POMPEU FABRA UNIVERSITY

MASTER THESIS

The Point of No Return in Action Cancellation: Deciphering Its Influence on the Human Sense of Agency via Real-Time Brain-Computer Interfaces

Author:
Hamed GHANE

Signature:

Supervisor:
Prof. Salvador SOTO FARACO

Signature:

A thesis submitted in fulfillment of the requirements
for the degree of Master of Science

in

Brain and Cognition

Center for Brain and Cognition
Department of Information Technology and Communications.

July 12, 2023

"When I raise my arm, my arm goes up. And the problem arises:
What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?"

— Ludwig Wittgenstein, Proposition 614, The Philosophical Investigations

"Above the sea's great sweep,
view beacons beyond the leap,
to find the path, you have to keep.

Atop the sky's grand steep,
look at the far light in the deep,
guidance from yonder we keep. "

— By the help of OpenAI's GPT-4 (2023)

Statement of Contribution

This section clarifies the contributions of each individual involved in this thesis, according to the CRediT (Contributor Roles Taxonomy) guidelines.

- **Hamed Ghane**: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing - Original Draft Preparation, Writing - Review & Editing, Visualization, Project Administration.
- **Salvador Soto Faraco**: Conceptualization, Resources, Writing - Review & Editing, Supervision, Project Administration, Funding Acquisition.
- **Xavi Mayoral**: Resources (BCI setup and experimental equipment provision), Technical Support.
- **OpenAI's GPT-4**: Software (Code Debugging), Writing - Review & Editing.

POMPEU FABRA UNIVERSITY

Abstract

Center for Brain and Cognition

Department of Information Technology and Communications.

Master of Science

The Point of No Return in Action Cancellation: Deciphering Its Influence on the Human Sense of Agency via Real-Time Brain-Computer Interfaces

by Hamed GHANE

Human volition, composed of intention and agency as subjective experiences, is often considered one of the most distinguishing characteristics of our species. Despite extensive debate around intention due to its implications on free will, the concept of agency has not received equivalent attention. The seminal work of Schultze-Kraft et al., 2016 advanced our understanding of voluntary actions by quantifying the 'point of no return' in movement initiation. This crucial temporal threshold is reached approximately 200 ms before an action, beyond which it becomes almost impossible to cancel the imminent action. Our research expands on this concept by investigating how the 'point of no return' influences our post-action sense of agency. By utilizing a Brain-Computer Interface and machine learning-based techniques for imminent action detection, we propose a novel hypothesis: our sense of agency regards an imminent action as 'initiated' once the 'point of no return' is crossed, which in turn shapes our perception of agency following an action. Our experimental design consists of two stages—a preparatory stage for data collection and classifier training, and a main real-time experiment designed to test the hypothesis in real-time. This setup enables us to study how the timing of an action's outcome relative to the 'point of no return' impacts the perceived sense of agency. This thesis, however, only covers the first stage of the experimental design, namely the data collection and classifier training. Our findings could potentially open up new avenues of research into the intricacies of human agency and volition, and may have significant implications for designing human-machine interface technologies aimed at enhancing user's perceived control and interaction.

Acknowledgements

The first words of gratitude are for my son, **Roham**, a little man with a big heart. His patience and love throughout this lengthy process have been an immeasurable source of strength and motivation. Thank you, my sweetheart.

To my parents, without their unwavering and unconditional support and faith in me, none of this would have been possible at all.

I am sincerely grateful to my supervisor, Salvador Soto Faraco. I appreciate not only his guidance throughout this process but also his valuable suggestion to explore this captivating topic, which set the course of this enriching journey.

Special thanks to Xavi Mayoral, who was always there, at the CBC lab, to lend a hand when I faced challenges and couldn't make any progress.

My gratitude also goes out to Irene Vigué-Guix for her time in familiarizing me with the BCI setup used in her experiment. Furthermore, I am appreciative of the people from the Multisensory Research Group, particularly Alice Drew and Angela Marti for their generous help and assistance.

I would also like to appriciate those classmates from the CBC2023 class who participated in this research. Their willingness to participate and contribute has been truly valuable to this study.

Lastly, a special acknowledgment for ChatGPT-4. By extending the definition of friend, it stands as the most patient, knowledgeable, and helpful friend I have ever had. Its assistance during this journey has been invaluable.

Contents

List of Figures

List of Tables

List of Abbreviations

SoA	**S**ense **o**f **A**gency
BCI	**B**rain **C**omputer **I**nterface
RP	**R**eadiness **P**otential
LRP	**L**ateralized **R**eadiness **P**otential

To the duty that accompanies the privilege of Intellect…

1. Introduction

1.1 Voluntary Action and Human Agency: Bridging Philosophy, Neuroscience, and Human-Brain Interface

Throughout history, the nature of human action and our abilities as agents have been central themes in philosophy. The concept of agency is intrinsically linked to the idea of human freedom and the extent to which individuals can exert control over their actions. This fundamental inquiry, dating back to the early days of philosophical thought, has driven investigations into human consciousness and self-awareness (Leibniz, 1991).

Leibniz's "Theodicy" exemplifies this line of inquiry, posing the problem of necessity and questioning whether humans are truly free agents or merely subject to the determinism of a pre-established order (Leibniz, 1985). This understanding of agency presupposes individuals' ability to identify themselves as agents capable of initiating and controlling actions. Such self-awareness is fundamental to human nature, permitting purposeful behavior and decision-making that reflects individuals' intentions, desires, and beliefs.

Transitioning from traditional philosophical debates into the realm of contemporary neuroscience research, as depicted in Figure 1.1 , Haggard, 2017 delineates voluntary action as entailing two distinct subjective experiences: intention and the sense of agency (SoA). The concept of intention primarily refers to the cognitive processes underlying the planning and initiation of actions. It enters into the vast and often contested discussion about human freedom in decision-making and is closely tied to the concept of 'free will', a term we consciously avoid here to prevent getting entangled in complex philosophical debates. While intention is typically probed via introspective methods, it has been empirically investigated through various experimental designs as well.

Contrarily, SoA deals with the subjective experience of controlling one's actions and perceiving oneself as the author of those actions. Unlike intention, the SoA is often studied retrospectively, examining how individuals perceive their actions after their occurrence. This distinction helps us better channel our research efforts to explore the unique aspects of each concept and ensure that our investigations remain grounded within their respective theoretical frameworks.

Focusing on SoA, our research aims to delve deeper into the subjective experience associated with the execution of a voluntary action. The exploration of SoA holds promise in illuminating the human condition, our freedom, our responsibilities, and our control over actions. Probing the neural and cognitive mechanisms underpinning SoA, we aspire to demystify the processes enabling us to experience ourselves as agents. In doing so, we inch closer to solving one of the "sphinxes of science," as termed by Leibniz, 1985.

FIGURE 1.1: Subjective components of a voluntary action.

The SoA also takes center stage in the realms of ethics and law, particularly in the context of emerging technologies like brain-computer interfaces (BCIs). Deciphering responsibility for actions performed through these devices raises urgent questions. Seminal works by Cornelio et al., 2022 and Wen and Imamizu, 2022 elucidate SoA's role in human-machine interactions and underscore its importance in devising user-centric and ethically sound technologies. These investigations collectively highlight SoA's potential to bolster human performance while grappling with the ethical complexities spurred by technological advancement.

1.2 Investigation Parameters: Defining SoA and Actions

This section elaborates on the specific definitions and parameters that guide our investigation of the SoA and voluntary action. We will discuss the distinct interpretation of SoA adopted for this thesis and provide a definition for action, focusing on the particular type of voluntary action under study.

1.2.1 Narrowing the Concept of SoA

In line with Haggard's perspective (Haggard, 2017), our investigation of SoA is more specifically oriented. Whereas some definitions interpret SoA as a broad subjective feeling of capability to act (or self-efficacy), we take a more restricted approach. We center our exploration on the experiential component of SoA, which emerges before, during, and after an actual muscular movement. Therefore, throughout this thesis, SoA refers to the subjective experience related to the execution of a particular motor action.

1.2.2 Defining Action and Its Types

Clarity necessitates a precise definition of action and the identification of the specific type of action our thesis focuses on.

Actions form the core of our everyday experiences and are fundamental for achieving our goals. It's argued that our brains have evolved primarily to meet the needs of action, rather than solely focusing on cognitive functions (Llinas, 2002). Human actions can generally be divided into two categories: stimulus-dependent reactions

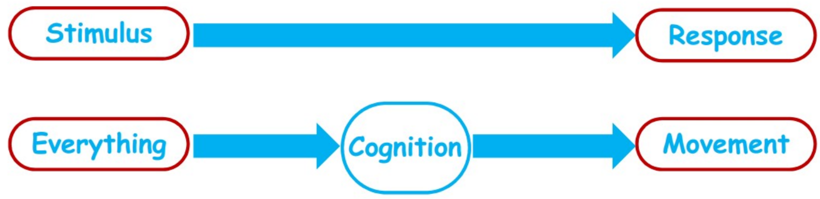

FIGURE 1.2: Two types of actions; **Above**: Stimulus-dependent and **below**: non[aparent]-stimulus-dependent

and stimulus-independent voluntary actions. Reactions are direct responses to external stimuli, while voluntary actions are internally initiated, uninfluenced by external stimuli, and accompanied by unique subjective experiences like the sensation of personal choice (Haggard, 2017).

Most daily actions lie on a continuum between stimulus-dependent and stimulus-independent actions. For instance, answering a ringing phone is voluntary (stimulus-independent) as we decide to answer, and reactional (stimulus-dependent) as the decision is triggered by the sound (Khalighinejad, 2022). The line between internal (voluntary) and external (stimulus-dependent) agency is crucial but challenging to delineate.

This challenge stems from the difficulty of defining an 'agent.' Linked to the concept of 'self,' the definition of an agent has been a contentious philosophical subject for centuries. Furthermore, the complex relationship between the agent and their environment, as well as the intricate processes governing decision-making and action initiation, add to the complexity.

Haggard, 2016 highlights that neuroscience often defines voluntary action by exclusion, characterizing it as not caused by stimuli. The focus is less on what provokes these actions and more on what doesn't—namely, stimuli.

For this thesis, we adopt a pragmatic criterion for differentiating between voluntary and involuntary actions. We deem an action voluntary if no apparent, discernible external stimulus perceived by an average human triggers the action, something like what has been ilustrated in Figure 1.2. This approach provides a working definition of voluntary actions, enabling us to investigate SoA in a more structured and focused manner.

This pragmatic approach enables us to explore the complexities of the SoA while recognizing the limitations imposed by the lack of a definitive boundary between the internal and external aspects of an agent. By examining the factors that influence human behavior and the mechanisms that underlie voluntary actions, we can gain valuable insights into the SoA and its potential implications for autonomy and responsibility.

1.3 Measuring the SoA: Methods and Challenges

The SoA is a subjective experience, making it complex to measure. Due to its multifaceted nature, a variety of methodologies are employed to capture its nuances. Two primary methods exist: explicit and implicit.

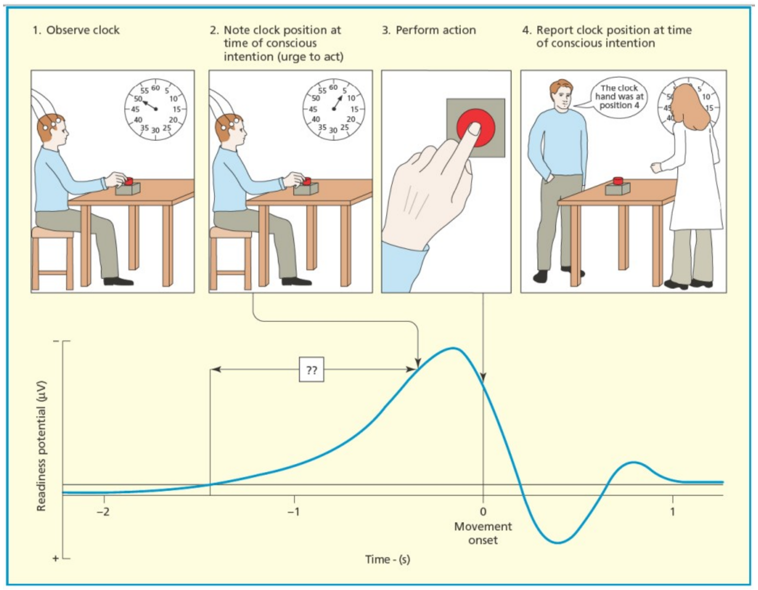

FIGURE 1.3: The classic Libet-style experiment; from Haggard, 2008 .

Explicit methods often involve self-report tasks, such as answering yes/no questions or rating the level of control over an event. These methods provide a direct and straightforward approach to gauging an individual's perceived agency. Despite their simplicity, explicit methods may sometimes be influenced by factors such as mood, expectations, or prior experiences, which could potentially affect their precision.

On the other hand, implicit methods involve evaluating the differences in perception and behavior between situations when individuals feel a SoA versus when they do not, or between situations when individuals experience strong versus weak agency. These methods, being less susceptible to the individual's momentary mental states, offer a more nuanced understanding of SoA.

1.3.1 Implementing a BCI Setup for SoA Measurement: Embracing the Libet Paradigm

Our primary objective is to establish a firm basis for the investigation. This involves the setup and configuration of a Brain-Computer Interface (BCI) system in our lab, which aligns with the implementation of a Libet-style experimental platform (Figure 1.3).

In this paradigm, participants perform self-paced actions, such as pressing a button, while their brain activity is monitored using electroencephalography (EEG). This design enables us to examine the relationship between conscious intention, action execution, and the perception of control, thereby uncovering crucial insights into the neural underpinnings of the SoA.

Given the time constraints and the intricacies involved in setting up the BCI, the current research will primarily focus on using explicit methods to measure the SoA. Despite the potential limitations of these methods, they offer an effective means of exploring our research questions and establishing an initial proof of concept.

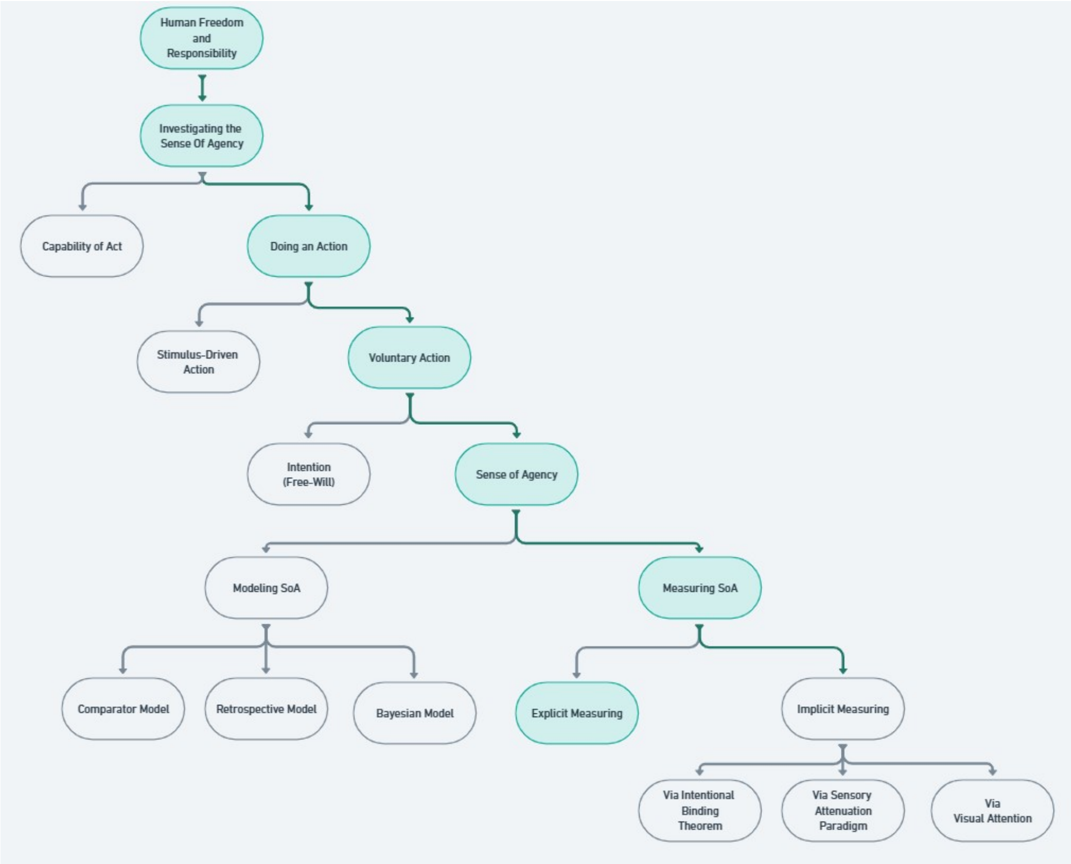

FIGURE 1.4: A roadmap depicting our journey from initial idea to our current thesis focus.

1.4 Review of the Journey So Far

In summary, our exploration began with the concept of SoA, leading us to investigate agency arising from motor actions. We proceeded to define actions and introduce their two primary types: voluntary and stimulus-dependent actions. Recognizing the challenge of establishing a clear boundary between these categories, we addressed the difficulty in defining the agent and the notion of self. Despite the complexities involved, we adopted a practical, though somewhat simplistic, definition of voluntary actions. Then we introduced implicit and explicit measurement of SoA, each with their pros and cons. Finally, we selected a Libet-style experimental design for our study.

The path that has led us to the current thesis topic is shown in Figure 1.4, encapsulating the key aspects discussed in the introduction section.

2. Hypothesis and Experimental Design

2.1 Readiness Potential and Intention Detection

Determining intentionality has always held immense interest in the fields of BCIs, neuroscience, and cognitive psychology, with the Readiness Potential (RP) as a focal neural signal in this pursuit. As our research aims to explore the SoA associated with actual actions, rather than the mere capability of performing actions, understanding and detecting RP is critical. As previously stated, we have chosen the Libet-style experiment as our platform to capture the subject's intention prior to action execution. While the famous Libet experiment, described in Libet et al., 1983, wasn't the first to detect RP, its intelligent design spotlighted the association of RP with voluntary movement, igniting profound debates around the concept of free will both in the scientific and philosophical communities. This section will unpack the complexities, challenges, and debates surrounding the real-time detection of RP and its role in intention detection. The groundwork laid here serves to contextualize our subsequent exploration of the transition from traditional RP-centric methods to contemporary machine learning approaches in intention detection.

2.1.1 Readiness Potential; What is it?

The Readiness Potential (RP), also known as Bereitschaftspotential (BP), is a neural signal associated with voluntary movement and has been utilized to challenge the concept of free will. First discovered by Kornhuber and Deecke, 2016 (original paper: Kornhuber and Deecke, 1965), the RP is identified through an analysis of electroencephalogram (EEG) data collected during experiments involving spontaneous or self-initiated actions. By time-locking EEG recordings to the onset of movement and averaging them, a gradually increasing negative electrical potential becomes apparent preceding the movement's initiation.

The most significant breakthrough in understanding the RP and its implications came with the famous Libet experiment. According to Benjamin Libet's findings, the RP begins 550 ms before the act, but human subjects become aware of their intention to act 350-400 ms after the RP starts, yet 200 ms before the motor act. Hence, Libet proposed that the initiation of a voluntary act appears to start in the brain unconsciously, well before the person consciously knows he wants to act, potentially challenging the conventional understanding of free will (Libet, Freeman, and Sutherland, 1999).

Despite these findings, Libet did not reject the concept of free will entirely. He proposed the possibility of a conscious veto, wherein our conscious intention could override or stop the unconscious initiation of an action (Libet, Freeman, and Sutherland, 1999). An interval of approximately 150ms between the conscious will (W)

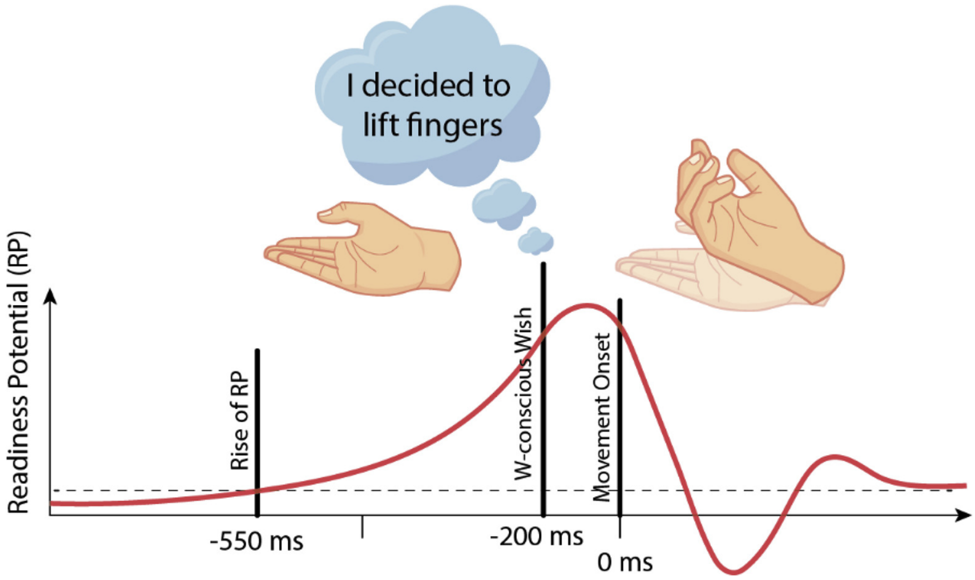

FIGURE 2.1: Temporal profile of the RP in Libet experiment, form Jamali, Golshani, and Jamali, 2019

and the actual movement allows for such a veto. However, only about 100ms of this window is practically available, as the final 50ms before muscle activation is the time for the primary motor cortex to activate the spinal motor nerve cells, and during this time, the act proceeds to completion with no possibility of stopping it (Libet, Freeman, and Sutherland, 1999). The temporal profile of the RP is illustrated in Figure 2.1.

Libet's interpretation of RP and its implications for free will continue to be discussed in contemporary research, and the concept of the conscious veto has been particularly controversial. Some scientists have argued that the veto, like the conscious intention, could itself be a consequence of some preceding unconscious neural activity. A breakthrough study by Schultze-Kraft et al., 2016 further elucidated this debate by investigating the timing window for this conscious veto, providing an in-depth examination of when we can effectively veto self-initiated movements (Schultze-Kraft et al., 2016). This work, alongside others, underlines that conscious veto might influence brain activity in ways we do not yet fully understand (Haggard, 2008). Nevertheless, the concept of a veto window where conscious intention can override unconscious initiation of action remains a cornerstone in the study of voluntary action and the neuroscience of free will.

2.2 The Shift to Machine Learning: Rationale and Preliminary Studies

Recognizing the complexities and ambiguities associated with traditional readiness potential detection, depicted in Appendix A, our research pivots towards employing machine learning. This transition is motivated substantially by the groundbreaking work of Schultze-Kraft et al., 2016, who successfully harnessed machine learning classifiers to capture subjects' intentions rather than solely attempting to detect readiness potentials. Their studies provide a strong foundation, exhibiting the promising potential of machine learning in this research realm.

In addressing the inherent challenges associated with traditional methodologies—specifically the debates surrounding the buildup or onset of readiness potential and the question of causality between RP detection and imminent action—machine learning emerges as a practical tool. It redirects the focus towards the pragmatic aim of imminent action detection, a target better suited to the strengths of machine learning classifiers.

From a technical perspective, machine learning confers significant advantages. Instead of depending on a constrained set of electrodes to seek readiness potentials—an endeavor fraught with difficulties—machine learning allows for the use of a broader set of electrodes for classification. This approach could potentially enhance the reliability of imminent action detection, while possibly providing a deeper exploration into neural activities associated with the SoA.

Moreover, the intrinsic value of data science comes to the fore with the application of machine learning, as it can potentially unearth meaningful patterns in the vast volume of EEG data that are not yet understood. It allows us to leverage the rich, high-dimensional data provided by EEG, extracting and utilizing complex, possibly nonlinear patterns that may be vital for reliable and robust intention detection. This represents a marked departure from traditional methodologies that typically rely on predefined features derived from our current understanding of brain activity, which, while valuable, may not encapsulate the full story.

In the following sections, we will delve into the specifics of our machine learning approach, detailing our experimental design, hypotheses, and the possible implications of the study.

2.3 Building Upon the Foundation: Revisiting the Study by (Schultze-Kraft et al., 2016)

The influential work of Schultze-Kraft et al., 2016 profoundly deepened our understanding of voluntary actions. They introduced the concept of a 'point of no return' in movement initiation, which is a temporal threshold following the onset of neural preparation but preceding action execution. Beyond this 'point of no return,' approximately 200 ms before the action, it becomes impossible to cancel a planned motor action.

Their findings not only echo but also refine Libet's notion of a conscious veto. Libet had proposed that a voluntary act could be canceled within a window of approximately 150 ms before the movement. However, the results from Schultze-Kraft et al., 2016 provide a crucial clarification: vetoing a movement becomes impossible beyond the 200 ms threshold before movement onset.

Interestingly, this 'point of no return' discovered by Schultze-Kraft et al. closely aligns with Libet's proposed 'W' time, when the awareness of the intention to act emerges. According to Libet, the 'W' time occurs roughly 200 ms before the action (see Figure 2.1), which suggests a paradoxical situation: the decision to cancel an action might occur at a point when we are not even fully conscious of the decision to act. This compelling observation opens up an intriguing area of exploration into the mechanisms underlying our subjective experience of intention and agency.

While Schultze-Kraft et al.'s work focused on pre-action neural processes and their relation to the cancellation of an action, our research aims to extend this concept

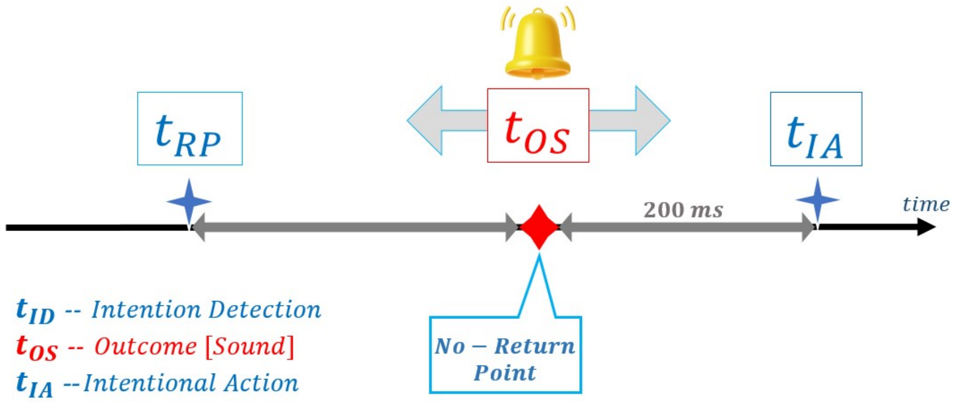

FIGURE 2.2: Rough timeline of the main experiment.

further. We propose to explore how the timing of this 'point of no return' influences our SoA following an action.

Inspired by Schultze-Kraft et al.'s findings, the following sections of our research journey build upon their groundbreaking study. We aim to delve deeper into the temporal dynamics of voluntary actions, the SoA, and how these can be better understood through machine learning-based imminent action detection methods. This shift away from traditional readiness potential detection methodologies promises to open up new possibilities for real-time intention recognition.

2.4 Formulating the Hypothesis

Grounding our research on the findings of (Schultze-Kraft et al., 2016), we put forth the following hypothesis related to the 'point of no return' and its impact on the perception of agency:

We hypothesize that the point of no return, occurring approximately 200 ms before a voluntary action, serves as a temporal marker that significantly shapes our post-action sense of agency. Specifically, once this point is crossed, the impending action is perceived as initiated in our subjective experience, thereby evoking a SoA even before the physical occurrence of the action. Consequently, if the outcome of the action is presented within this 200 ms window, subjects will still perceive a strong SoA, despite the physical action not having occurred. In contrast, if the outcome is presented before the point of no return, even though the intention to act has been formed, the perceived SoA significantly decreases.

This hypothesis suggests that our subjective SoA considers an action as 'initiated' once the point of no return has been passed, rather than at the moment of physical action execution. Therefore, it not only provides a retrospective perspective on the role of the point of no return in shaping our SoA, but also offers additional evidence in support of the existence of this critical temporal threshold, by demonstrating its effect on the SoA, rather than solely its role in the veto mechanism.

A rough schematic representation of the sequence and timing of these events, including the moment of intention detection, the point of no return, the time of generated outcome, and the moment of actual action, is illustrated in Figure 2.2.

2.5 Expectations

According to the hypothesis that we made, we anticipate that crossing the Point of no return, in terms of the timing of the outcome -which in our experiment is sound-will significantly alter the SoA experienced by the subject.

2.6 Experimental Design

The overarching aim of our experiment is to explore how the timing of an action's outcome, relative to the moment of decision, affects a person's perceived SoA. To facilitate this, we designed an experiment leveraging machine learning-based classifiers for detecting a participant's intention to perform an action in real time. The choice of classifier plays a pivotal role in this process, which will be elucidated in a following subsection.

Our experimental design comprises two main stages: Preparatory Stage: Collecting Subject Data and Classifier Training and the main Real-Time experiment.

In the Preparatory stage, we gather EEG data while participants spontaneously perform button press actions. Here, each button press synchronously triggers a sound, simulating a natural action-outcome relationship. The collected data serve two critical purposes: Training participant-specific classifiers for real-time BCI operation, and identifying the most discriminative EEG channels for each participant.

During the Real-Time stage, the classifiers trained in the preparatory stage are deployed to detect the participants' intention to press the button in real time. Upon intention detection, we trigger an auditory outcome at varying time intervals. This manipulation allows us to assess the effect of timing of the outcome relative to the action on the participants' perceived SoA over the sound.

2.6.1 Classification Algorithm Selection

Within the context of the real-time BCI experiment, one of the crucial steps following the EEG data acquisition involves extracting relevant features from the data and feeding them into a classifier. The chosen classifier, trained on these features, plays a critical role in discerning the current state of the subject. The classifier makes this determination by distinguishing between two main classes:

- **'Idle' or 'No Intention' Class:** This class is assigned when no significant pattern suggestive of an impending action is detected in the EEG data. In essence, when the classifier assigns this label, it indicates that no imminent action is expected from the subject.
- **'Pre-movement' or 'Intention' Class:** This class is assigned when the EEG data exhibits a pattern that is indicative of an impending voluntary action. An assignment of this class by the classifier signifies that an action from the subject is imminent.

The ability to accurately and swiftly classify these states is of paramount importance for the successful operation and performance of the real-time BCI experiment. Swift classification ensures minimal latency between the subject's intent to act and the response of the BCI, which is critical for maintaining a natural and seamless interaction in the experiment

There are numerous options for classifier selection when dealing with a binary classification problem, among which, the Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) classifiers are two common choices, particularly suitable for real-time BCI systems due to their combination of accuracy, robustness, and computational efficiency:

1. **Support Vector Machine (SVM) Classifier:** SVM is a powerful and widely-used supervised learning method that can handle both linearly separable and non-linearly separable data by using appropriate kernel functions. However, SVM requires solving a quadratic programming problem, which can be computationally expensive, especially for large datasets.

2. **Linear Discriminant Analysis (LDA) Classifier:** LDA is a supervised linear classification technique that assumes linear separability and equal covariance matrices for classes. It seeks to maximize the between-class variance while minimizing the within-class variance. LDA has a closed-form solution, which means that it can be solved directly, making it computationally efficient.

Inspired by the work of (Schultze-Kraft et al., 2021) and specially (Schultze-Kraft et al., 2016) we have opted to utilize a Regularized Linear Discriminant Analysis (RLDA) with automatic shrinkage as our classifier. This choice is motivated by several key factors:

1. **Improved performance:** RLDA incorporates regularization, which helps reduce overfitting and enhances generalization capabilities. This leads to more robust classification performance compared to standard LDA, particularly when dealing with limited training data or high-dimensional feature spaces.

2. **Automatic shrinkage:** The automatic shrinkage feature in RLDA allows for an optimized balance between the standard LDA and a purely diagonal discriminant analysis. This optimization is achieved by estimating the optimal shrinkage intensity from the data itself, leading to better classification results without the need for manual tuning of hyperparameters.

3. **Computational efficiency:** Although RLDA with automatic shrinkage adds some extra computational load compared to traditional LDA, it still maintains a high level of computational efficiency, making it suitable for high-speed applications.

By selecting RLDA with automatic shrinkage as our classifier, we aim to leverage its enhanced performance and adaptability, while maintaining the computational efficiency required for successful real-time BCI implementation.

2.6.2 Feature Extraction

Feature extraction is a pivotal step in EEG data classification, determining the selection of informative, independent, and most notably, discriminative features. The noisy, high-dimensional nature of EEG data often necessitates a transformation of the original signals into a more suitable feature space for achieving accurate classification results. Consequently, feature selection significantly influences classifier performance. The subsequent sections detail the feature selection process for identifying the specific classes of interest in this experiment.

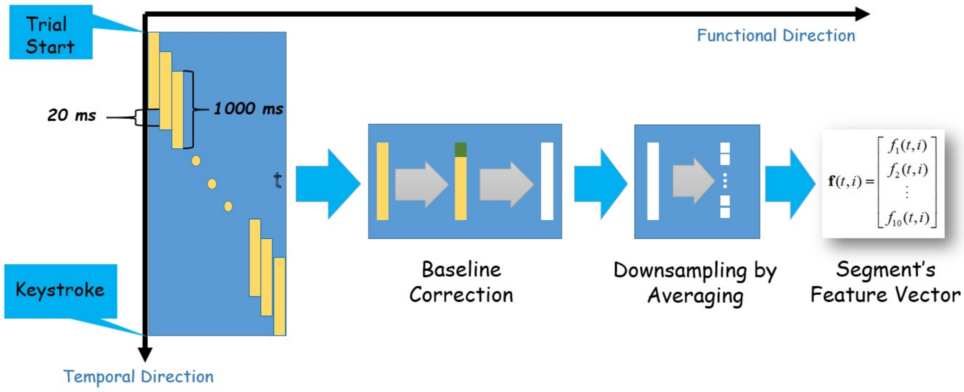

FIGURE 2.3: Feature Extraction Process. **Baseline Correction**: subtracting the average signal value in the first 100 ms (from -1000 ms to -900 ms) from all the data points in the entire segment. **DownSampling**: averaging the values in consecutive 100 ms intervals, resulting in 10 temporal features per segment and channel. **Segment's Feature Vector**: Concatenate the temporal features across all channels to form a feature vector for each segment.

Feature Selection

Our real-time BCI experiment design's feature selection approach draws extensively from the work of Schultze-Kraft et al., 2021 and Schultze-Kraft et al., 2016. In line with their studies, the feature selection process employs a downsampled version of each data segment. Initially, each segment undergoes baseline correction concerning its first 100 ms. Following this, data downsampling is achieved by averaging the values within consecutive 100 ms intervals, generating 10 temporal features per segment and channel. Thus, the chosen features comprise temporal averages of the initial EEG data within each 100 ms interval, an effective strategy for capturing the pre-movement brain activity integral to the success of this real-time BCI experiment design.

Extract the Feature Vector

Building on the aforementioned feature selection approach, the following steps outline the feature extraction process:

1. Continuously stream and store the EEG data in a buffer at the trial's commencement.
2. Select a 1000-ms-long segment, updated continuously with incoming data in 20 ms steps, ending at the current time point.
3. Execute baseline correction on each segment by subtracting the average signal value in the first 100 ms (from -1000 ms to -900 ms) from all data points within the entire segment.
4. Downsample the baseline-corrected segments by averaging the values within consecutive 100 ms intervals, yielding 10 temporal features per segment and channel.

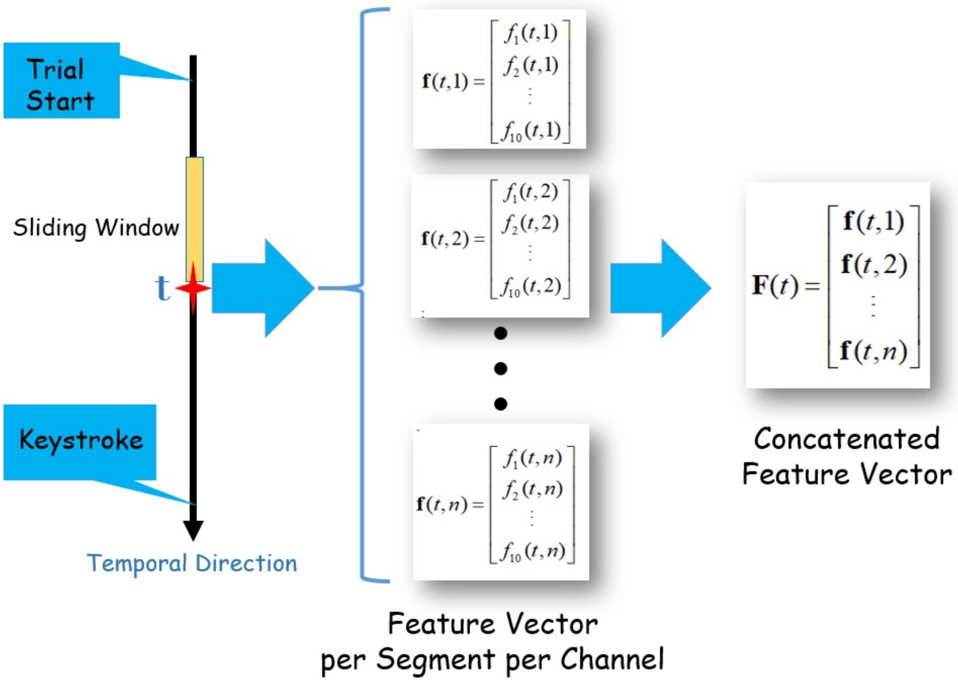

FIGURE 2.4: Feature Concatenation Process.

5. Concatenate the temporal features across all selected channels to formulate a comprehensive spatiotemporal feature vector for each segment.

Adhering to these steps ensures that the extracted feature vectors efficiently capture the pre-movement brain activity in real-time. These feature vectors can then be used as input to the classifier, allowing for accurate classification of the EEG data in this real-time BCI experiment.

As illustrated in Figure 2.3, the general process of feature extraction during the online experiment can be visualized. At each moment, with a resolution of 20 ms, a feature vector is obtained for each segment and channel after performing baseline correction and downsampling. The resulting feature vector for each segment and channel can be represented as follows:

$$\mathbf{f}(t,i) = \begin{bmatrix} f_1(t,i) & f_2(t,i) & \cdots & f_{10}(t,i) \end{bmatrix}^T$$

where $\mathbf{f}(t,i)$ is the feature vector at time t for channel i , and $f_1(t,i)$, $f_2(t,i)$, ..., $f_{10}(t,i)$ are 10 temporal features extracted from each segment of channel i at that specific time. This feature vector will be continuously updated as new data is streamed into the buffer, allowing for real-time classification of the EEG data in the BCI experiment.

The next step in the process is to concatenate all feature vectors from each channel in order to create a single comprehensive feature vector for each segment. By combining the features from all channels, the classifier can utilize both the spatial information across channels and the temporal features extracted from each segment, leading to improved classification performance.

The concatenated feature vector for a segment can be represented as follows:

$$\mathbf{F}(t) = \begin{bmatrix} \mathbf{f}(t,1) & \mathbf{f}(t,2) & \cdots & \mathbf{f}(t,n) \end{bmatrix}^T$$

where $\mathbf{F}(\mathbf{t})$ is the concatenated feature vector at time t and n is the number of selected channels and $\mathbf{f}(t,i)$, $1 \leq i \leq n$, as stated before, is the feature vector at time t for channel i. The number of selected channels for classification is n. This concatenated feature vector will be continuously updated every 20 ms, as new data is streamed into the buffer, allowing for real-time classification of the EEG data in the BCI experiment. This contamination process is shown in Figure 2.4.

Thus, at any given time, there are $10n$ distinct features fed into the classifier, where n is the number of channels and each channel contributes 10 temporal features.

2.6.3 Preparatory Stage: Collecting Subject Data and Classifier Training

The preparatory phase of the experiment is an integral and commonplace stage in creating accurate and robust BCIs, particularly in machine learning-based classifications as demonstrated in studies like Schultze-Kraft et al., 2021 and Schultze-Kraft et al., 2016. This stage involves two pivotal components: the collection of individual subject data and the training of the classifier based on this data.

During this stage, participants are asked to engage in a voluntary action, specifically, pressing a button. This self-paced action generates a computer-produced sound. For enhanced congruency between the action and its corresponding outcome, a box containing the button and the speaker is provided to participants (see Figure 2.5). As participants perform these voluntary actions while their EEG signals are being recorded, we collect the required data for the 'Pre-movement' or 'Intention' class. In addition, periods of inactivity captured during this phase serve to provide the data needed for the 'Idle' or 'No Intention' class. The data collected during these activities encapsulates the unique patterns associated with each class for every individual, thereby providing the foundation for subject-specific classifier training. The aim is to tailor the classifier so it can accurately differentiate between these two classes for each participant.

Moreover, the preparatory phase plays a crucial role in the optimal selection of EEG channels. This step aims to identify the most informative and discriminative EEG channels for each participant, optimizing the classifier's performance. By focusing on the channels that carry the most relevant information for each individual, we can enhance the overall performance of the BCI system.

The following subsections will delve into the specifics of these steps, detailing the process involved in collecting subject data, training the classifier, and selecting the optimal EEG channels.

To reiterate:

- The **Voluntary action** in this context is the participant pressing a button, at a time of his or her own choosing.
- The **Outcome** is the computer-generated sound as a result of the action.

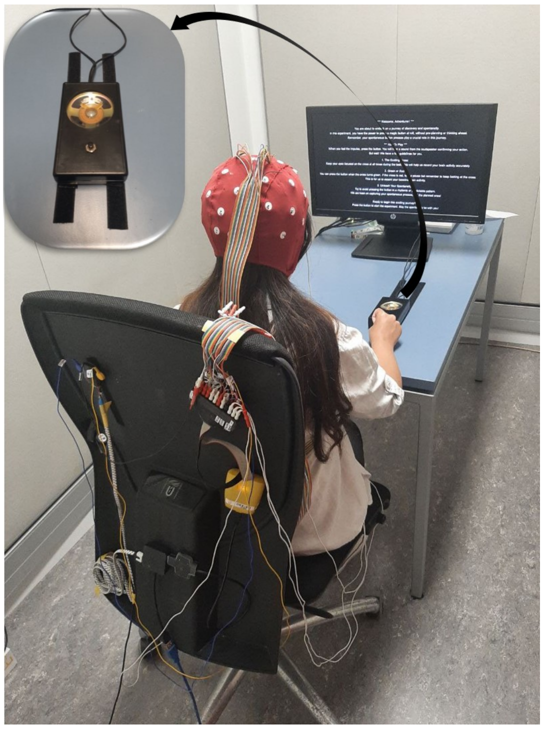

FIGURE 2.5: A participant in the resting state, ready to press the button on the box that also contains a speaker.

Gathering Data and Subject-Specific Classifier Training

In order to train the classifier to accurately distinguish between 'Idle' or 'No Intention' and 'Pre-movement' or 'Intention' states, a substantial amount of subject data needs to be collected. This is achieved during a dedicated preparatory stage of the experiment where participants are engaged in a series of trials designed to elicit the required brain activity patterns.

This phase of the experiment which consists of at least two blocks, each lasting for about 8-10 minutes, depending on the number of button presses that occur can be summarized as follows:

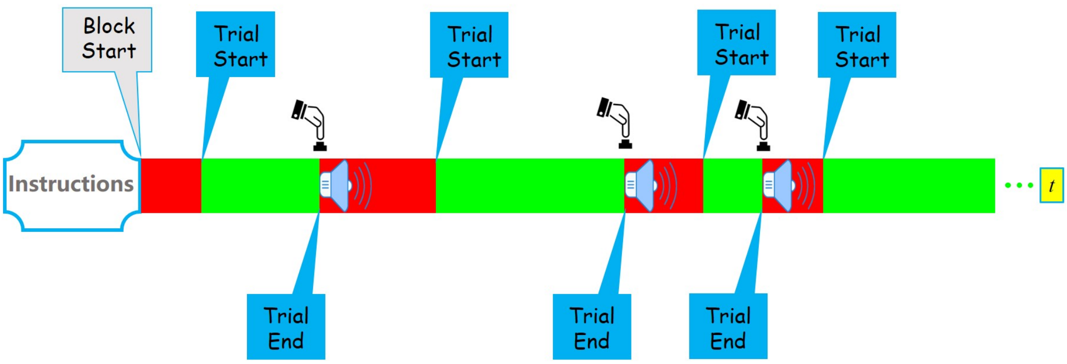

FIGURE 2.6: Timeline of the Prepratory Stage Experiment

1. **Resting State**: Depending on whether the participant is left-handed or right-handed, a box containing a button and a speaker is affixed on the table on the side of the participant's dominant hand (see Figure 2.5). Participants are instructed to place their dominant hand in a rest position on the box and be ready to press the button in a manner that involves minimal movement, maintaining the same physical pattern for pressing the button throughout the experiment.
2. **Instruction**: Participants are instructed to fixate on a cross in the middle of the screen that appears at the start of each block until the end of the block. This measure helps minimize eye movement artifacts during recording.
3. **Block Start**: Each block commences in a continuous manner, take 8-10 min. After pressing the button at the end of the instructions, a red fixation cross appears for a fixed time (2.5 sec), followed by a green fixation cross, signaling the start of the first trial.
4. **Voluntary Action**: Participants are required to perform a voluntary button press whenever they wish without any time constraint. The only restriction is that the action should not be taken while the fixation cross is red, i.e., during the inter-trial intervals which include the classifier-driven action inhibition period.
5. **Trial Start**: Each trial starts when the fixation cross turns green.
6. **Trial End**: Each trial ends with the participant's voluntary button press. The color of the fixation cross changes to green and the next trial begins. During this phase, the brain activity data collected will be used to train the classifier.
7. **Inter-Trial Interval**: Between every two trials, there is an inter-trial interval (ITI) composed of a random delay, t_{iti}, averaging 2 second, and the time required for playing the outcome sound initiated by the participant's button press. This structured interval prevents the participants from falling into a fixed, rhythmic pattern of button pressing. The fixation cross remains red during this interval, indicating that the participant should not press the button.
8. **Block End**: Once the participant has performed a sufficient number of button presses or after 8 minutes (whichever comes first), and no later than 10 minutes, a break page appears on the screen signaling the end of the block. Participants can take this time to rest before they initiate the next block by pressing the button. This structure ensures that the block does not extend beyond a set limit, preventing potential fatigue and maintaining the quality of data collected.

This protocol outlines the procedure for collecting subject data during the preparatory stage of the experiment. The data collected through this process will subsequently be used for training the classifier in a subject-specific manner.

Optimal EEG Channel Selection

For effective differentiation between the 'Idle' or 'No Intention' and 'Pre-movement' or 'Intention' states, the selection of the most discriminative EEG channels for each participant is vital. This selection process constitutes the second part of this prepratory phase. The EEG data collected during the 'Gathering Data and Subject-Specific Classifier Training' phase is analyzed using various feature extraction and channel ranking methods. By selecting the most informative channels unique to each individual, the performance of the classifier can be optimized, enhancing the accuracy and robustness in distinguishing between the two classes during the main experiment.

However, in the landscape of research, there are varying approaches to channel selection. Some studies choose a set of pre-established electrodes, typically those that cover areas of the brain known for their strong correlation with intentional neural activity, such as the motor area. On the other hand, some researchers opt for an individualized approach, selecting the channels that yield the most significant impact on the classification for each participant. This latter approach typically happens after an evaluation of the trained classifier.

In our study, we adopt an individualized approach and select channels based on their correlation with each subject's actions. We utilize 17 passive electrodes positioned to cover the primary and pre-motor cortex, where the RP typically appears, with additional electrodes scattered around to capture more global brain activity. This approach aligns with our primary objective, which is to identify the channels that are most closely correlated with the subjects' actions, rather than focusing on understanding the nature of the RP itself.

2.6.4 Main Real-Time Experiment

In the main real-time experiment phase, we employ a setup similar to the preparatory stage, but with some critical modifications. Firstly, EEG data are collected in real-time and instantaneously fed into the classifier, which has been specifically trained for each participant. This process allows us to detect the neural activity related to the intention to press the button (the 'Pre-movement' class) in real-time. Secondly, the outcome - a computer-produced sound - is initiated at various time points across different trials, based on the classifier's output. A third key modification in this phase is that, at the end of each trial, participants rate their SoA for that specific trial on a scale. This direct measure provides valuable insight into how participants perceive their control over the timing of the outcome sound, relative to their voluntary button press.

Through this tailored experimental setup, we aim to investigate how the temporal relationship between a voluntary action, its outcome, and the detection of the intention to act influences the perceived SoA. This approach aligns with our research hypothesis, probing into the role of the 'point of no return'—the moment around 200 ms before a voluntary action when the intention to act has become irreversible—in shaping the perception of agency. By introducing variability in the timing of outcomes and collecting participants' SoA, we anticipate generating a rich data set to shed light on these intricate relationships.

The main experimental design, which consists of at least two blocks, each lasting for about 8-10 minutes, depending on the number of button presses that occur can be summarized as follows:

1. **Resting State**:Depending on whether the participant is left-handed or right-handed, a box containing a button and a speaker is affixed on the table on the side of the participant's dominant hand. Participants are instructed to place their dominant hand in a rest position on the box and be ready to press the

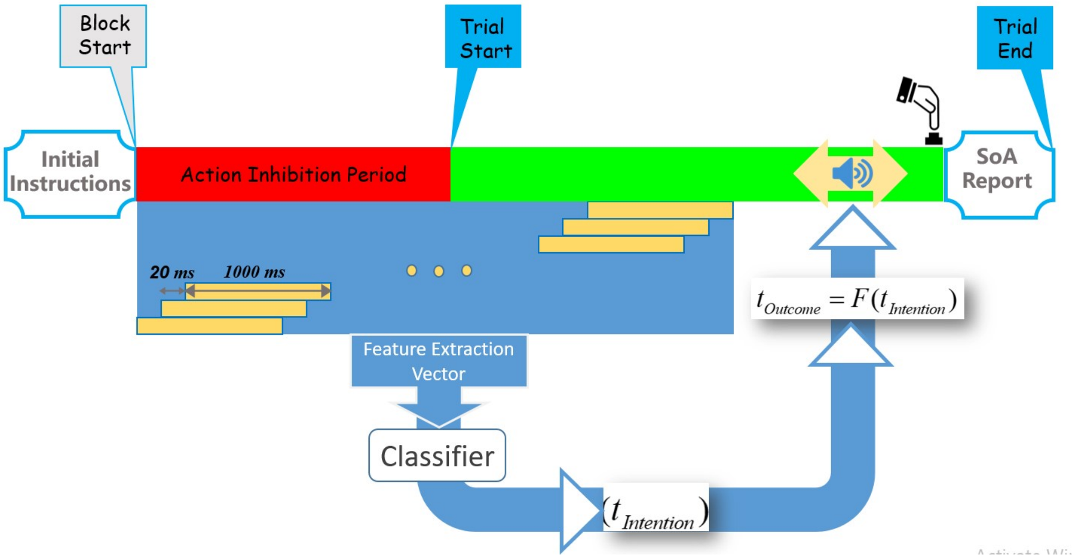

FIGURE 2.7: Timeline of the Time Experiment; Only for First Trial

button in a manner that involves minimal movement, maintaining the same physical pattern for pressing the button throughout the experiment.

2. **Instruction**: As in the preparatory stage, participants are instructed to fixate on a cross in the middle of the screen throughout each block to minimize eye movement artifacts during recording. At the end of the instruction page, they are asked to press the button to start the experiment.

3. **Block Start**: Each block commences in a continuous manner, similar to the preparatory stage. After pressing the button at the end of the instructions, a red fixation cross appears for a fixed time (2.5 sec), followed by a green fixation cross, signaling the start of the first trial.

4. **Voluntary Action**: Participants are required to perform a voluntary button press whenever they wish without any time constraint. The only restriction is that the action should not be taken while the fixation cross is red, i.e., during the inter-trial intervals which include the classifier-driven action inhibition period.

5. **Trial Start**: Each trial starts when the fixation cross turns green.

6. **Real-time Classification**: From the moment the experiment starts, the last 1000 ms window of the participant's EEG data is continuously fed into the trained classifier in 20 ms steps. The classifier output determines the participant's state, i.e., either in an 'Idle' or an 'Intention' state. If the classifier output classifies an 'Intention' state, the outcome sound is initiated at a varying time in different trials before the actual action.

7. **Outcome**: The outcome is a simple computer-generated sound that is intended to occur synchronously with the actual action, i.e., the button press. However, it's important to note that this synchrony is natural rather than literal; technical limitations and inherent processing delays mean that there will be always a

minor delay between the button press and the initiation of the sound. This delay mimics the kind of synchrony typically expected in real-world interactions between mechanical actions and their outcomes

8. **SoA Explicit Report**: After the participant presses the button or the outcome sound occurs, whichever happens later, they are asked to rate their SoA for the specific trial on a scale. A page for this rating appears on the screen, which participants should answer using the mouse provided for this purpose.

9. **Trial End**: Each trial ends once the participant has submitted their SoA assessment.

10. **Inter-Trial Interval**: Following the SoA report, another button press prompts the appearance of a red fixation cross for a random delay with an average duration of 1.5 sec. This inter-trial interval prevents participants from falling into a fixed, rhythmic pattern of button pressing and guarantees a 1000 ms feature extraction temporal window. The fixation cross remains red during this interval, indicating that the participant should not press the button. Afterward, it turns green, signaling the start of the next trial.

11. **Block End**:Once the participant has performed a sufficient number of button presses or after 8 minutes (whichever comes first), and no later than 10 minutes, a break page appears on the screen signaling the end of the block. Participants can take this time to rest before they initiate the next block by pressing the button.

The Figure 2.7 illustrates the timeline for the first trial following the start of the block. This provides a visual representation of the process outlined above, showing the sequence of events as they unfold in real-time.

3. Results, Discussion and Future Plan

3.1 Some Extra Practical Considerations and Challenges

Over the course of our research journey, due to time constraints, our primary focus was on the preparatory stage of our experimental design. This stage serves as a proof of concept, indicating that we are on the correct path towards fulfilling our research objectives. This thesis does not include results from the main real-time experiment phase. However, the ensuing sections provide a detailed discussion of the findings from our preparatory stage, their implications, and outline future plans for completing the subsequent stages of this experiment.

EEG Recordings

During the initial phase of our study, we consistently gathered EEG data via 17 passive electrodes (F3, Fz, F4, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, Oz). These were strategically positioned in line with the globally recognized 10-20 system. The exact montage of these electrodes is illustrated in Figure 3.1 To track horizontal ocular movements, we deployed an external electrode, and for the possibility of offline re-referencing, two additional electrodes were affixed to the left and right mastoids. For the online reference, we employed the AFz electrode and designated the right mastoid as the grounding electrode. This data was meticulously recorded with the aid of the ENOBIO 20 5G system, which operates at a sampling rate of 500 Hz, along with a medical-grade touch-proof adapter. Both these tools were sourced from Neuroelectrics, based in Barcelona, Spain.

More About Outcome Sound

We prepared an outcome sound, which is a simple, computer-generated tone produced using PsychToolbox. After several iterations, we settled on a quasi-periodic sound. It was created from a linear combination of four frequencies: 500 Hz, 700 Hz, 900 Hz, and 1000 Hz. The sound had a duration of 500 ms and a sampling rate of 22100 Hz.

Timing of Events

Accurately capturing the timing of events was an essential aspect of our experimental setup. In particular, we needed to precisely track the timing of each button press. To ensure this, we employed a multi-method approach.

To accurately capture and mark each button press event within the EEG data, we utilized the TCP/IP protocol, as facilitated by the Neuroelectrics Instrument Controller (NIC). This system enables the real-time transmission of each button press event to

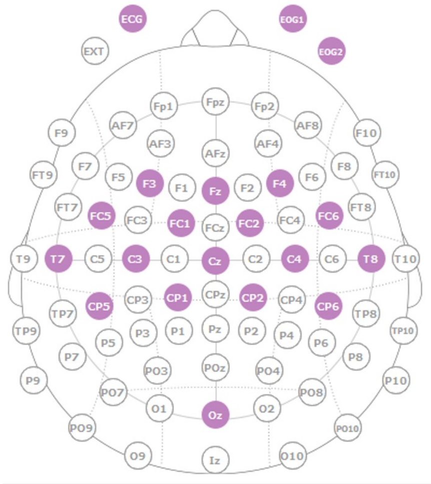

FIGURE 3.1: EEG Electrode Montage for Prepratory Stage

the host PC, where it is annotated as a marker on the corresponding EEG data. The NIC software ensures synchronization between the EEG data and the event markers, even in the context of inherent system latencies. Thus, every button press is accurately timestamped in the EEG data stream, providing us with a reliable record of participant actions in relation to their brain activity.

In addition to the use of TCP/IP protocol for button press event tracking, we employed a second method to further increase the precision of timing. We incorporated a dedicated Transistor-Transistor Logic (TTL) circuit, as offered by the Enobio system, into our setup. This circuit is designed to synchronize the delay from the moment a button press is captured to the moment it is recorded on the host PC's EEG data. This is achieved by aligning this delay with the time difference between the capture of EEG signals by the Enobio electrodes and their subsequent recording on the host PC via the NIC interface.

The functionality of this TTL circuit is based on the reception of a TTL signal that NIC can display through one of its EEG channels using a dedicated TTL adapter. This approach requires the presence of a parallel port on the computer running the external software that controls the experiment. By configuring the connections appropriately, a TTL input signal can be sent and recognized whenever a TTL pulse is dispatched. The adapter's TTL input is set to recognize signals between 3.3V and 5V. Once received, these signals are synchronized with the EEG data stream and displayed on the corresponding EEG channel. However, it is important to note that these signals are not registered as markers in the penultimate column of the .easy files - the detection and offline analysis of the TTL pulses is the responsibility of the user.

Additionally, we included a system for the experimenter to visually monitor each button press. Using a different parallel port on the same computer, each button press activated an LED to flash, serving as a visible signal to the experimenter outside the room. Furthermore, the total number of button presses was continuously displayed on the main PC. This information allowed us to adapt the experiment in real-time. For instance, if we did not reach the desired number of button presses, we could extend the current block without interrupting the participant or add another extra block.

Upon comparing the TCP/IP and TTL methods, we observed that the delay introduced by the TCP/IP protocol was always under 26 ms, which equates to less than 13 samples at the 500 Hz frequency of the EEG device. Given the 20 ms steps of our sliding feature extraction window, our system effectively achieved a temporal resolution greater than 40 ms. Thus, the delay introduced by the TCP/IP protocol is unlikely to significantly affect our results.

After careful consideration and facing some challenges with the TTL circuit, we adopted the TCP/IP method as our main approach for capturing the time of button presses, for this preparatory stage.

This meticulous attention to timing accuracy underpins our objective to continuously monitor and classify neural signals, thereby reliably detecting the participant's intention to press the button.

3.2 Results

We collected data from 5 participants (3 Male and 2 Female; age: 22-29) for the Preparatory Stage of the experiment. The subsequent steps were focused on transforming this collected EEG data into a labeled version suitable for classifier training, using a series of processes implemented via custom MATLAB scripts

During the Preparatory Stage, the continuous EEG data underwent a two-stage transformation into a format suitable for classifier training. In the first stage, the raw EEG data was segmented into trials based on participants' button presses, resulting in trial lengths of several seconds each. In the second stage, these trials were further divided into 1000 ms sliding windows. Following the segmentation, we assigned each window to either an idle or pre-movement state, shaping the temporal structure within the EEG data.

Once segmented, each 1000 ms window was then transformed into a feature space. We applied a baseline correction and then averaged the values within each 100 ms interval of the window, producing a set of temporal features for each segment and channel. Each set of features was subsequently labeled according to its respective class, resulting in a labeled feature set. This set formed the foundation for our supervised learning approach, where the classifier was trained to identify and discriminate between patterns associated with each class.

Once the EEG data was successfully transformed into a labeled feature set, we employed this data to train a classifier. This training process was unique, particularly due to our use of Linear Discriminant Analysis (LDA) classifier, and the way we automated the process of finding the optimal gamma (regularization) parameter. We accomplished this by implementing a leave-one-out cross-validation scheme, iteratively training the LDA model for each fold across a range of gamma values. The gamma that provided the highest cross-validation accuracy was then chosen as the best gamma and used to train the LDA model. This automated process significantly reduced the need for manual tuning and ensured an optimal and generalizable model.

The performance of the classifier was not only assessed through the mean cross-validation accuracy but also through the weights assigned to each EEG channel by the LDA model. This analysis gave us valuable insights into the relative importance of each channel for distinguishing between the 'Idle' and 'Pre-movement' classes.

Through this strategic and automated approach to data processing and classifier training, we efficiently transformed the raw EEG data into insightful analysis of the participant's brain activity. This process laid the groundwork for the real-time stage of the experiment, where the trained classifier will be employed for real-time BCI operation. It also informed us about the most discriminative EEG channels for each participant, which would be instrumental for improving the accuracy and efficiency of the real-time BCI operation.

The primary accuracy of the trained classifier for each participant, based on the data gathered from their performance during the preparatory stage, and the time it takes for the classifier to be trained is shown in Table 3.1.

TABLE 3.1: Mean accuracy, training time and optimal selected channels for Classifiers Trained on Different Participants

Participant Number	Mean Accuracy	Training Time (Sec)	Most Influential Channels
P1	0.91195	65.693	T7, C3, C4, CP1, Fz
P2	0.9016	61.8430	T7 , C3, CP6, FC5, FC6
P3	0.9151	64.9135	CP1, T8, C3 , C4, F3
P4	0.9336	71.0107	T8, Fz, C4 , Cz, FC5
P5	0.8723	69.1088	F3, C3, T7, Cz, C4

The classifier accuracies presented in Table 3.1, obtained from each participant during the Preparatory Stage of our experiment, demonstrate promising results. The high mean accuracy values, computed through a leave-one-out cross-validation scheme, indicate that the classifier performs well in distinguishing between the 'Idle' and 'Pre-movement' states. In this process, for each participant's EEG data, one sample was left out in each iteration, and the classifier was trained on the remaining data and tested on the left-out sample. This was repeated until each sample had been used for testing exactly once. It is important to note that while these accuracy estimates provide valuable insights, they do not necessarily reflect the real-time classification performance that will be achieved in the later stages of the experiment. This is due to the potential influence of various factors not captured in the cross-validation process. Nevertheless, the successful performance in the preparatory stage builds confidence for the upcoming real-time experiment phase.

3.3 Future Plan

We are presently in the process of coding a simulation for the real-time experiment, to evaluate the performance of participant-specific classifiers in a real-time context, albeit in an offline mode. To achieve this, we are generating overlapped feature vector segments, identical to those that will be used in the real-time experiment. The objective is to determine the earliest point in time at which the classifier can detect the intention class and the corresponding accuracy.

For each participant-specific classifier, we aim to generate a graph that indicates the average accuracy of imminent button press detection at a 1-second period leading up to the actual button press. If we manage to identify an imminent button press a sufficient time prior to the actual event with satisfying accuracy, we can then proceed to the next step – the main real-time experiment.

In terms of the main real-time experiment, it is crucial to emphasize that the preparatory stage, including data collection, classifier training, and the aforementioned simulated real-time experiment, should be performed for each participant prior to initiating the main real-time experiment. We aim to conduct all these stages in a single session.

This strategy is justified by the data-dependent nature of our technique. The collected data, and consequently, the performance of the trained classifier, could be influenced by various factors, including the timing of the data collection. Therefore, by training the data-based classifier and employing it in the same session, we aim to minimize the impact of time-related variability on the model's performance. This approach will enhance the reliability of our results and ensure the efficiency of the real-time BCI operation.

3.4 Discussion and Conclusion

The satisfactory mean accuracies of the participant-specific classifiers, trained with data collected in the preparatory stage, provide a promising start for our experiment. However, transitioning from this preparatory phase to the real-time experiment introduces new challenges, with timing emerging as a critical aspect.

In our real-time experimental design, the main task is to estimate the timing of an imminent action as soon as the intention is detected from a participant (indicated by the classifier output). We aim to initiate the outcome sound prior to the actual action, with the exact initiation point varying within the experiment.

We acknowledge that the performance measured during the preparatory stage might not fully reflect the performance during the real-time experiment due to the dynamic and real-time nature of the data in the latter. To gain a more accurate estimation of the expected accuracy in real-time, we plan to simulate the real-time experiment, as outlined in our future plans. This simulation will allow us to investigate at what point within a segment the classifier output reaches a level of confidence that satisfies our requirements.

Achieving a high-confidence classifier output at least 300 ms before the actual action is particularly crucial. This lead time allows for the initiation of the outcome sound within a reasonable time frame (i.e., from -250 ms to +250 ms relative to the actual action time). This timing constraint is an essential factor for testing our hypothesis and achieving our experiment's desired results.

Our journey of investigating SoA through a neuroscientific approach and setting up a BCI system for it has proven to be an ambitious yet rewarding endeavor. Each stage of the process has deepened our knowledge and understanding of this complex field, presenting both challenges and invaluable learnings. We look forward to the upcoming real-time experiment phase with eagerness and optimism, hopeful that our work will contribute significantly to the broader understanding of SoA and pave the way for further discoveries in neuroscience and BCI technology.

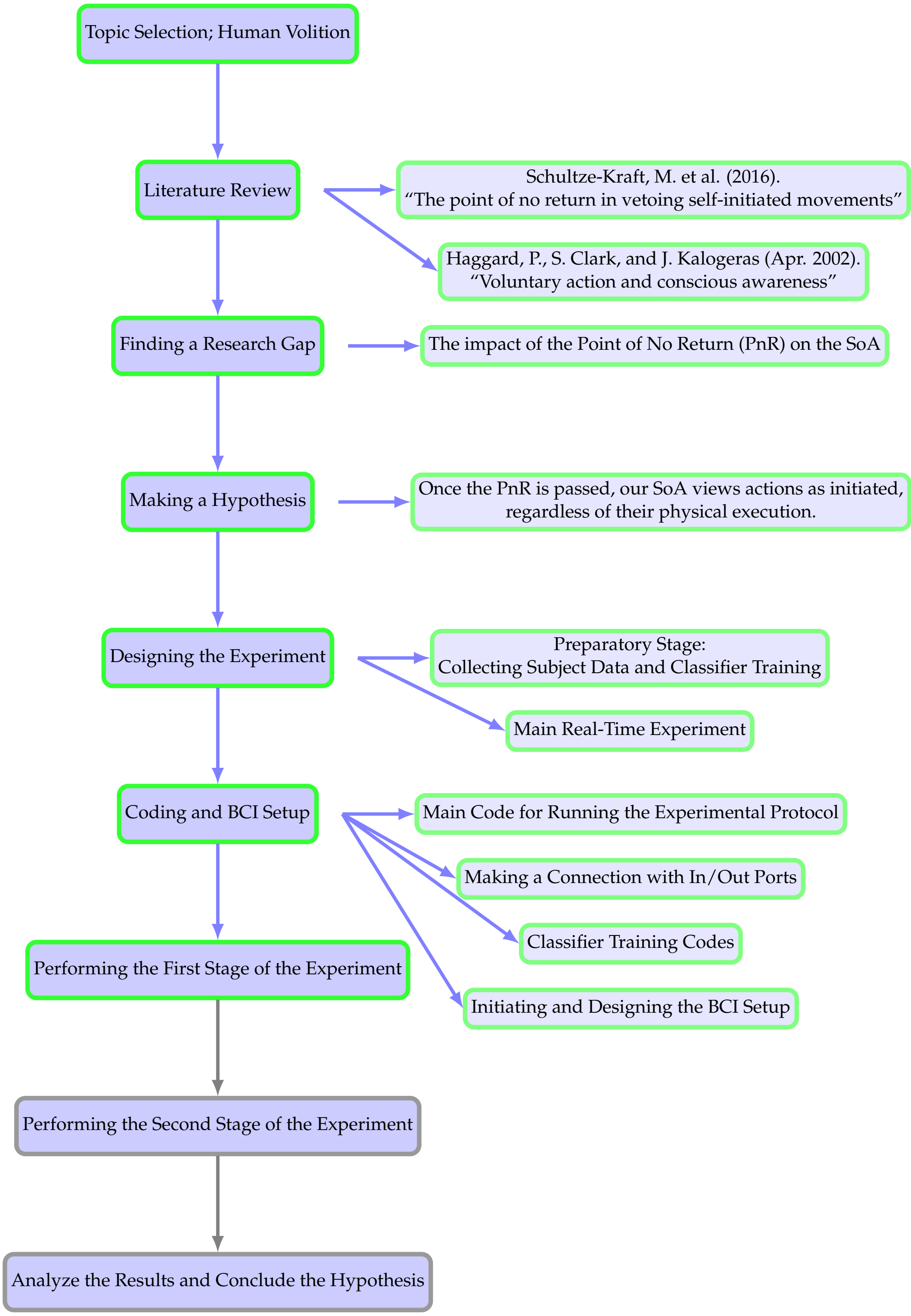

FIGURE 3.2: Flow Diagram of the Thesis Stages: The stages of the thesis are represented as blocks with associated sub-blocks. The blocks outlined in green are stages that have been completed, while the blocks outlined in gray are still in progress. 'PnR' stands for for 'Point of No Return'.

Bibliography

Abou Zeid, E. and T. Chau (2015). "Electrode fusion for the prediction of self-initiated fine movements from single-trial readiness potentials". In: *International journal of neural systems* 25.4, p. 1550014. URL: https://doi.org/10.1142/S0129065715500148.

Abou Zeid, E., A. Rezazadeh Sereshkeh, and T. Chau (2016). "A pipeline of spatio-temporal filtering for predicting the laterality of self-initiated fine movements from single trial readiness potentials". In: *Journal of Neural Engineering* 13.6, p. 066012. URL: https://doi.org/10.1088/1741-2560/13/6/066012.

Blakemore, S., D. M. Wolpert, and C. D. Frith (1998). "Central cancellation of self-produced tickle sensation". In: *Nature Neuroscience* 1.7, pp. 635–640. URL: https://doi.org/10.1038/2870.

Cornelio, P. et al. (2022). "The Sense of Agency in emerging technologies for human–computer integration: A review". In: *Frontiers in Neuroscience* 16. URL: https://doi.org/10.3389/fnins.2022.949138.

Garipelli, G., R. Chavarriaga, and J. d. R. Millan (2011). "Single trial recognition of anticipatory slow cortical potentials: The role of spatio-temporal filtering". In: *Fifth International Conference on Neural Engineering*. Cancun, Mexico.

Haggard, P. (2008). "Human volition: Towards a neuroscience of will". In: *Nature Reviews Neuroscience* 9.12, pp. 934–946. URL: 10.1038/nrn2497.

Haggard, P. (May 2016). "Volition, Agency, Responsibility: Cognitive Mechanisms of Human Action". The Jean-Nicod Conferences, Paris Institute for Advanced Studies, France.

Haggard, P. (Apr. 2017). "Sense of Agency in the human brain". In: *Nature Reviews Neuroscience* 18, pp. 196–207. URL: https://doi.org/10.1038/nrn.2017.14.

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

Jamali, Mohammad, Mehdi Golshani, and Yousef Jamali (2019). "A proposed mechanism for mind-brain interaction using extended Bohmian quantum mechanics in Avicenna's monotheistic perspective". In: *Heliyon* 5.7, e02130. ISSN: 2405-8440. DOI: https://doi.org/10.1016/j.heliyon.2019.e02130. URL: https://www.sciencedirect.com/science/article/pii/S2405844019357901.

Khalighinejad, N. (2022). "Neural mechanisms underlying voluntary action control in the human brain". PhD thesis. Institute of Cognitive Neuroscience, University College London. URL: https://discovery.ucl.ac.uk/id/eprint/1539885/1/Khalighinejad_Nima_PhD%20thesis.pdf.

Kornhuber, H. H. and L. Deecke (1965). "Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale". In: *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, pp. 1–17.

Kornhuber, H. H. and L. Deecke (July 2016). "(English Translation) Brain potential changes in voluntary and passive movements in humans: readiness potential and reafferent potentials". In: *Pflugers Archiv: European journal of physiology* 468, pp. 1115–1124. URL: https://doi.org/10.1007/s00424-016-1852-3.

Krauledat, M. et al. (2004). "Improving speed and accuracy of brain-computer interfaces using readiness potential features". In: *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, pp. 4511–4515. URL: https://doi.org/10.1109/IEMBS.2004.1404253.

Leibniz, G. W. (1985). *Theodicy: Essays on the goodness of God, the freedom of man, and the origin of evil*. Trans. by E. M. Huggard. (Original work published 1710). Open Court.

Leibniz, G. W. (1991). *Monadology*. Trans. by N. Rescher. (Original work published 1714). University of Pittsburgh Press.

Lew, E. et al. (2012). "Detection of self-paced reaching movement intention from EEG signals". In: *Frontiers in Neuroengineering* 5. URL: https://doi.org/10.3389/fneng.2012.00013.

Libet, B., A. Freeman, and K. Sutherland, eds. (1999). *The Volitional Brain: towards a neuroscience of free will*. Imprint Academic, p. 298. ISBN: 0-907845-11-8. URL: http://dx.doi.org/10.1136/jnnp.70.5.709c.

Libet, B. et al. (1983). "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act". In: *Brain* 106.3, pp. 623–642. URL: https://doi.org/10.1093/brain/106.3.623.

Llinas, R. R. (2002). *I of the vortex: From neurons to self*. MIT press.

Schultze-Kraft, M. et al. (2016). "The point of no return in vetoing self-initiated movements". In: *Proceedings of the National Academy of Sciences of the United States of America* 113, pp. 1080–1085. URL: https://doi.org/10.1073/pnas.1513569112.

Schultze-Kraft, M. et al. (2021). "Suppress Me if You Can: Neurofeedback of the Readiness Potential". In: *eNeuro* 8. URL: https://doi.org/10.1523/ENEURO.0425-20.2020.

Schurger, A. et al. (2021). "What is the readiness potential?" In: *Trends in Cognitive Sciences* 25, pp. 558–570. URL: https://doi.org/10.1016/j.tics.2021.04.001.

Wen, W. and H. Imamizu (2022). "The Sense of Agency in perception, behaviour and human–machine interactions". In: *Nature Reviews Psychology* 1, pp. 211–222. URL: https://doi.org/10.1038/s44159-022-00030-6.

A. Implicit Methods for Measuring SoA

A recent review paper by (Wen and Imamizu, 2022) provides a comprehensive list of the primary methods used to implicitly measure the SoA. To provide a more comprehensive understanding, we will discuss three such methods in detail in the following sections.

A.0.1 Intentional Binding

To address the challenge of explicitly measuring the SoA (SoA), (Haggard, Clark, and Kalogeras, 2002) developed an innovative approach for implicitly assessing it. Participants observed a rotating clock hand on a screen while making judgments about the timing of their actions and the resulting effects. The experiment consisted of four blocks, with baseline conditions in which participants performed a voluntary action or heard a beep and judged the timing of either event. In operant conditions, participants performed voluntary key presses that produced a beep after 250ms, and they judged the timing of the key press or the subsequent tone.

Comparing the perceptual time of action and outcome in baseline and operant conditions revealed that the perceptual times of action and outcome tended to converge when participants' key presses resulted in a tone (operant condition). This phenomenon was named "Intentional binding." When voluntary movements were replaced by TMS-induced twitches, a 'repulsion' effect occurred, with an increased perceived interval between action and outcome. (Haggard, Clark, and Kalogeras, 2002) suggested that the brain mechanism responsible for action-outcome binding may be crucial for a normal SoA.

A.0.2 Sensory Attenuation

Sensory attenuation is a phenomenon suggests that individuals have an inherent understanding of their own actions, and that this understanding influences the perception of sensory feedback. A classic example of sensory attenuation is the reduced ticklishness people experience when they touch themselves, compared to when the touch is applied by someone else. This difference in perception is thought to be closely related to the SoA (Blakemore, Wolpert, and Frith, 1998).

To assess sensory attenuation, researchers often use tactile stimuli and a force-matching task, where participants feel a force (pressure) and reproduce an equivalent pressure on themselves (self-touch) or via a pressure device (external touch). Sensory attenuation can be calculated as the ratio or difference between reproduced forces in the external touch and self-touch conditions. This method provides an indirect measure of SoA, as self-produced effects are less "surprising," and the degree of surprise

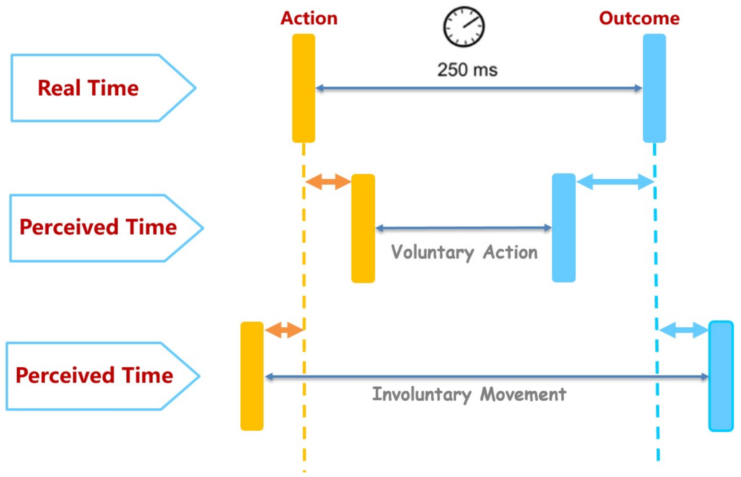

FIGURE A.1: Temporal schematic of intentional binding effect.

elicited by an effect can serve as an indicator of the SoA over that effect (Wen and Imamizu, 2022).

A.0.3 Visual Attention (Mainly from (Wen and Imamizu, 2022))

Measuring the SoA (SoA) implicitly through visual attention involves analyzing how attention is drawn to controllable objects or events. When control over an object or event is ambiguous, attention is automatically attracted to the controllable aspects. Studies show that reaction time to identify avatars whose motions match participants' body movements is shorter, highlighting the salience of SoA in allocating visual attention. Neural measures of attention, such as steady-state visual evoked potential (SSVEP), examine visual attention towards controllable and uncontrollable objects. Research demonstrates that control over objects automatically attracts visual attention. Attention to controllable objects also improves their visual detection, shortening detection time as control increases.

Control detection tasks are useful for investigating SoA in ambiguous environments for both healthy participants and those with impaired decision-making processes. SSVEPs are primarily modulated by top-down selective attention, while event-related potentials (ERPs) serve as indexes of early bottom-up attention. Larger ERP amplitudes occur when events are rare and salient, and they are suppressed when events are self-triggered by a participant (agency condition) compared to computer-triggered (non-agency condition). The attraction of attention to a loss or gain of SoA depends on whether control has already been established in an environment. When control is uncertain, detecting control attracts attention to controllable objects. Conversely, when SoA is expected, the absence of SoA is perceived as unexpected and attracts more attention than conditions with SoA.

B. Real-Time RP Detection Challenges

The real-time detection of Readiness Potential (RP) presents significant challenges due to the complexity and subtlety of brain signals, as well as the interpretive complexities associated with RP's nature. Accurately isolating and identifying the RP amidst ongoing neural activity demands precision and technological finesse, while the inherent variability within the data adds to the difficulty. Furthermore, differing scientific opinions about the nature of RP signal complicates the detection process. Addressing these challenges to enhance RP detection reliability is a fundamental step towards a deeper understanding of human volition.

Readiness Potential; It's Amplitude and Shape

As the RP signal is an order of magnitude weaker than the noise, usually averaging many (>30) trials has been necessary to reveal the RP and examine its properties. This makes its online detection problematic.

Literature review As (Abou Zeid and Chau, 2015; Abou Zeid, Rezazadeh Sereshkeh, and Chau, 2016) and (Lew et al., 2012) have highlighted, the detection of RP in online single trials is challenging due to its nature as a slow cortical potential (SCP) close to zero-frequency. The presence of the RP in single trials is often elusive, as it occurs concurrently with task-unrelated brain activity. This makes the RP typically invisible in single trials and difficult to identify. However, it has been demonstrated ((Garipelli, Chavarriaga, and R. Millan, 2011)) that narrow band filtering of EEG signals in the 0.1–1 Hz range can help to enhance the detection of the RP, making it more discernible despite its elusive presence in single trials.

Readiness Potential; False Positive

The challenge of false positives in RP detection affects the reliability and validity of the obtained results. Researchers attempt to mitigate this issue by employing refined feature extraction and optimized signal preprocessing techniques to enhance detection accuracy and deepen the understanding of voluntary actions. As a result, it remains uncertain whether the (averaged) RP merely represents a de-noised version of the signal present in each trial or an artifact of trial averaging. This uncertainty highlights the potential advantages of utilizing classification methods based on multi-channel signal input and assigning the entire data set to a state or class of pre-movement, as it appears to offer a more reliable approach compared to relying solely on the traditional readiness potential.

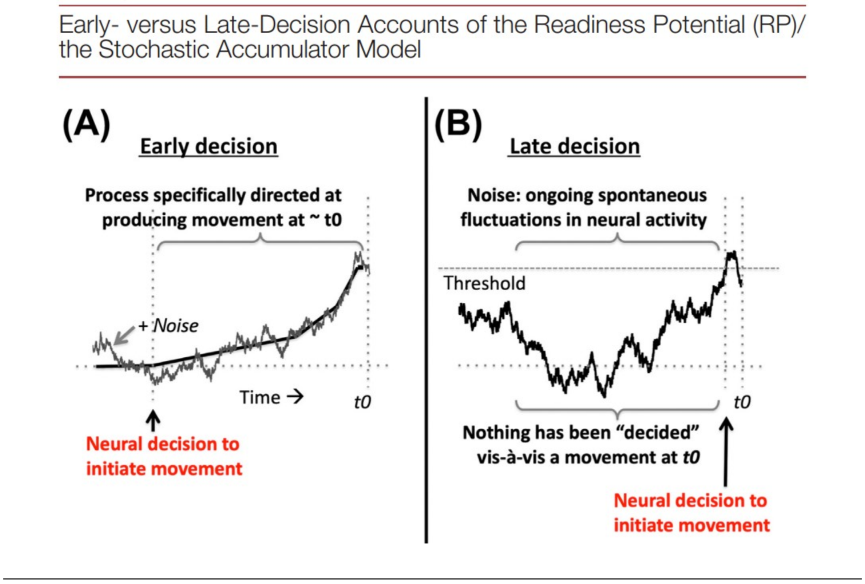

FIGURE B.1: RP Onset or Buildup: Deciphering Movement Causes (from: (Schurger et al., 2021))

Readiness Potential; Onset or Buildup?

An intriguing question to consider is: Should the onset of RP or the crossing of a threshold during its buildup be regarded as the primary cause for voluntary movement? This inquiry delves into the nuances of the neural processes underlying voluntary actions and could offer valuable insights into the mechanisms governing our movements.

Readiness Potential; Eye movement Artefact

A significant challenge in detecting RP signals within BCI systems is the interference caused by eye movements. In particular, horizontal eye movements can produce lateralized low-frequency shifts in EEG channels that closely resemble lateralized readiness potentials (LRPs, which are a type of readiness potential that can be observed in preparation for a lateralized movement), which may substantially impact the detection of low-amplitude RP signals. This issue raises concerns that the BCI classifier could be vulnerable to EOG artifacts or even be controlled by left versus right eye movements ((Krauledat et al., 2004)). To counteract these effects, researchers typically implement computational compensatory preprocessing techniques and provide participants with behavioral instructions, such as maintaining fixation during the experimental setup. In the following section, we will conduct a literature review on this topic to gain a deeper understanding of the challenges and potential solutions related to eye movement artifacts in RP detection.

C. Strategies for Optimal Classification in Real-Time BCI

Classifying 'Idle' or 'No Intention' and 'Pre-movement' or 'Intention' states in a real-time BCI setup demands the implementation of strategic approaches to ensure the classification process's accuracy and robustness. Since intention detection via classification fundamentally involves RP signals, the strategies are often similar to those outlined in Section **??**, Real-Time RP Detection Challenges.

C.0.1 Gaze-Stabilized Trial Initiation

To reduce the influence of eye movement artifacts on intention detection, presenting a central fixation point at the start of each trial is often employed. As explored in section 2.1.2, this practice has been implemented in several studies including (Lew et al., 2012) and (Schultze-Kraft et al., 2016) among others. This ensures that eye movement artifacts that can significantly affect classification accuracy are minimized.

C.0.2 Defining a Feature Extraction Temporal Window

The establishment of a precise sliding time window is crucial for continuously analyzing neural activity over time. The update rate of this window, as shown in (Schultze-Kraft et al., 2016) with a 10 ms resolution, must balance the classifier's precision, computational capacity, and data accessibility. This technique enables continuous monitoring and classification of neural signals.

C.0.3 Classifier-Driven Action Inhibition Period

Several studies have emphasized the significance of an action inhibition period in their experimental designs. This period typically ranges from 1 to 2 seconds, depending on the feature extraction temporal window, and provides the classifier sufficient time to extract features and accurately detect the states.

For instance, (Schultze-Kraft et al., 2016) maintained a 2-second gap between the start of the trial and the onset of movement. Participants were instructed to wait about 2 seconds before pressing a button with their right foot after a "go signal" appeared on the screen. Similarly, in (Schultze-Kraft et al., 2021), participants were instructed to wait approximately 2 seconds before pressing a pedal after the start of a trial, signaled by a white circle appearing on the screen. (Lew et al., 2012) instructed subjects to initiate the movement whenever they wished, but not before 2 seconds after the presentation of an auditory cue. (Krauledat et al., 2004) instructed subjects to press keys with their index and little fingers in a deliberate order and at a pace of approximately 2 seconds. This action inhibition period ensures the classifier can effectively analyze the data and differentiate between the movement and idle states, contributing to the overall performance of the intention detection algorithm.

By implementing these strategies, we aim to ensure reliable and accurate classification between 'Idle' or 'No Intention' and 'Pre-movement' or 'Intention' states, thereby enhancing the performance and success rate of real-time BCI experiments.