions, volition, intended [motor] actions, self-paced actions; we also checked all studies on the full-text level that mentioned any term in their abstract that might be used as a synonym to these terms to be as inclusive as possible and avoid missing out on any relevant studies) or free won't or mentioned Libet et al.'s (1983) implications or results, and (d) did not investigate samples consisting entirely of participants with clinical conditions.

In addition, we checked the full text for the following criteria: Studies were eligible for inclusion if (e) participants performed a task in which they were asked to freely choose the time of a given motor action (i.e., focused on the *when*-component of intentional action, Brass & Haggard, 2008), (f) they reported at least one of the following times: the time of the conscious intention to move relative to the actual onset of movement (Intention) or the time of the awareness of the onset of movement relative to the actual onset of movement (Awareness). We also checked the "Participants" subsection to exclude any study samples consisting entirely of participants with clinical conditions.

Data from unpublished studies were included if (a) participants performed a task in which they were asked to freely choose the time of a given motor action, (b) participants were asked to report at least one of the following times: the time of their conscious intention to move relative to the actual onset of movement (Intention) or the time of their awareness of the onset of movement relative to the actual onset of movement (Awareness), and (c) the sample did not consist solely of participants with clinical conditions.

2.2. Search Strategy

The search strategy consisted of four parts. First, we searched the two electronic databases PubMed and Web of Science on the 29 and 30 April 2019. Searches were forward searches for articles that cited the landmark study by Libet and colleagues (1983) or a theoretical paper by Benjamin Libet discussing the 1983 experiment and its results (Libet, 1985). Both are seminal papers by Benjamin Libet on the topic, each cited more than 3,000 times according to Google Scholar. We repeated these searches on 29 and 30 December 2019 to retrieve articles that were published after the first search. On the same occasion, we additionally searched the EBSCO database. Only the 1983 article was available on PubMed and only the 1985 article was available in the EBSCO database at the time of the literature search.

The first author screened all results for eligibility on the abstract level. All articles eligible on the abstract level were then retrieved and rigorously checked for the eligibility criteria on the full-text level by the first author and a research assistant. Discrepancies were discussed and settled within the research team. If relevant statistical information was missing, we contacted the authors and asked for this information (e.g., the onset of the readiness potential or the standard deviation of the time reports).

Second, we issued calls for unpublished data via the EEGLAB mailing list (Delorme & Makeig, 2004; http://www.sccn.ucsd.edu/eeglab) and the list server of the biopsychology and neuropsychology section of the German Psychological Society (DGPs). As preregistered, we additionally contacted the Society for Neuroscience (SfN), the Cognitive Neuroscience Society (CNS), and the Federation of European Neuroscience Societies (FENS). The SfN and CNS did not respond to our request and the FENS declined to share our call with their members. Third, we harvested references from a qualitative review on research in the Libet tradition (Saigle et al., 2018) and from all studies included in the meta-analysis. Fourth, we asked all authors of included studies we had corresponded with during the project for additional published or unpublished data.

2.3. Study Coding

The first author and a research assistant extracted information concerning publication characteristics, sample characteristics, study characteristics and statistics from the written study reports. Interrater reliability was examined using the intra-class correlation for continuous moderators (ICC[1,1]; Shrout & Fleiss, 1979) and Cohen's Kappa for categorical moderators (Cohen, 1968). Interrater reliability for the study coding was high by common standards (Cicchetti, 1994), $\kappa = 0.93$, mean ICC(1,1) = 0.95.

2.3.1. Study characteristics

We coded a large number of study characteristics, of which we focus on only a subset in this article. A full list of all coded variables and the complete coding sheet are available in the associated OSF project.

2.3.1.1. Instructions for reporting awareness

The wording of the instructions to report the time of one's subjective awareness of the onset of movement varied across studies. Whether the *time of the movement itself* (e.g., the time of the button press) or the *time of the onset of movement* (e.g., the time when participants first felt that their finger moved) were reported might have influenced the reported times.

2.3.1.2. Instructions for reporting intention

The wording of the instructions to report the time of one's conscious intention to move varied across studies. Whether the time of the *urge* to move, the time of the *intention* to move, the time of the *wanting* to move, or the time of the *decision* to move were reported might have influenced the reported times.

2.3.1.3. Method of time reporting

Some researchers asked participants to report the times of interest via a mouse click on the clock face, via a keyboard, or verbally. We examined whether the modality of the time reports systematically influenced the results.

2.3.1.4. Mode of recall

The participants in the Libet experiment reported the times of interest in two different modes: the absolute mode of recall and the order mode of recall (Libet et al., 1983). These different modes might be a source of systematic variance in the time reports.

2.3.1.5. Monitoring instrument

Lying in an fMRI scanner or wearing an EEG cap might distract participants. Thus, their time reports might systematically differ compared to those of participants in experiments without fMRI or EEG monitoring.

2.3.1.6. Number of trials

Reporting their conscious intention to move and the time of their awareness of the onset of movement should be challenging to participants in Libet-style experiments because they are asked to pay attention to processes they do not usually pay attention to. Hence, participants might need some trials to get used to the task. This might result in more precise time reports when more trials are completed.

2.3.1.7. Order of time reports

Some researchers controlled the order in which participants reported the times of interest (Dominik et al., 2017; Sanford et al., 2020). The temporal pattern these researchers found substantially differed from the temporal pattern found by Libet and colleagues.

2.3.1.8. Time measurement

Some studies used a randomized stream of letters instead of a clock as the method of time measurement (e.g., Haynes, 2011; Parés-Pujolràs et al., 2019). We examined whether the time measurement used systematically influenced the results.

2.3.1.9. Type of action

Some actions are more complex than others. Complex actions (e.g., writing a string of letters) might differ from simpler actions (e.g., pressing a key) regarding the time and the nature of the associated intention. Furthermore, different types of actions might cause different tactile experiences. Hence, the corresponding times of subjective awareness of the onset of movement might differ.

2.4. Meta-analytic procedure

2.4.1. Effect size

We pre-registered converting all effect sizes concerning the time differences mentioned in the research questions into Hedges' g_{av} as suggested by Lakens (2013). There are two reasons why we deviated from this plan. First, none of the publications included in this meta-analysis report standardized effect sizes concerning the time differences outlined in the research questions. They exclusively report unstandardized effect sizes (i.e., time differences in milliseconds). Second, calculating Hedges' g_{av} and its variance requires an estimation of the correlation between the time measurements within one sample. Both Hedges' g_{av} and its variance are highly dependent on the value of this estimation. However, none of the publications included in this meta-analysis reported this correlation.

For these reasons, we decided to calculate mean raw time differences (*D*) for each study as effect sizes and used the corresponding standard error means (*SEM*) as estimates of the standard deviations of the distributions of these time differences. Thus, we followed the logic for dependent sample statistics (Borenstein et al., 2009). The interpretation of these mean raw time differences is straightforward and perfectly suits our research questions. Furthermore, it is in alignment with the scientific debate on Libet-style experiments, which relies on these mean raw time differences (e.g., Gomes, 1998; Walter, 2011).

To compute the time difference between the onset of unconscious brain activity and the conscious intention to move (RQ 1) and the time difference between the conscious intention to move and subjective awareness of the onset of movement (RQ 5) and their corresponding standard deviations, we needed an estimate of the correlation between multiple time measurements within one sample. We calculated this correlation from the raw data by Lush and colleagues (2016). This data is published on the OSF and is the only publicly available set of behavioral data from the studies included in the present meta-analysis. This calculation yielded $r = .399$ for the correlation between the conscious intention report and the report of the subjective awareness of the onset of movement within Lush et al.'s sample. It serves as an estimate for the correlation between multiple time measurements within one sample. We then conducted sensitivity analyses for this estimate with $r = .2$ and $r = .6$ instead of $r = .399$ to examine the influence of the magnitude of this correlation on the results. Varying the magnitude of this correlation did not meaningfully change the results as reported below. The results of this sensitivity analysis can be found in Supplementary Table S1.

2.4.2. EMG onset correction

Different definitions of the onset of a movement are used in the literature. The majority of studies define the actual onset of movement as the moment in which an apparatus (i.e., a computer) recognizes the movement onset (e.g., a mouse click, Douglas et al., 2015; Lau et al., 2007). In the Libet experiment (Libet et al., 1983) and a minority of other studies (e.g., Rigoni et al., 2013; Sirigu et al., 2004), the actual onset of a movement is defined as the onset of an EMG signal that records the activity of muscles responsible for the movement. We used this latter definition in the present meta-analysis for two reasons. First, the detection of muscle movements is arguably the more valid indicator compared to the detection of, for example, a button press that necessarily trails the beginning of the muscle movement. Second, this approach secures best comparability of the meta-analytical results and those reported by Libet et al. (1983).

An EMG recognizes the movement onset approximately 60 ms earlier than an apparatus (Dominik et al., 2018). The weighted mean reported delay in the articles included in the present meta-analysis is 63.09 ms ($M = 59.73$ ms) and the reported delays range between

40.6 ms (Sirigu et al., 2004) and 76.2 ms (Haggard et al., 2002). Thus, we corrected the reported times from studies that defined the actual onset of a movement as the moment in which an apparatus recognizes the movement onset by 60 ms. Further, we ran sensitivity analyses with 20 ms, 40 ms and 80 ms instead of 60 ms for the time difference between the onset of unconscious brain activity and the actual onset of movement (RQ 2), the time difference between the conscious intention to move and the actual onset of movement (RQ 3), and the time difference between subjective awareness of the onset of movement and the actual onset of movement (RQ 4).

The majority of the times reported in the included articles had to be corrected, because these studies did not use EMG, but apparatuses to detect the onset of movement. Hence, varying the magnitude of the correction results in a linear transformation of the absolute values of the corrected time differences. Importantly, the temporal order of the time differences and their relative differences are unaffected by the magnitude of the correction. The results of the sensitivity analyses for the EMG onset correction can be found in Supplementary Table S2. The time difference between the onset of unconscious brain activity and the conscious intention to move (RQ 1) and the time difference between the conscious intention to move and subjective awareness of the onset of movement (RQ 5) rely on two reported times each. Therefore, no EMG onset correction was necessary in these cases.

2.4.3. Data synthesis

We computed summary effects using intercept-only random-effects models following the robust variance estimation (RVE; Hedges et al., 2010) approach. The RVE approach accounts for dependencies originating from studies that produce several estimates. These estimates can arise either from the same individuals or from clusters of studies that are not independent (e.g., a series of studies carried out in the same laboratory or by the same researcher). Therefore, the RVE approach does not require statistical independency. Second, it does not require knowledge about the covariance structure of the dependent estimates. Third, it does not make any assumptions about the underlying sample distributions of the effect sizes. Thus, the RVE approach is well-suited for the analysis of within-subjects designs such as those prevalent in the present set of studies.

We used the small-sample RVE approach as proposed by Fisher and Tipton (2015), as there were less than 40 studies included in each of the analyses. Following existing recommendations (Tanner-Smith & Tipton, 2014), we decided to set the weights of the effect sizes to weights that adequately account for the type of dependencies that are most prevalent in the data set (i.e., correlated effects in the present case). As an estimate for the within-study correlation of the effect sizes, we used the default setting of $\rho = .8$ (Fisher & Tipton, 2015). Additionally, we ran sensitivity analyses with $\rho = 0$, $\rho = .2$, $\rho = .4$, $\rho = .6$, and $\rho = 1$ for the five research questions. The results of the sensitivity analyses for the choice of the value of ρ can be found in Supplementary Table S3. Varying the magnitude of this correlation did not meaningfully change the results as reported below.

We report two measures for the heterogeneity of the effect sizes: τ and I^2 (Borenstein et al., 2009). First, τ is an estimate of the between-study standard deviation of the true effect sizes. It is an estimate of the true heterogeneity of effects that uses the same scale as the effect size index used in the analysis (i.e., milliseconds in the present meta-analysis; Fisher & Tipton, 2015). Thus, τ reflects the absolute amount of true effect size variation and says nothing about the proportion of the overall variation that is due to true effect size variation and not merely to sampling error. The I^2-statistic is a proportion that provides a different information. It is a descriptive statistic for the ratio of true heterogeneity to total variance in the observed effect sizes. Thus, the I^2-statistic reflects the estimated proportion of the variance that reflects true variance resulting from real effect size differences and not from sampling error (Higgins et al., 2003).

In addition to τ and I^2, we computed the 95% prediction intervals and the 95% confidence intervals for the estimated mean effect sizes. The 95% prediction intervals estimate the range in which 95% of the true effect sizes of randomly sampled studies (sampled from the universe of studies included in this meta-analysis) would fall (Borenstein et al., 2017; Riley et al., 2011). The prediction intervals are based on the between-study variance of the true effect sizes. The 95% confidence intervals indicate the precision of the estimates and are based on between-study variance of the true effect sizes and sampling error (Borenstein et al., 2009). In 95% of cases the mean effect size falls inside the confidence interval. Thus, these two intervals provide different information. Prediction intervals address the real dispersion of effect sizes while the confidence intervals quantify the accuracy of the means (Borenstein et al., 2009).

2.5. Moderator Analyses

In the case of continuous characteristics, we conducted moderation analyses by performing meta-regressions using the RVE approach (Hedges et al., 2010). In the case of categorical characteristics, we calculated the RVE intercept-only random-effects model for each subgroup. We then compared the resulting estimates via *Z*-tests (Borenstein et al., 2009). Here, we deviated from our preregistered approach as a maximum of two subgroups with at least $k = 5$ studies could ultimately be identified per categorical characteristic. Hence, the *Z*-test approach we followed here is straightforward and mathematically equivalent to our preregistered approach (Borenstein et al., 2009).

2.6. Publication bias

Publication bias occurs when studies that did not produce the desired outcomes are published less often than studies that did produce the desired outcomes (Fanelli, 2012; Franco et al., 2014). The result is a published body of literature that is not representative of the population of all completed studies. Statistically significant results are easier to publish than non-significant results. Publication bias emerges because authors are less likely to submit "failed" studies for publication, and if they do, reviewers and editors are less likely to endorse publication compared to studies that "worked" (e.g., Egger et al., 1997; Fanelli, 2010; Sterling, 1959; Sterling et al., 1995; Tramer et al., 1997). Studies that produced significant results are likely to have produced larger effect sizes, because for any given sample size, a larger effect is more likely to be significant than a smaller effect size. As a result, in a literature with publication bias, studies with larger effect sizes are over-represented. Strong publication bias can severely distort meta-analytic effect size estimates (Friese & Frankenbach, 2020).

2.6.1. Bias detection

To detect publication bias, we employed two different methods: Funnel plot and Egger's regression test.

2.6.1.1. Funnel plot

A funnel plot displays the relationship between studies' effect sizes and their precision. The effect size (here: the raw mean time difference) is plotted on the x-axis; study precision (i.e., the standard error of the study effect size) is plotted on the y-axis. Small standard errors are depicted at the higher end of the y-axis. The 1.96 standard error lines, indicating the 95% confidence intervals, are depicted as well. As a result, the funnel plot resembles a triangle or a funnel that has been turned upside down, centered on the estimated meta-analytic effect size (e.g., Light & Pillemer, 1984; Sterne & Egger, 2001). In the case of no publication bias, the effect sizes should scatter symmetrically around the estimated mean effect, with 95% of the effect sizes lying within the funnel (Borenstein et al., 2009). However, funnel plots tend to overestimate a possible bias when a random-effects model is used (i.e., when true heterogeneity of effect sizes is expected), as is the case in the present meta-analysis (Lau et al., 2006).

2.6.1.2. Egger's regression test

Egger's regression test examines whether there is a statistically significant relationship between a study's effect size and the study's precision (Egger et al., 1997; Sterne & Egger, 2005). To investigate this relationship, a weighted random-effects meta-regression is computed in which a study's effect size is regressed on the study's standard error. A significant slope indicates the presence of small-study effects and potential publication bias (Sterne & Egger, 2005).

This logic behind Egger's regression test can be extended to the RVE model used in the present meta-analysis by computing an RVE meta-regression with the dependent effect sizes at hand. Previous meta-analyses have already extended the logic of Egger's regression test to the RVE model (Coles et al., 2019; Friese et al., 2017). Nonetheless, caution is warranted in interpreting the results of this test as it suffers from low statistical power when the number of studies is small (Kromrey & Rendina-Gobioff, 2006) and the performance of the RVE extension of Egger's regression test has only been investigated in between-participant designs (Rodgers & Pustejovsky, 2020).

2.6.2. Bias correction

An extensive recent simulation study indicates that correcting for publication bias is a difficult task, since none of the current methods is consistently convincing (Carter et al., 2019). Therefore, we corrected for publication bias using two methods: Trim and Fill (Duval & Tweedie, 2000a, 2000b) and the Precision Effect Estimation with Standard Error (PEESE, Stanley & Doucouliagos, 2014). We then interpreted the results of these methods in light of recent simulation studies (Carter et al., 2019; Stanley & Doucouliagos, 2015; Stanley & Doucouliagos, 2016). Note that these simulation studies focused on two- (or more) group designs or regression coefficients of observational studies rather than studies using within-designs as in in the present meta-analysis. Hence, caution is warranted in interpreting the results of the publication bias analyses exclusively in light of the results of these simulation studies.

2.6.2.1. Trim and Fill

The Trim and Fill method (Duval & Tweedie, 2000a, 2000b) examines (a)symmetry of effect sizes in a funnel plot. The rationale is that asymmetry may be the result of studies that were in fact conducted, but never reported, leading to systematic omissions in the funnel plot. The Trim and Fill algorithm first removes extreme studies until the funnel plot is symmetric, yielding an adapted overall effect size. The adapted effect size estimate is interpreted as a more accurate (i.e., unbiased) true effect size estimate. The algorithm then imputes effect sizes of hypothetical studies (mirror images of the trimmed studies) to estimate the correct variance of the overall distribution of studies. We used the same sets of effect sizes as in the computation of the funnel plots to implement the Trim and Fill method.

The Trim and Fill method suffers from a few limitations. First, simulation studies show that Trim and Fill may adjust for publication bias when factually none exists. Conversely, it may adjust insufficiently in cases of strong publication bias (Moreno et al., 2009; Terrin et al., 2003). Second, Trim and Fill only slightly reduces bias and shows high false positive rates in the presence of publication bias combined with a small to zero true effect size (Carter et al., 2019).

2.6.2.2. Precision Effect Estimation with Standard Error (PEESE)

PEESE computes a meta-regression in which the squared standard error of the effect size is used as a predictor of the effect size. In the case of a significant relationship between these two parameters, publication bias is assumed to be present. The intercept of this regression is interpreted as the effect size of a (hypothetically “perfect”) study with a (squared) standard error of zero. PEESE uses this intercept as a bias-corrected estimate of the true effect size (Stanley & Doucouliagos, 2014).

PEESE was developed for traditional meta-analysis, but the logic can be extended to the RVE model (Coles et al., 2019; Friese et al., 2017). Nonetheless, caution is warranted in interpreting the results of this method because it has not yet been validated in the RVE context. Also, as a method based on linear regression, PEESE requires a large number of studies to yield adequate power and reliable results.

The methods used here assume that the effect sizes are statistically independent. Including dependent effect sizes (e.g., stemming from multiple ways to measure an outcome) violates this assumption (Lipsey & Wilson, 2001). Some studies contribute more than one effect size to the same research question. For the regression-based methods, we control for these dependencies in the same way as we did in the main analyses: by using the RVE approach (Hedges et al., 2010). In cases in which one study contributed more than one effect size to the same funnel plot or Trim and Fill analysis, we calculated the mean of these effect sizes and the associated standard errors. In this way, each set of effect sizes used for computation of the funnel plots or Trim and Fill analyses satisfies the independency assumption. Nonetheless, the results have to be interpreted with caution as they are based on effect size summaries and serve only as a first (visual) assessment of publication bias.

2.7. Weighted-least-squares meta-regression estimator

Regression-based methods to detect and correct for publication bias and small-study effects (e.g., Egger's Regression test and PEESE) suffer from low statistical power when the number of studies is small (Kromrey & Rendina-Gobioff, 2006; Stanley, 2008). Therefore, publication bias and small-study effects might be present even in cases where these methods indicate the contrary. It has been shown that the weighted-least-squares (WLS) meta-regression estimator yields less biased estimates than the random-effects estimator in the case of publication bias and small-study effects (Carter et al., 2019; Stanley & Doucouliagos, 2015). However, this has only been shown for two- (or more) group designs or regression coefficients of observational studies (Stanley & Doucouliagos, 2015; Stanley & Doucouliagos, 2016). Thus, even in the case of publication bias and small-study effects, the WLS estimator should not per se be seen as outperforming the random-effects estimator given the dependencies due to the within-designs in the present meta-analysis that the RVE model accounts for. We computed WLS meta-regression estimates for the five research questions by computing an ordinary least squares regression of the standardized mean time differences (i.e., *raw mean difference/SEM*) against the study precision (i.e., *1/SEM*) with no intercept (Stanley & Doucouliagos, 2015). Then, we compared the results of this computation to the RVE intercept-only random-effects model's mean effect sizes in light of the results of our publication bias analyses.

Analyses that rely on the RVE model (Hedges et al., 2010) suffer from inflated Type I error rates if the degrees of freedom fall below four (Fisher & Tipton, 2015; Tipton, 2015). Thus, the computed p-value underestimates the real p-value in these cases. As a result, the inferential statistics cannot be trusted if the usual α-level of $\alpha = .05$ is applied. To solve this issue, we used a higher standard of evidence for rejecting the null hypothesis by adjusting the α-level to $\alpha = .01$ in cases in which the degrees of freedom fell below four.

All analyses were computed using the R statistical programming environment (R Core Team, 2020) with the following packages: grid, irr (Gamer et al., 2019), lm.beta (Behrendt, 2014), metafor (Viechtbauer, 2010, 2016), readxl (Wickham & Bryan, 2019), robumeta (Fisher et al., 2017), and tidyverse (Wickham et al., 2019).

3. Results

3.1. Study identification

In the first forward searches of the Web of Science and PubMed databases, 1,093 and 200 articles were identified, respectively. Together with the Libet at al. articles from 1983 and 1985, this yielded a total of 1,114 articles after removal of duplicates. In the second forward search of these databases, 73 additional articles (53 after removal of duplicates) were identified. The forward search of the EBSCO database revealed 447 articles. In sum, these searches yielded 1,614 articles. Reference harvesting of the included articles delivered four additional articles. Reference harvesting of the qualitative review by Saigle and colleagues (2018) delivered another two articles.

The calls for unpublished data and the correspondence with authors in the context of this project did not reveal any unpublished data. However, we received one more publication through correspondence with another author. Hence, a total of 1,621 articles were screened. We excluded 1,483 articles after checking them against the eligibility criteria on the abstract level. This yielded 138 articles to be checked against the eligibility criteria on the full-text level. We excluded 101 articles after checking them against the eligibility criteria on the full-text level. This yielded a total 37 articles and $k = 43$ studies, $m = 150$ effect sizes with a total sample size of $N = 804$ included in the meta-analysis (see Figure 2). The average sample size was $M = 18.70$ ($Mdn = 14$, $SD = 15.18$). Publication dates ranged from 1983 to 2019 ($Mdn = 2012$). A list of the articles that were excluded after checking them against the eligibility criteria on the full-text level, including each article's reason for exclusion, is available in the associated OSF project.

3.2. Main Analyses

3.2.1. Time difference between the onset of unconscious brain activity and the conscious intention to move (RQ 1)

The following analysis is based on $m = 27$ effect sizes from $k = 6$ studies. The RVE intercept-only random-effects model's mean effect size for the time difference between the onset of unconscious brain activity and the conscious intention to move was $D = -479$ ms, CI_{95} [-614 ms, -344 ms], $t(4.19) = -9.67$, $p < .001$. Hence, the onset of unconscious brain activity preceded the conscious intention to move by on average 479 ms. The 95% prediction interval was PI_{95} [-842 ms, -116 ms]. The estimated between-study variance in the true effect sizes was τ = 121 ms. Around two-thirds of the variance in observed effect sizes was estimated to reflect true differences in effect sizes, I^2 = 67.11%. This indicates a moderate heterogeneity according to common conventions (Higgins et al., 2003). The forest plot for RQ 1 is shown in Figure 3.

3.2.2. Onset of unconscious brain activity relative to the actual onset of movement (RQ 2)

The following analysis is based on *m* = 21 effect sizes from *k* = 6 studies. The onset of unconscious brain activity preceded the actual onset of movement by on average 698 ms, CI_{95} [-987 ms, -409 ms], PI_{95} [-1474 ms, 79 ms], $t(4.82) = -6.29$, $p = .002$. The estimated between-study variance in the true effect sizes was τ = 257 ms. More than 90% of the variance in observed effect sizes was estimated to reflect true differences in effect sizes, I^2 = 91.37%. This value indicates a high heterogeneity, according to common conventions (Higgins et al., 2003). The forest plot for RQ 2 is shown in Supplementary Figure S1.

3.2.3. Intention to move relative to actual onset of movement (RQ 3)

The following analysis is based on *m* = 38 effect sizes from *k* = 33 studies. The conscious intention to move was reported on average 122 ms before the actual onset of movement, CI_{95} [-156 ms, -89 ms], PI_{95} [-257 ms, 12 ms], $t(29.18) = -7.50$, $p < .001$. The estimated between-study variance in the true effect sizes was τ = 64 ms. More than 90% of the variance in observed effect sizes was estimated to reflect true differences in effect sizes, I^2 = 90.97%. This value indicates a high heterogeneity (Higgins et al., 2003). The forest plot for RQ 3 is shown in Supplementary Figure S2.

3.2.4. Awareness of the onset of movement relative to the actual onset of movement (RQ 4)

The following analysis is based on *m* = 38 effect sizes from *k* = 33 studies. Subjective awareness of the onset of movement was reported on average 13 ms after the actual onset of movement, CI_{95} [1 ms, 24 ms], PI_{95} [-31ms, 56ms], $t(23.85) = 2.30$, $p = .030$. The estimated between-study variance in the true effect sizes was τ = 21 ms. Around 80% of the variance in observed effect sizes was estimated to reflect true differences in effect sizes, I^2 = 78.09%. This value indicates a high heterogeneity in the set of effect sizes (Higgins et al., 2003). The forest plot for RQ 4 is shown in Supplementary Figure S3.

3.2.5. Time difference between the conscious intention to move and subjective awareness of the onset of movement (RQ 5)

The following analysis is based on m = 26 effect sizes from k = 23 studies. The time of the conscious intention to move preceded the time of subjective awareness of the onset of movement by on average 134 ms, CI_{95} [-156 ms, -111 ms], PI_{95} [-214 ms, -53 ms], $t(17.49)$ = -12.42, $p < .001$. The estimated between-study variance in the true effect sizes was τ = 37 ms. More than two-thirds of the variance in observed effect sizes was estimated to reflect true differences in effect sizes, I^2 = 70.41%. This value indicates a moderate heterogeneity (Higgins et al., 2003). The forest plot for RQ 5 is shown in Supplementary Figure S4. The results of the main analyses are illustrated in Figure 4.

3.3. Moderator Analyses

Due to the small amount of data available, only a handful of the initially planned moderator analyses could be computed. These moderator analyses could only be computed for the time difference between the conscious intention to move and the actual onset of movement (RQ 3), the time difference between subjective awareness of the onset of movement and the actual movement (RQ 4) and the time difference between the conscious intention to move and subjective awareness of movement (RQ 5). For RQ 1 and RQ 2, there was not enough data available to conduct moderator analyses. The results of the moderator analyses for the categorical characteristics are depicted in Table 1. Supplementary Table S4 provides an overview of how the examined categorical moderators are distributed over the various times of interest.

3.3.1. Study characteristics

3.3.1.1. Instructions for reporting awareness

Descriptively, participants who were instructed to report when they realized they had moved their finger reported earlier subjective awareness of the onset of movement (D = 9 ms, CI_{95} [-17 ms, 36 ms], k = 14, m = 17) than participants who were instructed to report the time of their button or key press (D = 13 ms, CI_{95} [-0 ms, 27 ms], k = 17, m = 19). The difference was not significant (Z = -0.28, p = .781).

Participants who were instructed to report when they realized that they had moved their finger reported a significantly larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who were instructed to report the time of their button or key press (D = -143 ms, 95% CI [-204 ms, -83 ms], k = 10, m = 13, D = -129 ms, CI_{95} [-152 ms, -107 ms], k = 12, m = 12, respectively, Z = -2.37, p = .018).

3.3.1.2. Instructions for reporting intention

Descriptively, participants who were instructed to report their *urge* to move reported earlier times for the conscious intention to move (D = -160 ms, CI_{95} [-250 ms, -70 ms], k = 8, m = 10) than participants who were instructed to report their *intention* to move (D = -98 ms, CI_{95} [-146 ms, -50 ms], k = 17, m = 17). This difference was not significant (Z = -1.41, p = .158).

Participants who were instructed to report their *urge* to move reported a significantly larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who were instructed to report their *intention* to move (D = -153 ms, CI_{95} [-237 ms, -69 ms], k = 5, m = 7, D = -124 ms, CI_{95} [-159 ms, -90 ms], k = 12, m = 12, respectively, Z = -4.61, p < .001).

3.3.1.3. Method of reporting time

Participants who used either a mouse or a keyboard for their time reports reported significantly earlier times for the conscious intention to move (*D* = -131 ms, CI_{95} [-175 ms, -87 ms], *k* = 18, *m* = 22) than participants who reported the times verbally (*D* = -56 ms, CI_{95} [-130 ms, 17 ms], *k* = 6, *m* = 6; *Z* = -2.12, *p* = .034). Descriptively, participants who used either a mouse or a keyboard for their time reports reported earlier subjective awareness of the onset of movement (*D* = 17 ms, CI_{95} [4 ms, 29 ms], *k* = 19, *m* = 21) than participants who reported the times verbally (*D* = 28 ms, CI_{95} [-41 ms, 96 ms], *k* = 5, *m* = 7). This difference was not significant (*Z* = -0.50, *p* = .618).

Descriptively, participants who used either a mouse or a keyboard for their time reports reported a larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who reported the times verbally (*D* = -119 ms, CI_{95} [-139 ms, -98 ms], *k* = 13, *m* = 15, *D* = -108 ms, CI_{95} [-243 ms, 26 ms], *k* = 3, *m* = 3, respectively). This difference was not significant (*Z* = -1.66, *p* =.097).

3.3.1.4. Monitoring instrument

Descriptively, participants who were monitored either with an EEG or an fMRI reported earlier times for the conscious intention to move (*D* =-154 ms, CI_{95} [-201 ms, -107 ms], *k* = 18, *m* = 22) and earlier times for subjective awareness of the onset of movement (*D* = 11 ms, CI_{95} [-2 ms, 23 ms], *k* = 14, *m* =16) than participants who were not monitored (*D* = -95 ms, CI_{95} [-143 ms, -47 ms], *k* = 15, *m* = 16; *D* = 13 ms, CI_{95} [-5 ms, 31 ms], *k* = 19, *m* = 22, respectively). None of the differences was significant (*Z* = -1.90, *p* = .057; *Z* = -0.24, *p* = .809, respectively).

Participants who were monitored with either an EEG or an fMRI reported a significantly larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who were not monitored (*D* = -155 ms,

CI_{95} [-194 ms, -115 ms], k = 11, m = 13, D = -120 ms, CI_{95} [-151 ms, -90 ms], k = 12, m = 13, respectively, Z = -6.26, p < .001).

3.3.1.5. Type of action

Participants performing a button or key press reported descriptively earlier times for the conscious intention to move (D = -144 ms, CI_{95} [-193 ms, -94 ms], k = 19, m = 20) compared to those performing a mouse click (D = -99 ms, CI_{95} [-138 ms, -59 ms], k = 11, m = 14). Furthermore, mean reported times for subjective awareness of the onset of movement were descriptively earlier for button or key presses (D = 10 ms, CI_{95} [-4 ms, 24 ms], k = 22, m = 23) than for mouse clicks (D = 24 ms, CI_{95} [1 ms, 47 ms], k = 9, m = 10). Both differences were not significant (Z = -1.56, p = .118; Z = -1.28, p = .199, respectively).

Participants who performed a button or key press reported a significantly larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who performed a mouse click (D = -147 ms, CI_{95} [-181 ms, -114 ms], k = 14, m = 15, D = -108 ms, CI_{95} [-125 ms, -90 ms], k = 8, m = 9, respectively, Z = -8.50, p < .001).

3.3.1.6. Number of trials

The number of trials moderated the reported times for the conscious intention to move (b = -2.05, t(10.9) = -2.24, p = .047) but neither the times reported for subjective awareness of the onset of movement (b = -0.17, t(7.69) = -0.56, p = .591) nor the time difference between the conscious intention to move and subjective awareness of the onset of movement (b = -0.37, t(6.86) = -0.69, p = .513).

3.4. Publication bias

3.4.1. Funnel plots and Trim and Fill analyses

We computed five funnel plots (one for each of the research questions) to display the relationship between study precision and effect size. The funnel plots after the Trim and Fill analyses are depicted in Figure 5. Only the funnel plot and Trim and Fill analysis for the time difference between the onset of unconscious brain activity and the conscious intention to move (RQ 1) yielded evidence for small-study effects and potential publication bias. Visual inspection revealed that one effect size fell outside the interval in which 95% of studies are expected for any given level of precision. The effect sizes were asymmetrically distributed around the summary effect. The Trim and Fill method (see below) indicated that three effect sizes on the bottom right side (two around -300 ms and one just above zero) were missing. The Trim and Fill analysis yielded a bias-corrected effect size estimate of D = -422 ms, CI_{95} [-490 ms, -354 ms], z = -12.17, p < .001. This estimate is less negative than the uncorrected effect size (D = -479 ms). Neither the visual inspection of the funnel plots nor the Trim and Fill analyses for the remaining research questions (RQ 2 – RQ 5) yielded evidence for the presence of small-study effects or publication bias.

3.4.2. Egger’s regression test

The results of Egger’s regression test for the five research questions are shown in Table 2. Egger’s regression tests yielded no evidence for the presence of small-study effects or publication bias except for the conscious intention to move relative to the actual onset of movement (RQ 3, k = 33, m = 38), b_{se} = -1.60, SE = 0.62, t(12.6) = -2.57, p = .024.

3.4.3. PEESE

PEESE was computed in the RVE model extension for the five research questions. The results are shown in Table 2. None of the analyses yielded evidence for the presence of small-study effects or publication bias.

3.5. WLS

As preregistered, we also computed the WLS meta-regression estimates for the five research questions. It has been shown that the WLS estimator outperforms random-effects estimators in case of publication bias in combination with a zero or small-true effect size

(Carter et al., 2019). This is not the case in the present meta-analysis. Most of our publication bias analyses yielded no evidence for small-study effects or publication bias. Furthermore, the main analyses yielded significant, robust mean effect sizes. Additionally, the RVE approach that we used in our random-effects models is well suited to account for the dependencies of effect sizes that stem from the same primary study. Thus, we report the WLS meta-regression estimates in Table 2 as preregistered, but do not interpret them further. Instead, we continue to rely on the effect sizes that the RVE model yielded.

4. Discussion

In Libet et al.'s (1983) pioneering experiment, participants freely decided when to start an instructed movement, and reported when they felt the intention to move and when they became aware of their movement. An EMG recorded when the movement actually started. This experiment has had a tremendous impact on scientific and lay discussions about conscious causation and the existence of free will, particularly because Libet and colleagues found unconscious brain activity well before participants reported their conscious intention to move. The present meta-analysis provides the first quantitative review investigating the robustness and moderators of the temporal pattern found by Libet and colleagues. Using the robust variance estimation approach — which effectively handles dependencies among effect sizes — we found a temporal pattern that is largely consistent with the one found by Libet and colleagues: The onset of unconscious brain activity preceded the conscious intention to move by around 479 ms (CI_{95} [-614 ms, -344 ms]), which in turn preceded subjective awareness of the onset of movement by around 134 ms (CI_{95} [-156 ms, -111 ms]). Other than in the original experiment by Libet et al. (1983), the subjective awareness of the onset of movement was reported on average 13 ms after (not before) the actual onset of movement (CI_{95} [1 ms, 24 ms]).

In addition to the identified temporal pattern, another major contribution of the present meta-analysis was to elucidate the amount of evidence that has been gathered since the publication of the seminal original experiment almost 40 years ago. Somewhat surprisingly, the evidence base is remarkably thin. The five research questions considered together, we were able to locate 43 studies, delivering 150 effect sizes, based on a total sample size of N = 804. The number of studies addressing each research question ranged from k = 6 to k = 33, the number of effect sizes from m = 21 to m = 38, and the total sample size from N = 53 to N = 641.

The amount of available data was especially small for the question of whether the onset of unconscious brain activity precedes the conscious intention to move, and if so, by how long. It is this time difference in particular that is crucial for many skeptical positions on free will (e.g., Bargh, 2008; Wegner, 2002). For this difference, we could procure data from k = 6 studies delivering m = 27 effect sizes based on only N = 53 participants (!). This meager data base stands in sharp contrast to the huge influence the Libet experiment has had on both scientific thinking in various disciplines and public discourse. It is worth noting that we identified 6 more studies that investigated brain activity in combination with participants' intention or awareness reports but did not report the onset time of the brain activity. We asked the authors for the onset times or the raw data but did not receive either.

The scarcity of evidence concerning the onset of brain activity was also associated with a large degree of uncertainty around the estimate of 698 ms before the onset of actual movement. This was attested to by the wide 95% confidence and prediction intervals (see Figure 4) and the considerable heterogeneity, particularly when compared to the estimates for intention and awareness. Despite this uncertainty about the mean time delay, the meta-analytic results clearly suggest that – based on the available evidence – the onset of unconscious brain activity precedes the conscious intention to move. The most intensively debated finding of Libet et al.'s (1983) study was thus meta-analytically confirmed.

One may wonder why the uncertainty around the estimate of the onset of brain activity was so much larger compared to the estimates for intention and awareness. Unfortunately, the available data was too scarce to conduct formal moderator hypotheses, but one reason for the large variability around the mean estimate for the onset of brain activity was that various methodological factors contributing to the exact time estimation have varied across the six studies (e.g., use of the readiness potential or lateralized readiness potential, different ways of determining the exact onset, Dominik et al. 2018; Libet et al., 1983). It will be up to future research to investigate how these methodological factors in determining the onset of brain activity affect the time estimates.

The meta-analytic temporal pattern was largely in line with the pattern in the original Libet experiment, with one noteworthy deviation. While Libet and colleagues found the reported awareness of the movement to precede the actual movement by 85 ms, in our meta-analysis the reported awareness trailed the actual movement onset by on average 13 ms. Note that the meta-analytic result fits intuition better than the Libet finding: One cannot be subjectively aware of an event before it actually happens. From this perspective, one may even wonder why the difference is not more pronounced than 13 ms on average. One reason that may contribute to the deviation from the Libet experiment is that Libet et al.'s findings were based on only five participants compared to $N = 579$ participants from $k = 33$ studies in the meta-analysis. The Libet findings are therefore necessarily less precise and more subject to random fluctuation. The meta-analysis likely estimates any true effect more precisely than a single, small study. Concerning the question why the difference is not even larger, one reviewer suggested that possibly the conscious mind projects slightly into the future, mentally simulating the present moment slightly ahead of reality. Although the present data do not directly speak to this possibility, it would be a worthwhile endeavor for future research to examine this intriguing possibility.

Although we coded a host of theoretically derived and methodological potential moderating variables, we could only conduct a handful of the initially planned analyses due to the small amount of data available. In total, we found six moderator effects. For example, participants who were instructed to report when they realized that they had moved their finger reported a significantly larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who were instructed to report the time of their button or key press. Similarly, participants who were instructed to report the time of the *urge* to move reported a significantly larger time difference between the conscious intention to move and subjective awareness of the onset of movement than participants who were instructed to report the time of the *intention* to move. These two moderator effects indicate that the reported time difference between the conscious intention to move and subjective awareness of the onset of movement is sensitive to the exact wording of the instructions. This result expands on the previous finding that instructions to report the start of the movement yielded earlier report times than instructions to report the end of the movement (Pockett & Miller, 2007). Of note, all six significant moderator effects were of a rather technical nature and do not question the overall Libetian pattern and its interpretation. Potentially more substantial moderators for the debate around the Libetian pattern could not be analyzed due to the scarcity of the data available (e.g., the order of the intention and awareness reports, Dominik et al., 2018).

The results of the Libet experiment (Libet et al., 1983) have been interpreted as questioning the causal role of conscious mental states in behavior (e.g., Roediger et al., 2008; Pockett, 2006, cf. Mele, 2009) and the existence of free will (e.g., Bargh, 2008; Wegner, 2002). What can the present meta-analysis contribute to these debates? The present work did not address methodological or philosophical questions about the suitability of the Libet paradigm to speak to these issues. Rather, it asked how confident we can be about any claims that are based on the empirical findings originating from this paradigm. Researchers may then consider how these findings square with their respective theoretical positions. In terms of the amount of available data, the results are rather sobering. After almost 40 years, the evidence base of the most crucial of Libet and colleagues' findings appears thinner than anticipated in light of the substantial scientific work that built on it. Basing far-reaching theoretical positions on such a small amount of data seems premature. This conclusion is not meant to question the trustworthiness of the research summarized in the present meta-analysis. Rather, it points out the surprisingly large uncertainties and calls for prudence in deducing positions and theories that rely on the robustness and generalizability of the pattern.

4.1. How to move forward?

How can future research move forward based on what we now know about the evidence in the Libet tradition? First, the available evidence is riddled with problems that impair robustness. It is scarce, heterogeneous, built on tiny studies, and contains only a single full replication of the Libet experiment (Dominik et al., 2018). The field urgently needs high-powered, preregistered direct and conceptual replications and variations of the Libet experiment to provide more robust data on the temporal pattern. Special attention should be given to potential moderators that might be crucial for the validity of Libet-style experiments (e.g., the order of intention and awareness reports; Dominik et al., 2017; Sanford et al., 2020).

Second, besides assuring the robustness of findings in the context of the original paradigm, future research should further develop alternative approaches suited to address similar questions as the Libet experiment that address shortcomings of the original operationalization in Libet-style experiments (see Wolpe & Rowe, 2014, for an overview of approaches that were partially inspired by the Libet experiment). One example of these approaches is the intentional binding paradigm (Haggard et al., 2002), which promises to shed light on human agency (Moore & Obhi, 2012). As a relative measure, intentional binding successfully addresses limitations of the original Libet paradigm, such as subjective biases regarding the time reports.

4.2. The role of the preceding brain activity

When interpreting the data from experiments in the Libet tradition, the brain activity that precedes self-initiated actions assumes a pivotal role. However, the properties and meaning of this brain activity are increasingly scrutinized. A deeper understanding of the role of this brain activity will be important for future work drawing on the Libet paradigm.

The classical view interprets the readiness potential as a specific causal precursor of voluntary actions (e.g., Libet et al., 1982; Travers et al., 2020). It is this interpretation that many influential interpretations of the Libetian pattern presuppose as being true (e.g., Libet, 1999; Wegner, 2002). However, a growing number of findings challenges this interpretation. First, some research suggests that the movement initiation in Libet-style experiments depends on the crossing of a threshold due to stochastic fluctuations of neural signals rather than a specific causal precursor (e.g., Schurger et al., 2012; for a discussion see Brass et al., 2019; for a review see Schurger et al., in press).

Second, several studies question both, the sufficiency and the necessity of the readiness potential for voluntary movements. Concerning sufficiency, in one study, readiness potentials also occurred in the absence of subsequent movements (Alexander et al., 2016). Two studies speak to the issue of necessity. One study replicated the observation that the readiness potential preceded arbitrary, purposeless decisions of no further importance to participants (e.g., when to move a finger). It did not, however, precede actions that arose from deliberate, purposeful decisions that participants cared about (Maoz et al., 2019). Hence, this finding questions the generalizability of the temporal pattern typically found in Libet-style experiments to deliberate, meaningful decisions. The second study used a Libet-type setup and found that only around two thirds of the voluntary button presses were preceded by a slow negative potential shift (i.e., a readiness potential, Jo et al., 2013). The other third, in contrast, was preceded by a slow positive potential shift. Averaging the EEG signal across many trials yields the readiness potential. According to Schmidt and colleagues (2016), this pattern of results is best accounted for by assuming that there are continuous fluctuations of slow cortical potentials and that the initiation of voluntary actions more likely occurs during negative phases of these fluctuations, but sometimes also during positive phases. Together, the work discussed in this paragraph suggests that the readiness potential may neither be sufficient nor necessary for voluntary actions.

4.3. Strengths and Limitations

The present meta-analysis features several strengths, but also limitations. Starting with the strengths, we note, first, that the present work was preregistered, followed best-practice recommendations for the reproducibility of meta-analyses (Lakens et al., 2016) and strove to meet highest standards of openness and transparency. For example, we closely followed our preregistered procedures and reported and justified all deviations. To facilitate future updates of this work, we made all data, code, a list of all coded characteristics, the coding sheet, supplementary analyses, and further materials available in the associated OSF project.

Second, using the robust variance estimation approach, we employed state-of-the-art meta-analytic techniques that allow for the analyses of dependent effect sizes that are ubiquitous in research in the Libet tradition. Third, we thoroughly addressed the prevalent problem of publication bias using several methods to detect and potentially correct for small-study effects and publication bias.

On the side of limitations, we again stress that the available evidence for each of our five research questions was slim, particularly for the most crucial questions involving brain activity. Although there is nothing meta-analysts can do about slim evidence bases, less evidence obviously allows for less confident conclusions. Far-reaching conclusions based on the available evidence seem premature.

Second, while the performance of the publication bias methods used for the data from two- (or more) group designs has been the object of systematic investigation (e.g., Carter et al., 2019; Stanley & Doucouliagos, 2015; Stanley et al., 2017), we are not aware of systematic simulation results concerning the performance of these methods in the case of dependent data that are prevalent in the present meta-analysis. Therefore, the respective results should be interpreted with caution.

5. Conclusion

The present random-effects meta-analysis yielded a temporal pattern that is largely consistent with the one found by Libet and colleagues (Libet et al., 1983). Surprisingly, the evidence base is remarkably thin. This is especially true for the crucial time difference between the onset of unconscious brain activity and the conscious intention to move ($k = 6$ studies, $m = 27$ effect sizes, $N = 53$). Additionally, there is a high degree of uncertainty associated with the onset of this brain activity. Thus, even after almost 40 years, some of Libet et al.'s findings appear more fragile than anticipated in light of the substantial scientific work that built on them.

Acknowledgements

We would like to thank Roy Baumeister and an anonymous reviewer for their valuable comments that helped to improve this article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

Conceptualization: M.N.B., J.W., and M.F.; Data Curation: M.N.B.; Design: M.N.B., J.W., and M.F.; Formal Analysis: M.N.B.; Investigation: M.N.B.; Project Administration: M.N.B. and M.F.; Supervision: J.W. and M.F.; Visualization: M.N.B.; Writing - original draft: M.N.B.; Writing - review & editing: M.N.B., J.W., and M.F.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Alexander, P., Schlegel, A., Sinnott-Armstrong, W., Roskies, A. L., Wheatley, T., & Tse, P. U. (2016). Readiness potentials driven by non-motoric processes. *Consciousness and Cognition*, *39*, 38–47. https://doi.org/10.1016/j.concog.2015.11.011

*Baek, K., Doñamayor, N., Morris, L. S., Strelchuk, D., Mitchell, S., Mikheenko, Y., Yeoh, S. Y., Phillips, W., Zandi, M., Jenaway, A., Walsh, C., & Voon, V. (2017). Impaired awareness of motor intention in functional neurological disorder: Implications for voluntary and functional movement. *Psychological Medicine*, *47*(9), 1624–1636. https://doi.org/10.1017/S0033291717000071

Bargh, J. (2008). Free will is un-natural. In J. Baer, J. Kaufmann, & R. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 128–154). Oxford University Press.

Baumeister, R. F. (2005). *The cultural animal: Human nature, meaning, and social life*. Oxford University Press.

Baumeister, R. F. (2008). Free will in scientific psychology. *Perspectives on Psychological Science*, *3*(1), 14–19. https://doi.org/10.1111/j.1745-6916.2008.00057.x

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., & Vohs, K. D. (2011). Do conscious thoughts cause behavior? *Annual Review of Psychology*, *62*(1), 331–361. https://doi.org/10.1146/annurev.psych.093008.131126

Beckermann, A. (2008). *Gehirn, Ich, Freiheit. Neurowissenschaften und Menschenbild.* [Brain, self, freedom. Neurosciences and anthropology.]. mentis.

Behrendt, S. (2014). *lm.beta: Add standardized regression coefficients to lm-objects* (Version 1.5-1) [Software package]. https://CRAN.R-project.org/package=lm.beta

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Wiley.

Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I² is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Brass, M., Furstenberg, A., & Mele, A. R. (2019). Why neuroscience does not disprove free will. *Neuroscience & Biobehavioral Reviews*, *102*, 251–263. https://doi.org/10.1016/j.neubiorev.2019.04.024

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *Neuroscientist*, *14*(4), 319–325. https://doi.org/10.1177/1073858408317417

Braun, M. N. (2021a, February 22). Schematic illustration of the results of the Libet experiment. Retrieved from osf.io/2f57j

Braun, M. N. (2021b, June 03). Meta-Analysis of Libet-style experiments main results. Retrieved from osf.io/jnbt5

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*(2), 115–144. https://doi.org/10.1177/2515245919847196

Caruso, G. D. (Ed.). (2013). *Exploring the illusion of free will and moral responsibility*. Lexington Books.

*Caspar, E. A., & Cleeremans, A. (2015) Free will: Are we all equal? A dynamical perspective of the conscious intention to move. *Neuroscience of Consciousness*, *2015*(1), 1–10. https://doi.org/10.1093/nc/niv009

Chomsky, N. (2006). *ali g interviews noam chomsky* [Interview]. http://www.youtube.com/watch?v=fOIM1_xOSro

Christensen, J. F., Di Costa, S., Beck, B., & Haggard, P. (2019). I just lost it! Fear and anger reduce the sense of agency: A study using intentional binding. *Experimental Brain Research*, *237*(5), 1205–1212. https://doi.org/10.1007/s00221-018-5461-6

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220. https://doi.org/10.1037/h0026256

Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin*, *145*(6), 610–651. https://doi.org/10.1037/bul0000194

Deecke, L., Grözinger, B., & Kornhuber, H. H. (1976). Voluntary finger movement in man: Cerebral potentials and theory. *Biological Cybernetics*, *23*(2), 99–119. https://doi.org/10.1007BF00336013

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

*Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., & Procházka, R. (2017). Libet's experiment: Questioning the validity of measuring the urge to move. *Consciousness and Cognition*, *49*, 255–263. https://doi.org/10.1016/j.concog.2017.01.017

*Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., & Procházka, R. (2018). Libet's experiment: A complex replication. *Consciousness and Cognition*, *65*, 1–26. https://doi.org/10.1016/j.concog.2018.07.004

*Dominik, T., Dostál, D., Zielina, M., Šmahaj, J., Sedláčková, Z., & Procházka, R. (2018). EEG data and introspective reports from the Libet's experiment replication. *Data in Brief*, *20*, 2040–2044. https://doi.org/10.1016/j.dib.2018.09.066

*Doñamayor, N., Strelchuk, D., Baek, K., Banca, P., & Voon, V. (2018). The involuntary nature of binge drinking: Goal directedness and awareness of intention. *Addiction Biology*, *23*(1), 515–526. https://doi.org/10.1111/adb.12505

*Douglas, Z. H., Maniscalco, B., Hallett, M., Wassermann, E. M., & He, B. J. (2015). Modulating conscious movement intention by noninvasive brain stimulation and the underlying neural mechanisms. *Journal of Neuroscience*, 35(18), 7239–7255. https://doi.org/10.1523/JNEUROSCI.4894-14.2015

Duval, S., & Tweedie, R. (2000a). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal*, *315*(629), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLoS ONE*, *5*(4), Article e10068. http://doi.org/10.1371/journal.pone.0010068

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, *90*(3), 891–904. https://doi.org/10.1007/s11192-011-0494-7

Fischborn, M. (2016). Libet-style experiments, neuroscience, and libertarian free will. *Philosophical Psychology*, *29*(4), 1–9. https://doi.org/10.1080/09515089.2016.1141399

Fisher, Z., & Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. *ArXiv:1503.02220 [Stat]*. http://arxiv.org/abs/1503.02220

Fisher, Z., Tipton, E., & Hou, Z. (2017). *robumeta: Robust Variance Meta-Regression* (Version 2.0) [Software package]. https://cran.r-project.org/package=robumeta

Focquaert, F., Glenn, A. L., & Raine, A. (2013). Free will, responsibility, and the punishment of criminals. In T. A. Nadelhoffer (Ed.), *The future of punishment* (pp. 247–274). Oxford University Press.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Friese, M., & Frankenbach, J. (2020). P-hacking and publication bias interact to distort meta-analytic effect size estimates. *Psychological Methods*, *25*(4), 456–471. https://doi.org/10.1037/met0000246

Friese, M., Frankenbach, J., Job, V., & Loschelder, D. (2017). Does self-control training improve self-control? A meta-analysis. *Perspectives on Psychological Science*, *12*(6), 1077–1099. https://doi.org/10.1177/1745691617697076

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). *irr: Various coefficients of interrater reliability and agreement* (Version 0.84.1) [Software package]. https://CRAN.R-project.org/package=irr

*Ganos, C., Asmuss, L., Bongert, J., Brandt, V., Münchau, A., & Haggard, P. (2015). Volitional action as perceptual detection: Predictors of conscious intention in adolescents with tic disorders. *Cortex*, *64*, 47–54. https://doi.org/10.1016/j.cortex.2014.09.016

Gholipour, B. (2019). A famous argument against free will has been debunked. *The Atlantic.* https://www.theatlantic.com/health/archive/2019/09/free-will-bereitschaftspotential/597736/

Gilden, L., Vaughan, H. G. Jr., & Costa, L. D. (1966). Summated human EEG potentials with voluntary movement. *Electroencephalography and Clinical Neurophysiology*, *20*(5), 433–438. https://doi.org/10.1016/0013-4694(66)90100-3

*Giovannelli, F., Mastrolorenzo, B., Rossi, A., Gavazzi, G., Righi, S., Zaccara, G., Viggiano, M. P., & Cincotta, M. (2016). Relationship between impulsivity traits and awareness of motor intention. *European Journal of Neuroscience*, *44*(7), 2455–2459. https://doi.org/10.1111/ejn.13359

Gomes, G. (1998). The timing of conscious experience: A critical review and reinterpretation of Libet's research. *Consciousness and Cognition*, *7*(4), 559–595. https://doi.org/10.1006/ccog.1998.0332

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience*, *9*(12), 934–946. https://doi.org/10.1038/nrn2497

*Haggard, P., Cartledge, P., Dafydd, M., & Oakley, D. A. (2004). Anomalous control: When 'free will' is not conscious. *Consciousness and Cognition*, *13*(3), 646–654. https://doi.org/10.1016/j.concog.2004.06.001

*Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385. https://doi.org/10.1038/nn827

*Haggard, P., & Cole, J. (2007). Intention, attention and the temporal experience of action. *Consciousness and Cognition*, *16*(2), 211–220. https://doi.org/10.1016/j.concog.2006.07.002

*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

*Haggard, P., Martin, F., Taylor-Clarke, M., Jeannerod, M., & Franck, N. (2003). Awareness of action in schizophrenia. *NeuroReport*, *14*(7), 1081–1085. https://doi.org/10.1097/01.wnr.0000073684.00308.c0

*Haggard, P., Newman, C., & Magno, E. (1999). On the perceived time of voluntary actions. *British Journal of Psychology*, *90*(2), 291–303. https://doi.org/10.1348/000712699161413

Haynes, J.-D. (2011). Decoding and predicting intentions. *Annals of the New York Academy of Sciences*, *1224*(1), 9–21. https://doi.org/10.1111/j.1749-6632.2011.05994

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. https://doi.org/10.1002/jrsm.5

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hirose, S., Nambu, I., & Naito, E. (2018). Cortical activation associated with motor preparation can be used to predict the freely chosen effector of an upcoming movement and reflects response time: An fMRI decoding study. *NeuroImage*, *183*, 584–596. https://doi.org/10.1016/j.neuroimage.2018.08.060

*Jo, H.-G., Habel, U., & Schmidt, S. (2019). Role of the supplementary motor area in auditory sensory attenuation. *Brain Structure and Function*, *224*(7), 2577–2586. https://doi.org/10.1007/s00429-019-01920-x

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

*Jo, H.-G., Hinterberger, T., Wittmann, M., & Schmidt, S. (2015). Do meditators have higher awareness of their intentions to act? *Cortex*, *65*, 149–158. https://doi.org/10.1016/j.cortex.2014.12.015

*Jo, H.-G., Wittmann, M., Borghardt, T. L., Hinterberger, T., & Schmidt, S. (2014). First person approaches in neuroscience of consciousness: Brain dynamics correlate with the intention to act. *Consciousness and Cognition*, *26*, 105–116. https://doi.org/10.1016/j.concog.2014.03.004

*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

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. [Brain potential changes in voluntary and passive movements in humans: Readiness potential and reafferent potentials]. *Pflügers Archiv für Gesamte Physiologie*, *284*, 1–17. https://doi.org/10.1007/BF00412364

Krieghoff, V., Waszak, F., Prinz, W., & Brass, M. (2011). Neural and behavioral correlates of intentional actions. *Neuropsychologia*, *49*(5), 767–776. https://doi.org/10.1016/j.neuropsychologia.2011.01.025

Kromrey, J. D., & Rendina-Gobioff, G. (2006). On knowing what we do not know: An empirical comparison of methods to detect publication bias in meta-analysis.

Educational and Psychological Measurement, *66*(3), 357–373. https://doi.org/10.1177/0013164405278585

*Lafargue, G., & Duffau, H. (2008). Awareness of intending to act following parietal cortex resection. *Neuropsychologia*, *46*(11), 2662–2667. https://doi.org/10.1016/j.neuropsychologia.2008.04.019

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Lakens, D. D., Hilgard, J., & Staaks, J. J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BioMed Central Psychology, 4*, Article 24. https://doi.org/10.1186/s40359-016-0126-3

Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). Evidence based medicine: The case of the misleading funnel plot. *British Medical Journal*, *333*, 597–600. https://doi.org/10.1136/bmj.333.7568.597

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

*Lau, H. C., Rogers, R. D., & Passingham, R. E. (2006). On measuring the perceived onsets of spontaneous actions. *Journal of Neuroscience*, *26*(27), 7265–7271. https://doi.org/10.1523/JNEUROSCI.1138-06.2006

*Lau, H. C., Rogers, R. D., & Passingham, R. E. (2007). Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience*, *19*(1), 81–90. https://doi.org/10.1162/jocn.2007.19.1.81

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*(4), 529–539. 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., 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. https://doi.org/10.1016/0013-4694(82)90181-x

*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): The unconscious initiation of a freely voluntary act. *Brain*, *106*(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research.* Harvard University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Sage.

List, C. (2019). *Why free will is real.* Harvard University Press.

*Lush, P., Naish, P., & Dienes, Z. (2016). Metacognition of intentions in mindfulness and hypnosis. *Neuroscience of Consciousness*, *2016*(1), 1–10. https://doi.org/10.1093/nc/niw007

Maoz, U., Yaffe, G., Koch, C., & Mudrik, L. (2019). Neural precursors of decisions that matter – an ERP study of deliberate and arbitrary choice. *eLife*, *8*, Article e39787. https://doi.org/10.7554/eLife.39787.001

Mele, A. R. (2009). *Effective intentions: The power of conscious will*. Oxford University Press.

Mele, A. R. (2014). *Free: Why science hasn't disproved free will.* Oxford University Press.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097

*Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, *17*(1), 136–144. https://doi.org/10.1016/j.concog.2006.12.004

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*(1), 546–561. https://doi.org/10.1016/j.concog.2011.12.002

Moreno, S. G., Sutton, A. J., Ades, A. E., Stanley, T. D., Abrams, K. R., Peters, J. L., et al. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BioMed Central Medical Research Methodology*, *9*, Article 2. https://doi.org/10.1186/1471-2288-9-2

*Moretto, G., Schwingenschuh, P., Katschnig, P., Bhatia, K. P., & Haggard, P. (2011). Delayed experience of volition in Gillles de la Tourette syndrome. *Journal of Neurology, Neurosurgery and Psychiatry*, *82*(12), 1324–1327. https://doi.org/10.1136/jnnp.2010.221143

Nahmias, E. (2005). Agency, authorship, and illusion. *Consciousness and Cognition, 14*(4), 771–785. https://doi.org/10.1016/j.concog.2005.07.002

Nahmias, E. (2015). Why we have free will. *Scientific American*. https://www.scientificamerican.com/article/why-we-have-free-will/

Overbye, D. (2007). Free will: Now you have it, now you don't. *The New York Times*. https://www.nytimes.com/2007/01/02/science/02free.html

*Parés-Pujolràs, E., Kim Y.-W., Im, C.-H., & Haggard, P. (2019). Latent awareness: Early conscious access to motor preparation processes is linked to the readiness potential. *Neuroimage*, *202*, 116–140. https://doi.org/10.1016/j.neuroimage.2019.116140

*Penton, T., Thierry, G. L., & Davis, N. J. (2014). Individual differences in attributional style but not in interoceptive sensitivity, predict subjective estimates of action intention.

Frontiers in Human Neuroscience, *8*, Article 638. https://doi.org/10.3389/fnhum.2014.00638

Pereboom, D. (2001). *Living without free will*. Cambridge University Press.

Pockett, S. (2006). The neuroscience of movement. In S. Pockett, W. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior?* (pp. 9–24). MIT Press.

*Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition*, *16*(2), 241–254. https://doi.org/3389/fnhum.2014.0063810.1016/j.concog.2006.09.002

R Core Team (2020). *R: A language and environment for statistical computing* [Software]. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

*Rigoni, D., Brass, M., Roger, C., Vidal, F., & Sartori, G. (2013). Top-down modulation of brain activity underlying intentional action and its relationship with awareness of intention: An ERP/Laplacian analysis. *Experimental Brain Research*, *229*(3), 347–357. https://doi.org/3389/fnhum.2014.0063810.1007/s00221-013-3400-0

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*, Article d549. https://doi.org/10.1136/bmj.d549

Rodgers, M. A., & Pustejovsky, J. E. (2020). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*. Advance online publication. https://doi.org/10.31222/osf.io/vqp8u

Roediger, H. L., Goode, M. K., & Zaromb, F. M. (2008). Free will and the control of action. In J. Baer, J. Kaufmann, & R. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 205–225). Oxford University Press.

*Rossi, A., Giovannelli, F., Gavazzi, G., Righi, S., Cincotta, M., & Viggiano, M. P. (2018). Electrophysiological activity prior to self-initiated movements is related to impulsive personality traits. *Neuroscience*, *372*, 266–272. https://doi.org/10.1016/j.neuroscience.2018.01.011

Saigle, V., Dubljević, V., & Racine, E. (2018). The impact of a landmark neuroscience study on free will: A qualitative analysis of articles using Libet and colleagues' methods. *AJOB Neuroscience*, *9*(1), 29–41. https://doi.org/10.1080/21507740.2018.1425756

Sanford, P., Lawson, A. L., King, A. N., & Major, M. (2020). Libet's intention reports are invalid: A replication of Dominik et al. (2017). *Consciousness and Cognition*, *77*, 1–16. https://doi.org/10.1016/j.concog.2019.102836

Schlosser, M. E. (2012). Free Will and the Unconscious Precursors of Choice. *Philosophical Psychology*, *25*(3), 365–384. https://doi.org/10.1080/09515089.2011.622366

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

Schurger, A., Hu, P. „Ben", Pak, J., & Roskies, A. L. (in press). What is the readiness potential? *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2021.04.001

*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*, *109*(42), E2904–E2913. https://doi.org/10.1073/pnas.1210467109

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037//0033-2909.86.2.420

*Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., & Haggard. P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neuroscience*, *7*(1), 80–84. https://doi.org/10.1038/nn1160

Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*(5), 543–545. https://doi.org/10.1038/nn.2112

Stafford, T. (2015, August 7). Why do we intuitively believe that we have free will? *BBC Future*. https://www.bbc.com/future/article/20150806-why-your-intuitions-about-the-brain-are-wrong

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effect in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, *70*(1), 103– 127. https://doi.org/10.1111/j.1468-0084.2007.00487.x

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

Stanley, T.D., & Doucouliagos, H. (2015). Neither fixed nor random: Weighted least squares meta-analysis. *Statistics in Medicine*, *34*(13), 2116–2127. https://doi.org/10.1002/sim.6481

Stanley, T.D., & Doucouliagos, H. (2016). Neither fixed nor random: Weighted least squares meta-regression. *Research Synthesis Methods*, *8*(1), 19–42. https://doi.org/10.1002/jrsm.1211

Stanley, T.D., Doucouliagos, H., & Ioannidis, J. P. A. (2017). Finding the power to reduce publication bias. *Statistics in Medicine*, *36*(10), 1580–1598. https://doi.org/10.1002/sim.7228

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance— or vice versa. *Journal of the American Statistical Association*, *54*(28), 30–34. https://doi.org/10.2307/2282137

Sterling, T. D., Rosenbaum, W. L., & Weinkam, J. J. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice-versa. *The American Statistician*, *49*(1), 108–112. https://doi.org/10.1080/00031305.1995.10476125

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*, *54*(10), 1046–1055. https://doi.org/10.1016/S0895-4356(01)00377-8

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 99-110). Wiley.

Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies*, *75*, 5–24. https://doi.org/10.1007/BF00989879

Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional binding without intentional action. *Psychological Science*, *30*(6), 842–853. https://doi.org/10.1177/0956797619842191

*Tabu, H., Aso, T., Matsuhashi, M., Ueki, Y., Takahashi, R., Fukuyama, H., Shibasaki, H., & Mima. T. (2015). Parkinson's disease patients showed delayed awareness of motor intention. *Neuroscience Research*, *95*, 74–77. https://doi.org/10.1016/j.neures.2015.01.012

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*(1), 13–30. https://doi.org/10.1002/jrsm.1091

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, *22*(13), 2113–2126. https://doi.org/10.1002/sim.1461

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, *20*(3), 375–393. https://doi.org/10.1037/met0000011

Tramer, M. R., Reynolds, D. J., Moore, R. A., & McQuay, H. J. (1997). Impact of covert duplicate publication on meta-analysis: A case study. *British Medical Journal*, *315*, 635–640. https://doi.org/10.1136/bmj.315.7109.635

Travers, E., Khalighinejad, N., Schurger, A., & Haggard, P. (2020). Do readiness potentials happen all the time? *NeuroImage*, *206*, 116286. https://doi.org/10.1016/j.neuroimage.2019.116286

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

*Tsakiris, M., & Haggard, P. (2003). Awareness of somatic events associated with a voluntary action. *Experimental Brain Research*, *149*(4), 439–446. https://doi.org/10.1007/s00221-003-1386-8

van Inwagen, P. (2008). How to think about the problem of free will. *The Journal of Ethics*, *12*(3–4), 327–341. https://doi.org/10.1007/s10892-008-9038-7

Vargas, M. (2013). *Building better beings: A theory of moral responsibility*. Oxford University Press.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W. (2016). *metafor: Meta-analysis package for R* (Version 1.9-9) [Software package]. https://cran.r-project.org/package=metafor

*Vinding, M. C., Jensen, M., & Overgaard, M. (2014). Distinct electrophysiological potentials for intention in action and prior intention for action. *Cortex*, *50*, 86–99. https://doi.org/10.1016/j.cortex.2013.09.001

*Walsh, E., Kuehn, S., Brass, M., Wenke, D., & Haggard, P. (2010). EEG activations during intentional inhibition of voluntary action: An electrophysiological correlate of self-control?. *Neuropsychologia*, *48*(2), 619–626. https://doi.org/10.1016/j.neuropsychologia,2009.10.026

Walter, H. (2011). Contributions of Neuroscience to the Free Will Debate: From random movement to intelligible action. In R. Kane (Ed.), *Oxford Handbook of Free Will* (515–529). Oxford University Press.

Wegner, D. (2002). *The illusion of conscious will*. MIT Press.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Wickham, H. & Bryan, J. (2019). *readxl: Read Excel files* (Version 1.3.1) [Software package]. https://CRAN.R-project.org/package=readxl

Wolpe, N., & Rowe, J. B. (2014). Beyond the "urge to move": Objective measures for the study of agency in the post-Libet era. *Frontiers in Human Neuroscience*, *8*, Article 450. https://doi.org/10.3389/fnhum.2014.00450

Table 1

Results of the Moderator Analyses for the Categorical Characteristics

Research Question	Moderator		Mean Effect and Inference Statistics								Test of Moderation			
			D (ms)	LL (ms)	UL (ms)	*t*	*df*	P_t	*k*	*m*	D_{Diff} *(ms)*	SE_{Diff} *(ms)*	*Z*	p_Z
Intention (RQ 3)	Instructions													
		Urge	-160	-250	-70	-4.25	6.62	.004	8	10				
											-62	44	-1.41	.158
		Intention	-98	-146	-50	-4.36	15.20	< .001	17	17				
	Method of Time Reporting													
		Mouse/Keyboard	-131	-175	-87	-6.32	14.61	< .001	18	22				
											-75	35	-2.12	.034
		Verbally	-56	-130	17	-1.98	4.95	.105	6	6				
	Monitoring Instrument													
		EEG/fMRI	-154	-201	-107	-6.97	15.53	< .001	18	22				
											-59	31	-1.90	.057
		None	-95	-143	-47	-4.26	13.43	< .001	15	16				
	Type of Action													
		Button/Key press	-144	-193	-94	-6.16	16.56	< .001	19	20				
											-45	29	-1.56	.118
		Mouse click	-99	-138	-59	-5.80	7.69	< .001	11	14				
Awareness (RQ 4)	Instructions													
		Movement	9	-17	36	0.81	8.59	.441	14	17				
											-4	14	-0.28	.781
		Button/Key press	13	-0	27	2.15	12.61	.052	17	19				
	Method of Time Reporting													
		Mouse/Keyboard	17	4	29	2.91	10.95	.014	19	21				
											-11	22	-0.50	.618
		Verbally	28	*-41*	*96*	*1.29*	*2.97*	*.288*	5	7				

	Monitoring Instrument													
		EEG/fMRI	11	-2	23	1.96	6.36	.095	14	16				
											-2	10	-0.24	.809
		None	13	-5	31	1.56	15.60	.138	19	22				
	Type of Action													
		Button/Key press	10	-4	24.1	1.47	17.42	.161	22	23				
											-14	11	-1.28	.199
		Mouse click	24	1	47	2.86	4.28	.043	9	10				
Awareness minus Intention (RQ 5)	Instructions													
		Movement	-143	-204	-83	-5.58	7.25	< .001	10	13				
											-14	6	-2.37	.018
		Button/Key press	-129	-152	-107	-12.96	9.18	< .001	12	12				
	Instructions													
		Urge	-153	*-237*	*-69*	*-6.75*	*2.39*	*.013*	5	7				
											-28	6	-4.61	< .001
		Intention	-124	-159	-90	-8.08	9.70	< .001	12	12				
	Method of Time Reporting													
		Mouse/Keyboard	-119	-139	-98	-14.30	5.88	< .001	13	15				
											-10	6	-1.66	.097
		Verbally	-108	*-243*	*26*	*-3.48*	*1.99*	*.074*	3	3				
	Monitoring Instrument													
		EEG/fMRI	-155	-194	-115	-9.57	6.12	< .001	11	13				
											-34	5	-6.26	< .001
		None	-120	-151	-90	-8.92	9.98	< .001	12	13				
	Type of Action													
		Button/Key press	-147	-181	-114	-9.63	11.36	< .001	14	15				
											-40	5	-8.50	< .001
		Mouse click	-108	-125	-90	-16.99	4.21	< .001	8	9				

Note. The results of the moderation analyses for the categorical characteristics are shown. *D* (ms) = effect size in milliseconds. LL (ms) = lower limit of the 95% confidence interval (CI) in milliseconds. UL (ms) = upper limit of the 95% CI in milliseconds. t = t-value associated with the effect size in the same row testing whether the effect size differs significantly from zero. *df* = associated small-sample-adjusted degrees of freedom. p_t = p-value associated with t and *df* in the same row. k = number of studies that contributed to the respective moderator level. m = number of effect sizes on the respective moderator level.

Test of Moderation: D_{Diff} *(ms)* = mean difference between the effect sizes of the two moderator levels in milliseconds. SE_{Diff} *(ms)* = standard error of the mean difference in milliseconds. Z = Z-value associated with the effect sizes of the two moderator levels in the same row, testing whether these effect sizes significantly differ. p_z = p-value associated with Z in the same row. Note that when degrees of freedom fall below four (italics), the computed p-value tends to underestimate the real p-value (Fisher & Tipton, 2015; Tipton, 2015). Hence, in these cases, the α-level should be adjusted to $\alpha = .01$ to avoid drawing inaccurate conclusions.

Table 2

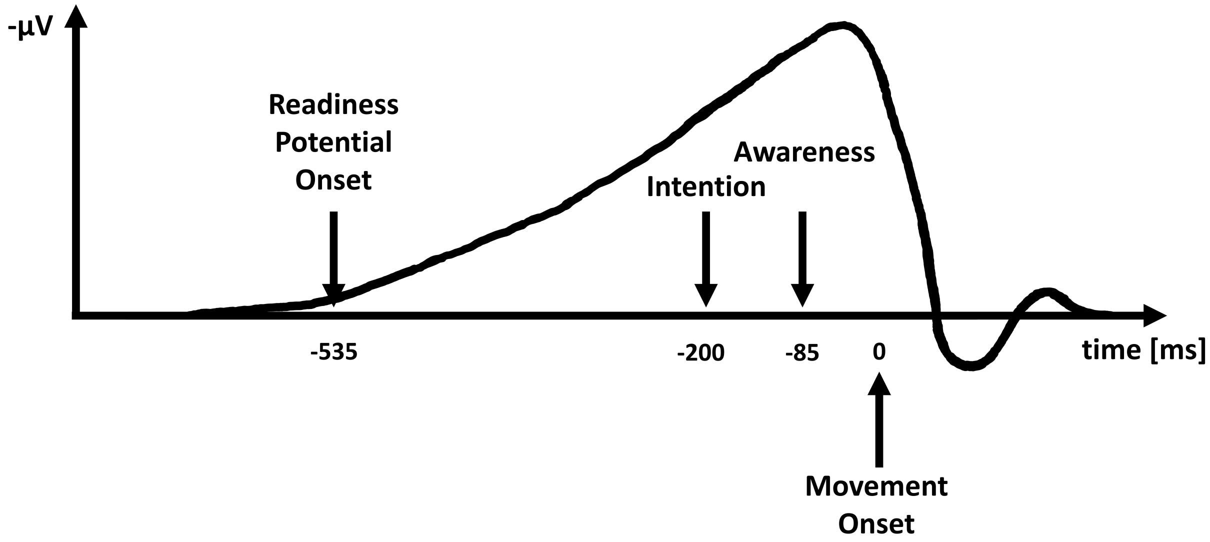
Results of the Publication Bias Analyses

Research Question			Egger's Regression Test					PEESE					WLS				
	k	*m*	b_{SE}	*SE*	*t*	*df*	*p*	b_1	*SE*	*t*	*df*	*p*	D_{WLS} *(ms)*	*SE (ms)*	*t*	*df*	*p*
Onset Brain Activity minus Awareness (RQ1)	6	27	-1.37	1.21	-1.14	2.24	.362	-0.01	0.01	-0.98	1.65	.447	-434	44	-9.95	26	< .001
Onset Brain Activity (RQ 2)	6	21	-2.54	1.44	-1.77	2.00	.220	-0.01	0.01	-1.20	1.48	.388	-415	59	-7.06	20	< .001
Intention (RQ 3)	33	38	-1.60	0.62	-2.57	12.6	.024	-0.01	0.01	-1.67	5.83	.147	-89	12	-7.40	37	< .001
Awareness (RQ 4)	33	38	-0.36	0.37	-0.97	13.7	.350	-0.00	0.00	-0.47	6.90	.656	18	4	4.87	37	< .001
Intention minus Awareness (RQ 5)	23	26	-0.83	0.48	-1.73	8.64	.118	-0.00	0.00	-1.32	5.45	.240	-122	9	-14.35	25	< .001

Note. *k* = number of studies. *m* = number of effect sizes. b_{SE} = slope of the RVE-model meta-regression of effect sizes on their standard errors. b_1 = slope of the RVE-model meta-regression of effect sizes on their squared standard errors. D_{WLS} *(ms)* = effect size estimated by the WLS meta-regression in milliseconds. *SE* = standard error associated with the respective effect size. *t* = *t*-value associated with the respective effect size. *df* = degrees of freedom. *p* = *p*-value associated with respective *t* and *df*.

Figure 1

Illustration of the Temporal Pattern found in the Libet Experiment (Libet et al., 1983)



Note. The onset of the readiness potential precedes the conscious intention to move (Intention) by around 335 ms, which in turn precedes subjective awareness of the movement (Awareness) by around 115 ms. Furthermore, subjective awareness of the movement precedes the actual onset of movement by around 85 ms. The amplitude of the readiness potential (black line) is depicted as -µV on the y-axis (Braun, 2021a).

Figure 2

PRISMA Flowchart of the Literature Search and Study Coding

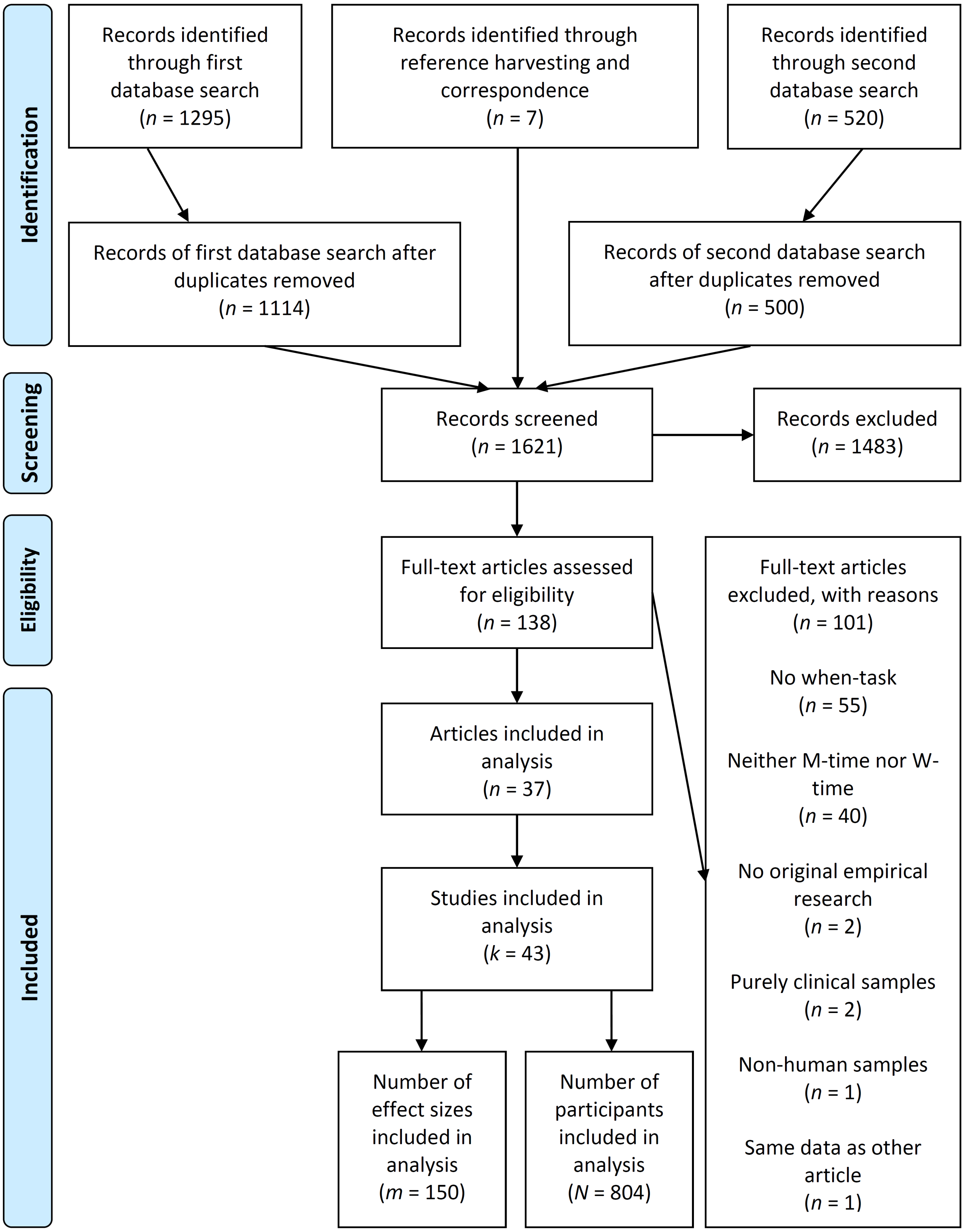
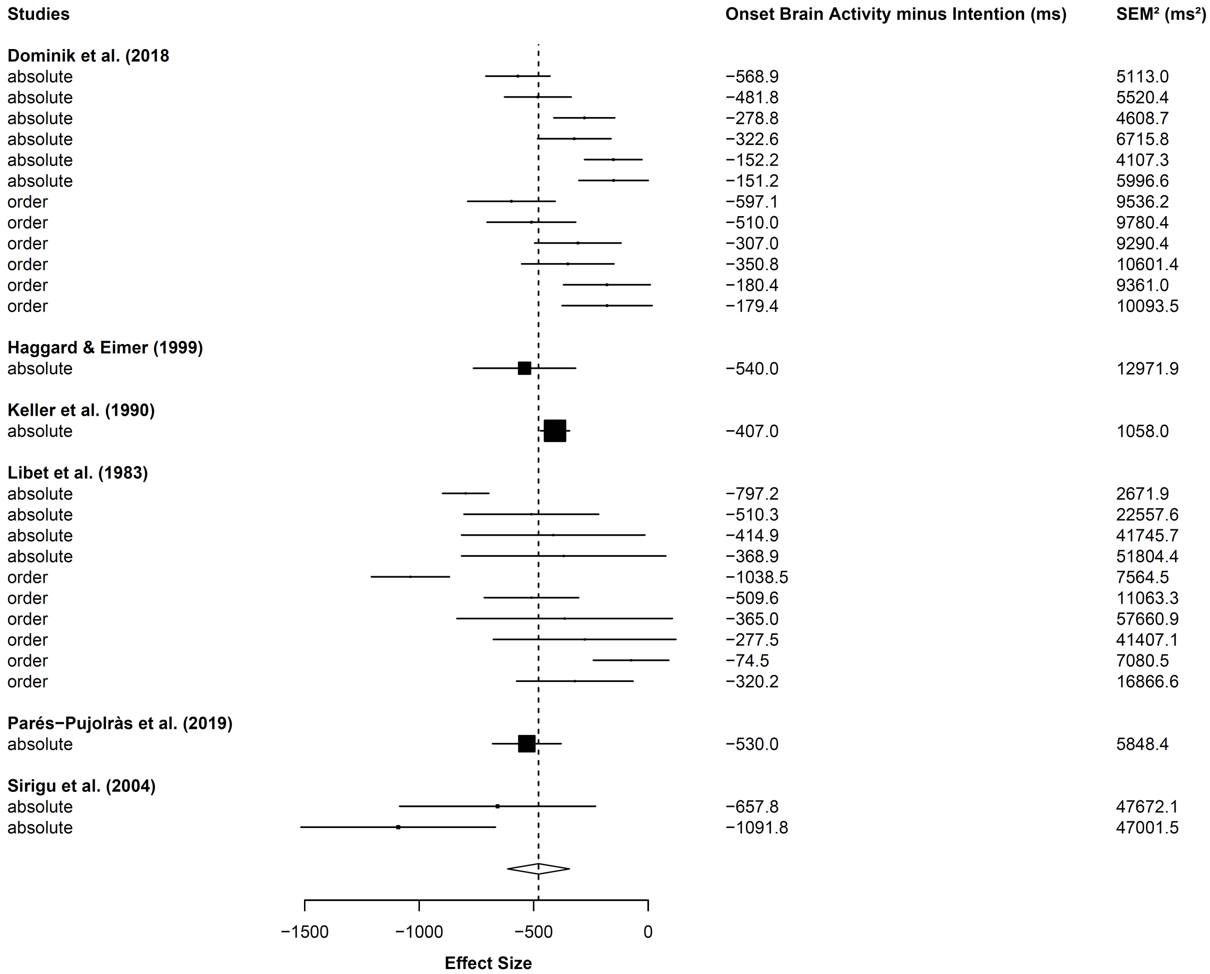


Figure 3

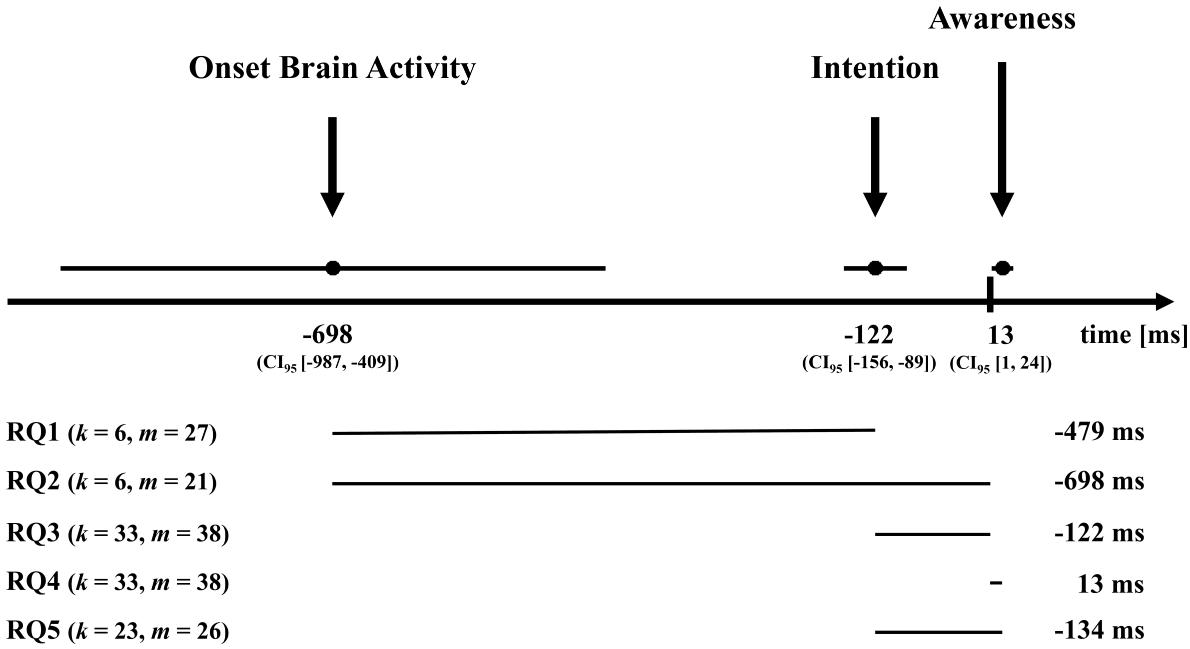
Forest Plot for the Time Difference between the Onset of Unconscious Brain Activity and the Conscious Intention to Move (RQ 1)



Note. The black boxes depict the mean effect size (in milliseconds) associated with the study in the same row. The size of the box corresponds to the weight assigned to the respective effect size: The higher the weight, the bigger the box. The horizontal lines depict the corresponding 95% confidence intervals (CI). The white diamond depicts the 95% CI around the overall effect size (D = -479 ms) yielded by the RVE intercept-only random-effects model computed in the main analysis. SEM² (ms²) = squared standard error mean of the mean effect size associated with the study in the same row in squared milliseconds. absolute = absolute mode of recall. order = order mode of recall.

Figure 4

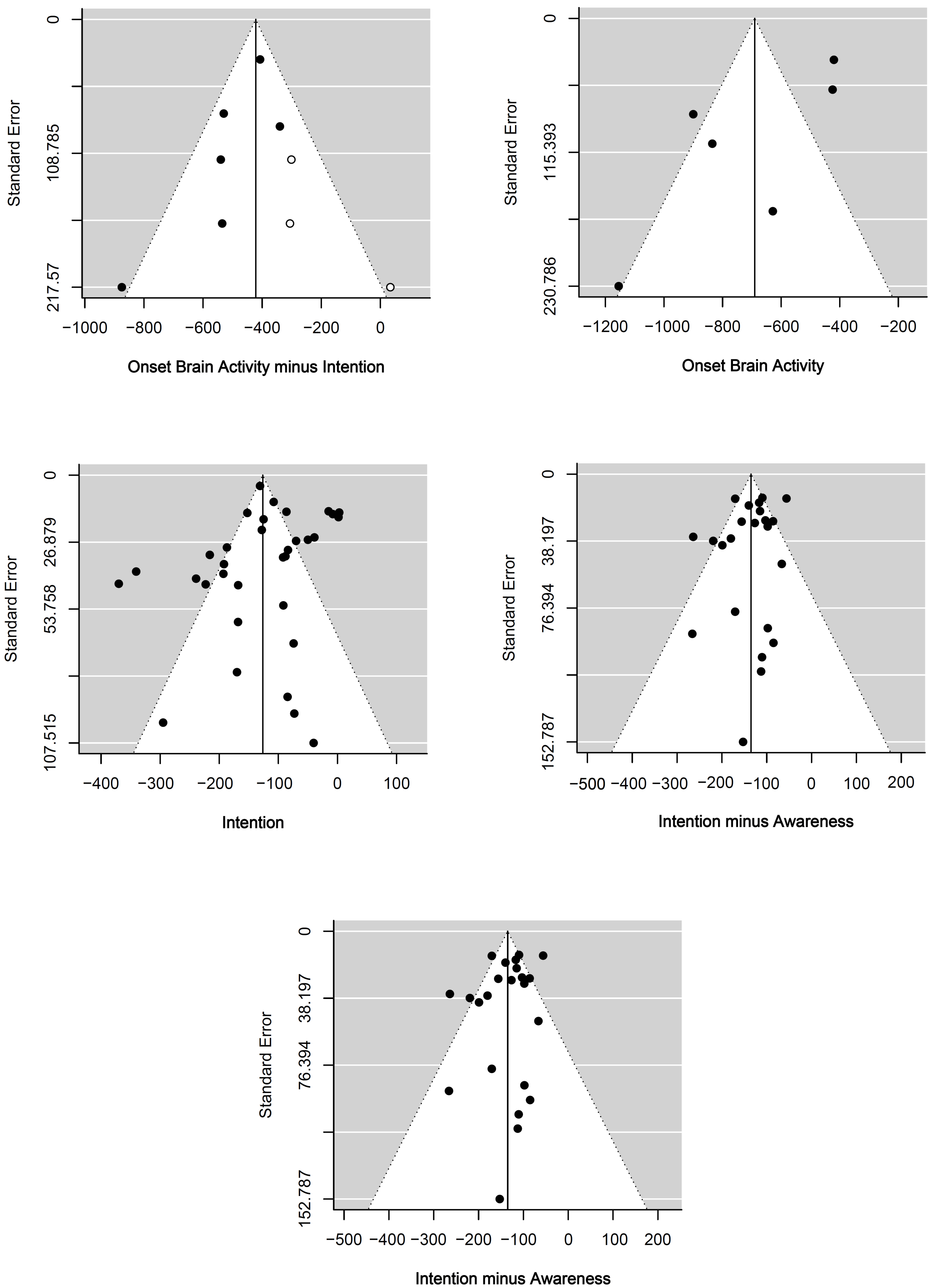
Temporal Pattern yielded by the Main Analyses of the respective Times relative to the actual Onset of Movement



Note. The black dots depict the effect size while the horizontal bars above the x-axis depict the 95% confidence intervals. The onset of unconscious brain activity (Onset Brain Activity) precedes the conscious intention to move (Intention), which in turn precedes subjective awareness of the onset of movement (Awareness). Subjective awareness of the onset of movement trails the actual onset of movement (vertical line at 0 ms). This pattern is largely in line with the temporal pattern found in the Libet experiment (see Figure 1). The horizontal bars below the x-axis visualize the time differences addressed by the five research questions and the values on the right denote the respective mean time differences estimated in the main analyses (Braun, 2021b).

Figure 5

Funnel Plots for the five Research Questions after Trim and Fill Analyses



Note. Effect sizes were only asymmetrically distributed in the funnel plot for the time difference between the onset of unconscious brain activity and the conscious intention to move (RQ 1). The respective Trim and Fill analysis indicated three missing effect sizes (white-filled dots). Effect sizes were relatively symmetrically distributed in the funnel plots for the remaining research questions. The respective Trim and Fill analyses did not indicate missing effect sizes.