NEURAL DYNAMICS UNDERLYING PERCEPTUAL DECISION-MAKING: AN EEG STUDY USING BIOLOGICAL MOTION DIRECTION DISCRIMINATION TASK

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

NEUROSCIENCE

By

Osman Çağrı Oğuz

July 2025

Neural Dynamics Underlying Perceptual Decision-Making: An EEG Study Using Biological Motion Direction Discrimination Task
By Osman Çağrı Oğuz
July 2025

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Burcu Ayşen Ürgen(Advisor)

Zahide Pamir Karatok

Didem Kadıhasanoğlu

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

NEURAL DYNAMICS UNDERLYING PERCEPTUAL DECISION-MAKING: AN EEG STUDY USING BIOLOGICAL MOTION DIRECTION DISCRIMINATION TASK

Osman Çağrı Oğuz
M.S. in Neuroscience
Advisor: Burcu Ayşen Ürgen
July 2025

Perceptual decisions constitute the core of most interactions with the environment and are modeled with the basic assumption that noisy evidence is accumulated over time until a certain bound is reached. The centro-parietal positivity (CPP), has been shown to reflect accumulation-to-bound dynamics in human EEG studies. Although CPP has been validated in various tasks, biological motion perception has not been investigated thoroughly in the context of perceptual decision-making. Here we employed a design similar to widely employed random-dot motion tasks, where the stimuli were presented in a continuous visual stream. In particular, scrambled motion control stimuli were transitioned into target point-light displays of human walkers in a two-alternative forced choice task of reporting walking direction. It has been shown that CPP displays characteristics expected by the hidden decision variable occurring through evidence accumulation dynamics, including a build-up rate scaling with the evidence strength and reaching a common level near decision reports. Moreover, CPP differentiates between correct and invalidly responded trials, shows predictive qualities for reaction times, and covaries with the drift rate parameter estimates after drift-diffusion model fitting. In addition, investigation of the pre-stimulus attention levels and biological-motion specific signals revealed that pre-target attention might exert its influence on the decision-making process with the enhanced quality of higher-level visual encoding following the target stimulus. As a result, our study extended the previous findings on CPP into the biological motion literature and provided an overview of the neural dynamics for perceptual decisions.

Keywords: Perceptual decision-making, drift-diffusion models, centro-parietal positivity, biological motion perception.

ÖZET

ALGISAL KARAR VERME SÜREÇLERİNİN ALTINDA YATAN NÖRAL DİNAMİKLER: BİYOLOJİK HAREKET YÖN AYRIMI DENEYİYLE BİR EEG ÇALIŞMASI

Osman Çağrı Oğuz
Nörobilim, Yüksek Lisans
Tez Danışmanı: Burcu Ayşen Ürgen
Temmuz 2025

Algısal kararlar çevreyle olan etkileşimlerimizin temelinde yer alır. Bu karar süreçleri, temel varsayımı gürültülü duyusal kanıtın belirli seviyeye kadar biriktirilmesi olan modeller aracılığıyla incelenmektedir. Algısal kararların altında yatan dinamikler insanlarda da EEG aracılığıyla gözlemlenmiş ve bu olaya ilişkin potansiyel centro-parietal pozitiflik (CPP) olarak isimlendirilmiştir. Bu varsayımlar ve CPP çeşitli deneylerle doğrulansa da biyolojik hareket algısı bağlamında henüz geniş kapsamlı incelenmemiştir. Bu çalışmada, ekranda süregelen bir görsel uyaranın olduğu ve hedef uyaranlara kesintisiz bir şekilde geçiş yapıldığı bir tasarım kullanılmıştır. Bu deney dizaynı, yaygın olarak kullanılan rastgele nokta hareketi deneylerinden esinlenilmiştir. Bu bağlamda, karıştırılmış hareketi takip eden nokta-ışık uyaran şeklinde insan hareketi gösterimleri kullanılmıştır. Katılımcıların iki alternatifli zorunlu seçim görevinde yürüme yönünü tayin etmeleri beklenmiştir. CPP'nin kanıt birikiminden ortaya çıkan karar değişkenini yansıttığı, duyusal kanıtın gücüyle ölçeklenen bir artış oranı ve karar bildirme zamanına yakın bir ortak genlik seviyesine ulaşmasıyla gösterilmiştir. Ek olarak, CPP deney içerisindeki doğru ve geçersiz cevaplar arasında ayrım göstermiş, karar süresini tahmin edebilmiş ve difüzyon-sürüklenme modeli çıktılarından sürüklenme oranı parametresiyle pozitif yönde ilişkilendirilmiştir. Ayrıca, hedef uyaran öncesi dikkat seviyeleri ve biyolojik hareket algısına özgü sinyaller incelenmiştir. Hedef öncesi dikkatin, takip eden gelişmiş görsel işleme arabulucuğuyla karar verme süreci üzerindeki etkisi görülmüştür. Sonuç olarak, çalışmamız CPP hakkındaki önceki bulguları biyolojik hareket literatürüne genişletmiş ve algısal kararların altında yatan farklı dinamikleri ortaya çıkarmıştır.

Anahtar sözcükler: Algısal karar verme, difüzyon sürüklenme modeli, centro-parietal pozitiflik, biyolojik hareket algısı.

Acknowledgement

I would especially like to thank my advisor, Prof. Burcu Ayşen Ürgen, for the guidance and ongoing faith in me during both my undergraduate and graduate education. I would also like to express my gratitude for being a member of CCNLAB with amazing people in such a safe and supportive environment, and I am thankful to Burcu hocam for this opportunity. I am looking forward to being a collaborator with her in the future. I thank Berfin Aydin for experimental coding and data collection, and Şeyda Evsen for her support and insightful comments on the data analysis.

I would also like to express my gratitude to my thesis committee members, Dr. Zahide Pamir Karatok and Dr. Didem Kadıhasanoğlu, for their dedicated time and constructive feedback. Furthermore, I am thankful to The Scientific and Technological Research Council of Turkey (TÜBİTAK) for the BIDEB scholarship.

Finally, I would like to acknowledge how lucky I am to have family members and friends who believe in me more than I believe in myself, who are there for me when I fall, and who renew my faith that I can overcome any challenge. I am especially grateful to my mother, Canan Oğuz, for all the hard work she put into our education, and to my father, Bülent Oğuz, for all the opportunities they have provided me. I would also like to thank my siblings for their emotional support and my dear friend, Ilgaz Şenyüz, for all the hours we spent together working on our projects. Last but not least, I also want to thank my family in Ankara, Nil Eymen Yılmaz, for patiently listening to and supporting me, for our lovebird Cacık and our cat Kopernik, for bringing joy to our lives.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Perceptual Decision-Making

Perceptual decisions have key importance for most interactions with the environment. Its functional significance includes the detection and interpretation of objects and navigation in space. Imagine a common everyday situation, seeing a pedestrian approaching you while walking on a sidewalk. In order to avoid collision, you must decide how fast and in which direction to move based on the sensory information coming from various sources, such as motion trajectory, body posture, or even emotional expression in potentially threatening situations. In case the view is obscured by the surrounding irrelevant objects or if you have limited attentional resources at the time, reaching perceptual decisions becomes more challenging. This example illustrates the importance of investigating perceptual decision making as the brain should quickly perceive and evaluate complex, noisy sensory information.

To understand how the brain transforms noisy sensory input into evidence for decision alternatives, researchers developed computational models under the umbrella term of sequential sampling models (SSMs). These models share the

basic assumption, known as accumulation-to-bound [1], and view the sensory evidence as being noisy and variable. To reach reliable decisions, the optimal system should reduce the variability resulting from the noisy inputs. In this respect, at the core of these models lies the idea that sensory evidence is gradually integrated over time until a certain threshold is reached. Once this boundary is crossed, a choice is made, and a response that conveys the decision report is initiated. The integration view has been challenged with the claims that non-integration strategies may also give rise to similar predictions regarding the behavioral patterns. However, the opposing views are addressed by devising different variants of the random-dot-motion (RDM) task, which is the most common paradigm for studying simple perceptual decisions, to disentangle the proposed alternatives. These displays accompany moving clouds of dots in which some manipulated proportion is moving in the same direction (i.e., signal dots) while the rest is moving randomly (i.e., noise dots). Therefore, RDMs come with ease of difficulty manipulation by changing the signal-to-noise ratio within the displays. As a result, non-integration strategies are falsified and further support for the integration strategy is provided [2]. Applications of SSMs are not limited to simple perceptual decisions. They are extended across a range of domains, including object categorization [3], lexical decision-making [4], value-based choices [5], and social decisions [6].

More excitingly, the expectations from these models regarding the evolving decision variable have been extensively supported by neurophysiological findings [7]. It has been shown by the monkey studies that single neuron firing rates in task-relevant brain areas that control saccades (e.g., frontal eye field - FEF) may show the very pattern of rising to threshold [8]. Other studies with rhesus monkeys also demonstrated dynamics predicted by integration to threshold models in interconnected areas. Specifically, neurons in the lateral intraparietal area (LIP) [9] and superior colliculus [10] show a steeper increase in spike rates with higher sensory evidence while reaching a common level during response regardless of sensory evidence quality. In humans, the task of finding such neural correlates of the decision variable (DV) unfolding temporally is more challenging

due to constraints related to the use of non-invasive techniques with lower resolution. In this respect, several attempts have been successful through careful experimentation to isolate decision signals. Electroencephalography (EEG), with its high temporal precision to capture the timely dynamics, became the natural candidate. Particularly, researchers designed experiments in which the target RDM display is subtly transitioned from a continuous stream of visual stimulation to prevent the large bursts of visual activity from occipital areas [11]. This approach has been fruitful in finding a signal showing properties of an abstract decision variable. As such, relevant event-related potential (ERP) is found and has been named as Centro-Parietal Positivity (CPP). This ERP is found over centro-parietal areas, beginning approximately 300ms following the stimulus onset as a positive-going potential while its rising slope scales with sensory evidence and reaches a common bounded level at response in various difficulty conditions [12]. It has also been found that CPP does not rely on specific sensory modalities, is motor-independent by showing ramping activity even when no decision report is required, and can predict RT. Moreover, it differentiates incorrect from correct trials and can even be manipulated in the middle of the trial by changing the evidence for decision alternatives [11] [13]. Some researchers also found that CPP can differentiate between the response alternatives by showing amplitude modulation in memory or perceptual tasks [14] and even reflect subjective evidence and confidence levels using a sequential vibrotactile comparison task on stimulus frequencies [15]. CPP's scalp topography and temporal qualities are similar to those of a well-established decision-related signature, P300. Some studies endeavored to integrate these two lines of research and found that P300 can reliably account for the variations in reaction time (RT) and scales with evidence strength, in the direction that higher slopes are associated with faster reports and easier conditions. As for the peak amplitude, the differentiation with respect to different RT bins has not been observed, which is interpreted as reflecting the common action-triggering threshold [16]. However, the CPP time window may or may not coincide with those of P300's established temporal qualities (occurring around $300-600$ms with respect to stimulus onset), depending on the tasks with prolonged stimulus viewing conditions. In those cases, CPP is analyzed according to response-locked waveforms [17].

1.2 Modeling Perceptual Decisions

The most widely employed model within the family of SSMs is the drift-diffusion model (DDM) [18]. This model maps the decision process to functionally and cognitively different underpinnings while accounting for behavioral data. Relying on a single measure of RT and accuracy may sometimes lead to misleading conclusions. For example, a participant with fast responses may appear superior based on RT alone, even if those responses reflect impulsivity and accompany low accuracy. On the other hand, a participant with a high number of correct responses may appear superior based on considering accuracy alone, despite adopting an overly cautious strategy leading to slower decisions (i.e. longer RTs). By investigating the two core behavioral indices together under a unified process of evidence accumulation (EA), DDM enables a more nuanced understanding of the phenomena related to decision strategies, such as speed-accuracy trade-offs [19]. For the two-alternative forced choice tasks, the core assumptions of the DDM include stochastic evidence accumulation reflecting internal and external noise in the decision process. If noise outweighs the signal in some trials due to its stochastic nature, the end result will be reaching the bound for the incorrect response alternative. More specific assumptions of DDM include that integration takes place in time continuously, as an extended version of the discrete random walk process [18]. Also, DDM postulates a single accumulator, meaning that the evidence favoring one option is automatically conceptualized as evidence opposing the other [20].

DDM has four essential parameters to jointly account for RT and accuracy, each having cognitive interpretability and variability in some settings more than others. *Drift rate (v)* reflects the average rate of evidence accumulation and the difficulty manipulations target this variable, with higher drift rates observed in easier conditions. It has been suggested that if the experimental manipulations were targeting the difficulty of the trial, the corresponding changes in RT distributions should be captured by different drift rates rather than adjustments in other main parameters, with the exception that the participants knew the upcoming trial's difficulty beforehand [21]. The decision process terminates when it crosses

boundary separation (a), which determines the amount of evidence required before committing to one of the alternatives. The speed or accuracy instructions are reflected in the narrowing or widening of the thresholds, respectively. Evidence accumulation begins from *starting point (z)*, which represents the preliminary bias between the two alternatives. In case of differently valued response alternatives or when the stimulus probability is manipulated, the starting point lies closer to the bound representing the favored option. However, when the bounds are determined as accuracy-coding (upper bound is correct and lower bound is incorrect), the models do not include a free starting point parameter and instead fix it as the halfway between the bounds [22], [3]. Lastly, *non-decision time (t0)* reflects the sum of sensory encoding and motor delays, which are the processes assumed prior to the evidence accumulation starts and after boundaries are crossed, respectively. It has been shown that especially these central parameters of DDM have high test-retest reliability, which further encourages the usage as a diagnostic tool [23]. Moreover, the model's main parameters have been shown to generalize and have powerful cognitive interpretation across domains, including simple perceptual decisions, value-based or lexical decisions, memory experiments, and even clinical applications [19].

In addition to the essential parameters, there is a diffusion coefficient that reflects within-trial variability in drift rate, although this term is often fixed to a constant value for estimation of other parameters. The core model is further extended by adding between-trial variabilities in drift-rate, starting point [24], and non-decision time [25], which is sometimes named as the Full Diffusion Model [19]. However, these random between-trial variabilities are cautiously treated and challenged by the notion that they lead to model overcomplexity without offering interpretable outputs [26], or detrimental to parameter recoveries for cognitively most interesting ones [27]. Despite the established theoretical foundation with supportive findings and mathematical tractability for static main parameters throughout the trial, some researchers claim that parameters change in time [28]. For instance, the time-dependent drift rate stems from the idea of a gain for sensory processing that is reflected in elevated drift rates as processing time increases [29]. Another influential perspective was on the adjustments

regarding decision criteria, which challenges the traditional fixed boundary view by suggesting a collapsing bound phenomenon [30]. This mechanism leads to flexible criteria for the required amount of evidence, and in turn, gives rise to timely decision reports. However, these approaches have not found comprehensive support from modeling human data despite support from monkey studies [20]. Yet, if the decisions are needed to be made within the imposed short deadlines, there is evidence supporting urgency-gating [31], where the accumulated evidence is manipulated by the urgency signal that increases in a time-dependent manner to prompt responses [32]. Furthermore, although DDM assumes discrete stages between visual encoding, evidence accumulation, and motor response selection, this traditional view is sometimes challenged by different formulations. Some influential theories suggest that there is a leakage between perceptual or decision systems and the motor system [33]. As for the visual encoding stage in non-decision times, there may be at least some overlapping time windows where both encoding and accumulation of evidence take place rather than being fully discrete [34]. Keeping all these challenging views in mind, DDM still provides a powerful tool to draw inferences for using estimated parameters for psychometric evaluations rather than theory development due to its established success in capturing behavioral data with cognitive plausibility [20].

In addition to being powerful in algorithmic and computational layers of processing, the underlying processes should be biologically realizable within the brain to reach reliable inferences. Connecting with Marr's level of analysis [35], without demonstrating *how* the system might realize the underlying and mathematically proven instantiations will provide only a limited part of the whole picture. Therefore, it is crucial to support DDM with neurophysiological findings. Indeed, previous research has provided reliable counterparts of the DDM processing stages from the brain. It has been shown using both perceptual and memory tasks that when the single-trial data is partitioned into two groups according to rising slope of CPP, DDM returned lower drift rates in low CPP slope than high CPP slope trials [14]. Another influential perspective is to estimate the model parameters, then to connect with the neural measures. In this respect, it has been reported that in easier trials potentially accompanying cautionary strategy,

both elevated boundary separation estimates and CPP amplitudes at response is observed [36]. More precisely, neural signatures can be used as a validation for the hidden dynamics predicted by the DDM. As such, using complex tasks where the evidence is interrupted or changes during the course of the trial, CPP showed close agreement with the simulated decision variable resulting from the parameter estimates of DDM [37]. Additionally, CPP waveforms are found to closely follow the simulated decision variable until response, and even the dynamics of post-response evolution of DV when it is formulated by gradual decay back to baseline [16]. In case of differently valued response alternatives, some researchers also found drift rate bias to favor the highly valued alternative. As a neural correlate of this effect, CPP showed the turnaround effect, characterized by an initial rise towards the high-valued bound but later turning to the correct but low-valued region in decision-variable space [13]. Regarding other DDM components, N200 peak latencies are used successfully to mark the onset of accumulation following the visual encoding stage that sequentially precedes it [38]. As for the decision boundary parameter, some researchers informed the parameter space with the contralateral motor-related neural markers at decision time [39]. In addition, the starting point parameter has been associated with motor oscillations preceding target onset, which is thought to reflect perceptual expectations [40]. These studies collectively suggested that DDM parameters not only have cognitive interpretability but also neural plausibility.

Lastly, another important decision is the fitting method employed for DDM. Among the approaches for fitting DDM to behavioral data, the Chi-Square (CS) based fitting is one of the most recognized methods. CS-based approach requires RT to be separated into several bins rather than modeling the full distribution. The CS-based approach requires less time to fit the data, has demonstrated robustness to outlier RTs and reliable parameter estimates if the experiment includes a sufficiently large number of trials [23]. Another widely employed fitting method is based on maximum-likelihood (ML-based) and proven to be more robust in terms of parameter estimation and recovery with a relatively lower number of trials compared to the CS-based approach. However, the pre-processing step to

remove outliers should be carefully done [25] if one employs an ML-based procedure. On the other hand, as with the recent increases in available computational resources, researchers were attracted by the hierarchical Bayesian approaches to model fitting [41]. In this methodology, low trial numbers are effectively handled with the parameter regularization incurred by the group-level hyperparameters. However, Bayesian approach comes with another challenges regarding the time required for the fitting procedure and the interpretability of the output because of returning full posterior distribution rather than single estimates. Because each fitting procedure having their costs and benefits, the final decision for which methodology to be employed should be considered in tandem with the computational resources and data quality [42].

1.3 Biological Motion Perception

The most general description of biological motion (BM) refers to the movement patterns of living entities. However, in the relevant literature, the term biological motion perception typically denotes the human ability to perceive and interpret human movement patterns [43]. This remarkable perceptual capacity constitutes the core of social interactions and is regarded as the hallmark of social cognition [44]. Given its critical role in interpreting intentions, detecting threats, coordinating actions, maintaining social relationships, and resolving social conflicts, it is unsurprising that humans exhibit exceptional sensitivity to BM — an ability likely shaped by evolutionary pressures to survive in complex social environments (see [45], [46], for reviews). There are several key brain areas found to be supporting BM perception, including premotor cortex, potentially due to its role in action understanding through multisensory neurons [47] and posterior superior temporal sulcus (pSTS), especially in the right hemisphere [48], but see [49]. The area pSTS has key importance due to demonstrating reliable activation not only for whole-body motion displays, but also with respect to motion sequences of specific parts of the body and even the sounds of footsteps [46].

One of the most influential tools in BM research is the point-light display

(PLD) paradigm introduced by Gunnar Johansson [50]. PLDs consist of a limited number of dots attached to the major joints of the human body, effectively eliminating visual cues related to form (texture, color, facial features, or attractiveness). Despite reduced detail, PLDs convey sufficient information to evoke a vivid perception of a human figure, particularly when the dots are seen in motion. Research shows that observers perceive BM from PLDs in as little as 150 milliseconds, highlighting the efficiency and automaticity of BM processing [45]. PLDs have been used to study various perceptual and cognitive processes, including action recognition, emotion detection [51], and even intention inference [52]. Importantly, studies employing PLDs have revealed selective impairments in populations with social cognition deficits. For example, individuals with autism [53], schizophrenia [54], or broader social processing difficulties [55] often show impaired performance specific to recognizing human PLDs, further emphasizing the relevance of PLDs in both basic and clinical research. The power of PLDs also attracts researchers in related fields, such as facial expression perception, where the created point-light stimuli reflect facial kinematics [56].

To examine the temporal characteristics of the neural activity in response to BM displays, EEG provides a powerful tool. Activities in response to PLDs as early as around 150ms [57] have been observed, though some researchers suggest that early components might reflect general integration of form and motion rather than BM-specific processes [58]. However, a relatively late negativity around 300- 600 ms time windows reported in posterio-occipital regions is also observed, which is thought to reflect the semantic processing specific to biological motion for higher-level action understanding. As such, the peak amplitude observed for this component is modulated by the attention devoted to motion rather than another task, and also by biological motion point-light displays compared to their scrambled counterparts [59], [60]. Additionally, some studies reported hemispheric selectivity for the component of interest, as the effect is more pronounced in right hemispheric regions [58], such as electrodes PO8, P6, and P8. The source localization endeavors revealed that the likely generators for this negativity are biological-motion specific posterior superior temporal sulcus [61], or the known regions for general motion processing, such as right superior temporal gyrus and

V5 [57].

PLDs also offer flexibility for experimental control with ease of manipulation. These manipulations include spatial scrambling to distort configural information [62], distortion of configuration by creating upright or inverse walkers [63], and even temporally distorting the phase [64]. Different techniques have been useful in analyzing the relevant contributions of form and motion information embedded in PLDs that underlie the remarkable perceptual ability of BM perception. Across different manipulations, scrambled motion is one of the most employed and effective types of noise mask [65]. This mask is created by randomizing the starting points of the dots, but keeping the same motion trajectory as that of the ones depicting human joints. Connecting with the idea of preventing large visual evoked potentials that potentially attenuate decision-related neural markers, it would be ideal to use scrambled motion stimuli prior to meaningful biological motion displays. It is therefore convenient to adjust the perceptual decision-making paradigms utilizing RDMs to biological motion tasks involving PLDs. Building on this, biological motion stimuli provide a useful framework to investigate whether the properties of a domain-general abstract decision variable and its underlying temporal mechanisms generalize to more socially relevant tasks and settings.

1.4 Current Study

The current study incorporates PLDs depicting human walkers to investigate perceptual decision-making using EEG. In this way, the purpose of this study is to extend the previous literature on decision-making by introducing socially meaningful stimuli. The design has been inspired by one of the pioneering studies on CPP by Kelly & O'Connell [12], in which the target RDM stimuli with varying difficulty are presented in a continuous visual stream. In particular, in the current study, scrambled motion displays are utilized before the biological motion targets as a control stimulus to ensure an ongoing visual stream and are followed by PLD walkers in different levels of noise to manipulate task difficulty. The participants are tasked with reporting the walking direction of the PLDs. One may argue that

this design requires detection of human form prior to the evidence accumulation process for direction discrimination, which is not generally common in perceptual decision-making studies [66]. However, seamless transition from scrambled to biological motion displays is utilized to prevent visually evoked potentials from contaminating the decision-related signal.

In our recent paper [67], using the same data, we have demonstrated that the latency of the CPP peak amplitude is delayed in the BM direction discrimination task compared to the previous findings. Yet, the characteristics expected from a decision signal were replicated by demonstrating a rising activity until the decision report. Despite the observed patterns, the expected qualities of an abstract decision variable still need to be thoroughly investigated*ü, such as the relationship of CPP with RT or whether CPP can differentiate between correct and incorrect trials. Furthermore, in this study, the behavioral data is fitted to DDM to better understand the relationship between CPP slope and estimated drift rates. As for the task-specific neural markers, this study also incorporates measures of biological-motion related EEG signatures (e.g., posterior-occipital negativity in relatively late time windows), pre-stimulus attentional markers (e.g., posterior-occipital alpha band spectral amplitude), and motor preparation levels before the stimulus presentation (e.g., mu-beta band spectral amplitude). It is also aimed to uncover the relationships between these neural markers to provide a more complete picture of how perceptual decisions are made.

We believe that it is promising to introduce BM perception in perceptual-decision-making literature, with such an experimental design allowing the investigation of a neural representation of an abstract decision variable. As will be detailed in the discussion, both fields have the potential to inform and benefit from each other. That is, perceptual decision-making literature can be informed by examining task-related signatures. As for the BM perception, the relative contributions from different information sources (e.g., form and motion) can be better disentangled by temporally inspecting the neural representation of the accumulated decision variable.

Chapter 2

Methods

Before explaining the methodological details, it should be noted that the data from this study was used for the undergraduate thesis [68] and our previously published paper last year [67].

2.1 Participants

EEG and behavioral data are collected from 16 participants (8 female; age range: 18–40 years; mean age: 23.18) with no history of neurological or psychiatric disorders. All participants had normal or corrected-to-normal vision and were not taking any psychiatric medication at the time of data collection. The study was approved by the Human Ethics Board of Bilkent University. Informed consent was obtained from all participants after they were briefed on the experimental procedures. To minimize potential confounds on EEG and behavioral measures, participants were required to have adequate sleep the night before the session, which was ensured by a pre-screening questionnaire form.

2.2 Task and Procedure

Each trial began with a 2-second fixation period during which participants were instructed to focus on the center of the screen. The target stimuli were point light biological motion displays, a widely used paradigm to convey human form and motion dynamics. Each display consisted of 13 dots representing the major joints and the head, overlaid with a variable number of randomly moving noise dots. The task difficulty was manipulated by introducing four levels of noise. Each target display was preceded by a scrambled motion display in which the spatial configuration of the 13 joint dots was randomized, eliminating the perception of biological form. Scrambled displays lasted for 3, 5, or 8 seconds to prevent the predictability of target onset (see Fig. 2.1). Transitions from scrambled to coherent motion were implemented seamlessly to minimize transient visual evoked potentials that could interfere with EEG signals of interest, following best practices outlined by Kelly & O'Connell (2013) [12].

The task is to report the walking direction of a PLD walker depicted in a sagittal view of a human figure as if walking on a treadmill, moving either leftward or rightward. The stimulus remained on screen until the participant responded, but for a maximum of 2 seconds. Each gait cycle spans 700 ms; hence, up to three full gait cycles could be presented if participants used the entire response window. Participants were instructed to report the walker's direction as soon as they were confident, using designated keys on the keyboard with their corresponding hand (i.e., reporting rightward-moving walker with the index finger of the right hand and vice versa).

Participants completed 480 trials in total, presented in random order. Walker direction, noise level, and pre-stimulus duration were fully counterbalanced across trials. Before the main experiment, participants underwent a practice session to familiarize themselves with the task. All stimuli were displayed on a screen, capturing 6.5 degrees of visual angle, viewed from a distance of 57 cm in a light-attenuated room.

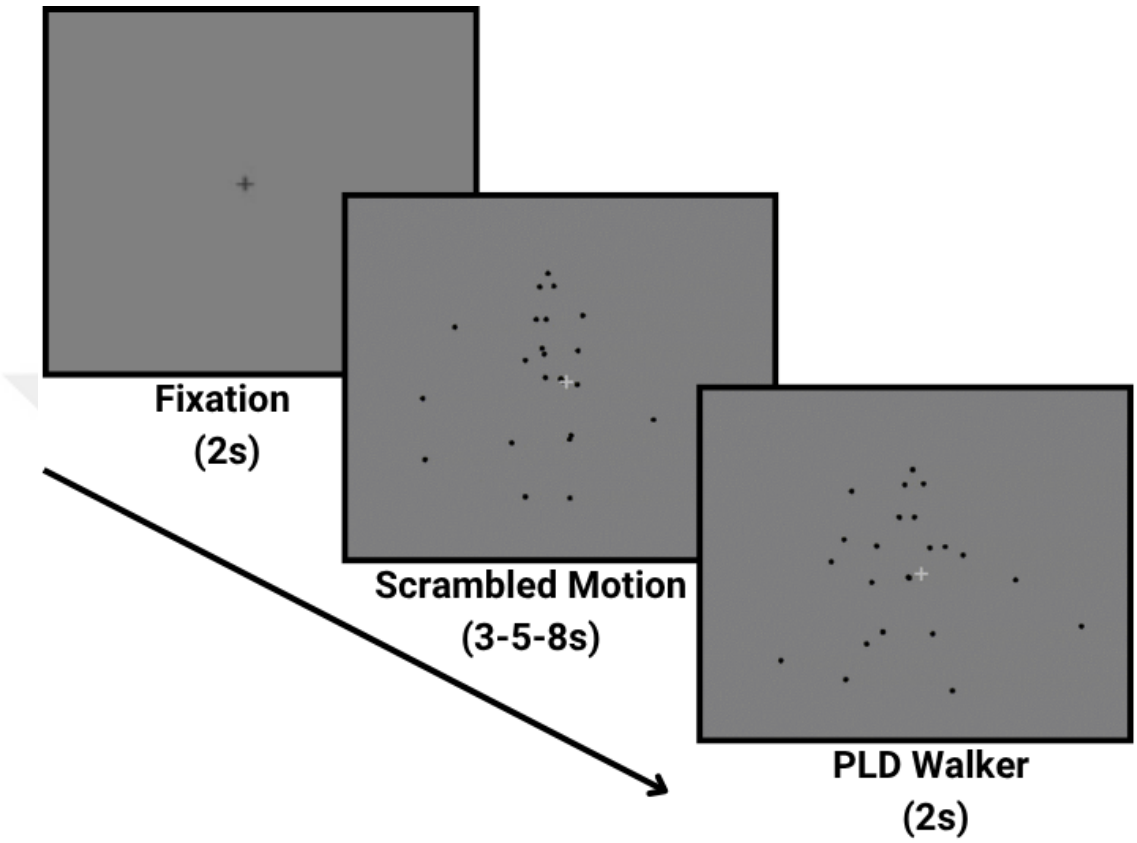

Figure 2.1: An example trial with a walking direction of right and noise level of 10. Each trial began with a 2-second fixation period, followed by varying durations of scrambled motion displays, then seamlessly transitioned to point-light display walkers that were displayed on the screen until response and a maximum of 2 seconds.

2.3 Behavioral Analysis

2.3.1 Reaction Time (RT)

For the reaction time (RT) analysis, independent variables are experimental manipulations for the duration of scrambled motion before the stimulus (hereafter named as pre-stimulus duration, accompanying 3 levels - 3, 5, or 8 seconds) and the number of added noisy dots to target walker stimuli (hereafter named as noise level, accompanying 4 levels from easiest to most difficult - 10, 20, 35, 55). The combination of two manipulations creates 12 different conditions for which

subject-wise mean RTs are calculated. Thus, the resulting data has 192 rows, with 4 columns: subject, noise level, pre-stimulus duration, and mean RT. Repeated measures ANOVA is employed to investigate each experimental manipulation's effect, while noise level and pre-stimulus duration are within-subject factors. Assumptions related to normality, sphericity, and residuals are also checked and corrected for if necessary. In case of significant main effects or interaction, a post-hoc test is also employed to reveal pair-wise differences between the conditions. The main effect of the noise level is expected to be seen in such a way that RT decreases in lower noise conditions, reflecting its relationship with the difficulty of the task. Additionally, the main effect of pre-stimulus duration is expected with the pattern of RT decrease in higher pre-stimulus durations due to increased preparedness for the upcoming target.

To precisely estimate the relationship of single-trial RT with single-trial neural measures, two different procedures are followed. Firstly, based on each participant's single-trial RTs, fast and slow decisions were quantified with a median split within each noise level. Incorporating the neural markers detailed below as dependent variables, separate two-way repeated-measures ANOVAs with noise level and RT category as within-subject factors were employed. As for the neural markers measured prior to stimulus onset, a median split procedure is additionally done within each subject and each pre-stimulus duration. This is because before the target onset, it is only the pre-stimulus duration that is explicitly manipulated, rather than the task difficulty via noise level. Therefore, additional repeated-measures ANOVAs are conducted where the neural measures before the target onset as the dependent variable, and the within-subject factor of RT category is determined based on the other within-subject factor, pre-stimulus duration.

In case of significant main effects for RT category, a further test is conducted to more precisely identify the relationship between neural markers of interest and RT regardless of the experimental manipulations. That is, the single-trial data from each participant was z-scored and then aggregated across subjects. This step ensures the removal of the potential differences across participants. For instance, one participant might follow an overly cautious strategy throughout the experiment that hinders the statistical effects from being observed as a result

of their data differing appreciably from others. In this aggregated and z-scored RT data, it is now feasible to categorize more precisely due to the increased number of trials. Hence, RT is divided into 6 bins within each experimental manipulation of interest rather than simply 2 as in the previous analysis. Then, the neural markers are regressed against 6 RT bins separately for each level of the experimental factor. To ensure that any significant effects are not due to chance, a permutation test with 10.000 replications is performed, where in each of the replications, trials were randomly assigned to one of the RT bins.

2.3.2 Accuracy

To investigate accuracy in relation to the experimental manipulations, the major imbalance between the two main response categories (correct: 6347, incorrect: 352) resulted in the failure to account for incorrect responses when a binomial generalized linear model with a logit link is employed. More specifically, all the model predictions were 'correct' responses. To address this, a Chi-square test was utilized, which also offered the flexibility to include misses (no response during the target presentation) and premature responses (responses given before the target stimulus) as additional response categories. It is noteworthy to include misses and premature responses, since errors in this particular design result not only from incorrect responses. In particular, the lack of anticipation about the timing of the upcoming trial resulted in premature responses, while the problems with detecting seamless transitions led to misses. As a result, 0 was encoded as incorrect (N = 352), 1 as correct (N = 6347), 2 as miss (N = 857), and 3 as premature (N = 98) responses. This way, the contingency tables are created. There is also an R package that implements Chi-square post-hoc testing, which was used in the case of significant main effects [69]. For the difficulty manipulation, it is expected to see significantly more correct responses in lower noise conditions, while more misses and incorrect responses are expected in higher noise conditions. As for the pre-stimulus duration, lower waiting intervals accompanying significantly more misses are expected. This is due to unanticipated target stimulus transitions occurring relatively early that might prevent the detection of a PLD walker. On the other

hand, significantly more premature responses are expected following longer pre-stimulus durations due to false alarms caused by the increased preparedness and potentially growing urgency to respond as soon as possible.

2.3.3 Drift Diffusion Model Fitting

Drift Diffusion Model fitting is done with the fast-dm-30 software developed by Voss and colleagues [70]. Because of the RT discretization caused by the experiment code (getting RTs in 35-ms steps), more robust fitting methods that require full RT distribution (e.g., maximum-likelihood or Bayesian approaches) were not employed. Hence, a Chi-square-based (CS) fitting procedure is used. CS-based fitting methods often divide RTs into several bins. Traditionally, 6 bins are used, corresponding to [0.1, 0.3, 0.5, 0.7, 0.9] percentiles. A small random jitter with a magnitude on the order of 10^{-6} seconds drawn from a uniform distribution is added to single-trial RTs to ensure fine-grained binning. The model selection procedure is done in a nested hierarchical manner, for which the individual datasets are partitioned with respect to noise level, and the free parameters included are allowed to vary between those levels. The goodness of fit measures (e.g., CS fit values) are noted for each participant and each model configuration. For model comparison, the Chi-Square table's critical value at a 0.99 significance level is evaluated as in [71], for which the degrees of freedom (df) is calculated as follows:

$$df = (bins \times (2) - 1) \times (conditions) - free\_params \tag{2.1}$$

Accordingly, in the case of using 6 bins as conventionally done, using noise level to partition the data (4 levels), and supposedly having different drift rates for each noise level, one boundary separation and one non-decision time parameter while the starting point is fixed at halfway between the bounds, the equation will result in a df of 38. On the other hand, a simpler model where the only difference is that the drift rate does not differ according to noise levels results in a df of 41. As illustrated, within the model structures divided according to noise levels, as the number of free parameters increases, the df decreases. According to the above example, the critical value is 11.45 when df difference is equal to 3 at a significance level of 0.99. As a result, one should demonstrate that the goodness-of-fit

measure should differ more than that critical value to favor the complex model over the simple one with fewer parameters. Since the trial numbers are relatively lower than suggested in the literature for reliable estimation, starting-point and drift-rate variabilities are not included. The remaining essential parameters of bound, non-decision time, drift rate, and optional non-decision time variability are evaluated in different configurations depending on their relation with the experimental condition.

After establishing the superior model structure, cumulative density functions are investigated to compare model predictions and empirical data. Furthermore, a parameter recovery study is conducted to give as much reliable interpretation [72] as possible by simulating 1000 data sets that have the same number of trials as the empirical data sets. The resulting datasets are fitted with the exact same model structure using the CS-based methods, and from the resulting fits, 95% confidence intervals for parameter estimates are compared to those of the results from empirical fits.

2.4 EEG Data Collection & Analysis

EEG data were recorded using a 64-channel system (Brain Products GmbH, Germany) at a sampling rate of 1000 Hz. Electrode placement followed the international 10–20 system. The ground electrode was placed at AFz, and FCz served as the online reference. Electrode offsets were maintained below 25 kΩ throughout data acquisition to ensure signal quality.

2.4.1 Preprocessing

Preprocessing is conducted individually using MATLAB and the EEGLAB toolbox [73]. First, the data were downsampled to 500 Hz for computational efficiency. The ECG channel was removed, and a bandpass filter was applied with a high-pass cutoff at 0.1 Hz and a low-pass cutoff at 50 Hz. Data were then re-referenced

to the average of all electrodes. Break periods were removed using the ERPLAB toolbox [74]. Subsequently, automatic bad channel detection was performed using EEGLAB's built-in functions. The default rejection criteria were as follows: (1) flat signal for more than 5 seconds, (2) standard deviation of high-frequency noise exceeding 4, and (3) correlation with neighboring channels below 0.8. Channels identified as bad were removed and later interpolated following independent component analysis (ICA). ICA decomposition was performed using the Picard algorithm [75], optimized for speed and robustness. Artifactual components were identified and removed using the ICLabel plugin [76], which classifies components based on probabilistic estimates across categories such as brain, eye, and muscle. Particularly, the components labeled "eye" or "muscle" with a probability greater than 90% were removed.

Epoching was carried out from -2000 to 1200 ms relative to stimulus onset for stimulus-locked analyses. For response-locked analysis, epoching interval spans -1200 to 200 ms relative to response onset, with time 0 defined by the keypress. Epoch rejection was conducted using semi-automated procedures, including visual inspection, identification of improbable data, and detection of abnormally distributed epochs. To ensure consistency between the analyses, the same epochs were rejected in both stimulus-locked and response-locked waveforms. It is important to convert the resulting EEG data from epoching to current source density (CSD), to prevent contamination of the activity from neighboring channels and isolate decision-signals, since some components such as frontocentral negativity may serve another function for motor execution during response yet can have significant impact on CPP waveforms [16] [77]. Therefore, CSD transformation using spline interpolation is utilized. Before computing averaged event-related potentials (ERPs), data were binned according to noise or pre-stimulus duration levels. Furthermore, stimulus-locked ERPs were subjected to baseline correction using a -200 to 0-ms time window relative to stimulus onset, whereas no baseline subtraction was applied for the response-locked analysis.

Additionally, it is aimed to investigate the neural markers in correct versus incorrect or missed trials. For this reason, incorrect responses and misses were combined into a single category to ensure a sufficient number of trials within

each difficulty condition. Hereafter this category is named as invalid trials for clarity. The averaged invalid trial ERPs for each noise level and participant are calculated. This is followed by weighted-averaging across noise levels, ensuring that the resulting signal is not heavily affected by the trial number imbalance across different conditions. Then, the ERPs of interest that will be detailed below were compared between invalid and correct trials.

2.4.2 Centro-Parietal Positivity (CPP) Analysis

The evidence accumulation process in DDM is assumed to reflect both perceptual categorization and response selection. Although in the latter case, motor signals such as readiness potential are used to index accumulation dynamics, CPP is generally used in the first case [22]. Therefore, CPP is the best-suited neural marker to analyze in our particular design for walking direction discrimination.

CPP was computed by averaging the signal across electrode CPz and its two closest neighbors, CP1 and CP2. To demonstrate whether CPP reflects the evidence accumulation dynamics predicted by the hidden decision variable in the sequential sampling models (SSM) framework the following hypotheses were tested, similar to previous research [14]:

1. CPP amplitude and slope scale with task difficulty, with the expectation that steeper slopes and larger amplitudes will be seen in easier trials (i.e., lower noise levels).

2. CPP amplitudes do not show differences across noise levels at the time of decision report, in line with the common-threshold assumption.

3. CPP slope changes with RT, with shallower slopes associated with longer and steeper slopes with shorter RTs.

4. CPP differentiates correct versus invalid trials, with larger amplitudes observed in correct trials.

5. The variations of CPP slope across subjects and difficulty levels will be related to variations observed in drift rate estimates from DDM, in a positively correlated manner.

To test the assumptions related to CPP amplitude from hypotheses (1) and (2), grand-average stimulus- or response-locked waveforms were investigated. ANOVA with false-discovery rate (FDR) correction using a conventional procedure by Benjamini & Hochberg performed within the fMUT package [78]. For this analysis, the centro-parietal electrodes CP1, CPz, and CP2 are included. Additionally, a cluster-based permutation test was applied within the same package to assess amplitude differences in a broader posterior region. The posterior and posterior occipital electrodes are included in this analysis due to observed characteristic profile during the candidate time interval ($300 - 800$ms), and previous studies employing some of those channels to quantify CPP. A combination of these two tests led to the identification of spatially and temporally contiguous electrode clusters showing significant amplitude modulation across experimental conditions. Informed by the results, single-trial CPP mean amplitudes were extracted from baseline-corrected, stimulus-locked waveforms within the significant time windows. CPP mean amplitudes are then subjected to repeated-measures ANOVA utilizing noise level as within-subject factor, and if appropriate, pair-wise comparisons with a T-test are followed on aggregated measures.

To test the assumptions related to CPP slope in hypotheses (1) and (3) and to determine the time intervals for significant CPP slope differences between the noise levels, both stimulus-locked and response-locked waveforms were analyzed using a sliding window regression approach. These windows have a 100-ms length and are shifted in time with 10-ms steps. For each time window and participant, the CPP slope was estimated by fitting a linear regression line to the 100-ms-long ERP signal across noise levels. These slope values were then regressed against levels of difficulty. Specifically, increasing noise levels were hypothesized to decrease the CPP slope due to increased difficulty. Informed by the results for significant time windows, single-trial CPP slopes were estimated from response-locked waveforms as the slope of a linear line fitted, starting from the first time window

center reaching significance, and ending at 150ms following that [77]. As for the repeated-measures ANOVA analysis incorporating noise level and RT category as within-subject factors, it is expected to see steeper CPP build-up with fast RTs. In addition, this pattern is also expected to be seen in more precise analysis using a permutation test and regression against 6 RT bins.

To test hypothesis (4), stimulus-locked CPP waveforms were calculated by weighted averaging across noise levels for each subject. From these waveforms, the mean amplitude values in the same time window determined by the FDR-corrected ANOVA procedure are extracted. Then, the CPP mean amplitudes from correct and invalid trials are compared using a one-tailed paired-samples T-test, with the alternative hypothesis being that the measurement from correct trials is greater than that of invalid trials.

Lastly, to test the relationship between drift rate estimates and CPP more closely, as outlined in hypothesis (5), mean CPP slopes within each subject and each noise level are calculated. According to the winner DDM model structure, the resulting mean CPP slope and drift rate estimates are tested with a one-tailed Pearson's correlation test, with the alternative hypothesis being that those two measures are positively correlated. In case of significance, to ensure the positive relationship is not driven solely by simply partitioning the data into different difficulty conditions, the effect of noise level is regressed out from both measures. The remaining residuals from both measures were subjected to linear mixed model analysis with a random effect of subject.

2.4.3 Biological Motion Related Negative ERP (N2-bm) Analysis

Based on visual inspection of the scalp topographies and in line with prior literature, three right-hemisphere channels—PO8, P6, and P8—were identified as sites of maximal negativity during relatively late time windows ($300 - 400$ms).

Informed by the literature, this signal is hereafter named as the second negative time interval biological motion-related ERP (N2-bm, for short) for the rest of the paper. To examine whether activity at these sites was modulated by noise level, an FDR-corrected ANOVA procedure was applied on stimulus-locked grand-averaged waveforms using the specified channels. It is expected to see that the N2-bm amplitude scales with evidence strength, as lower noise levels are associated with more negative activity. This finding would support the hypothesis that easier stimuli accompany better visual encoding for higher-level BM perception. The significant time windows as a result of this procedure were used to calculate the single-trial N2-bm peak negativities in baseline-corrected stimulus-locked waveforms. These single-trial measures are then analysed with a repeated-measures ANOVA having noise level and RT category as within-subject factors. It is expected to see significantly larger N2-bm peak negativities in faster RTs.

In addition, to test if N2-bm peak negativity is different between correct and invalid trials, the weighted averaging procedure is followed. Mean peak negativities for correct and invalid trials are then compared by means of a one-tailed paired-samples T-test. It is expected to demonstrate statistically larger peak negativity in correct compared to invalid trials as a reflection of enhanced visual encoding quality.

2.4.4 Attentional and Motor-Related Neural Markers

The fluctuations in alpha-band activity were interpreted as indices of attentional engagement, with lower power reflecting increased attentional engagement [12]. In that research, it is reported that alpha-band activity levels right before the stimulus onset are linked with RTs and CPP slope, in a way that higher values of alpha power are related to longer RTs and slower CPP build-up. Additionally, it has been suggested that attention might modulate relatively late biological motion-specific negative ERPs [60]. Therefore, alpha band activity from $8-13$Hz within $[-700, 0]$ms time window with respect to target stimulus onset is examined

across five posterior-occipital electrodes spanning from PO7 to PO8 using the newtimef() function in EEGLAB. Event-related spectral perturbations (ERSPs) are returned as the output from this procedure and then averaged to obtain single-trial measures. To enhance precision, the Morlet wavelet was convolved over a broader window from -1000 to 0 ms, providing an additional temporal buffer. The characteristics of the Morlet wavelet were defined as three cycles at the lowest frequency, with a scaling factor of 0.5 to balance temporal and frequency resolution.

Previous studies reported that pre-stimulus motor preparation levels may elevate depending on manipulating the expectancy, which is associated with the starting point parameter being closer to the decision boundaries [40]. Furthermore, in a neurally-informed modeling study, motor preparatory activity reflected in mu-beta band spectral amplitudes has been used to constrain the starting points before the evidence accumulation starts [39]. In this respect, to test if motor preparatory activity is significantly related to experimental manipulations or RT, mu-beta band spectral power is calculated using fronto-central sites, C3 and C4. Apart from the channels employed, the remaining procedural details (e.g., wavelet properties, time window) are the same as explained for the alpha-band analysis. The resulting ERSPs were integrated separately for each channel, and the mean spectral amplitudes were stored in a 3D matrix with dimensions corresponding to subject (s), trial (i), and channel (C3 or C4). A lateralized motor preparation index was computed by subtracting activity in the contralateral motor cortex from that in the ipsilateral cortex, based on the direction of the upcoming walker stimulus. Although the output is termed as ERSPs for both attention and motor markers, baseline correction was applied in neither of the analyses regarding pre-stimulus activities. This is because any selection for the baseline would be arbitrary, as there are no experimentally induced changes in the epoching time intervals while participants are expected to passively view scrambled motion displays. Moreover, applying an arbitrary baseline correction could introduce challenges related to interpretation since there are 3 different levels of pre-stimulus durations. Therefore, it is more appropriate to refer to these measures as dB-power rather than ERSPs, as ERSPs typically denote activity

relative to reference windows.

Following this step, two separate repeated-measures ANOVAs on each of the single-trial pre-stimulus measures are conducted. In both of them, the subject is the random factor. In the first RMAnova, one of the within-subject factors was pre-stimulus duration and in the second, noise level. Due to relatively long RTs observed in our task compared to simple perceptual decisions spanning several hundred milliseconds, the expectations regarding the relationship between pre-stimulus activity and RT are not as clear as those of CPP and N2-bm.

Lastly, lateralized readiness potential (LRP) is calculated, which has previously been found to be modulated by evidence strength and reaching a common level at response, although lagging behind that of CPP's temporal profile [12]. Additionally, some researchers found that the readiness potential, rather than CPP, reflects the evidence accumulation process [22]. To calculate LRP, response-locked epochs with 8 bins are created, for which each noise level is partitioned into two, depending on the walker's direction (right or left). From resulting ERPs, LRP was computed as the difference between contralateral and ipsilateral activity with respect to the hand used for the response on each trial. The electrodes used for LRP computation were C3 and C4, corresponding to the left and right motor cortices, respectively [79]. After LRPs are calculated, sliding regression analysis is used as in CPP slope analysis to see if the build-up rate of LRP differs across noise levels. It is expected to see that significant windows marked as a function of difficulty to be closer to response onset than of CPP. This temporal lag would further suggest CPP's role in representing the modality-independent abstract decision variable by temporally preceding that of the motor preparation component.

2.4.5 Relationship Between Neural Markers

The remarkable performance and sensitivity observed in response to BM perception [45] lead to questioning whether attentional resources are needed at all for this excellent ability. However, previous research showed that attention is still

required for active processing in dual-task settings [80], or when the sensory information is degraded and in higher-level tasks such as intention understanding [81]. On the other hand, CPP slope has already been linked positively with attentional engagement levels preceding the stimulus onset [12]. Hence, it is noteworthy to examine the relationship of N2-bm and CPP with pre-stimulus alpha power.

To quantify the extent of the relationship between the neural markers, a series of correlation analyses was conducted. The single-trial data are z-scored within each subject, and averaged within each combination of experimental manipulations. This normalization procedure aimed to remove between-subject variability and met the assumptions of correlation analyses regarding independence. It is expected to see a significant relationship between N2-bm or CPP with the alpha power from the pre-stimulus duration. In other words, higher quality visual encoding and enhanced evidence accumulation are expected when the trials accompany already elevated attentional engagement before target PLDs. As such, lower alpha power is expected to be associated with larger N2-bm peak negativity and higher CPP slopes or mean amplitudes.

Following correlation analyses, single-trial data are investigated to quantify if the relationship between pre-stimulus attention (i.e., alpha power) and evidence accumulation (i.e., CPP) is mediated by the high-level processing of biological motion (i.e., N2-bm). To precisely investigate this hypothesis, a Bayesian multilevel analysis is employed. This approach allows for the simultaneous estimation of multiple hierarchical regression models while accounting for subject-level variation. Particularly, two models are used to accommodate this analysis within the mediation framework. In the first model (mediator model), N2-bm peak negativity is the outcome variable, and the neural predictor is alpha power. For the second model (outcome model), CPP mean amplitude is the outcome while N2-bm peak negativity and alpha power are the predictors. Both models also include both experimental manipulations (i.e., pre-stimulus duration and noise level) as categorical predictor variables. Lastly, random-effect is specified as for each subject. By incorporating several potential sources of variability, the estimated parameters for the indirect and direct effects will reflect a more fine-grained

relationship between the neural markers than simple correlation on averaged measures. To jointly fit these two models, the brms package in R [82] is utilized with 4 independent Markov chains, each run for 4000 iterations, of which the first 1000 were the warm-up to ensure sampling efficiency. Results give the indirect effect as the multiplication of regression coefficients from alpha to N2-bm (in the mediator model) and N2-bm to CPP (in the outcome model). As for the direct effects, the regression coefficient from alpha to CPP (in the outcome model) is used. Lastly, the total effect is the sum of direct and indirect effects. It should be noted that for the analyses including alpha power, data from 13 subjects were used due to interpolation of critical channels for the remaining 3 out of 16 participants.

Chapter 3

Results

3.1 Behavioral Results

Repeated-measures ANOVA, having RT as the dependent variable, revealed that for the within-subject noise level factor, the sphericity assumption is violated, which is addressed by Greenhouse-Geisser correction. Although the Shapiro–Wilk test indicated a deviation from normality in the residuals ($W = 0.974$, $p = .001$), visual inspection of the Q–Q plots suggested that the violation was minor. Given that the test statistic was close to 1 and considering the known robustness of repeated-measures ANOVA to mild violations of normality, the analysis proceeded. The main effect of the noise level was significant after correction ($F(1.82, 27.32) = 270.76$, $p < .001$, $\eta^2_G = .40$). There was also a significant main effect of the pre-stimulus duration ($F(2, 30) = 13.78$, $p < .001$, $\eta^2_G = .035$). Moreover, the interaction between these factors was also significant ($F(6, 90) = 3.08$, $p < .01$, $\eta^2_G = .012$) (see Fig. 3.1).

Post-hoc pairwise comparisons were conducted using estimated marginal means (EMMs), controlling for multiple comparisons using a Bonferroni adjustment. It is revealed that for the noise level, all pairs have significantly different mean RTs from each other (all $p < 0.001$), with lower RTs observed in easier

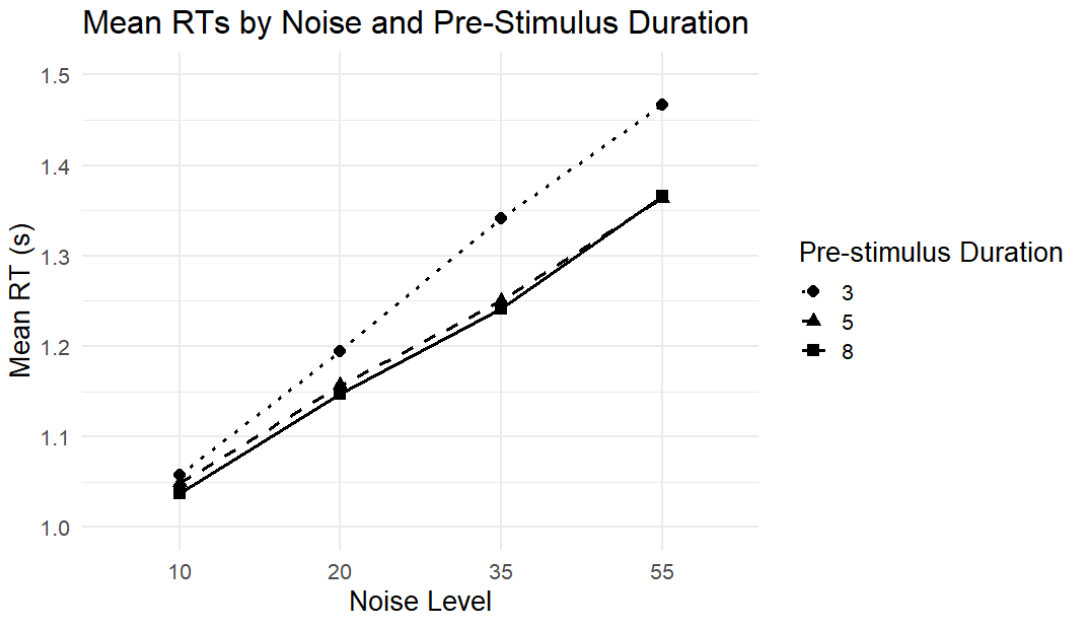

Figure 3.1: Mean RTs across the combination of experimental manipulations are shown on the y axis. The X axis shows different noise levels, while different pre-stimulus durations (in seconds) are shown as separate lines.

conditions and vice versa. This provided support for a successful experimental manipulation of difficulty. On the other hand, higher RTs are observed in the pre-stimulus duration of 3 seconds than in the 5 or 8-second durations ($p < 0.01$ for both comparisons). This pattern might be due to problems with detecting the transition from scrambled motion to PLD walkers in short durations, where preparation is limited. To inspect the significant interaction more closely, the main effect of one experimental manipulation was investigated within each level of the other. This analysis revealed that, especially for the higher noise levels (35 and 55), the differences between the pre-stimulus duration of 3 compared to 5 or 8 seconds become significant, so that significantly faster RTs are observed after longer waiting times (all $p < 0.05$).

Different response categories (correct, incorrect, misses, premature responses) across experimental manipulations are analyzed with a chi-square test. There was a significant main effect of pre-stimulus duration ($\chi^2(6) = 214.21$, $p < .001$, $Cramer'sV = 0.118$) and noise level ($\chi^2(9) = 961.03$, $p < .001$, $Cramer'sV = 0.205$). Following significant main effects, post-hoc pair-wise comparisons are

conducted between combinations of experimental manipulations. First, the contingency table with a combination of pre-stimulus duration (3) and noise level (4) is formed, followed by calculating the standardized residuals. The resulting post hoc comparison table, along with the standardized residuals and their statistical significance, is shown below (see Table 3.1). Inspection of the table revealed that lower noise levels are associated with more correct but fewer misses and incorrect responses. In contrast, higher noise levels show the opposite pattern. These trends suggest that difficulty is successfully manipulated. For the pre-stimulus duration, longer waiting periods were associated with significantly more premature responses, especially if it was coupled with higher difficulties, as the standardized residuals for 8.35 and 8.55 showed a significant effect. This effect might stem from the increased urge to respond as soon as possible imposed by the increased timing of scrambled motion. In addition, 8.35 and 8.55 conditions are associated with significantly more incorrect responses, whereas this is not seen at 3.35 or 3.55, respectively. This intriguing pattern points to the evidence accumulation process terminating at the wrong decision boundary, which may also be explained by the increased urgency imposed by longer waiting intervals. More specifically, growing urgency might prevent sufficient evidence accumulation, resulting in either internal or external noise outweighing the signal. This claim is further supported by the observed pattern in mean RTs, where the significantly lower RTs (hence limited time for evidence accumulation) are observed in higher pre-stimulus durations at higher noise levels. However, for the shorter pre-stimulus durations, the expectation of seeing more misses was not confirmed. This is potentially due to the effect of noise level becoming more powerful to account for the missed trial counts compared to the pre-stimulus duration.

3.1.1 DDM Results

Overall, the model where the drift rate depends on the noise level (v_noise) is favored by most participants according to goodness-of-fit measures. In this model, the starting point is fixed halfway between the bounds due to accuracy coding. The remaining 6 free parameters include four drift rates for each noise level, one

Table 3.1: Standardized residuals for each response type across a combination of experimental conditions, where the first number denotes prestimulus duration, the second denotes noise level.

Condition	Incorrect (0)	Correct (1)	Miss (2)	Premature (3)
3.10	**-4.80***	**10.09***	**-8.05***	-2.27
5.10	**-4.42***	**9.91***	**-8.21***	-1.91
8.10	**-3.82***	**9.00***	**-8.32***	0.31
3.20	**-3.82***	**4.62***	-1.91	**-3.01***
5.20	-2.85	**7.38***	**-6.37***	-1.54
8.20	-1.83	**6.01***	**-6.86***	2.53
3.35	-0.63	**-3.83***	**-6.05***	-2.99
5.35	1.91	0.87	-2.16	-0.43
8.35	**4.49***	-1.34	**-3.19***	**5.10***
3.55	0.90	**-17.42***	**21.26***	-3.01
5.55	**6.05***	**-13.41***	**12.01***	-0.06
8.55	**8.83***	**-11.90***	**5.73***	**7.30***

Note: Asterisks (*) denote the conditions where post-hoc comparisons based on standardized residuals produced significance ($p < .05$). Values are rounded to 2 decimal places for readability. The overall effect size for the Chi-Square post-hoc analysis, Cramer's V, is 0.234.

boundary separation, and one non-decision time. The table below shows the subject-wise parameter estimates for the v_noise model (see 3.2). The mean CS-fit value (64.87) is just above the critical value of 53.38 at 38 degrees of freedom and 0.05 significance. However, this does not necessarily mean that there is a significant misfit given the overly-conservative nature of the Chi-Square approach [83]. Moreover, the differences are mostly driven by the resulting extremely large fit value for subject 12 (226.26). As such, when that participant is excluded, the mean CS-fit value becomes very close to the critical value (M=54.11).

Drift rate estimates across noise levels and subjects are compared with repeated-measures ANOVA. However, due to violations in parametric assumptions, a non-parametric alternative, the Friedman rank sum test, is used. This test returns a significant effect of noise level on drift rate ($\chi^2(3) = 48.00$), $p < .001$). For the following pair-wise comparisons, Exact Test is used, which revealed that drift rates for noise level 10 and 35, noise level 10 and 55, and noise level 20 and

Table 3.2: Estimated drift diffusion model parameters and resulting chi-square fit values for each subject (**a** is boundary separation, **t0** is non-decision time, and **v_i**s are the estimated drift rate parameters for noise level i). CS Fit values represent the goodness of fit measures for each participant. The final row shows the mean parameter estimates across participants (standard deviation).

Subj	**a**	**t0**	**v_55**	**v_35**	**v_20**	**v_10**	**CS Fit**
1	5.30	0.33	2.98	3.64	4.40	5.31	32.37
2	4.03	0.50	1.99	2.12	2.47	2.79	21.52
3	2.22	0.73	0.92	1.79	2.56	3.06	104.33
4	2.43	0.52	1.70	2.23	3.18	3.95	19.81
5	1.84	0.73	0.62	1.01	1.73	2.52	65.73
6	2.07	0.63	0.56	1.23	1.54	2.24	84.58
7	6.37	0.10	2.39	2.61	2.68	3.09	15.89
8	2.95	0.55	1.72	2.25	2.48	2.93	41.64
9	2.15	0.63	0.90	1.89	2.88	3.98	64.03
10	6.71	0.38	2.70	3.15	3.77	4.32	22.59
11	1.86	0.64	0.80	1.18	2.18	3.51	95.10
12	2.08	0.92	0.39	0.52	1.41	2.20	226.26
13	2.96	0.59	1.97	2.36	2.85	3.68	24.75
14	2.10	0.55	0.77	1.67	2.17	3.83	58.41
15	2.29	0.81	0.54	1.48	1.61	2.29	93.23
16	2.43	0.59	1.25	1.82	2.22	3.33	67.73
M(SD)	3.11(1.56)	0.58(0.19)	1.39(0.81)	1.94(0.77)	2.51(0.78)	3.31(0.83)	64.87(50.66)

55 significantly differ (all p values $< .001$) after Bonferroni correction. Therefore, lower noise levels, hence easier conditions, are associated with higher drift rates.

After getting parameter estimates and deciding on a winning model based on a nested comparison of goodness-of-fit measures, it is important to check whether the empirical data are in line with the model predictions. Visual inspection of a plot accompanying the empirical cumulative density function (CDF) collapsed across noise levels and its predicted counterpart showed a decent overlap (see 3.2). It should be noted that step-like increases in the empirical CDF are due to a lack of precision in getting single-trial RTs as a result of 35-ms discretization. Another important check is whether the estimated parameters of model fitting to empirical data can be reliably recovered from the generated datasets using those exact parameter configurations. For that purpose, 1000 datasets for each participant that match the trial numbers as much as possible to those of different noise levels are generated, and the resulting parameter estimates are noted. However,

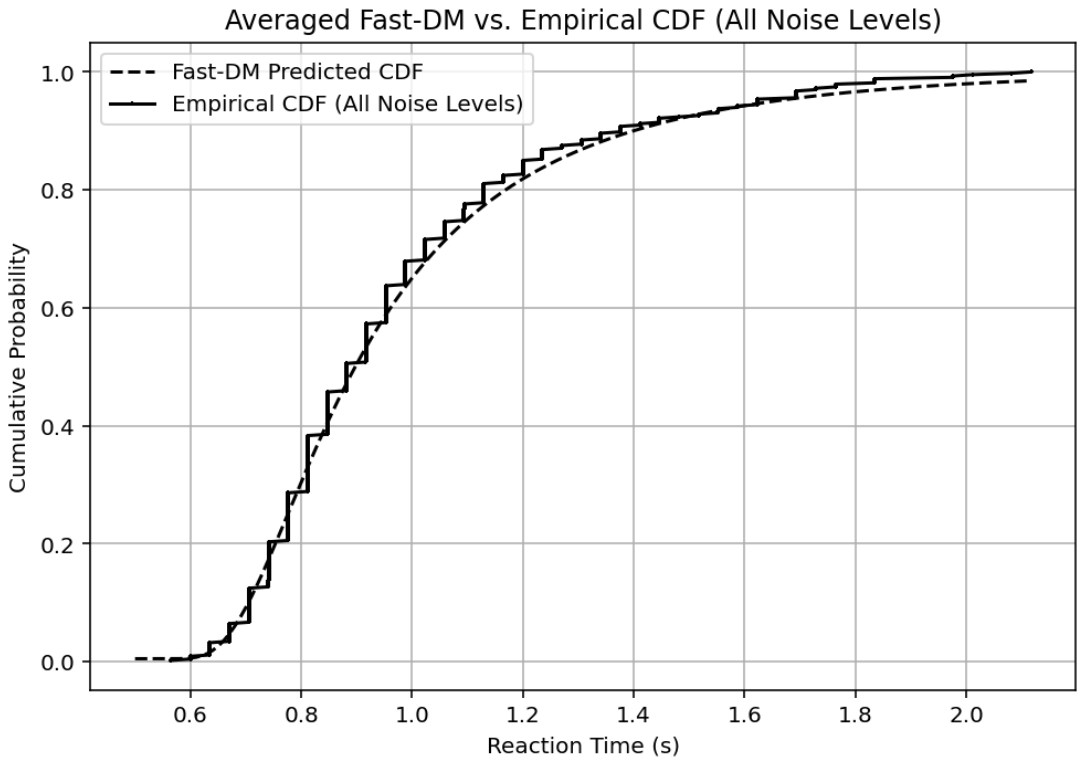

Figure 3.2: Cumulative Density Functions (CDFs) for subject 4, averaged across all noise levels. The solid line represents the CDF from empirical data, and the dashed line shows the predicted CDF from parameter estimates after model fitting with Fast-DM. A visual inspection revealed a decent overlap between model predictions and empirical data.

estimated parameters were not recovered successfully when evaluated at 95% CIs. This might be caused by the CS-based fitting method being less successful in parameter recovery, especially in low trial numbers. Indeed, in this study, the trial numbers are lower than the previous studies and best practices outlined [23] after excluding misses and partitioning the individual datasets according to the experimental conditions. Yet another potential reason might be longer RTs observed in our task compared to simple perceptual decisions often investigated with DDM, leading to inherently less reliable parameter estimates.

3.2 EEG Results

3.2.1 CPP

First, whether there exist amplitude differences in the stimulus-locked CPP signal across experimental conditions of pre-stimulus duration or noise levels is investigated. For the false discovery rate (FDR) corrected ANOVA, the channels of interest are CP1, CPz, and CP2. In those electrode sites, activity has been found to differ significantly between noise levels starting at around 380 ms and consistently continuing up to 830 ms following the stimulus onset (see Supplementary Figure A.1). Similarly, a cluster-based permutation test showed comparable results for the effect of noise level. In addition, when the response-locked CPP waveforms were investigated using the same procedure, there were no significant differences in amplitude with respect to difficulty levels revealed near 100ms of the decision reports (see Supplementary Figure A.2). This finding is in line with the assumptions that CPP scales with evidence strength and reaches a common level at response, as expected from an abstract decision variable. The significant time window that spans from $380 - 830$ ms informed the single-trial analysis to calculate the mean amplitude of the CPP. Using those single-trial measures, repeated-measures ANOVA showed a significant main effect of noise level ($F(3, 45) = 15.79$, $p < .001$, $\eta^2_G = .06$), where the lower noise levels are associated with higher CPP mean amplitudes. The results of pair-wise T-test applied on aggregated single-trial measures produced significant differences between noise levels 10 and 35 ($p = 0.034$), 10 and 55 ($p < 0.001$), 20 and 55 ($p = 0.002$), and 35 and 55 ($p = 0.049$) after Bonferroni correction. As for the pre-stimulus duration, the significant activity differences across CPP channels are not as consistent as in the case of noise level. Rather, this modulation is observed in more parietal scalp sites (e.g., P1, P2 channels) (see Supplementary Figure A.3).

Another important property of CPP is having a build-up rate that scales with evidence strength. In this respect, a running window-based linear regression was applied across subjects to estimate the rising slope of the stimulus-locked CPP waveforms for each difficulty level. Significant differences in rising slope emerged

between noise levels starting from the $280-380$ ms window and extending to the $370-470$ ms window. These significant windows were determined under the assumption that lower noise levels (i.e., easier conditions) are associated with steeper CPP build-up. The grand-averaged CPP waveforms for each difficulty level are presented, with the centers of the significant time windows marked along the time axis (see Fig. 3.3). Response-locked waveforms were also analyzed using a similar sliding-window procedure as in the stimulus-locked analysis, whereas the time window of interest spanned from -1200 ms to 0 ms relative to the decision report. The same hypothesis as for the relationship between difficulty and the rising slope of the CPP (i.e., lower noise associated with steeper build-up) is used, and FDR correction for multiple comparisons is applied. It is revealed that the significant differences in slope emerged beginning at approximately -600ms and continued up until around -330ms before the response. Response-locked grand-averaged CPP waveforms for each difficulty level are presented with the significant time window centers marked below (see Fig. 3.4). The marked onset of slope differentiation thus informed the single-trial analysis on estimating CPP slopes, which is taken from a window of 150ms length, from -550 to -400ms.

For each subject, weighted averaged CPP mean amplitudes across noise levels are compared between invalid versus correct trials using the time window of $380-830$ms in stimulus-locked waveforms. One-tailed paired-samples T-test result was significant ($t(15) = 7.504$, $p < .001$, Cohen's d $= 1.876$). Therefore, it is concluded that CPP mean amplitudes can differentiate between the correct and invalid trials as expected from a neural representation of a decision variable.

In addition, to connect the drift rate with the CPP rising slope, the mean single-trial CPP slopes were calculated for each subject and each noise level, A one-tailed Pearson correlation revealed a significant positive relationship ($r(62) = .42$, $p < .001$, 95% $CI[0.24, 1.00]$). To more precisely estimate the relationship between these two measures, the effect of noise is regressed out. This way, it is ensured that the significant relationship is not the result of partitioning the datasets according to different noise levels. A linear mixed model analysis on the resulting residuals with a random effect of subject revealed that the positive relationship still persists ($r(62) = .32$, $p < .01$, 95% $CI[0.12, 1.00]$). These results

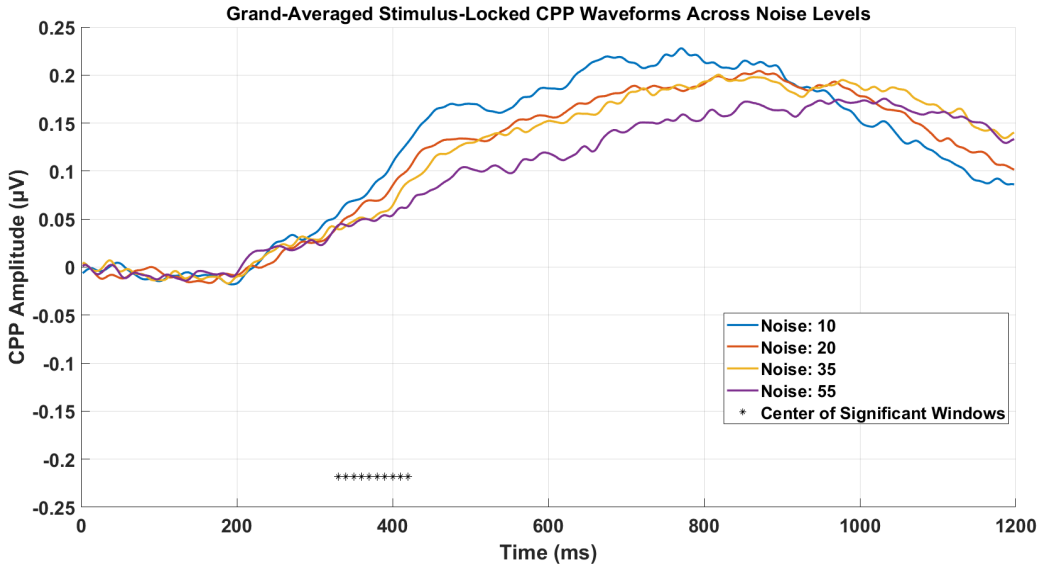

Figure 3.3: Stimulus-locked grand-averaged CPP (N = 16) for each noise level. Asterisks along the bottom of the plot mark the centers of significant time windows as a result of linear regression, assuming that lowering the noise leads to a steeper rise in the signal. Note that while determining the significant windows, false discovery rate correction is used to prevent Type I errors.

suggest that the variations in CPP slope is related to variations in drift rate, further supporting the properties of an abstract decision variable.

3.2.2 N2-bm

FDR-corrected ANOVA procedure is also applied to N2-bm waveforms taken from parieto-occipital channels, including PO8, P8, and P6 based on the inspection of scalp maps and stimulus-locked waveforms (see Fig.3.5). As a result, it has been statistically validated that around 290ms onwards and up until around 510ms, N2-bm activity is modulated by the noise level. Informed by the output of this procedure, single-trial peak negativities of N2-bm are measured in the same interval. Using single-trial measures and applying repeated-measures ANOVA demonstrated a significant main effect of noise ($F(1.88, 28.26) = 17.78$, $p < .001$,

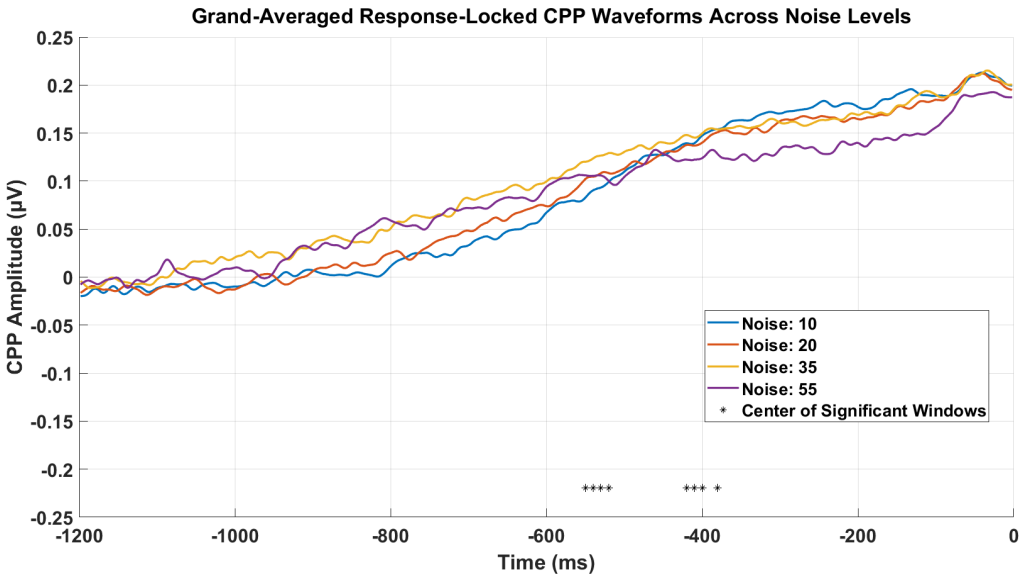

Figure 3.4: Response-locked grand-averaged CPP (N = 16) for each noise level. Asterisks along the bottom of the plot mark the centers of significant time windows as a result of linear regression, assuming that lowering the noise leads to a steeper rise in the signal. Note that while determining the significant windows, false discovery rate correction is used to prevent Type I errors.

$\eta^2_G = .041$). As a next step, pair-wise comparisons with a T-test on the aggregated single-trial measures revealed significant differences in N2-bm peak negativity between noise levels 10 and 20 ($p = 0.049$), 10 and 35 ($p < 0.01$), 10 and 55 ($p < 0.001$), and 20 and 55 ($p < 0.01$) after Bonferroni correction. Therefore, significantly larger N2-bm peak negativities are observed in lower noise conditions. In other words, N2-bm peak negativity scales with sensory evidence strength, as larger negativities potentially reflect enhanced higher-level perceptual processing. On the other hand, there was no modulatory effect of pre-stimulus duration observed in FDR-corrected ANOVA, suggesting that it is the stimulus strength rather than the interval preceding the stimulus onset that drives the amplitude differences.

Additionally, weighted average N2-bm waveforms across noise levels are compared between invalid and correct trials. Specifically, peak N2-bm negativities in a $290 - 510$ms time window are compared with a non-parametric one-tailed Wilcoxon Signed-rank due to a violation of the normality assumption in the

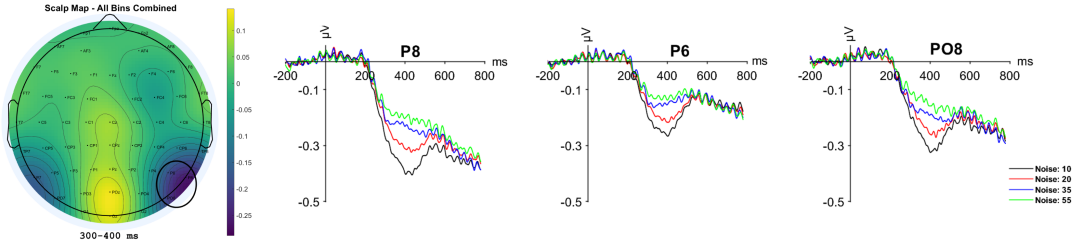

Figure 3.5: Stimulus-locked grand-averaged scalp map for the mean activity over $300 - 400$ms is shown on the left panel. Next to the scalp map, stimulus-locked grand-averaged waveforms across noise levels are shown for the N2-bm specific channels: P8, P6, PO8 in order, where the larger peak negativities are observed in lower noise levels.

paired-samples T-test. The alternative hypothesis was that the measurement from correct trials is greater in magnitude than that from invalid trials. The result was significant ($z = -3.516$, $p < .001$) and perfectly consistent across subjects ($W = 0.000$, Rank-Biserial Correlation = -1.000).

3.2.3 Attentional and Motor Neural Markers

Greenhouse-Geisser corrected results showed that alpha power significantly differs across pre-stimulus durations ($F(1.19, 14.24) = 12.57$, $p = .002$, $\eta^2_G = .013$) and noise levels ($F(1.92, 23.04) = 3.51$, $p = .048$, $\eta^2_G = .0016$). The pair-wise comparisons using averaged single-trial estimates revealed that all pairs significantly differ between pre-stimulus durations (all $p < .05$), with longer intervals leading to lower alpha power. This finding may reflect that attentional engagement is becoming increasingly higher with longer waiting intervals before the stimulus, potentially due to growing uncertainty for the timing of the upcoming target PLDs. For the noise level, the significant main effect is found to be driven by the differences between the highest and lowest noise levels ($p = .038$). That is, lower alpha power is observed in the most difficult condition compared to the easiest. This might be due to how the scrambled motion stimuli are created as a control for target BM displays by accompanying the same number of dots. Therefore, the ambiguity of the stimulus shown on the screen during the pre-stimulus interval

may lead to elevated attentional engagement to prevent missing the transition to target PLDs. The same procedure is applied to pre-stimulus motor preparatory signals indexed by the mu-beta band power at C3 and C4 channels as the contralateral minus ipsilateral activity with respect to the upcoming trial. However, none of the within-subject factors showed significance. This might be due to a lack of explicit cueing regarding the upcoming stimulus timing compared to previous research.

To quantify whether the LRP build-up rate differs across noise levels and the timing of this differentiation, similar to running regression analysis for CPP response-locked waveforms, the slopes are investigated. The expectation was to see significant time windows where the faster build-up rate of LRP is to be seen in lower noise levels. There are only two 100-ms window centers marked in the output, -230 and -220ms, satisfying this presumption. The centers being closer to response onset than CPPs (-550 to -360ms) is in line with the view that CPP reflects an abstract decision variable preceding the motor preparatory activity [12]. However, there were no significant centers that survived after FDR correction for multiple comparisons. This might be due to prolonged RTs in our task that may interfere with the build-up dynamics of motor preparatory activity.

3.2.4 Relationship Between Neural Markers

According to averaged neural measures across experimental conditions following within-subject z-scoring, a significant negative correlation was observed between alpha power and N2-bm peak negativity (r = -0.18, p =.027, 95% CI [-0.33, -0.02]). On the other hand, a significant positive correlation was found between alpha power and CPP mean amplitude ($r = 0.21$, $p = .008$, 95% CI [0.06, 0.36]) or CPP slopes ($r = 0.23$, $p = 0.003$, 95% CI [0.08, 0.38]). This positive relationship pattern is in contrast with the findings of Kelly & O'Connell (2013), where lower alpha band activity, interpreted as indicating higher attentional engagement, precedes greater CPP slopes. Lastly, there was a significant negative correlation between N2-bm peak negativity and CPP mean amplitude (r = -0.27, $p < .001$,

95% CI [-0.41, -0.12]). This result suggests that enhanced visual encoding (larger N2-bm peak negativity) is linked with enhanced evidence accumulation.

Following correlation analyses, the possibility of visual encoding quality, indexed by N2-bm, to be the mediator for the relationship between pre-stimulus attention (alpha) and the evidence accumulation process (CPP) is investigated. In this context, Bayesian multilevel modeling in the mediation framework is utilized. Before reporting Bayesian multilevel model outputs, convergence diagnostics were examined, including $\hat{R}$ values indexing whether the chains are converged and effective sample sizes (ESS), where higher numbers reflect reduced autocorrelation between the estimated effects. $\hat{R}$ values are generally regarded as ideal if it is less than 1.1 or, even more conservatively, 1.02 [3]. All $\hat{R}$ values found to be equal to 1.00, suggesting chain convergence and hence reliable posterior estimates. On the other hand, effective sample size (ESS) should be as high as possible to index how many samples are used to calculate the posterior for that parameter [84]. The ESS values for all parameters are large enough to suggest that there is a sufficient amount of independent information present in MCMC samples. Therefore, after ensuring robust and stable estimates for the posterior distributions, indirect, direct, and total effects were computed by combining path coefficients from the two models.

The indirect effect of alpha power on CPP mean amplitude via N2-bm peak negativity was found to be consistently positive across samples (Median = 0.0009, 95% CI [0.0005, 0.0013]), as well as the direct effect of alpha power on CPP mean amplitude (Median = 0.0038, 95% CI [0.0014, 0.0063]. Lastly, the total effect combining both direct and indirect effects suggested a reliable positive relationship between alpha power and CPP mean amplitude (Median = 0.0048, 95% CI [0.0023, 0.0072]). Therefore, the combined results suggest that the relationship between pre-stimulus alpha power on CPP mean amplitude is at least partially accounted for by the mediator N2-bm peak negativity. In other words, the relationship between attentional fluctuations before the target onset and the decision-making process is partially mediated by higher-level visual encoding processes. It should be noted that despite the indirect, direct, and total effects estimated numerically as small, this is expected given the scales of the neural variables involved

(e.g., microvolt-level ERP amplitudes and log-transformed power). Additionally, the indirect effect calculation involves the multiplication of two regression coefficients, which inherently leads to smaller values.

3.3 Reaction Time & Neural Markers

None of the pre-stimulus markers significantly differed between RT categories, regardless of which experimental manipulation determined those categories. Relatively longer RTs observed in our task might have prevented the activity before the stimulus onset from having any significant influence. Additionally, specific to contralateral preparatory motor activity, the non-significant result might be due to unpredictability of the upcoming stimulus timing as no cues are involved in our experimental design, contrary to Kelly et al. (2021) [39].

For the neural measures calculated following the stimulus onset, N2-bm peak negativity from stimulus-locked and CPP slope from response-locked waveforms are of interest. According to repeated-measures ANOVA analysis, both N2-bm peak negativity ($F(1, 15) = 11.94$, $p < .01$, $\eta^2_G = .03$) and CPP slope ($F(1, 15) = 48.98$, $p < .001$, $\eta^2_G = .26$) significantly differ across RT categories. Specifically, larger N2-bm peak negativity and steeper slopes are observed in fast compared to slow decisions.

To more thoroughly investigate the relationship between RT and neural markers, z-scored aggregated RTs are separated into 6 bins within the noise levels. Following that, either the CPP slope or the N2-bm peak negativity is regressed against RT bins for each one of the noise levels. The resulting slopes are statistically evaluated within the scope of a permutation test with 10000 iterations. As a result, it is revealed that the relationship between RT bins and CPP slope is significant and not due to chance. That is, the estimated regression coefficients were consistently and significantly negative across all noise levels (ranging from [-0.00011, -0.00017], all $p < .001$). This indicates that slower decisions were significantly associated with shallower CPP slopes even within difficulty levels (see

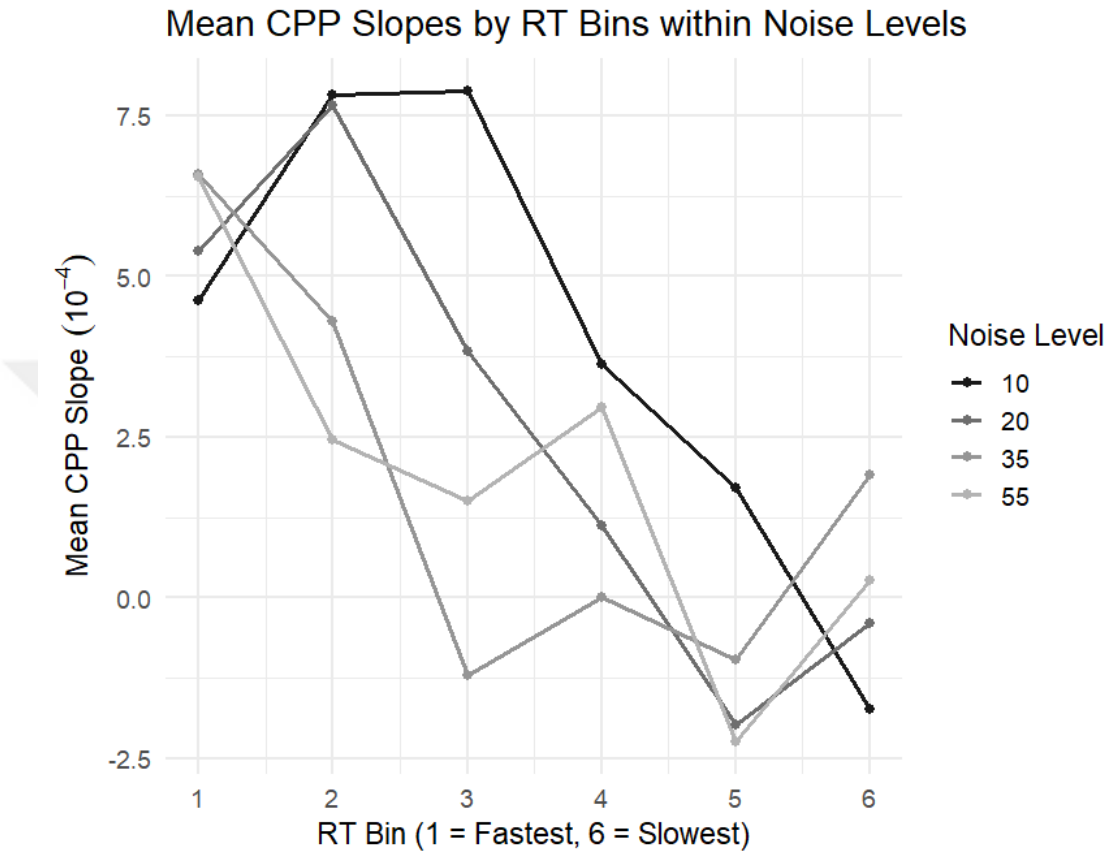

Figure 3.6: Mean CPP slopes are plotted across 6 reaction time (RT) bins on the x-axis, with separate lines representing each noise level condition. The y-axis values are scaled by a factor of 10^4 for improved readability and interpretability. It is demonstrated that CPP mean slopes are getting smaller in slower RT bins, as the downward trend is visible for all noise levels.

Fig. 3.6). However, this pattern is not observed for the N2-bm peak negativity. These combined results point to the distinction between the CPP and N2-bm in terms of which underlying decision process they may represent. In other words, although both measures are modulated by stimulus quality through noise level manipulation, it is the CPP slope that shows fine-grained RT-predictive qualities. Therefore, the claim that CPP reflects an abstract signal and N2-bm reflects the perceptual encoding quality of a biological motion stimulus is further strengthened.

Chapter 4

Discussion

4.1 General Findings

In this study, the aim is to investigate perceptual decision-making dynamics for biological motion perception. Given its established success in reflecting human motion information by effectively preventing potentially confounding form cues, PLDs are used as the stimuli. Informed by the previous research [12], our design enables seamless transition from scrambled motion displays to the target PLD walker, for which the participants are asked to report the walking direction. This design allowed us to examine the abstract decision signals without contamination by the large visually evoked potentials due to the sudden introduction of the visual targets. To manipulate difficulty, PLD walkers are embedded in noise dots that move randomly without conveying any meaningful human form or motion dynamics. In addition, temporal expectation regarding stimulus onset is prevented by the changing pre-stimulus duration. Behavioral results showed that an increase in noise level (i.e., difficulty) resulted in slower RTs with more misses or incorrect responses. Moreover, if the targets are preceded by increased pre-stimulus duration, faster decisions are observed. The effect of pre-stimulus duration also showed differences for the response categories. In particular, participants tend to respond prematurely with increasing pre-stimulus intervals. This

is potentially due to the imposing urgency-related processes to respond as soon as possible during increased waiting intervals, and ineffective preparation if the target stimulus appears after a short duration.

Fitting DDM to behavioral data revealed that (v_stim) model produced a superior fit when the data is separated by the noise level. It should be noted that the minimum trial number requirement is not satisfied when the datasets are partitioned into 12 levels according to the combination of noise level and pre-stimulus duration. Yet another nested comparison is also possible through partitioning the RT data from each subject according to pre-stimulus duration and allowing the free parameters to vary between its levels. In this nested comparison, the model in which non-decision time depends on the pre-stimulus duration (t0_prestim) provided a superior fit in most participants. Although not in the same nested structures, the comparison between (v_stim) and (t0_prestim) models revealed that (v_stim) produces a superior fit in 9 out of 16 participants. Specifically, v_stim has the df of 38, whereas the to_prestim has 28. As such, these models are evaluated at the critical value of 23.21 from the Chi-Square table with 10 df and at 0.99 significance level. Additionally, due to noise level producing larger effect sizes for significant main effect when inspected with behavioral metrics (RT and accuracy), (v_stim) has been selected as the winning model. After determining the winning model, it is revealed that the drift rate estimates are significantly higher in easier trials, accompanying a lower number of noisy dots. The visual inspection of the cumulative density functions produced by the empirical data and predicted by the model fitting revealed a decent overlap in participants who demonstrated better goodness-of-fit measures. However, parameter recovery is not as successful as comparing CDFs. Unsuccessful parameter recoveries may arise from the fact that ideal decisions to be investigated with DDM should span hundreds of milliseconds but not more than several seconds [19], as in our task. Nevertheless, DDM estimates provided informative indices for psychometric evaluations [20].

Neural measurements are taken through EEG, given its high temporal precision, to examine the underlying mechanisms occurring in time. Pre-processing is applied carefully to ensure a high signal-to-noise ratio in single-trial measures.

From the pre-processed datasets, single-trial measures of pre-stimulus attention fluctuation and motor preparation levels are calculated. Moreover, for the ERPs that follow stimulus onset, biological motion processing-related negativity (N2-bm) and CPP are calculated for each trial. It has been shown that pre-stimulus attention levels significantly differ between pre-stimulus durations and noise levels, as longer waiting intervals and higher difficulties lead to lower alpha power in posterior-occipital regions. This might be due to reduced suppression of scrambled motion processing in ambiguity, as when pre-stimulus durations are extended and stimuli become harder to evaluate with an increased number of moving dots. There were no significant relationships between RT and experimental manipulations observed with pre-stimulus motor preparation, despite previous research connecting it with starting point adjustments in the DDM framework. However, those studies have common manipulations, including changes in expectation or prior information [39] [40], in contrast to our experimental design, which includes no explicit cueing relative to the stimulus onset or stimulus properties.

For the ERP waveforms that are thought to reflect neural representation of the evidence accumulation process, CPP showed build-to-threshold dynamics, scaled with the evidence strength, covary with RT, and correlated with DDM drift rate estimates, which will be detailed below in a devoted section. An additional neural marker shown to display build-to-threshold dynamics scaling with the difficulty [12], namely lateralized readiness potential (LRP), is also investigated. The marked time points showing scaled slopes with evidence strength are found to temporally lag that of CPP. However, these time windows could not survive after FDR-correction. As for the N2-bm peak negativity, which is interpreted as reflecting visual encoding quality specific to higher-level processing of biological motion, larger negativities are observed in lower noise conditions. Moreover, N2-bm shows RT-predictive qualities to some extent, as larger N2-bm negativities are observed in faster decisions. Lastly, single-trial pre-stimulus attentional fluctuations are found to be significantly correlated with N2-bm peak negativity and CPP activity. A further test investigating whether the effect of attention on evidence accumulation is mediated by the visual encoding quality demonstrated at least a partial mediation, which has important implications that

will also be detailed in the devoted section.

4.2 CPP as a Neural Representation of Decision Variable

We endeavored to isolate the signal showing the decision variable characteristics that is integrated over time through evidence accumulation dynamics. This is challenging due to overlapping yet functionally distinct activities observed in the brain, for which the problem is even more of a concern if used non-invasive techniques, as in human decision-making tasks. Additionally, accumulation dynamics are also proposed to be at play not only in perceptual decision-making but in detecting signals [85], evaluating confidence, and even perceptual consciousness [86]. Keeping this caveat in mind, researchers should design careful experiments [87]. Informed by the previously suggested best practices in the field [12], we utilized an experimental design with biological motion stimuli during a continuous visual stream with varying difficulty.

Previous research has shown the decision variable properties from the centro-parietal areas (CPP), as the signal demonstrates build-to-threshold dynamics scaling with the evidence strength. By using socially meaningful stimuli, we have extended the prior findings on this account. Specifically, CPP amplitude significantly differed between the noise levels, with smaller activity observed in more difficult conditions. However, this modulation has not extended temporally to time windows closely preceding the decision report. This finding further supports a common threshold view. As for the build-up rate of CPP, it has been shown in both stimulus and response-locked waveforms that the CPP rising slope differs between the noise levels. In particular, we have found time windows where the steeper rises are observed in easier conditions. The CPP signal also reliably differentiated between correct and invalidly responded trials, by having larger mean amplitudes in trials accompanying correct responses. Another important property of a decision variable is that it covaries with RT. In this respect, when

the RTs are median-split within each participant and each noise level, the CPP slope has been shown to be steeper in fast decisions. More fine-grained analysis on RT by within-subject z-scoring and partitioning into 6 RT bins also revealed that the CPP slope is significantly steeper in faster decisions within each noise level. It is also worth connecting the CPP slope with the drift rate parameter in DDM, which reflects the speed of information uptake. When the relationship between these two measures is analyzed, it is shown that even after accounting for the effect of difficulty or subject-level variation, the significantly positive relationship still holds.

It should be noted that finding neural signals showing choice-predictive characteristics is not sufficient to disentangle their role in the decision process due to functionally different stages being involved (such as encoding of the stimulus, evidence accumulation, and motor response production). Despite N2-bm showing RT predictive characteristics in median-split analysis, this is not as precise as CPP for the 6 RT bins. Moreover, CPP activity is differentiated temporally from N2-bm, which peaked before CPP peak latency. There is another important issue of leakage between the perceptual/cognitive systems and motor signatures, which might falsely be interpreted as an abstract decision variable [33]. In this respect, LRP is also evaluated as it indexes motor preparation for the decision alternatives when mapped to specific effectors. However, no significant modulation by the evidence strength was observed in LRP after FDR correction. If one were to challenge using correction for multiple comparison for producing overly conservative results to prevent Type-I errors, even in that case, the marked windows still lag behind that of CPP. As a result, our findings provided further support for CPP to reflect the evolution of an abstract decision variable distinctively from the other processing stages underlying perceptual decisions, such as visual encoding and motor preparation.

It should also be noted that reaching a common threshold at response may not be the core property expected from a decision variable. For instance, a previous research utilizing single-trial EEG markers maximizing target stimulus discrimination showed that the amplitude of this neural marker is modulated by the difficulty near decision reports [66]. Additionally, it is difficult to find support

for the common threshold near decision reports where the speed is emphasized or during ambiguous stimulus presentations. Based on this perspective, CPP has been thought to index both the decision variable and the urgency signal growing over time in a multiplicative manner during those circumstances. This notion is further supported by the findings that CPP reaches high amplitudes even in the face of ambiguous sensory evidence, despite chance-level accuracy [17]. Another evidence supporting the urgency component comes from a monkey study investigating the area LIP, which shows evidence-independent rise for both choice alternatives modulated by the speed emphasis [88]. Although not found sufficient support in human studies, another mechanism that produces similar predictions to that of growing urgency is collapsing thresholds in time, which leads to loosening of the decision criteria in timely conditions [89]. Despite the caveats regarding common bound assumption, it is not counterintuitive that our results showed no amplitude modulation by the evidence strength near decision reports, given that the experimental design does not include any specific instructions within or across participants to emphasize speed over accuracy, and participants are given plenty of time to decide walking direction. However, our behavioral findings suggest a possible urgency mechanism is at play for the longer pre-stimulus durations, which may account for the observed patterns regarding fast RTs and more premature responses. Despite this plausible explanation, we could not explore this potential thoroughly in CPP waveforms following longer scrambled motion displays. This is due to a lack of distinct event markers at the response onset that are specific to the levels of pre-stimulus durations.

Some accounts claimed that motor preparation may at least partly coincide with the evidence accumulation process [90], and researchers model this process as secondary with leaky accumulation of the input coming from the decision process to reach the motor bound for triggering the response [91]. The sequential assumption of these processes is also targeted by another study using rats, which regards the decision process as a race between evidence accumulation and action-triggering mechanisms [92]. However, our study has not found any significant time windows where the build-up rate of LRP is modulated by the evidence strength. Therefore, our particular design without reliance on choosing between the motor

plans but rather focusing on perceptual categorization [22] may have prevented motor-related signals from showing similar dynamics in time, coinciding with that of CPPs.

4.3 The Role of Pre-stimulus Attention

It is difficult to reconcile our findings regarding the positive relationship between pre-stimulus alpha power and CPP with previously shown opposite patterns [12], particularly given the similarities between experimental designs. This discrepancy likely arises from task-specific properties modulating how alpha activity functions in pre-stimulus intervals. In our paradigm, the scrambled motion displays are shown before the PLD walkers to enable smooth transitions from control to the target stimuli. However, this choice might have increased the perceptual complexity by introducing ambiguous stimuli that closely resemble targets than typical random-dot-motion controls. As a result, participants may be more prone to false alarms during scrambled displays, especially in longer pre-stimulus durations. To prevent misinterpreting the irrelevant stimulus as the target to be decided upon, participants may have adopted a more cautious strategy. This may necessitate attentional disengagement during the scrambled motion period. This strategy could manifest as higher alpha power, which has been linked to functional inhibition of sensory processing [93] and higher perceptual thresholds [94]. From this perspective, increased alpha power may reflect successful suppression of the irrelevant visual input rather than fluctuations pointing to a lack of attentional engagement.

Supporting these claims, we observed that alpha power significantly varied with pre-stimulus duration, with longer durations associated with lower alpha-band activity. This pattern may arise from the temporal uncertainty about the stimulus timing and potentially reflect reduced confidence with increased duration to evaluate whether the visual stimulus is a task-irrelevant scrambled motion display or a task-related PLD walkers. The relationship between the neural markers further supports this interpretation. That is, when the processing of irrelevant

stimuli is successfully inhibited (higher alpha power), enhanced visual encoding quality of the targets (larger N2-bm peak negativity) is observed. Moreover, larger CPP mean amplitudes, pointing to information processing strength, are preceded by higher alpha power. Thus, in our task, higher alpha power may not necessarily reflect reduced attention, but rather a protective gating mechanism [95] that improves subsequent sensory encoding and decision-related processes.

To further understand these neural dynamics, it is demonstrated that there is at least partial modulation by N2-bm on the relationship between alpha and CPP. In our task, larger alpha power appears to have a functionally adaptive role by means of inhibition for the irrelevant scrambled motion processing. It has been previously argued that attention improves perceptual decisions by increasing the quality of perceptual encoding, thereby leading to more robust visual short-term memory traces, and hence more efficient information processing [96]. In the context of our design, attentional mechanisms act as suppressors for the processing of scrambled motion to preserve the acclaimed pipeline instantiated in visual and decision-making streams. As for the link between these two streams, previous research has proposed a gating mechanism, where the decision-related neurons showing accumulation to threshold pattern receive input from the visual encoding neurons [97]. This view postulates that visual encoding should reach a sufficiently high quality before the accumulation onset for its output to be fed into. In terms of the accumulation onsets, some researchers used the latency where the CPP signal reaches its 50% peak amplitude [98]. It is noteworthy that we observed a temporal overlap between higher-level visual encoding and accumulation onset latencies. In particular, 50% peak latency of CPP at 408ms and N2-bm peak negativity at 418ms is observed following the target onset in grand-average waveforms. Therefore, this close match in temporal dimension supports the outlined dynamics between visual encoding and the accumulation process.

4.4 Joint Models of Neural and Behavioral Data

Our study demonstrated that drift rate estimates from DDM are significantly and positively correlated with aggregated single-trial CPP slopes even after the effect of difficulty is regressed out. Although this approach improved the interpretability of CPP as reflecting the accumulation of an abstract decision variable, it is not the only way to connect neural and behavioral data. As such, initial work focused on improving the theoretical assumptions underlying the decision-making models informed by the neural characteristics. For instance, leaky competing accumulator models as an extension of classical DDM are inspired by the ideas of neurally plausible properties such as lateral inhibition between the neighboring neuron clusters and leaky integrate-and-fire dynamics [99]. A two-stage approach to find the neural correlates of model parameters also attracted the researchers. In this framework, first, the model parameters are estimated, and then their relationship with neural measures is evaluated by means of a subsequent correlation or regression analysis, as in our study. However, there are inherent limitations in this approach. Particularly, it does not allow enforcing constraints on the model structure or parameters through neural variations that have crucial information [100]. For instance, neural data may disentangle the roles of different parameters on the very same phenomenon, such as fast RTs. It might be that the accumulation process starting closer to the bound or enhanced drift rates may lead to faster decisions in some conditions than in others [101]. In addition, the 2-stage approach is not well-suited for studying single-trial data but rather aggregated measures. Furthermore, statistical constraints may also be reciprocal between the neural and behavioral parts, which further allows reliable inferences and interpretations regarding the underlying processes. Connecting with Marr's levels of analysis, showing computationally optimal and at the same time biologically realizable algorithms by a bidirectional link is crucial in understanding cognitive functioning [35].

In this respect, firing rates in FEF visual neurons were previously used directly as inputs to the accumulation process in a study with intriguing methodology by

Purcell and colleagues (2010) [97]. A combination of direct neural input and cognitive modeling suggested that there might be a gating mechanism with leakage between the visual stream responsible for encoding and the motor stream responsible for committing to a decision. Moreover, neural data can be as influential as to give explanatory power to empirically test or refine model assumptions [87]. In this regard, another study used the neural measures to directly constrain and validate a perceptual decision-making model [39]. In particular, motor preparation levels before the target stimulus onset determined the starting point parameter, whereas motor signals during response were used to constrain decision bounds. After establishing the model structure and parameter space informed by the neural data, the decision variable (DV) is simulated with the estimated parameters. These simulated DVs were then compared with CPP waveforms to validate the findings. It has been revealed that a neurally-informed model is superior in terms of capturing the behavioral data and CPP dynamics compared to conventional DDM with the same number of parameters. Recent advancements in joint modeling endeavors also give rise to mathematically more advanced methodologies. For instance, researchers utilized neural networks to simultaneously model behavioral and neural (including CPP) data in the context of DDM. This integrative approach was used to jointly describe EEG signals and behavioral outcome on single-trial level [90].

Connecting with our findings of task-dependent (N2-bm) and domain-general (CPP) neural signals, it would be noteworthy to evaluate the modeling of perceptual decisions within more advanced frameworks. For instance, similar to the direct input methodology used from FEF neurons, the sensory encoding quality extracted from the N2-bm signal can be directly fed into the evidence accumulation process in biological motion tasks. In addition, pre-stimulus alpha power may inform the parameters in the subsequent trial, as more successful suppression of scrambled motion may lead to elevated single-trial drift rates. On the other hand, within-trial fluctuations in the CPP signal may inform the diffusion coefficient parameter in the DDM framework as a reflection of internal noise. Furthermore, online evaluation of single-trial neural dynamics thought to underlie decisions may even inform brain-computer interface research. To illustrate,

in case of low-quality visual encoding or high attentional fluctuations, adaptive feedback may improve the performance [102]. Therefore, joint modeling of neural and behavioral data offers great promise for further research in both fine-grained understanding of perceptual decision-making dynamics and other technological applications.

4.5 Future Studies

The current study has important implications in revealing both domain-specific ERPs that reflect visual encoding quality and a domain-general abstract decision signal. It has been demonstrated that single-trial N2-bm peak negativities and CPP mean amplitudes are significantly correlated. Moreover, temporal investigation revealed that the onset of CPP and N2-bm peak latency roughly overlap. Future research should more extensively examine the hypothesis regarding the relationship between visual processing and evidence accumulation. In this context, it would be helpful to analyze connectivity between the centro-parietal areas and the right-lateralized posterior-occipital regions. Specifically, in case significant connectivity is found between the regions that coincides with the temporal window of interest (around 400ms), it would strengthen the claims for the notion that the decision-related accumulation process receives inputs from task-specific higher-level visual encoding areas, such as pSTS for biological motion perception tasks [61].

Our findings showed that if the scrambled motion is displayed for 5 or 8 seconds, the mean RTs are lower compared to 3 seconds. That is, the participants are responding more slowly at the lowest level of pre-stimulus duration. However, there were no significant differences found between 5 and 8 seconds. This pattern is interesting given the differences in seconds between these two categories. To more precisely investigate how urgency or attention-related processes might change with different pre-stimulus durations, future research might utilize a randomly jittered waiting time in a specified window, rather than using 3 fixed categories as in this study. This way, more insightful patterns might be observed

between all levels of pre-stimulus durations. This design choice may, in turn, lead to variations in single-trial neural measures related to evidence accumulation and attention.

A significant portion of previous research on biological motion processing is devoted to understanding the relative roles of form and motion cues underlying the remarkable performance in humans. Some accounts view local and relative motion dynamics as of great importance for BM perception, as PLDs are not as accurately and efficiently perceived without being seen in motion [103]. On the contrary, some researchers put great emphasis on form cues and revealed that BM perception can be realized even with static snapshots of PLDs, depending on the postural configuration [104]. There is also another line of research that underscores the integration of form and motion information into a coherent percept to explain neural and behavioral phenomena [65], [105], [106]. Nevertheless, the consensus is yet to be reached, and studying biological motion perception in a decision-making framework and with relevant markers offers great promise. Particularly, as outlined previously, fitting the behavioral data to DDM results in cognitively interpretable parameters. This approach is more fruitful in uncovering the underlying mechanisms than studying accuracy or RT alone in isolation. Additionally, investigation of the relevant experimental designs to address form, motion, or both cues offer another prospect to study how the decisions are unfolding temporally. Specifically, we have shown that research including abstract decision signals - CPP - can be extended to the biological motion domain using PLDs. Therefore, perturbations targeting different sources of information believed to be critical in biological motion perception may be reflected in CPP waveforms during decisions. For instance, disturbed form or motion information in biological motion detection tasks might lead to attenuated CPP amplitudes or reduced slopes, or even a turnaround effect that has been observed in some tasks when the evidence favors another alternative in the middle of the trial [13].

Lastly, although walking direction discrimination tasks are widely employed in BM perception studies, PLD walkers constitute only a limited part of the whole range of biological actions. PLDs offer great flexibility to convey other action types, including object manipulation, locomotion, or even communicative actions.

It has been thought that ERPs observed during the candidate time windows for N2-bm (i.e., around $300-600$ms following the stimulus onset) reflect higher-level visual processing of BM displays [57]. Hence, it would be worth examining how the tasks that requires semantically more complex processing (e.g., understanding intentions, socially communicative actions, etc.) will lead to changes in amplitude or temporal qualities of N2-bm. In addition, if a larger variation is expected to be seen with respect to more complex tasks, the relationship between N2-bm and CPP can be understood in a more precise manner.

4.6 Limitations

Several limitations should also be kept in mind to evaluate the findings of the study. Firstly, RT measures are not as precise as expected from robust fitting procedures for drift-diffusion modeling. In particular, due to problems in experiment coding, RTs are taken in 35ms steps, and this problem is targeted by adding an infinitesimal jitter and using Chi-Square (CS) based model fitting that does not require the full RT distribution, rather several bins. However, the CS-based approach provides limited power for reliable parameter estimations. Indeed, although visual inspection of empirical versus predicted cumulative density functions suggests overlap in some participants, the parameter recoveries were not successful in containing estimated values within their 95% CIs in most of the participants. Therefore, it should be noted that the parameter estimates are not robust enough to ground the claims on a strong basis for theory development [20]. More ideally, when the trial numbers are relatively lower, as in our design, the researchers are increasingly attracted to Bayesian hierarchical models [41] [107]. Considering problems related to parameter recovery in our study, future research with more precise single-trial RT measures should use a Bayesian approach, given the advantages in terms of providing full posterior distributions for quantifying uncertainty and parameter regularization through group-level hyperparameters.

On the other hand, although PLDs offer great flexibility and have established

usefulness, there are some inherent problems caused by the low-level representation of human motion patterns. We have interpreted the N2-bm as reflecting the visual encoding quality, and developed a view on the time-course of decisions based on the relationship between alpha power, N2-bm peak negativities, and CPP amplitudes. However, it might be that the walking direction can be deduced only by the local motion patterns of the feet [62], [108]. In that case, the global motion processing becomes limited, and higher-level processing is no longer required to extract walking direction. However, in our task, relying on local motion patterns of feet is only relevant after stimulus onset, as participants should detect the seamless transitions to targets from scrambled motion displays to produce correct responses. Otherwise, this would result in premature responses. In addition, inspection of mean RT (1200ms) revealed a substantial amount of time spanning several hundred milliseconds before the response, which is not suggestive of utilizing those low-level cues that should have produced faster RTs. Still, this is a noteworthy caveat that might be addressed in the future by designing experimental paradigms to prevent potential reliance on such cues as in [109] with the stimuli having distinct facing and movement directions to understand the global versus local processing.

4.7 Conclusion

It has been demonstrated that CPP is a domain-general neural representation of an abstract decision variable, showing the expected qualities in response to socially meaningful biological motion stimuli. Additionally, the rise of CPP is found to be delayed in our experimental paradigm compared to a similar design, potentially reflecting the additional perceptual demands imposed by higher-level visual stimuli than random dot kinematograms [12] and highlighting how CPP responds to task difficulty. In terms of task-specific neural markers, our findings suggest that N2-bm peak negativity indexes the quality of visual encoding for the biological motion stimulus. From this perspective, the observed relationship between N2-bm and CPP indicates that enhanced visual encoding of the targets is associated with a more robust evidence accumulation process. Additionally, the

temporal overlap between CPP's 50% and N2-bm's peak latencies is observed, which requires further investigation for the functional connectivity between the regions from which these signals originate. In this regard, future research could provide important insights regarding whether and how the visual system sends its output to the decision-making process. As for the effects of pre-stimulus attention on the subsequent trial performance, nuanced patterns are observed, necessitating investigations along with task-specific demands. Specifically, the findings are in line with the accounts for the functional inhibitory role of pre-stimulus attention for the processing of task-irrelevant scrambled motion displays. In relation to that, the trials accompanying enhanced visual encoding and evidence accumulation are found to be associated with successful suppression of scrambled motion before the targets. In conclusion, this study has extended the previous literature on perceptual decision-making by introducing point-light displays, opening ways for promising future research to explore the intricate neural dynamics underlying perceptual decisions and biological motion perception.

Bibliography

[1] C. Donkin, S. Brown, and A. Heathcote, "Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator," *Journal of Mathematical Psychology*, vol. 55, p. 140–151, Apr. 2011.

[2] G. M. Stine, A. Zylberberg, J. Ditterich, and M. N. Shadlen, "Differentiating between integration and non-integration strategies in perceptual decision making," *eLife*, vol. 9, p. e55365, Apr. 2020.

[3] J. Bolam, J. A. Diaz, M. Andrews, R. O. Coats, M. G. Philiastides, S. L. Astill, and I. Delis, "A drift diffusion model analysis of age-related impact on multisensory decision-making processes," *Scientific Reports*, vol. 14, p. 14895, June 2024.

[4] R. Ratcliff, P. Gomez, and G. McKoon, "A diffusion model account of the lexical decision task.," *Psychological Review*, vol. 111, no. 1, p. 159–182, 2004.

[5] M. Milosavljevic, J. Malmaud, A. Huth, C. Koch, and A. Rangel, "The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure," *Judgment and Decision Making*, vol. 5, p. 437–449, Oct. 2010.

[6] D. H. Arabadzhiyska, O. G. B. Garrod, E. Fouragnan, E. D. Luca, P. G. Schyns, and M. G. Philiastides, "A common neural account for social and nonsocial decisions," *Journal of Neuroscience*, vol. 42, p. 9030–9044, Nov. 2022.

[7] M. N. Shadlen and R. Kiani, "Decision making as a window on cognition," *Neuron*, vol. 80, p. 10.1016/j.neuron.2013.10.047, Oct. 2013.

[8] D. P. Hanes and J. D. Schall, "Neural control of voluntary movement initiation," *Science (New York, N.Y.)*, vol. 274, p. 427–430, Oct. 1996.

[9] J. D. Roitman and M. N. Shadlen, "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task," *The Journal of Neuroscience*, vol. 22, p. 9475–9489, Nov. 2002.

[10] G. D. Horwitz, A. P. Batista, and W. T. Newsome, "Representation of an abstract perceptual decision in macaque superior colliculus," *Journal of Neurophysiology*, vol. 91, p. 2281–2296, May 2004.

[11] R. G. O'Connell, P. M. Dockree, and S. P. Kelly, "A supramodal accumulation-to-bound signal that determines perceptual decisions in humans," *Nature Neuroscience*, vol. 15, p. 1729–1735, Dec. 2012.

[12] S. P. Kelly and R. G. O'Connell, "Internal and external influences on the rate of sensory evidence accumulation in the human brain," *Journal of Neuroscience*, vol. 33, p. 19434–19441, Dec. 2013.

[13] K. Afacan-Seref, N. A. Steinemann, A. Blangero, and S. P. Kelly, "Dynamic interplay of value and sensory information in high-speed decision making," *Current biology: CB*, vol. 28, pp. 795–802.e6, Mar. 2018.

[14] M. K. van Vugt, M. A. Beulen, and N. A. Taatgen, "Relation between centro-parietal positivity and diffusion model parameters in both perceptual and memory-based decision making," *Brain Research*, vol. 1715, p. 1–12, July 2019.

[15] J. Herding, S. Ludwig, A. Von Lautz, B. Spitzer, and F. Blankenburg, "Centro-parietal eeg potentials index subjective evidence and confidence during perceptual decision making," *NeuroImage*, vol. 201, p. 116011, Nov. 2019.

[16] D. M. Twomey, P. R. Murphy, S. P. Kelly, and R. G. O'Connell, "The classic p300 encodes a build-to-threshold decision variable," *European Journal of Neuroscience*, vol. 42, no. 1, p. 1636–1643, 2015.

[17] Y. Yau, T. Hinault, M. Taylor, P. Cisek, L. K. Fellows, and A. Dagher, "Evidence and urgency related eeg signals during dynamic decision-making in humans," *Journal of Neuroscience*, vol. 41, p. 5711–5722, June 2021.

[18] R. Ratcliff, "A theory of memory retrieval," *Psychological Review*, vol. 85, no. 2, p. 59–108, 1978.

[19] R. Ratcliff and G. McKoon, "The diffusion decision model: Theory and data for two-choice decision tasks," *Neural computation*, vol. 20, p. 873–922, Apr. 2008.

[20] C. Donkin and S. D. Brown, *Response Times and Decision-Making*, p. 1–33. Wiley, 1 ed., Mar. 2018.

[21] R. Ratcliff and R. Childers, "Individual differences and fitting methods for the two-choice diffusion model of decision making.," *Decision*, vol. 2, p. 237–279, Oct. 2015.

[22] K. K. Lui, M. D. Nunez, J. M. Cassidy, J. Vandekerckhove, S. C. Cramer, and R. Srinivasan, "Timing of readiness potentials reflect a decision-making process in the human brain," *Computational Brain Behavior*, vol. 4, p. 264–283, Sept. 2021.

[23] V. Lerche, A. Voss, and M. Nagler, "How many trials are required for parameter estimation in diffusion modeling? a comparison of different optimization criteria," *Behavior Research Methods*, vol. 49, p. 513–537, Apr. 2017.

[24] R. Ratcliff and J. N. Rouder, "Modeling response times for two-choice decisions," *Psychological Science*, vol. 9, no. 5, p. 347–356, 1998.

[25] R. Ratcliff and F. Tuerlinckx, "Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability," *Psychonomic Bulletin Review*, vol. 9, p. 438–481, Sept. 2002.

[26] G. Tillman, T. Van Zandt, and G. D. Logan, "Sequential sampling models without random between-trial variability: the racing diffusion model of speeded decision making," *Psychonomic Bulletin Review*, vol. 27, p. 911–936, Oct. 2020.

[27] V. Lerche and A. Voss, "Model complexity in diffusion modeling: Benefits of making the model more parsimonious," *Frontiers in Psychology*, vol. 7, Sept. 2016.

[28] T. Richter, R. Ulrich, and M. Janczyk, "Diffusion models with time-dependent parameters: An analysis of computational effort and accuracy of different numerical methods," *Journal of Mathematical Psychology*, vol. 114, p. 102756, June 2023.

[29] J. Ditterich, "Evidence for time-variant decision making," *European Journal of Neuroscience*, vol. 24, no. 12, p. 3628–3641, 2006.

[30] A. K. Churchland, R. Kiani, and M. N. Shadlen, "Decision-making with multiple alternatives," *Nature Neuroscience*, vol. 11, p. 693–702, June 2008.

[31] P. Cisek, G. A. Puskas, and S. El-Murr, "Decisions in changing conditions: The urgency-gating model," *Journal of Neuroscience*, vol. 29, p. 11560–11571, Sept. 2009.

[32] P. L. Smith and R. Ratcliff, "Modeling evidence accumulation decision processes using integral equations: Urgency-gating and collapsing boundaries," *Psychological review*, vol. 129, p. 235–267, Mar. 2022.

[33] J.-H. Song and K. Nakayama, "Hidden cognitive states revealed in choice reaching tasks," *Trends in Cognitive Sciences*, vol. 13, p. 360–366, Aug. 2009.

[34] J. L. McClelland, "On the time relations of mental processes: An examination of systems of processes in cascade," *Psychological Review*, vol. 86, no. 4, p. 287–330, 1979.

[35] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press, July 2010. Google-Books-ID: D8XxCwAAQBAJ.

[36] A. Boldt, A.-M. Schiffer, F. Waszak, and N. Yeung, "Confidence predictions affect performance confidence and neural preparation in perceptual decision making," *Scientific Reports*, vol. 9, p. 4031, Mar. 2019.

[37] C. Kohl, L. Spieser, B. Forster, S. Bestmann, and K. Yarrow, "Centroparietal activity mirrors the decision variable when tracking biased and time-varying sensory evidence," *Cognitive Psychology*, vol. 122, p. 101321, Nov. 2020.

[38] M. D. Nunez, A. Gosai, J. Vandekerckhove, and R. Srinivasan, "The latency of a visual evoked potential tracks the onset of decision making," *NeuroImage*, vol. 197, p. 93–108, Aug. 2019.

[39] S. P. Kelly, E. A. Corbett, and R. G. O'Connell, "Neurocomputational mechanisms of prior-informed perceptual decision-making in humans," *Nature Human Behaviour*, vol. 5, p. 467–481, Apr. 2021.

[40] F. P. d. Lange, D. A. Rahnev, T. H. Donner, and H. Lau, "Prestimulus oscillatory activity over motor cortex reflects perceptual expectations," *Journal of Neuroscience*, vol. 33, p. 1400–1410, Jan. 2013.

[41] T. V. Wiecki, I. Sofer, and M. J. Frank, "Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python," *Frontiers in Neuroinformatics*, vol. 7, Aug. 2013.

[42] C. E. Myers, A. Interian, and A. A. Moustafa, "A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences," *Frontiers in Psychology*, vol. 13, Dec. 2022.

[43] I. Thornton, "Biological motion: Point-light walkers and beyond," *Human body perception from the inside out*, p. 271–305, Jan. 2006.

[44] M. A. Pavlova, "Biological motion processing as a hallmark of social cognition," *Cerebral Cortex*, vol. 22, p. 981–995, May 2012.

[45] N. F. Troje and A. Basbaum, "Biological motion perception," *The senses: A comprehensive reference*, vol. 2, pp. 231–238, 2008.

[46] R. Blake and M. Shiffrar, "Perception of human motion," *Annual Review of Psychology*, vol. 58, p. 47–73, Jan. 2007.

[47] A. P. Saygin, S. M. Wilson, D. J. Hagler, E. Bates, and M. I. Sereno, "Point-light biological motion perception activates human premotor cortex," *The Journal of Neuroscience*, vol. 24, p. 6181–6188, July 2004.

[48] H. Peuskens, J. Vanrie, K. Verfaillie, and G. A. Orban, "Specificity of regions processing biological motion," *The European Journal of Neuroscience*, vol. 21, p. 2864–2875, May 2005.

[49] S. Gilaie-Dotan, R. Kanai, B. Bahrami, G. Rees, and A. P. Saygin, "Neuroanatomical correlates of biological motion detection," *Neuropsychologia*, vol. 51, p. 457–463, Feb. 2013.

[50] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Perception Psychophysics*, vol. 14, no. 2, p. 201–211, 1973.

[51] K. Alaerts, E. Nackaerts, P. Meyns, S. P. Swinnen, and N. Wenderoth, "Action and emotion recognition from point light displays: An investigation of gender differences," *PLOS ONE*, vol. 6, no. 6, p. e20989, 2011.

[52] V. Manera, B. Schouten, C. Becchio, B. G. Bara, and K. Verfaillie, "Inferring intentions from biological motion: A stimulus set of point-light communicative interactions," *Behavior Research Methods*, vol. 42, p. 168–178, Feb. 2010.

[53] R. Blake, L. M. Turner, M. J. Smoski, S. L. Pozdol, and W. L. Stone, "Visual recognition of biological motion is impaired in children with autism," *Psychological Science*, vol. 14, p. 151–157, Mar. 2003.

[54] J. Kim, M. L. Doop, R. Blake, and S. Park, "Impaired visual recognition of biological motion in schizophrenia," *Schizophrenia Research*, vol. 77, p. 299–307, Sept. 2005.

[55] Okruszek, "It is not just in faces! processing of emotion and intention from biological motion in psychiatric disorders," *Frontiers in Human Neuroscience*, vol. 12, Feb. 2018.

[56] C. Bidet-Ildei, A. Decatoire, and S. Gil, "Recognition of emotions from facial point-light displays," *Frontiers in Psychology*, vol. 11, 2020.

[57] A. I. Krakowski, L. A. Ross, A. C. Snyder, P. Sehatpour, S. P. Kelly, and J. J. Foxe, "The neurophysiology of human biological motion processing: A high-density electrical mapping study," *NeuroImage*, vol. 56, p. 373–383, May 2011.

[58] N. C. White, J. M. Fawcett, and A. J. Newman, "Electrophysiological markers of biological motion and human form recognition," *NeuroImage*, vol. 84, p. 854–867, Jan. 2014.

[59] M. Hirai, A. Senju, H. Fukushima, and K. Hiraki, "Active processing of biological motion perception: an erp study," *Cognitive Brain Research*, vol. 23, p. 387–396, May 2005.

[60] M. Hirai and K. Hiraki, "The relative importance of spatial versus temporal structure in the perception of biological motion: An event-related potential study," *Cognition*, vol. 99, p. B15–B29, Feb. 2006.

[61] D. Jokisch, I. Daum, B. Suchan, and N. Troje, "Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis," *Behavioural Brain Research*, vol. 157, p. 195–204, Feb. 2005.

[62] N. F. Troje and C. Westhoff, "The inversion effect in biological motion perception: Evidence for a "life detector"?," *Current Biology*, vol. 16, p. 821–824, Apr. 2006.

[63] M. Pavlova and A. Sokolov, "Orientation specificity in biological motion perception," *Perception Psychophysics*, vol. 62, no. 5, p. 889–899, 2000.

[64] S. M. Thurman and E. D. Grossman, "Temporal "bubbles" reveal key features for point-light biological motion perception," *Journal of Vision*, vol. 8, p. 28, Mar. 2008.

[65] M. Thirkettle, C. P. Benton, and N. E. Scott-Samuel, "Contributions of form, motion and task to biological motion perception," *Journal of Vision*, vol. 9, p. 28–28, Mar. 2009.

[66] M. G. Philiastides, H. R. Heekeren, and P. Sajda, "Human scalp potentials reflect a mixture of decision-related signals during perceptual choices," *Journal of Neuroscience*, vol. 34, p. 16877–16889, Dec. 2014.

[67] O. C. Oguz, B. Aydin, and B. A. Urgen, "Biological motion perception in the theoretical framework of perceptual decision-making: An event-related potential study," *Vision Research*, vol. 218, p. 108380, May 2024.

[68] O. C. Oguz, "Biological motion perception in the framework of perceptual decision-making: An event-related potential study," *(Unpublished Senior Thesis)*, 2023.

[69] D. Ebbert, *chisq.posthoc.test: A Post Hoc Analysis for Pearson's Chi-Squared Test for Count Data*, 2019. R package version 0.1.2.

[70] A. Voss, J. Voss, and V. Lerche, "Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30," *Frontiers in Psychology*, vol. 6, Mar. 2015.

[71] F. P. Leite and R. Ratcliff, "What cognitive processes drive response biases? a diffusion model analysis," *Judgment and Decision Making*, vol. 6, p. 651–687, Oct. 2011.

[72] N. J. Evans, J. S. Trueblood, and W. R. Holmes, "A parameter recovery assessment of time-variant models of decision-making," *Behavior Research Methods*, vol. 52, p. 193–206, Feb. 2020.

[73] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, p. 9–21, Mar. 2004.

[74] J. Lopez-Calderon and S. J. Luck, "Erplab: an open-source toolbox for the analysis of event-related potentials," *Frontiers in Human Neuroscience*, vol. 8, p. 213, 2014.

[75] P. Ablin, J.-F. Cardoso, and A. Gramfort, "Faster independent component analysis by preconditioning with hessian approximations," *IEEE Transactions on Signal Processing*, 2018. arXiv:1706.08171.

[76] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, p. 181–197, Sept. 2019.

[77] G. M. Loughnane, D. P. Newman, M. A. Bellgrove, E. C. Lalor, S. P. Kelly, and R. G. O'Connell, "Target selection signals influence perceptual decisions by modulating the onset and rate of evidence accumulation," *Current Biology*, vol. 26, p. 496–502, Feb. 2016.

[78] E. C. Fields and G. R. Kuperberg, "Having your cake and eating it too: Flexibility and power with mass univariate statistics for erp data," *Psychophysiology*, vol. 57, no. 2, p. e13468, 2020.

[79] M. Eimer, "The lateralized readiness potential as an on-line measure of central response activation processes," *Behavior Research Methods, Instruments, Computers*, vol. 30, p. 146–156, Mar. 1998.

[80] I. M. Thornton, R. A. Rensink, and M. Shiffrar, "Active versus passive processing of biological motion," *Perception*, vol. 31, p. 837–853, July 2002.

[81] J. Thompson and R. Parasuraman, "Attention, biological motion, and action recognition," *NeuroImage*, vol. 59, p. 4–13, Jan. 2012.

[82] P.-C. Bürkner, "brms: An r package for bayesian multilevel models using stan," *Journal of Statistical Software*, vol. 80, p. 1–28, Aug. 2017.

[83] R. Ratcliff, J. Love, C. A. Thompson, and J. E. Opfer, "Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses," *Child Development*, vol. 83, no. 1, p. 367–381, 2012.

[84] A. Heathcote, Y.-S. Lin, A. Reynolds, L. Strickland, M. Gretton, and D. Matzke, "Dynamic models of choice," *Behavior Research Methods*, vol. 51, p. 961–985, Apr. 2019.

[85] R. H. S. Carpenter, B. A. J. Reddi, and A. J. Anderson, "A simple two-stage model predicts response time distributions," *The Journal of Physiology*, vol. 587, no. 16, p. 4051–4062, 2009.

[86] M. Pereira, D. Perrin, and N. Faivre, "A leaky evidence accumulation process for perceptual experience," *Trends in Cognitive Sciences*, vol. 26, p. 451–461, June 2022.

[87] R. G. O'Connell, M. N. Shadlen, K. Wong-Lin, and S. P. Kelly, "Bridging neural and computational viewpoints on perceptual decision-making," *Trends in Neurosciences*, vol. 41, p. 838–852, Nov. 2018.

[88] T. Hanks, R. Kiani, and M. Shadlen, "A neural mechanism of speed-accuracy tradeoff in macaque area lip," *eLife*, vol. 2014, no. 3, 2014.

[89] A. Voss, V. Lerche, U. Mertens, and J. Voss, "Sequential sampling models with variable boundaries and non-normal noise: A comparison of six models," *Psychonomic Bulletin Review*, vol. 26, p. 813–832, June 2019.

[90] A. Ghaderi-Kangavari, J. A. Rad, and M. D. Nunez, "A general integrative neurocognitive modeling framework to jointly describe eeg and decision-making on single trials," *Computational Brain Behavior*, vol. 6, p. 317–376, Sept. 2023.

[91] S. Verdonck, T. Loossens, and M. G. Philiastides, "The leaky integrating threshold and its impact on evidence accumulation models of choice response time (rt)," *Psychological Review*, vol. 128, p. 203–221, Mar. 2021.

[92] L. Hernández-Navarro, A. Hermoso-Mendizabal, D. Duque, J. de la Rocha, and A. Hyafil, "Proactive and reactive accumulation-to-bound processes

compete during perceptual decisions," *Nature Communications*, vol. 12, p. 7148, Dec. 2021.

[93] B. F. Händel, T. Haarmeier, and O. Jensen, "Alpha oscillations correlate with the successful inhibition of unattended stimuli," *Journal of Cognitive Neuroscience*, vol. 23, p. 2494–2502, Sept. 2011.

[94] J. J. Foxe and A. C. Snyder, "The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention," *Frontiers in Psychology*, vol. 2, July 2011.

[95] J. M. Zumer, R. Scheeringa, J.-M. Schoffelen, D. G. Norris, and O. Jensen, "Occipital alpha activity during stimulus processing gates the information flow to object-selective cortex," *PLOS Biology*, vol. 12, no. 10, p. e1001965, 2014.

[96] P. L. Smith and R. Ratcliff, "An integrated theory of attention and decision making in visual signal detection," *Psychological Review*, vol. 116, no. 2, p. 283–317, 2009.

[97] B. A. Purcell, R. P. Heitz, J. Y. Cohen, J. D. Schall, G. D. Logan, and T. J. Palmeri, "Neurally constrained modeling of perceptual decision making," *Psychological Review*, vol. 117, no. 4, p. 1113–1143, 2010.

[98] D. Rangelov and J. B. Mattingley, "Evidence accumulation during perceptual decision-making is sensitive to the dynamics of attentional selection," *NeuroImage*, vol. 220, p. 117093, Oct. 2020.

[99] M. Usher and J. L. McClelland, "The time course of perceptual choice: the leaky, competing accumulator model," *Psychological Review*, vol. 108, p. 550–592, July 2001.

[100] J. J. Palestro, G. Bahg, P. B. Sederberg, Z.-L. Lu, M. Steyvers, and B. M. Turner, "A tutorial on joint models of neural and behavioral measures of cognition," *Journal of Mathematical Psychology*, vol. 84, p. 20–48, June 2018.

[101] B. M. Turner, L. van Maanen, and B. U. Forstmann, "Informing cognitive abstractions through neuroimaging: the neural drift diffusion model," *Psychological Review*, vol. 122, p. 312–336, Apr. 2015.

[102] M. D. Nunez, J. Vandekerckhove, and R. Srinivasan, "How attention influences perceptual decision making: Single-trial eeg correlates of drift-diffusion model parameters," *Journal of mathematical psychology*, vol. 76, p. 117–130, Feb. 2017.

[103] A. Casile and M. A. Giese, "Critical features for the recognition of biological motion," *Journal of Vision*, vol. 5, p. 6, Apr. 2005.

[104] R. Reid, A. Brooks, D. Blair, and R. van der Zwan, "Snap! recognising implicit actions in static point-light displays," *Perception*, vol. 38, no. 4, p. 613–616, 2009.

[105] M. Lappe, "Perception of biological motion as motion-from-form," *e-Neuroforum*, vol. 18, p. 67–73, Sept. 2012.

[106] J. O. Garcia and E. D. Grossman, "Necessary but not sufficient: Motion perception is required for perceiving biological motion," *Vision Research*, vol. 48, p. 1144–1149, Apr. 2008.

[107] A. Fengler, Y. Xu, K. Bera, A. Omar, and M. J. Frank, "HSSM: A generalized toolbox for hierarchical Bayesian estimation of computational models in cognitive neuroscience." In preparation, 2025.

[108] D. R. Saunders, J. Suchan, and N. F. Troje, "Off on the wrong foot: Local features in biological motion," *Perception*, vol. 38, p. 522–532, Apr. 2009.

[109] L. E. Miller and A. P. Saygin, "Individual differences in the perception of biological motion: Links to social cognition and motor imagery," *Cognition*, vol. 128, p. 140–148, Aug. 2013.

Appendix A

Supplementary Information

For the supplementary figures see A.1, A.2, A.3. Supplementary video on experimental stimuli depicting scrambled motion seamless transitioning to biological motion PLD walkers at the lowest noise level (noise:10) can be found here: https://osf.io/796fe/files/osfstorage.

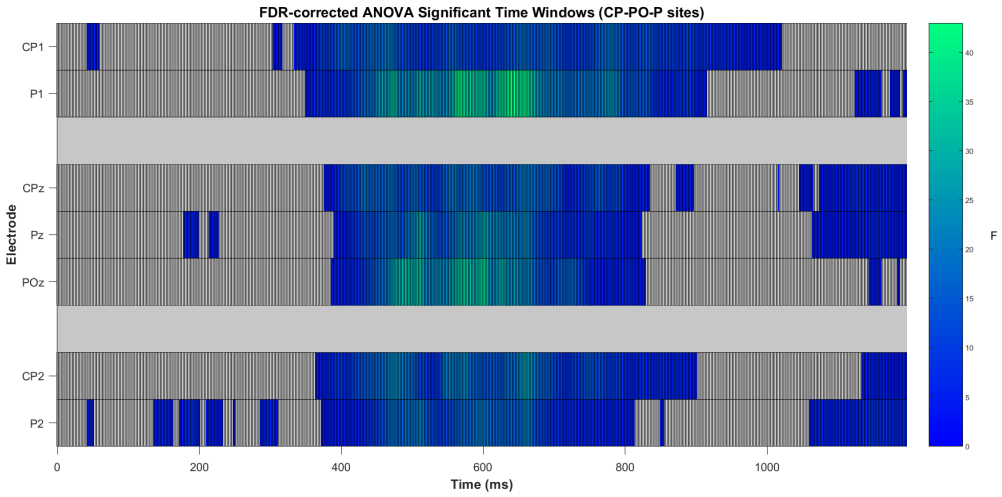

Figure A.1: According to false discovery rate-corrected ANOVA analysis, colored time points mark the significant amplitude differentiation with respect to noise level on stimulus-locked activity at CP-PO-P channels

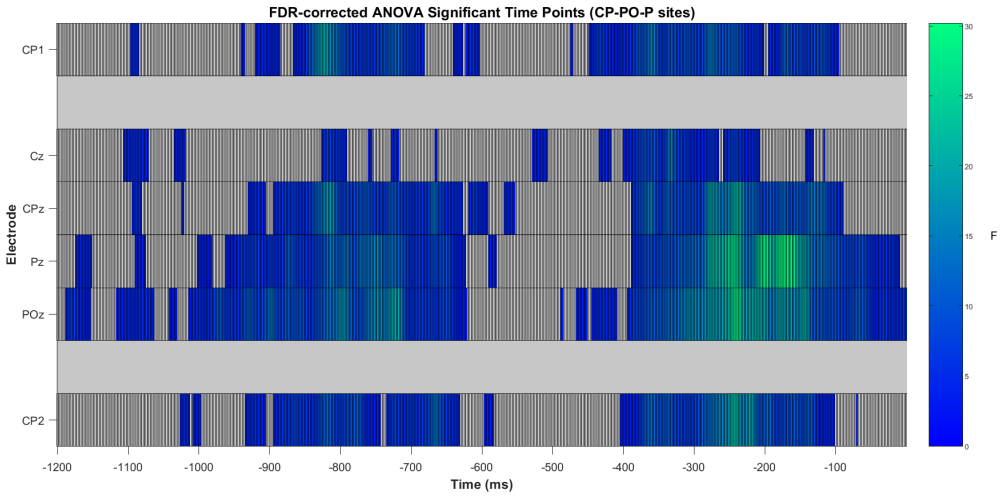

Figure A.2: According to false discovery rate-corrected ANOVA analysis, colored time points mark the significant amplitude differentiation with respect to noise level on response-locked activity at CP-PO-P channels

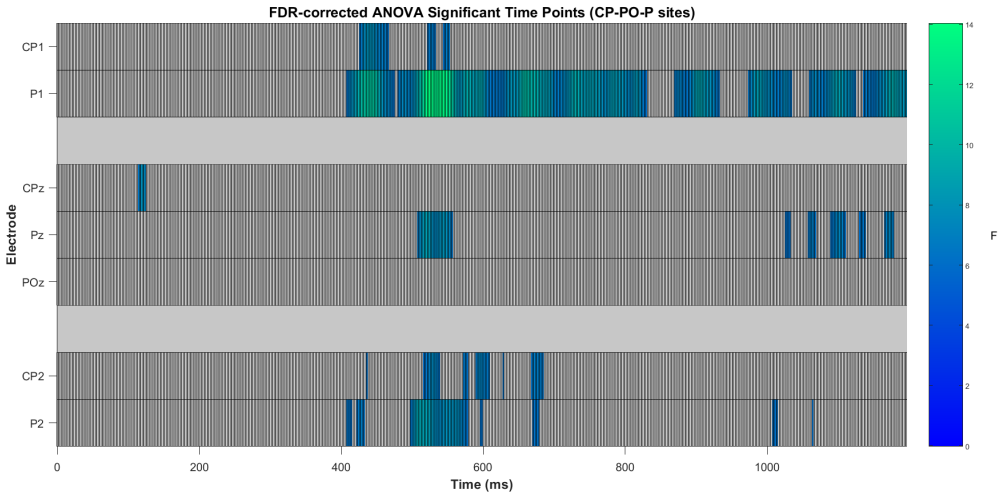

Figure A.3: According to false discovery rate-corrected ANOVA analysis, colored time points mark the significant amplitude differentiation with respect to pre-stimulus duration on stimulus-locked activity at CP-PO-P channels