UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Weber's Law and movement-correlates of decision confidence in human decision making

Íris Isabel Coimbra Damião

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Biofísica Médica e Fisiologia de Sistemas

Dissertação orientada por:
Prof. Dr. Alfonso Renart
Prof. Dr. Alexandre Andrade

ANO 2020

Acknowledgements

I would like to thank Juan Castiñeiras-de Saa and Mafalda Valente for all the incredible help in developing and finishing this work; to my supervisor, Alfonso Renart, for his guidance and for accepting me in his lab during this time, making it such a rich learning experience.

I thank all the lab members for showing me ”how work is done” and for teaching me, basically, all the little I know about neurosciences.

To all the persons in my life that in their own way showed their support, in particular, to Laara and Antonio, because if writing a thesis is already difficult, writing it during a pandemic felt 10 times harder.

Lastly, to my family and to my parents who supported all my education.

Abstract

Every day we collect information through our senses and use that information to make decisions. Thus, understanding how this sensory information is used to guide our behavior becomes extremely relevant, as perceptual decision-making is a central ability of human and animal behavior.

In particular, this work was dedicated to the study of two regularities observed in animal behavior in a context of sensory decision, promoting its extension to human behavior. The first, called TIED - time-intensity equivalence in discrimination - describes how reaction times involved in sensory intensity discrimination change as a function of the overall magnitude of the stimuli being discriminated. The identification of TIED in an experiment developed in rodents allowed to determine the specificities of a mechanism capable of justifying Weber's Law - old psychophysical regularity of the discrimination process. In this work, an adaptation of this experiment was carried out promoting the identification of TIED in human perception, thus extending not only its generality but allowing to determine an underlying mathematical mechanism for sensory discrimination.

The second behavior pattern studied in this work is the modulation of the speed used to indicate the response of a sensory decision. In particular, the identification of the movement's speed - vigor - as a function of sensory strength of evidence follows the same pattern as the level of confidence in a decision found by works dedicated to its study. In this way, a second human behavioral experiment was developed exposing a direct relationship between response movement speed and the level of confidence in the decision. Recent works devoted to the study of vigor in a decision-making context have exposed a relationship between vigor and the value of the reward obtained per unit of time - average reward rate - through the way it determines the computation of cost of time. Therefore, in this experiment a variable reward system and a discount protocol were also included. The purpose of this integration would be to understand how decision confidence, given its implicit relationship with vigor, is integrated into this suggested computation between reward, cost of time, and vigor.

Keywords: TIED, Weber's Law, Vigor, Confidence, Reward and Cost of Time.

Resumo

Várias vezes ao dia nós recolhemos informação através dos nossos sentidos e usamos essa informação para tomar decisões. Entender como essa informação sensorial é utilizada para guiar o nosso comportamento torna-se assim de extrema relevância, sendo esta não só uma capacidade central do comportamento humano, mas também, de todos os animais.

Este trabalho, em particular, dedicou-se ao estudo de duas regularidades observadas no comportamento animal num contexto de decisão sensorial, promovendo a sua extensão para o comportamento humano. A primeira, denominada TIED – time-intensity equivalence in discrimination – descreve como os tempos de reação envolvidos numa discriminação de dois estímulos sensoriais são influenciados em função da sua intensidade absoluta, ditando, mais especificamente, que o único efeito da intensidade absoluta dos estímulos a discriminar é um reescalonamento da unidade temporal deste mesmo processo de discriminação. A identificação do TIED numa experiência desenvolvida em ratos, importantemente permitiu encontrar respostas sobre as especificidades concretas do mecanismo capaz de justificar a Lei de Weber – antiga lei psicofísica que dita que a capacidade de discriminar entre dois estímulos sensoriais não depende do valor das suas intensidades absolutas, mas somente do rácio entre o valor destas intensidades. No entanto, sendo a Lei de Weber tão amplamente observada em diversas espécies e modalidades sensoriais, não se poderia concluir sobre a generalidade de um modelo que a procura explicar com base numa única experiência. Deste modo, neste trabalho, uma adaptação para humanos da experiência original foi desenvolvida, onde os sujeitos tinham de indicar qual o estímulo auditivo consideravam ser mais alto, aquele a ser apresentado no lado direito ou no seu ouvido esquerdo. Os seus resultados comprovaram a existência do TIED como regularidade psicofísica também presente na perceção humana, permitindo concluir que o mesmo modelo capaz de explicar o mecanismo que origina a Lei de Weber, cujas especificações do TIED sobre tempos de reação no processo de discriminação permitem determinar, também é o mecanismo matemático envolvido na perceção humana.

O segundo padrão de comportamento estudado neste trabalho consiste na modulação da velocidade utilizada para indicar a resposta de uma decisão sensorial. Em particular, a identificação que o padrão da velocidade do movimento – vigor – em função da força da evidência sensorial apresentado pelos ratos ao realizarem a experiência em cima mencionada, era muito semelhante ao padrão apresentado por valores explícitos de confiança numa decisão em função do seu grau de dificuldade, encontrado por trabalhos dedicados ao seu estudo. Esta observação, assim como conclusões próximas de outros trabalhos (Seideman et al., 2018) levaram-nos a sugerir a existência de uma relação de monotonia entre o vigor de uma resposta sensorial e o nível de confiança sobre a correção dessa mesma decisão. No entanto, e importantemente, os mais recentes estudos que visam entender porque várias espécies realizam as mesmas ações recorrendo de diferentes velocidades, têm relacionado o vigor de uma ação com a computação de um valor associado à passagem do tempo, denominado de custo do tempo. Segundo estas teorias normativas um sujeito decide mover-se mais rapidamente ou mais devagar de acordo com o valor computado para o custo da passagem do tempo, que por sua vez depende do valor da recompensa que foi adquirido com a

realização das mesmas ações no passado. Assim sendo, se um sujeito recebeu no passado uma recompensa maior, ele determina que o ambiente é mais rico, ou seja, que se for "preguiçoso" neste meio que lhe está a oferecer uma grande recompensa ele está a perder mais do que num ambiente que lhe oferece pouco, isto é, onde a taxa média de recompensa é mais baixa e por isso move-se mais rapidamente. Em suma, estas teorias normativas colocam o vigor de um movimento como sendo um equilíbrio entre o maior custo energético associado a uma ação mais rápida e o maior custo temporal associado a uma ação mais lenta, sendo o valor do custo temporal estimado a partir da taxa média de recompensa recebida no passado. No entanto, segundo estas teorias normativas o valor de confiança numa decisão não deveria influenciar o vigor de um movimento.

Este trabalho, procurou deste modo, não só expor a verdadeira relação entre vigor de uma resposta sensorial e o grau de certeza sobre essa decisão; mas também perceber como o valor da confiança poderia ser incluído nesta computação de custo temporal, propondo uma visão em que o custo temporal é influenciado não só pela recompensa recebida no passado, mas sobretudo pela recompensa oferecida no imediato, como proposto por muitos outros trabalhos que têm proposto esta teoria prospetiva da computação do custo temporal, opondo-se à computação retrospetiva acima explicada. Isto é, segundo esta norma prospetiva o custo temporal é mais elevado se o que nos é oferecido no imediato tem um maior valor, como é o caso de um rato com sede que é capaz de gastar mais energia (mover-se mais rápido) para obter água do que para obter comida, se soubesse à priori o que estava a ser oferecido pela sua ação. Ou como é o caso das amplas experiências realizadas em humanos de desconto temporal do valor de uma recompensa, em que é oferecida uma menor quantidade de dinheiro num futuro próximo ou uma maior quantidade de dinheiro num futuro mais longínquo. Sendo observado muitas vezes, que várias pessoas selecionam a opção que lhes permite receber uma menor quantidade de dinheiro no imediato, ilustrando bem como a passagem do tempo tem um custo, diminuindo o valor da recompensa.

Deste modo, se um entender o grau de confiança na correção da decisão sensorial tomada como o nível de certeza que a recompensa oferecida pela resposta correta será recebida, e a recompensa oferecida no imediato determinante deste valor temporal, estaríamos a incluir o valor de confiança na computação de custo do tempo. Deste modo, foi realizada uma experiência que procurou especificamente entender como o valor de uma recompensa (ao tornar este valor variável de trial para trial) e o grau de certeza numa decisão influenciam o vigor da resposta e o custo da passagem do tempo, ao incluir também um protocolo de desconto. Isto é, nesta experiência, sujeitos humanos tinham de ouvir um estímulo auditivo com várias dificuldades associadas, decidir sobre esse estímulo realizando um movimento para indicar a sua resposta e depois decidir quanto tempo estariam dispostos a esperar para receber, não apenas o valor da recompensa associado à decisão (indicado ao sujeito antes da apresentação do estímulo) mas pela possibilidade de receberem uma recompensa extra. Assim, foi possível medir não só a medir como o valor da recompensa oferecida influencia diretamente o vigor, mas também estaríamos a medir de forma direta como esta determina o custo associado à passagem do tempo, ao medirmos a quantidade de tempo que cada sujeito estaria disposto a esperar por uma recompensa extra.

Os resultados obtidos demonstraram que o grau de certeza na decisão a ser tomada está fortemente e diretamente relacionado com o vigor da resposta dada. No entanto, foram inconclusivos sobre o papel da recompensa nesse mesmo vigor e no custo do tempo, ao revelarem a recompensa como um valor não influenciador do vigor e da quantidade de tempo que um está disposto a esperar pela recompensa (custo do tempo). Deste modo, não foi possível concluir sobre como o valor da confiança pode ser incluído nas teorias normativas que têm sido propostas para descrever a seleção do vigor de um movimento. No entanto, é importante referir que estes resultados se devem sobretudo a problemas impossíveis de antecipar na construção da nossa experiência, tanto na maneira como nós definimos o sistema de recompensa

da tarefa, como na maneira como incluímos o protocolo de desconto na mesma, oferecendo no entanto, importantes referências para experiências futuras.

Palavras-chave: TIED, Lei de Weber, Vigor, Confiança, Recompensa, Custo do tempo.

Contents

List of Figures

List of Tables

List of Abbreviations

ABL Average Binaural Level

DC Decision Confidence

DDM Diffusion Decision Model

FT Fixation Time

ILD Interaural Level Difference

ITI Inter Trial Interval

JND Just-noticeable-difference

MS Movement Speed

MT Movement Time

RT Reaction Time

RTD Reaction Time Distribution

ROC Receiver Operating Characteristic

SAT Speed Accuracy Trade-Off

SDT Signal Detection Theory

SPL Sound Pressure Level

TIED Time-intensity Equivalence in Discrimination

WT Waiting Time

WL Weber's Law

2AFC Two alternative forced-choice

Chapter 1

Motivation and Aim

1.1 Motivation

Every day we are called upon to make several decisions. Whether in line at the cafeteria, in front of our closet, or wondering if we should run to catch the bus. We constantly identify potential options, estimate and compare their subjective value and ultimately make a choice. Decision making is unquestionably an essential part of our daily lives and, consequently, an important field of neuroscience. However, just like with many other processes our brain is capable of, a lot of the features of decision-making remain unanswered or not fully understood.

A big part of the decisions we make every day are based on the evidence we collect through our senses. For example, when we decide to switch lines at the supermarket because we see a shorter queue. Or when we decide if it is safe to cross the road. This process where sensory information is used to guide behavior toward the external world is called perceptual decision-making (Hauser and Salinas, 2014). To understand how these decisions are made, one has to understand how the physical characteristics of the stimuli are evaluated and integrated in accordance with the current goals of the internal state of the subject, and used to produce motor responses.

Historically, one of the firsts, if not the first, systematic set of studies on what we now call perceptual decision making were the experiments by Weber on the fist half of the 19th century on the properties of sensory discrimination made by human subjects on the sense of touch (Weber, E. H., 1834). Weber, not only qualitative assessed the subjects' accuracy in idetifying which two of the tactile sensations was stronger, but investigated how this level of accuracy depended on the overall magnitude of the two stimuli. He discovered that the minimal reliably detected difference between one stimulus and a fixed (standard) one - the just-noticeable-difference - was proportional to the magnitude of the standard. This quantitative regularity has turned out to be very general, and has been replicated for discriminations across all sensory modalities and in many species (Laming,1986; Link, 1992; Gescheider, 2013). Based on this subsequent work, one can reformulate Weber's law as stating that the probability of a correct discrimination between two stimuli (i.e if the stimuli are the same or different in their intensities) depends only on the ratio between their intensities. Even though this regularity is strongly established in the perceptual decision-making field, or psychophysics, no principled way had been identified to choose between its many proposed explanations (Pardo-Vazquez et al., 2018).

In the Renart group a perceptual decision-making experiment performed on rodents has been developed with the aim of finding some consensus regarding this question (Pardo-Vazquez et al., 2018). The experiment consisted in training rats to discriminate the relative intensity of sounds at the two ears at various absolute levels. Findings of the developed experiment not only corroborated the relationship that guides discrimination between two sensory stimuli (i.e Weber's Law), but revealed a new sensory regularity named Time-intensity equivalence in discrimination (TIED) describing the relationship between the intensity of a stimulus and the time required to make the perceptual decision. The identification of the TIED restricts the possible explanations to Weber's Law and suggests a biological plausible one that relies on a process of bounded accumulation of evidence, giving some closure to the question. However, the urge to prove the robustness of this finding and its pervasiveness to other species, required the adaptation of the performed experiment to humans. This project arises with this specific goal.

Additionally, in the original perceptual decision-making experiment performed with rats, another feature of the behavior was being largely observed - the easier the condition of the stimulus presented, the faster the animal moved to indicate its choice; that is, higher was his movement speed (also designated vigor) when the answer was correct. Surprisingly, when the animal was incorrect about its decision, it moved slower the easier the condition presented. This counterintuitive pattern found on the animals movement speed is, curiously, the same identified to be presented by response confidence (degree of certainty about the decision made) (Sanders et al., 2016). These observations seem to suggest a monotonic relationship between both of these behavioral variables. However, exposing this relationship by performing an experiment where both movement speed and decision confidence are measured has not been done yet. Working with human subjects has the advantage of allowing the direct collection of explicit measures of decision confidence, making it possible to reveal this relationship. The pertinence of approaching and looking for the true nature of this relationship, it is not only essential to strengthen the understanding of the link between the physical characteristics of the stimuli with the precepts they originate (decision confidence), but also to better understand the decision-making vigor that has been for long left aside in studies of perceptual decision-making.

Only recently studies on decision-making have been emphasizing the importance of vigor, by suggesting a close relationship between the speed of the decision and the subsequent response movement speed, with the cost associated with the passage of time (Niv et al., 2006; Guitar-Masip et al., 2011; Choi et al., 2014; Constantino & Daw, 2015; Summerside et al., 2018; Otto & Daw, 2019). That is, the idea that time comports a cost because it delays the acquisition of the expected reward and future possible rewards. Being this cost higher, higher is the value attributed by the subject to what is at 'stake'. However, different theories disagree in exactly how the value of cost of time is computed. Some suggest that the subject computes a value to how much each unit of time is worth by computing an average of 'how much' was received in the 'near past' by unit of time (Niv et al., 2006; Guitar-Masip et al., 2011; Constantino & Daw, 2015; Beierholm et al., 2013; Otto & Daw, 2019). Another branch of studies on vigor evokes that the cost of time is determined by what is being offered on the immediate (Kawagoe et al., 1998; Xu-Wilson et al., 2009; Hamid et al., 2015; Sackaloo et al., 2015; Summerside et al., 2018; Walton & Bouret, 2018). Another goal of this project is then to disentangle this relationship between cost of time, past reward, and future reward; connecting to what is observed by us and others, regarding the influence of decision confidence on vigor.

1.2 Aim

This project emerged with the specific aim of developing an experiment in humans, identical to one performed with rodents, to prove that the same results are present in other species, and therefore, concluding the universality of the TIED and the model which sets the mechanism behind Weber's Law. However, the results of movement speed in the original rodents task, left out in the study of Weber's Law, raised the idea and aim of pursuing another experiment, once humans can self-report their confidence levels in the decision being made. This creates the opportunity to measure the presented response vigor associated to the report of different confidence levels, analyzing the nature of their relationship. At the same time, it was not possible to develop an experiment to approach vigor, without having into consideration the real significant findings that have recently emerged from the study of this dimension of behavior. Therefore, this project is divided into two distinct experiments with specific and different goals, but which emerged from each other. The first with the direct goal of proving the existence of the TIED in the human sensorial discrimination process, revealing that the proposed mechanism to explain Weber's Law is transversal to multiple species. The second experiment, motivated by the analysis of the pattern presented by animals' movement speed when performing the first experiment, develops with the direct goals of:

1. Proving the movement speed - vigor- to be an implicit measure of decision-making confidence.

2. Addressing the impact of reward in the decision vigor. More precisely the theory that reward is reflected in how a subject values time, the cost of time. Which, by itself, according to recent experiments (Niv et al., 2006), determines the response vigor in a decision-making task.

1.3 Outline

Thus, this dissertation is divided into two major parts, each dedicated to one of the two experiments developed within its scope:

Part I is allocated to the study of Weber's Law and the TIED in humans and has the following outline:

Chapter 2 - Introduction: It provides a summary introduction to the Weber's Law and the two perceptual decision-making theories proposed to explain the process of differentiation between two sensory stimuli - the Signal Detection Theory and Sequential Sampling.

Chapter 3 - Weber's Law is the result of exact temporal accumulation of evidence: describes the experiment's design performed with rodents and its results, which allowed the identification of the TIED. It also contains a brief description of the model's main characteristics proposed to be the computational explanation of Weber's Law, not only necessary for the TIED to hold, but also sufficient to explain with high accuracy the rats' behavior.

Chapter 4 - Human Task: Looking for the TIED: describes the materials and methods employed to accomplish the rodents task's transformation into a human task.

Chapter 5 - Results and Discussion: is where the obtained results for the human adaptation of the task are shown and discussed.

Chapter 6 - Conclusions: focuses on the conclusions drawn from this work, also highlighting the importance of future experiments in other species and sensory modalities, in order to claim the TIED as an inherent psychophysical regularity of the brain.

Part II dedicates itself to the study of vigor and its relationship with decision confidence, reward, and direct link with cost of time.

Chapter 7 - Introduction: provides a summary of the main results found for the relationship between vigor and cost of time by recent studies; the resemblance between the pattern presented by rodents movement speed when performing the Weber's Law experiment and the patter found in decision confidence in other works; at last a brief summary of the experiments performed with temporal discounting conducted in humans. It concludes with the aims of this work.

Chapter 8 - A sound localization task in a changing environment with temporal discounting: describes the materials and methods employed to accomplish this work.

Chapter 9 - Results and Discussion: is where the main results obtained with the performance of this second experiment are shown and discussed.

Chapter 10 - Conclusions: focuses on what can be concluded from the obtained results and highlights its contributions for future works.

At last, **Chapter 11** summarizes the main conclusions of the developed work, that is, of both experiments.

Before moving to the report of the performed work, it is noteworthy that the elaboration and coding of both experiments, data collection, results processing and statistical analysis were all steps carried out by me, author of the thesis.

Part I

Weber’s Law and the TIED in humans

Chapter 2

Introduction

A constant challenge in neuroscience is to understand how the brain represents the external world through the stimulation of our diverse sensory systems and how it uses that information to guide behavior. The study of sensory perception started with E.H.Weber, in 1834 (Weber, E. H., 1834), when he reported that the just-noticeable-difference (JND) - the smallest difference between 2 stimulus magnitudes that enables them to be perceived as distinct - is not a mathematical difference but a ratio between the two stimuli intensities. This finding, which later established the Weber's Law (WL), has been, since its observation, extremely successful in describing human and animal behavior for a wide range of sensory stimuli. However, the mechanistic foundation that originates this law of perception remained, until recently, without a consensual explanation.

Different hypothesis have been suggested to explain how the absolute magnitude of the stimulus can be excluded during the process of sensorial discrimination. Early work focused mainly on discrimination and estimation of continuos stimuli, giving special attention to Signal Detection Theory as the explanation for this phenomenon. Later approaches brought out the concept of bounded accumulation of evidence and sequential sampling by taking into account the temporal structure of the stimulus. Nevertheless, despite being well established and widely observed, WL is not sufficiently restrictive to choose one of the theories over the other, as it is exclusively a statement about a single aspect of discrimination: its accuracy. Recently, some studies have highlighted the importance of Reaction Time (RT) as a key tool for comparing different models of WL (Teodorescu et al., 2016; Simen et al., 2016). Actually, it was Link who suggested a process of bounded accumulation of evidence based on RT observations during perceptual discrimination experiments (Link,1992). Although not so long ago, there was not enough clear empirical evidence that would allow an unambiguous identification of the mechanistic foundation of WL.

In order to reach a consensus on this question, the Renart Lab, which I now integrate as a student, built a model explaining the mechanistic foundation of WL in an auditory two-alternative-forced-choice decision making experiment (Pardo-Vazquez et al., 2018). In this experiment, the group studied how accurately rats could lateralize white noise for different overall levels of difficulty. The results showed that not only the accuracy of the task was level-invariant, obeying WL, but also, surprisingly, revealed that the shape of the reaction time distributions (RTD) was also level-invariant. In other words, the only behavioral effect of changes in the overall level of the sounds is a uniform scaling of time, that when applied makes the RTD overlap completely (Pardo-Vazquez et al., 2018). This evidence, denominated TIED - Time-intensity equivalence in discrimination- along with the demonstration that WL breaks for stimulus duration smaller than the average reaction time, suggest that WL is associated with a process of bounded accumulation of evidence (Pardo-Vazquez et al., 2018).

The aim of this first experiment developed within the scope of this thesis was to probe the generality of the TIED by testing if it was a regularity also present in human perception. Therefore, an adaptation of the rodents task was executed, which results revealed the TIED to be a behavioral regularity found in the human auditory perception system. These results point in the direction of the TIED as a signature inherent to the sensory perception, suggesting the universality of the proposed bounded accumulation of evidence model as the mechanistic explanation of WL.

2.1 Weber's Law

Weber's law (1834) states that the ability to perceive changes in magnitudes of stimuli is proportional to the magnitude. A simple and intuitive illustration is present in Figure 2.1. On both sides of the figure the lower squares contain 10 more dots than the upper one. However, the perception is different: on the left side one can clearly identify an increase in the dots intensity, that is, that the lower square contains more dots than the upper one. However, if we move our attention to the right side, even though the lower square has also 10 mores dots than the upper one, it is very hard to visualize this difference. This is a good illustration of how our perceptual ability varies proportionally with the intensity of the background, and that is exactly what Weber's Law stands for.

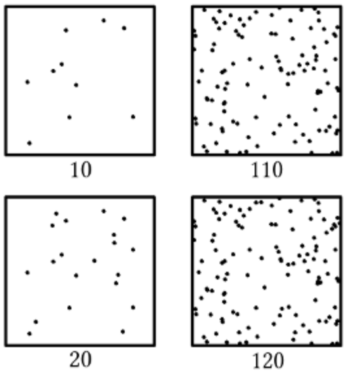

Figure 2.1: Illustration of Weber's Law. On each side the lower squares contains 10 more dots than the upper one. One the left side there is a clear perception that the number of dots increased, that is that the lower square has a higher number of dots. On the right side, the two squares look identical.

For a more formal understanding of the law, let us consider X as being a stimulus magnitude and $X + \Delta X$ the next greater magnitude that can just be distinguish from X, then WL states that

$$\frac{\Delta X}{X} = \Theta \tag{2.1}$$

where Θ is a constant for each sensory attribute (Laming D., 1986). According to the law, the only thing that matters upon deciding which stimulus is more intense, is the ratio between its objective magnitudes. Or in another way, when the intensities of two stimuli are scaled by any value k, the JND, ΔX, smallest change in stimuli that can be perceived, also scales by k, and consequently this ratio, Θ, remains constant across a wide range of stimuli intensities.

Later, Gustav Theodor Fechner came with a different understanding in an effort to relate the physical magnitude of a sensory stimulus and the magnitude of its associated subjective percept, envisaging that the interface between the two was based on a logarithmic transformation (Fechner G. T., 1860). By other words, Fechner believed that subjective stimulus intensity was a logarithmic function of stimulus intensity, that is $Subjective\ X = log(X)$. Therefore, he rewrote the WL in,

$$ln(X + \Delta X) - lnX = constant. \tag{2.2}$$

which maintains this idea that the JND increases with the initial intensity of the stimulus, X, being that a consequence of our logarithmic reading of the outside world. This relation remained known as Fechner's law and it was consensual at the time (Laming D., 2009). It led different scientists to look for logarithmic relationships in sensory pathways and to place a quite disproportionate emphasis on a finding by Hartline and Graham (Hartline and Graham, 1932), who were studying the effect of various intensities of light stimulation upon individual nervous discharges associated with photoreceptors. They found that the maximal frequency of discharge onset increased as the logarithm of luminance over about three log units (though the sustained rate of discharge, measured after 3,5s, followed a power law instead).

Later, in 1957, after dozens of experiments on direct scaling of stimuli, such as magnitude estimation, S.S.Stevens proposed a new law to relate sensation magnitude to stimulus intensity:

$$S = aX^{\beta} \tag{2.3}$$

where once again S stands for sensation magnitude and X for stimulus intensity (Stevens, 1957). Note, that both β and a constants depend on the sensory system in question. Even more important, the nature of the relation between sensation magnitude and the stimuli intensity is strongly influenced by the value of the exponent, β. So, this new law was able to explain positively or negatively accelerated responses as well as just linear ones. It was able to describe a big part of the responses observed in different sensory organs when stimulated with different intensity stimuli (Teodorescu et al., 2016).

Both Fechner and Stevens tried to understand how the magnitude of a stimulus is perceived by the brain in order to explain why Weber's Law holds for such a wide range of different sensory systems. They sought to find the dictionary that could explain how we measure outside stimuli. Today, it is known that both relationships found are extremely dependent on the configuration in which the two magnitudes to be distinguished are presented, and that is why they both can explain so accurately the data of their experiments. In reality, they both say very little about how the brain is able to correctly distinguish two stimuli intensities, or how sometimes it is not. What is the process that happens in the brain, that allows a difference in stimulus magnitude to be distinguished, and a smaller one not to? How is the JND kept constant regardless of the initial stimulus magnitude? How does our brain use this sensory information to guide behavior, to make a decision? Still in the 1950s and in relation to these questions, a new theory able to explain the properties of discrimination between two stimuli was emerging, the Signal-detection theory.

2.2 Signal-Detection Theory

A central concept in Fechner findings was the sensory threshold (Green and Swets, 1966). That is, the idea that if a subject detects a stimulus or identifies two stimuli as being distinct, it is because its intensity or the difference in intensities produces a sensory event of sufficient magnitude to exceed the sensory threshold. By other words, if the sensation of the stimulus is bigger than the threshold ($X > k$), then it is detected. Therefore, all the different experiments, carried by Fechner, to determine the difference between two signals necessary for the two to be just noticeably different, were also attempts to quantify this sensory threshold for different types of sensorial information.

However, even with slightly modifications (low-threshold theories (Green and Swets, 1966)), there was an important feature of the results obtained with the psychophysics experiments, that a simple threshold theory could not account for - the uncertainty. That is, the presentation of the same stimulus would not always produce a positive identification of the same, or that the same difference in intensity between two stimuli would not always make the subject perceive them as different. Actually, the value of the threshold was assumed by the criterion that 50% of the responses would correspond to the identification of the stimuli as having different magnitudes (Green and Swets, 1966). It is, therefore, with the ability to explain this inconsistency in responses that the Signal-detection theory (SDT) gained ground.

The general theory of signal detectability was developed most fully in the 1950's by mathematicians and engineers at the University of Michigan (Peterson et al., 1954) and at Harvard and Massachusetts Institute of Technology (Meter and Middleton, 1954). One of the first applications of the theory in psychology was the experiment, carried by Tanner and Swets , in the same year, 1954. Their experiment consisted in an yes-no visual task, where subjects had to indicate if there was a faint light being presented, or the background alone (Tanner and Swets, 1954). In addition, they inquired their subjects about the likelihood of the positive observation being in fact generated by the presence of a light. In this type of task there are two kinds of errors a subject might make:

1. Identify the presence of a signal when there is solely background noise - and let's consider this mistake to have a probability of occurrence α.

2. Identify no signal to be present, when there is one being presented - and here let's assume a probability of occurrence $1 - \beta$, where β would correspond to to the probability of a correct identification of the signal presence.

If the classical threshold theory is to be believed - if there is a magnitude level below which the stimulus has no effect and above which the stimulus is perceived and generates as output in our sensory system - then α should be a constant and β should depend only of the stimulus intensity (Laming D., 2009). However, as we observe by the results compiled in the ROC curve (Receiver-operating-characteristic), Figure 2.2-A , where the number of hits (correct identification of a light being present) are plotted against the number of false alarms (identification of a light to be present when there is background alone), for different likelihood ratios (each point represents a level of certainty about the presence of the stimulus), a given signal does not produce a consistent "yes" ("I detect it") or a consistent "no" ("I do not detect it"). On the contrary, what it is observed is that both probabilities - α and β - do not present a trivial and linear behavior, but a complex one, nicely, described by 2 curves.

The first person to suggest that a stimulus with the same intensity not always produces the same level of activity in the nervous system was Thurstone, in 1927, with his series of comparative judgment experiments (Thurstone, 1927). He introduced the idea that a a stimulation of any type produces a distribution of effects, with values of sensation that are more likely to occur, but that for various reasons, suffer fluctuations (Green and Swets, 1966) . Thurstone suggested that the distribution able to explain these random variations in the activity of the nervous system, was Gaussian and this became the base of the SDT. Going back to the experiment of Tanner and Swets, according to Thurstone's idea, the noise alone produces a normal distribution of activity scaled to unit variance, with mean, μ_N, as pictured in Figure 2.3. When there is a signal being presented, there is an increase in the activity, therefore the distribution moves to the right, to the mean μ_{SN}. The distributions are assumed to have the same variance, as a result of internal noise, characteristic of the sensory system in question, and eventual sources of external noise, which are considered to be constant for both $noise$ alone or $signal + noise$. Let's consider, as well, that the subject selects a criterion with value ξ_C, such that, if on any time the generated activity in the nervous system is $\xi \geq \xi_C$, then the subject will consider that there is a signal

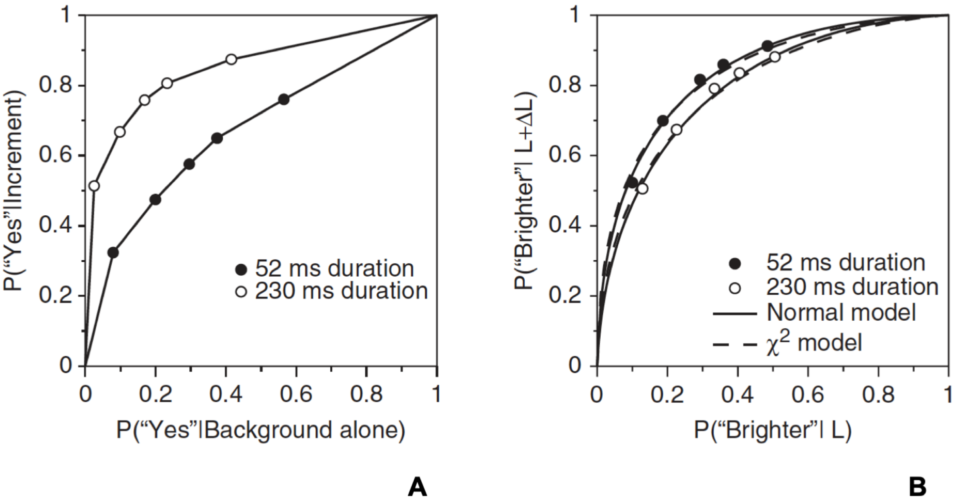

Figure 2.2: (a) ROC curve built with data from the experiment of Tanner and Swets (Tanner and Swets, 1954), where was asked to the subject to respond yes or no about the presence of a light (18% stronger than the background). The experiment was performed for 2 different durations of the stimulus presentation - 52 ms and 230ms. Each one of the dots represents the number of hits (correct identification of the light to be present) vs the number of false alarms (identification of a light to be present when was background alone), for different levels of likelihood. (b) ROC with the data from the experiment of Nachmias and Steinman, where subjects are asked to discriminate the brightness of two flashes of light with luminance L and $L + \Delta L$, being $\Delta L = 26\%$. Also here, subjects were asked to report the likelihood of their answer be indeed the most bright flash, by other words the level of confidence in their answer. Each dot corresponds to the number of hits - correctly indicate the most bright flash- versus false alarms - number of times subject incorrectly indicate the less bright flash as being the brightest. (This image was adapted from the book Inside Psychology: A Science Over 50 Years (Laming, 2009)).

being presented (in Tanner and Swets' experiment, that a light is being presented). On other hand, if the generated activity is $\xi < \xi_C$, the subject will assume to be noise alone. However, we see that both distributions overlap, that is, even above the criterion it is possible to find levels of activity that have origin in the observation of both signals, $noise$ alone and $signal + noise$. Which gives space to both mistakes to happen with probabilities α and $1 - \beta$, as mentioned above. More impressively, as ξ_C runs through its possible values, it generates a curvilinear relationship between α and β as the one observed in the ROC graph, Figure 2.2.

From the simple analysis of the Figure 2.3 one can easily understand that the discriminability of a signal depends both on the separation between the distributions and the spread of the same. That is why, is measured as function of d'

$$d' = (\mu_{SN} - \mu_N)/\sigma_N \tag{2.4}$$

Naturally, the relationship between α and β is also different for each value of d'.

The theory has been used to explain several results, inclusive other experiments characteristic to find JND's. For example, the Nachmias & Steinman experiment (Nachmias and Steinman,1965), where subjects had to indicate which one of two lights was the most bright. Their empirical data is also depicted in Figure 2.2 B. The theory accounts for all the properties of discrimination between two separate stimulus magnitudes, and therefore has been used to modulate several of the sensory precision related experiments (Banks, 1970; Laming, 1986; Clark, 1988; Verghese, 2001). But, more importantly for this work, the theory was able to give an explanation of WL, with a clear relationship between d' and the Weber's constant, above reported as θ (Thornton, 1969). However, it is noteworthy this theory was built on the fact that there is no absolute judgment related to perception, and sometimes we make mistakes. Also, for the first time a clear relationship between stimulus intensity and induced neural activity was able to

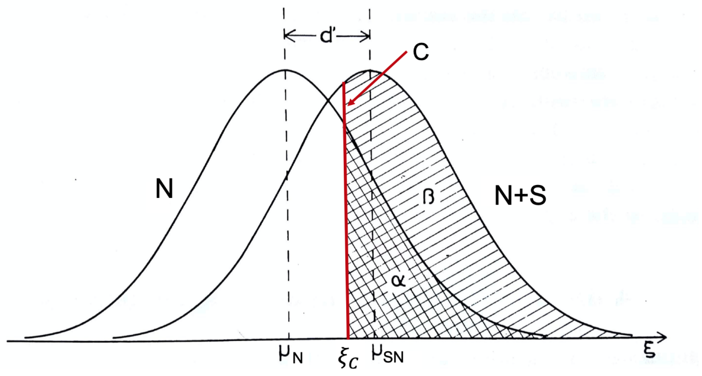

Figure 2.3: The normal signal-detection model of Tanner and Swets [7]. The presentations of noise alone N and a stimulus on top of the background $N + S$ are respectively represented by sample distributions, which are taken to be normal, of unit variance and separated by a difference in mean d'. The parameter d' measures the discriminability of the light signal presented compared to the noise of the background. The choice of criterion ξ_C reflects the subject's motivational bias and determines the two operating probabilities α and β. The probability α is the probability of incorrectly identifying the stimulus as being presented; β corresponds to the probability of correctly reporting the presence of the stimulus. (Image adapted from Sensory Analysis (Laming, 1986).

accurately explain perceptual behavior in animals (Laming, 2009). As a consequence, it was understood that sensation could not be measured in the sense both Fechner and Stevens envisaged (Laming, 2009). .

2.3 Sequential Sampling Theory

When analysing the results of Tanner and Swets in Figure 2.2 A, or even the results of Nachmias and Steinman in the same Figure B, there is a striking result not accounted by the previous introduced theory - the fact that when the stimulus is presented for a longer period of time, subjects make less mistakes. The sequential sampling theory in psychology and neuroscience arises from this observation, or better, as an attempt to unify the results of accuracy and response time of a decision-making task, in a single theory. Thus in 1960's, Stone (Stone, 1960) applies Wald theory (Wald, 1948) to explain reaction-times and the first step in this direction is given (Forstmann et al., 2016).

On the contrary to the theories previously mentioned, the Sequential sampling brings the idea that the internal representation of the external stimulus is not an unitary event, but a statistical process where successive samples of this stimulus are collected over time. In this way, a simple decision arises from repeated sampling of this external representation and comparison of some function of these samples to a criterion. Additionally, the evidence is considered to be noisy, either because the stimulus is a sequence of variable or noisy events, or because exists noisy coding in neural systems. Several different models constitute the totality of Sequential sampling theory. The ones introduced by Stone, in 1960, and Edwards in 1965 (Edwards, 1965) , later developed by Laming in 1968 (Laming, 1968) are part of a class denominated random-walk models (Ratcliff, 2001). In this class, the samples of the noisy signal are taken in discrete times and are added together to represent the evidence over time until an upper or lower bound is reached and the decision made. Later, Ratcliff in 1978 (Ratcliff, 1978), suggested that the process of sensory information acquisition was better viewed as a continuous- time rather than a discrete-time process. Therefore, the time steps were reduced to infinitesimals, becoming the model continuous in time and creating a new class of models, the Diffusion decision models (DDM). Other models and classes have been proposed since then (e.g. Accumulator models), although DDM have been gaining a lot of influence as models of the psychological and neural processes involved in decision-making (Ratcliff, 2001). The reasons are mainly 3:

1. These class of Sequential sampling models account for all the behavior data, namely accuracy, and the shapes and locations of the reaction-time distributions for both correct and incorrect responses.

2. They have been linked to neural processing for single cells and populations of neurons.

3. They have been successful in explaining decision making across wide domains of psychology such as aging, child development, various clinical populations, and animal species, often providing new interpretations of data.

Therefore, for a better understanding of this wide class of models and to clarify how it is connected with decision-making, in particular the distinction of 2 stimuli intensities the next paragraph will introduce what is consider to be the simplest version of the same.

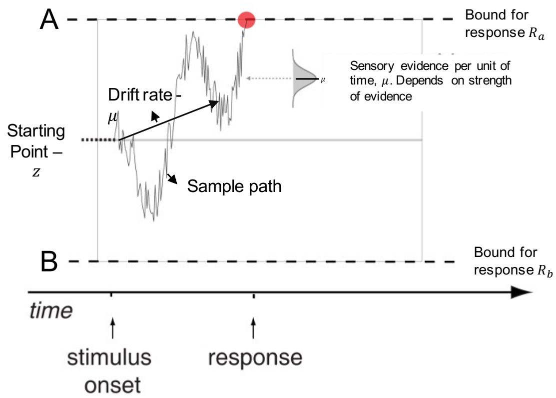

Figure 2.4: An illustration of a sample path of the accumulation of evidence underlying a perceptual decision. On each trial, evidence in favor of one alternative over another is accumulated as a function of time. For any particular stimulus strength, there is an accumulation of noisy evidence parameterized by the mean rate of accumulation, μ. A decision is made when the process reaches one of the bounds, R_a or R_b. (Image adapted from Mulder et al., 2012).

Let's consider the discrimination between stimuli S_a and S_b, where one is required to make a corresponding response R_a and R_b (standard design of a two-choice decision experiment). In parallel with the psychophysics experiments already mentioned, R_a would be stimulus A is more intense and R_b would correspond to B being more intense. As said, in the diffusion model, evidence is accumulated over time until an upper or lower bound or threshold (i.e criterion) is reached, (A or B) which triggers a response. The boundaries represent the amount of evidence that must be accumulated before a response is made. A single trial is illustrated in Figure 2.4. It shows the relative evidence of stimulus S_a over stimulus S_b as a function of time. A sample path from a single trial is shown by the jagged contour. It is, as observed, a noisy process; at each moment in time the evidence might point to one or the other of the 2 bounds. For this example trial, the accumulated evidence reaches the upper bound A and triggers a R_a response.

The ray from the origin illustrates the mean drift rate μ, which determines how fast the sample path reaches a bound. Also, even though it is not represented in Figure 2.4, there is another parameter σ, denominated diffusion coefficient, that conditions the behavior illustrated by the model. That is, at time t is assumed that the path would have a mean μt and a variance of $\sigma^2 t$. Regardless of the mathematical relation between μ, drift rate, and stimulus strength, generally all diffusion models associate a weak stimulus with small or even near zero values of μ. Therefore, a weaker stimulus originates a longer response time, leaving space for more mistakes. Contrarily, a stronger stimulus is associated with larger values of μ which originates smaller response times and less mistakes (Figure 2.5 B). In addition, bias towards one or other response can be represented by the relative values of the bounds A and B. That means, if there is no response bias, A=B (Figure 2.5 D). Also, the bound values control the speed-accuracy trade-off, being large values of the bound associated with slower responses and higher accuracies. Another important aspect is the starting point, that can also suffer variability, to account for trial-to-trial variability (Figure 2.5 C). It is as well, the variation of this starting point, or the bias towards one or another bound that justify the fact that sometimes mistakes have faster RTs.

In summary, the model has parameters for the bounds (A and B), the drift rate μ and the diffusion coefficient σ, and a starting point, z. All the effects that evoke changes in these parameters influencing RTs and chosen answers are well illustrated in the next figure.

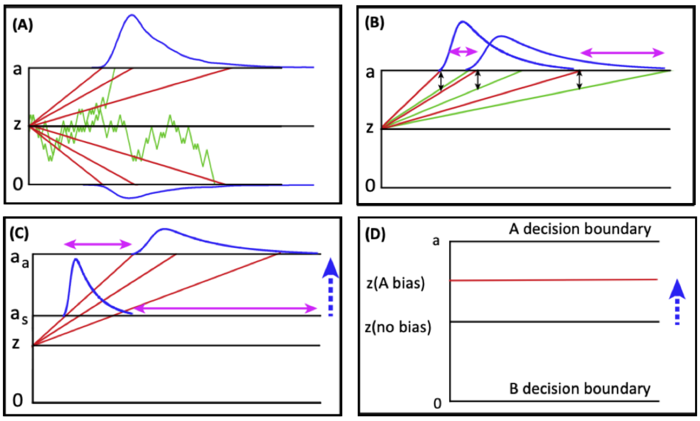

Figure 2.5: An Illustration of the Diffusion Model. (A) Shows two (irregular) simulated paths in the diffusion model (green). The blue curves represent response time distributions (RTD) for correct responses (top) and errors (bottom). The red lines represent the fastest, medium, and slowest responses. (B) Shows the effect of lowering drift rate by a fixed amount. The black double arrows show the effect on fast, medium, and slow average drift rates and the magenta arrows show the effect on the fastest and slowest responses from the blue RT distributions. There is a small change in the leading edge of the distribution and a large change in the tail. (C) Shows the effect of moving a boundary away from the starting point [a-speed (a_s) to a-accuracy (a_a), the blue broken arrow] to represent a speed–accuracy manipulation (both boundaries would move in a real experiment). The magenta arrows show the effect on the fastest and slowest responses from the blue RT distributions. There is a moderate change in the leading edge of the distribution and a large change in the tail. The difference in effects between (B) and (C) discriminates manipulations that change boundaries from manipulations that change drift rates. (D) Shows how a bias toward the A response can be modeled by a change in the starting point (blue broken arrow with the starting point moving from the black line to the red line). RT distributions change as in (C). (Image adapted from Ratcliff et al., 2016).

This type of models were, for the first time, able to explain the uncertainty among subjects answers but also, the different RTs associated to correct, incorrect answers, more difficult and easier conditions. They were able to accurately reproduce two dimensions of perceptual decision-making behavior: accuracy and response times. Thus, given its success, the theory has been highly used to sustain multiple behavioral observations. Although, different versions of the model address stimulus scaling in very different ways. By other words, the relation between μ and stimulus intensity varies from explanation to explanation. However, the most specific models, assume that the normalized drift rate is proportional to stimulus strength, that is, $\mu = kx$. Where the coefficient k is the sensitivity in this called proportional-rate diffusion model.

The fact that is not clear what is the contribution of the noise in the external stimulus or the noise in our neural systems to the process of Sequential sampling, made it hard to understand the exact role of overall stimulus magnitude in these models. As a consequence, classical problems such as WL have been left apart by the community studying perceptual decision making.

2.4 Where are we now?

Almost 200 years after the publication of Weber's law (1834) (Weber, 1834), and the extended replication of its results in diverse sensory systems (Oberlin, 1936; Gaydos,1958; Cornsweet and Pinsker, 1965; Indow and Stevens, 1966; Panek and Stevens, 1966; Leshowitz et al., 1968; Laming, 1986) for several species, a consensus has not be found among its many proposed explanations (Fechner, 1860; Laming, 1986; Link, 1992; Deco and Rolls, 2006; Teodorescu et al., 2016; Simen et al., 2015). It is a fact that Weber's law embodies a non-trivial computation, since sensory receptors and sensory neurons in the periphery encode absolute magnitude explicitly in the form of monotonic increases in firing-rate (Hartline and Graham, 1932). However, Weber's Law dictates that stimulus intensity is left aside in the discrimination process. Therefore, the question of how absolute magnitude can be excluded in the process remains unanswered.

As it as been presented above, most of the early work developed on sensory experience focused, mainly, on discrimination and estimation of different stimuli intensities, using SDT as the core explanation for most of the observed behaviors (Green and Swets, 1966). However, this model only accounts for accuracy, and that leaves aside interesting features of the sensory process, such as response time. Later, a different approach focused mainly on relating response time with accuracy in decision-making tasks originated a new theory, Drift diffusion model (Ratcliff, 1978). This new paradigm, which rests formally on the Sequential sampling framework, has been extremely successful in guiding the study of perception and decision-making theories (Ratcliff, 1978; Link, 1992; Brunton et al., 2013; Simen et al. 2015). Nevertheless, although it could account for RT in discrimination of two stimuli, this theory made it difficult to differentiate external and internal sources of noise. Consequently, being hard to identify the exact role of absolute stimulus magnitude in the theory combined with the success of the model, led problems and questions such as the ones in WL to remain not fully comprehended so far.

An important exception to this observed tendency of leaving WL back in the day, is the work of Stephen W. Link. In 1992 with his 'wave theory' (Link, 1992), Link reveals experimental evidences and a clear understanding that WL is the natural consequence of a process of bounded accumulation of evidence (Sequential Sampling Model) when the intensity of the stimulus is represented by the rate of a Poisson process. More recently, some studies have realized the importance of RT as a diagnostic tool for comparing different models of WL (Teodorescu et al., 2016). In this sense, a study of the effect of stimulus strength on the speed and accuracy of a perceptual decision (Palmer et al., 2005), revealed a close coupling between response time and accuracy according to a model of bounded accumulation of evidence, underling the importance of considering RT in the study of WL (although, WL does not embody any prediction of RT). Recently, a new study (Simen et al., 2015) with a set of elegant decision making experiments performed in visual, auditory and vibrotactil sensory systems went a step further, and revealed that the spike poisson-counting diffusion model, suggested by Link, predicts that overall stimulus level should rescale the reaction time distributions (RTD). In other words, Weber's consequence on RT in a two-choice perceptual decision should be exclusively a rescale in its distribution.

However, despite all these efforts, the empirical data available does not unambiguosly establish whether and how WL should be understood within a sequential sampling framework. Since, indeed, the most in-depth treatments still explain WL within a SDT. Consequently, an experiment able to gather the study of WL and its implications in RTs was necessary to finally establish the mechanistic foundation of WL. It is the case of the experiment developed in Renart's Lab performed in rodents that it was in the sequence of the work presented in this thesis, later transformed to be performed in humans. The details of this original work are given in the next section.

Chapter 3

Weber's Law is the result of exact temporal accumulation of evidence

So far, we have reviewed the most important findings in the study of perception since the publication of Weber's Law. Hopefully, it is now quite clear the importance of studying perceptual mechanisms, especially given the impact they have on our daily life and in the simplest animal behaviors. Therefore, with the goal of clarifying the mechanisms underlying WL, the research team developed an experiment where they trained rats to discriminate the lateralization of sounds for different overall levels (Pardo-Vazquez et al., 2018). The experiment brought some breakthroughs in the field of psychophysics and finally some consensus regarding the topic of WL. The next section starts by explaining in detail their approach to the study of WL followed by a short revision of their results and conclusions.

3.1 Experimental Design

Rodents and other species use inter-level differences (ILD) in the sound intensity received at both ears, caused by the acoustic shadow of the head, to localize sources in the horizontal plane (Schnuup and Nelken, 2011). In other words, through the comparison of the sound intensities received in both ears (Figure 3.1), rats can localize the relative position of a sound source:

$$ILD = SL_R - SL_L \tag{3.1}$$

Being SL_R and SL_L, the sound level (SL) at the right and left ear, respectively. As sound level is, by definition, defined as:

$$SL = 20log_{10}(P/P_0) \quad (dB\ SPL) \tag{3.2}$$

Where $P_0 = 20\mu Pa$ (Fastl and Zwicker, 1990). Then ILD is, consequently, given by:

$$ILD = 20log_{10}(P_R/P_L) \quad (dB\ SPL) \tag{3.3}$$

This equation of ILD shows that when subjects localize sound in the azimuthal plane, they are actually using the internal computation of the fraction between the pressure felt in both ears, which doesn't depend on the overall stimulus intensity as long as the ratio is kept constant. This fact makes sound localization the perfect experiment to address Weber's Law, as it is recruiting a circuit designed by the evolution for the purpose of comparing stimulus intensity (Pardo-Vazquez et al., 2018).

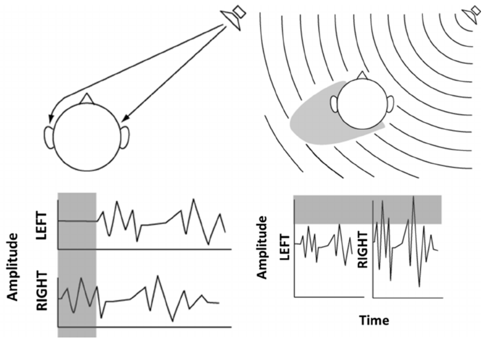

Figure 3.1: Illustration of what inter-aural differences are and how they are generated by the acoustic shadow of the head in the horizontal plane, evidencing how ILDs are used to identify sound source in the azimuthal plane. The amplitude plots reveal, as expected, that the sound amplitude is higher to the side where the sound source is. (Figure adapted from Sun et al., 2015).

Consequently, a sound localization task was developed in rodents with the possibility of studying the impact of the overall level of the sound - ABL:

$$ABL = (SL_R + SL_L)/2 \quad (dB\ SPL) \tag{3.4}$$

in rats' ability to discern the source of the sound for different ILDs, as well the RTs involved in the process. Additionally, WL is known to hold robustly for white noise discriminations (Miller, 1947; Laming, 1986; Stellmack et al., 2004), therefore the group trained rats to discriminate the lateralization of broadband noise bursts. In the task they varied ILDs while keeping the average binaural level (ABL) constant, and they did this for 3 different levels of ABL: 20, 40 and 60 dB SPL, Figure 3.2 C. Four different values of ILD were used, 1.5, 3, 4.5 and 6 dB.

The task consisted in indicating which side (left/right) the sound was located, which is the same as indicating the side where the sound is louder. For that, on each trial rats started by poking in the central port, Figure 3.2 A. After a variable fixation time (FT) the sound was played binaurally through custom-made headphones, until the animal left the central port or up to a maximum sound presentation of 6s. The moment of leaving the central port was the animal indication he had chosen between left and right. To indicate his response, the animal had to poke with the snout in the side correspondent to his decision. Naturally, both ports had the same distance from the central one and RT was considered to be the time necessary for the rat to make a decision, that is the amount of time between the sound onset and leaving the central port. The time between leaving the central port and reaching one of the lateral ports is denominated movement time (MT), and it will not enter the results of these experiments for their irrelevance to the study of WL. However, it will be quite important in the discussion and motivation of Part II of this thesis. For each correct response the animals were rewarded with water and incorrect responses penalized with a 10s timeout during which the rat was not able to start a new trial. A scheme of the multiple events in a trial is also illustrated in Figure 3.2 and for a deeper understanding of the task consult Pardo-Vazquez et al., 2018.

3.2 Results

The results of this experiment revealed that rat's discrimination accuracy is level-invariant, at least for broadband noise (Figure 3.3 A - psychometrics for different ABLs overlap and Figure 3.3 B - showing no significant difference in values of sensitivity (d') and criterion (c) associated to different ABLs). This means that the accuracy of ILD discriminations does not depend on ABL. A fact that confirms the rat's

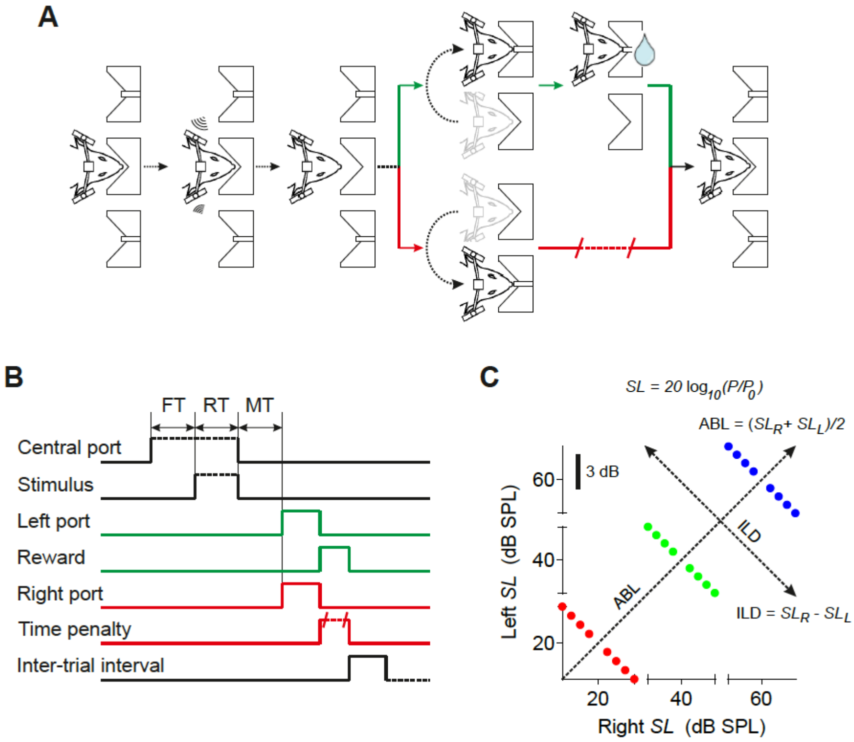

Figure 3.2: Task Structure and stimulus set - Rat's Experiment:**(A)** Schematic depiction of different task events. Rats were rewarded with water for making the correct choice and were punished with a time delay for making an error. **(B)** Time-line of relevant task events. FT, stands for fixation time; RT, reaction time; MT, movement time. **(C)** Stimulus set. All sounds were cosine-ramped broadband (5-20 kHz) noise bursts. The ABL (ILD) of a particular stimulus is given by the average (difference - by convention right minus left) of the intensity of the sound in dB SPL (sound level, SL) across both speakers. $P_0 = 20\mu Pa$ is the reference pressure of the SPL scale. (Image adapted from (Pardo-Vazquez et al., 2018).

ability to identify the source of the loudest sound does not depend on the overall intensity of the stimulus, but exclusively on the difference of sound level being presented to both ears. Thus, as expected, Weber's Law holds for auditory discrimination in rodents.

When analysing the results for the RTs, more difficult discriminations (i.e smaller ILDs) showed longer RTs (Figure 3.3 C) - speed-accuracy trade-off - which is a signature of bounded accumulation of evidence (Figure 2.5 B). This signature originates a typical right-skewed RTD (Figure 2.5), shape also found in the distributions of this experiment (Figure 3.3 D). Contrarily to the accuracy, RTs turned out to be dependent on the overall level of the stimulus - peak of the distributions in Figure 3.3 D is dislocated to the right when the ABL is smaller, which shows that discriminations involving overall quieter sounds (i.e lower ABLs) took longer on average. Result significant both at the individual and group level. However, when analysing in more depth these RTDs one observes that the shape of the distributions themselves are level-invariant. More specifically, if one rescales the distributions, one can observe they overlap almost perfectly and are identical for each difficulty and for all difficulties combined (Figure 3.3 E). This result, along with the level-invariance of discrimination accuracy, demonstrates that the sole effect of changes in the overall level of the sounds (changes in ABLs) is to change the effective unit of time of the sensory discrimination process, with a shorter effective unit of time for louder sounds (Pardo-Vazquez et al., 2018). In other words, the sensory discrimination process is independent of the overall stimuli magnitude - the probability of a correct response does not vary with ABL, the RTs follow robustly the same distribution. However, sensory discriminations happen faster (not differently) for louder sounds, that is, the effective unit of time is shorter for louder sounds. This observation was named Time-intensity equivalence in discrimination - TIED.

As mentioned, the speed-accuracy trade-off and the right-skewed shape found in the results suggest that the animals are solving the task by a process of bounded accumulation of sensory evidence. To test this prediction, the same task was performed while imposing a maximum sound duration smaller than the typical amount of time animals take to decide. Results revealed that the performance of the animals decreased significantly, being this difference more pronounced for the smaller ABLs, suggested as having a longer effective unit of time. Therefore, these results reinforce both the idea that WL is associated with a process of bounded accumulation of evidence and that WL occurs because multiplicative changes in stimulus intensity only modify the effective unit of time of this process.

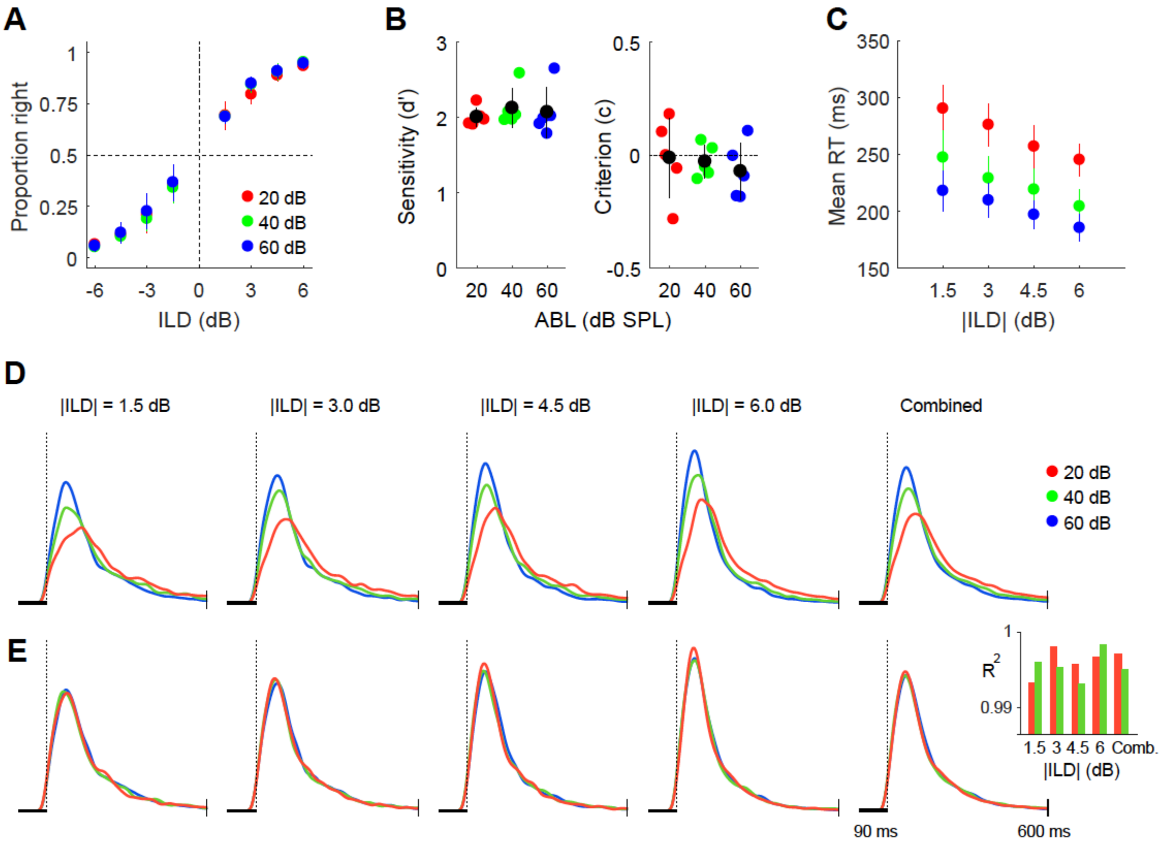

Figure 3.3: Analysis of behavior for all subjects in a level-invariant ILD discrimination task. **(A)** Choose-right probabilities as a function of ILD for each ABL separately (mean $\pm$ SD across rats).**(B)** Sensitivity (d', left) and (c, right) for each animal (mean $\pm$ SD across rats). **(C)** Reaction Time (mean $\pm$ SEM across rats) as a function of difficulty (ILDs) for each ABL separately.**(D)** RT distributions (RTDs) for the three ABLs are shown separately for each ILD, so for each difficulty, and combined across difficulties (right). For all RTDs, the dashed lines indicate when the RT becomes condition-dependent. Each RTD contains all data for that condition from all rats. **(E)** Rescaled RTDs uniformly in time to maximize the overlap of each RTD with that of the loudest sound (ABL = 60 dB SPL) (This figure was adapted from Pardo-Vazquez et al., where you can consult methods section to better understand the analysis performed to originate each one of these multiple plots.)

To test these predictions, the next step in this study was, naturally, to build a model with these characteristics and see how well it could reproduce the empirical data. Their work concludes that among a broad class of sequential sampling models, the only robust mechanism consistent with the temporal rescaling ($t\prime = \alpha t$ being t the time need to discriminate between a sound with RMS pressure at 2 ears of (P_R, P_L) and $t\prime$ a discrimination between the pressures (kP_R, kP_L) able to produce scale-invariant distributions, was a:

1. A model of constant decision bound with perfect accumulation of evidence - that is, no intrinsic decay of the decision variable across time.

2. A linear relationship between the variance of the evidence and its mean; and

3. A power-law relationship between physical stimulus intensity and sensory representation.

The condition (2) imposes that the statistics of the sensory evidence should behave like a Poisson process. As the spike counts statistics of a Poisson process are completely invariant if the rate and the count window are modified by multiplicative factors c and 1/c, respectively, condition (2) was necessary. Thus, scaling the rate of a Poisson process can always be compensated by a rescaling of time. This, combined with condition (1), implies that the probability of the decision variable hitting one of the bounds and the shape of the reaction time distributions are independent of the overall stimulus intensity, which is absorbed as the aforementioned time scaling. However, the TIED implies a scaling of the stimulus magnitude, not firing rates so an extra requirement was needed - condition (3). As a power-law is the most general transformation between stimulus magnitude and firing rate such that multiplicative changes in the intensity of sensory stimuli result in multiplicative changes in firing rate.

Remarkably, the implementation of a model with these characteristics, imposed by the TIED and level-independence accuracy constraints, was able to describe the rats performance with virtually no error, suggesting this mechanistic explanation of WL to be robust.

Chapter 4

Human Task - Looking for the TIED

The task explained in the previous section enabled the identification of the TIED as a crucial feature of the discrimination process and allowed the group to create an hypothesis for the mechanism underlying Weber's Law. However, the WL has been shown to be present in different species and for diverse sensory inputs. Therefore, if one wants to propose a mechanism able to explain WL, this has to be general and cannot be limited to the way a certain model solves a particular task. Wherefore, an expansion of the model to other species was essencial and this is where it starts the focus of this thesis. With the specific goal of finding some generality for the TIED and test the model proposed for different species, a similar task was performed in humans. This section will provide a detailed description of the experiment adapted from the rodents to humans.

4.1 Methodology

Like rodents, humans use ILDs generated by the acoustic shadow of the head to localize sound in the azimuthal plane (Celesia and Hickok, 2015). As a result, we aimed to maintain it as the most similar as possible to the original one when creating this task, since the former's design had already been successfully adapted to maximize the study of the contingencies in the process of discrimination with auditory stimuli. Therefore, an auditory two-alternative forced choice (2AFC) task was implemented for humans, where subjects had to indicate the side (left or right) they consider the sound to be located on.

To originate different perceptions of sound localization, we recurred once again of ILD's definition (Equation 3.3) and presented white noise with different intensities to each ear. In order to generate different levels of difficulty, we used 4 values of ILDs, although of smaller absolute magnitudes ($|ILD| = 0.3, 0.6, 1.2, 2.4$ dB) than the ones used in the rodents experiment, since human subjects revealed relative easiness in discriminating sound localization. These ILD values were presented for two different levels of overall stimuli magnitude (ABL = 40 and 60 dB SPL), so one could approach the effect of ABL in the discrimination of accuracy of human subjects.

One of the most important features of the rodents experiment, which allowed to establish TIED as a psychometric regularity, was its RT configuration. To accomplish the same in this experiment, we recurred of a gamepad (Figure 4.1 A). In the same way rats had to poke in the central port for as long as they needed to make a decision, subjects had to press both shoulder buttons (or bumpers) simultaneously until they had made their decision. To report their choice, they should release the shoulder button correspondent to the side considered to present a louder sound (Figure 4.1 A).

At last, to maintain subjects engaged in the experiment, we used a reward point system, rewarding subjects with 10 points for a correct response and penalizing them with a loss of 10 points for an incorrect response (Figure 4.1 C). Even though the points balance did not influence subjects' monetary compensation for each session, we considered this system effective in maintaining subject awareness of his/her performance.

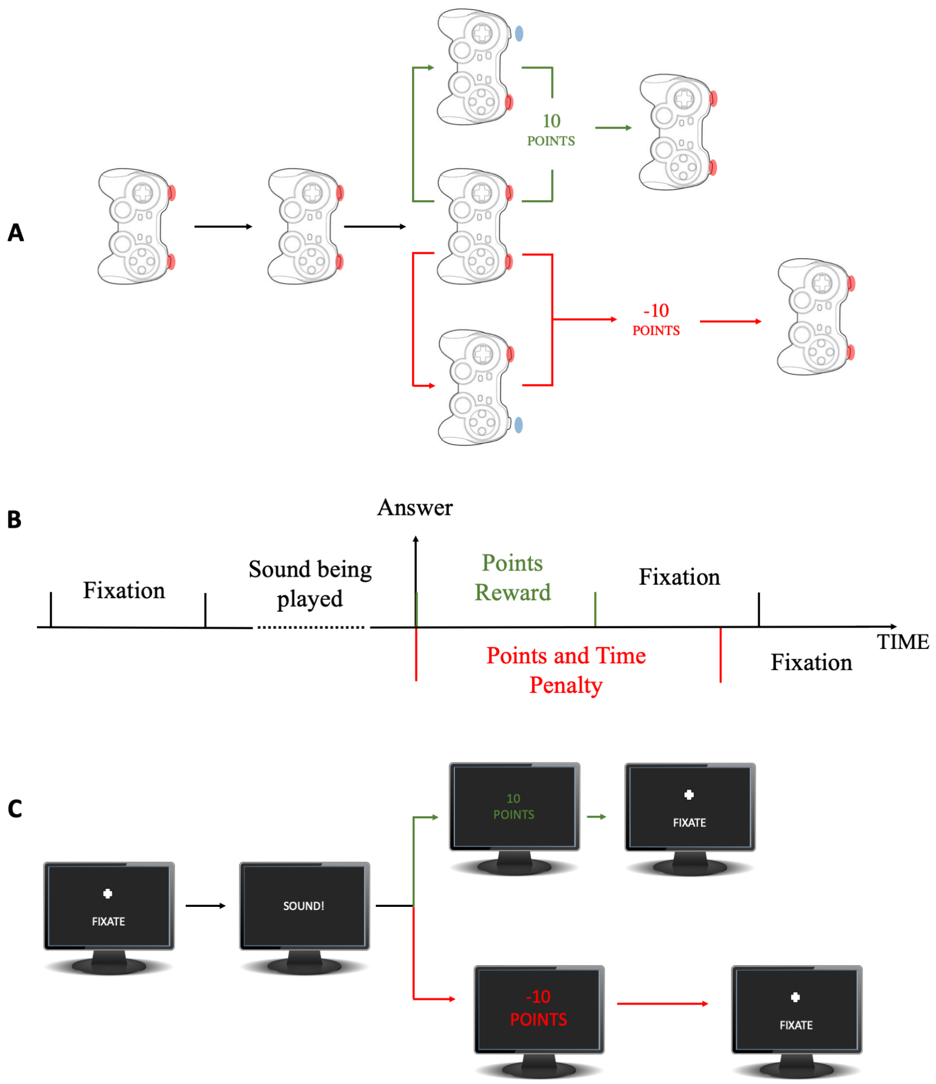

Figure 4.1: Task Structure and stimulus set - Human Experiment: **(A)** Schematic depiction of different trial gamepad events. The red circle represents trigger pressed, while blue circle means trigger unpressed. **(B)** Timeline of trial events. The events here represented are aligned with the representation of the same task events on the gamepad and screen. **(C)**Schematic depiction of how the different trial events were happening in the screen interface. For all the three schemes on the top part we have the events organized in time for a correct answer, while on the bottom the time task events associated to a wrong answer.

4.1.1 General parameters of the task

Subjects

All procedures were reviewed and approved by the Ethics Commission for human studies of the Champalimaud Centre for the Unknown, and all subjects signed a consent form also approved by the Ethics Committee. The experiment was performed by a total of 13 healthy human subjects, 8 males and 5 females, aged between 20 and 40 years old.

No subject had knowledge of any hearing impairment, and no abnormality was found in the analysis of the data. All subjects were naive about the aims of the task.However, 4 subjects were discarded by the following criteria. Two subjects showed an excessive bias (one of them more than 3 across-subject standard deviations away from the mean in criterion; the other more than 3 across-subject standard deviations away from the mean in the relative difference in average reaction time between left and right choices); and two other subjects clearly violated WL (relative variance in discriminability $\Delta d'/Mean(d')$ larger than 0.25).

Each session had approximately on average a duration of 20 minutes, with small fluctuations due to performance in the task. Subjects received 2.5 Euros per session regardless of performance. The first session corresponded to a training session, where subjects were tested for compatibility with the reaction time sound lateralization task and they were trained until they had reached 80$\%$ of accuracy for the two hardest conditions.

Material and Setup

The task was developed using the software Matlab version R2018a and the Psychtoolbox-3 (`http://psychtoolbox.org`) to produce the images on the screen and the auditory stimuli (in parallel with the sound card - Asus Xonar DX PCI Express 7.1 Audio Card). The different steps of each trial were signalized with visual information on a screen (Samsung SyncMaster S24B150). The auditory stimuli were delivered through a pair of over-the-ear headphones (Sennheiser HD 300 PRO) that were regularly calibrated. Subjects interacted with the behavioral software using a bimanual gamepad (Logitech Gamepad F310). The experiment took place in a small, closed room where subjects were free to sit comfortably in front of the screen, always guaranteeing equal access by both hands to the gamepad. The headphones were always carefully placed in the subject's head at the beginning of every session.

Instructions

When doing experiments with human subjects, instructions have to be carefully given otherwise the subject's behavior can be conditioned. As mentioned above, the subjects were simply instructed about the goal of the task - to identify on each trial if the sound playing was louder to the left or the right; that they had to initiate the trial by simultaneously pressing both shoulder buttons (or bumpers) of the gamepad, and that in a given moment the sound would start playing. To indicate their answer they were instructed to release the bumper of the side they considered to be the correct as soon as they knew the answer, with the understanding that there was no compensation for being fast, just correct. Subjects were notified about the number of points they could win for every correct answer and the number of points they could lose, for each incorrect response. It was guaranteed subjects understood all the instructions within the training session.

Auditory Stimuli

The percept of lateralization was created by presenting broadband noise (1 KHz to 15 KHz) with different intensities to each ear (creating an interaural level difference, or ILD). To generate the noise used we considered both the range of frequencies that humans can hear and also the intensities with which they feel comfortable (Cutnell and Johnson, 1998; Yost, 2000). Additionally, we selected broadband noise, since it is well know that WL holds robustly for discrimination with this type of sound, contrarily of pure tones (Stellmack et al., 2004; Laming, 2009). Previous studies also assured that the basic mechanisms responsible for WL are not only common to monaural but also binaural processing in humans, which allowed us to keep the same experimental design (Stellmack et al., 2004). The broadband noise was cosine-ramped (10ms) and independently generated for each ear (two different calibration files, one per speaker) and for each presentation using a sound card (Asus Xonar DX PCI Express 7.1 Audio Card) modulated at a sample rate of 192kHz.

The effect of loudness on sound localization was assessed by using different average binaural levels (ABLs), 40 and 60 dB SPL, where dB SPL is defined as in Equation 3.2. To be certain that the sound intensity being presented on each speaker was the desired one, the calibration of the speakers was performed. This was done through a styrofoam head and a Bruel & Kjaer Free-field 14 microphone, placed in front of the speaker 5 mm apart, in a soundproof box.

We used ILDs of 0.3, 0.6, 1.2 and 2.4 dB. Being the ILD of a particular stimulus given by the difference between the right and left speaker, (Equation 3.1, where SL_R and SL_L correspond to sound level in dB SPL for right and left speaker, respectively). Being the average ABL always kept constant: $ABL = (SL_R + SL_L)/2$ in each block of trials. Consequently, negative ILD values indicate higher sound intensity at the left hear; for example, an ILD of -2.4 dB for an ABL of 40 dB consisted in presenting the noise at 41.2 dB to the left and 38.8 dB to the right. While a positive ILD of 2.4 dB corresponded to the opposite situation, this is, 41.2 dB to the right and 38.8 dB to the left. For a better understanding of this methodology consult Figure 4.2.

It is also important to highlight that, by generating different ILDs in such a way, we ensured that the only cue used by the subjects to localize sound in the horizontal plane is ILD, and not also interaural timing differences (ITDs - differences in arrival time of a sound between two ears); as sounds in both headphone speakers were being presented exactly at the same time.

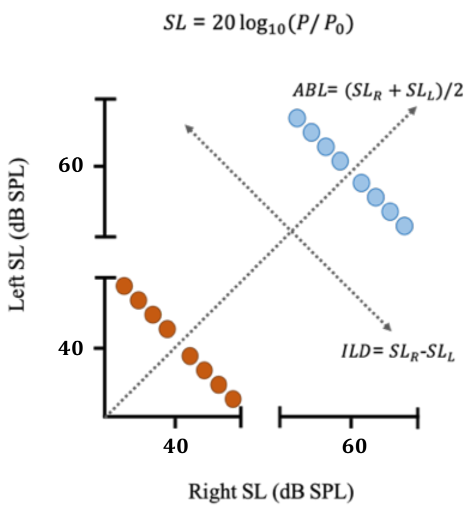

Figure 4.2: Illustration of the ABL and ILD distributions used in the human task. The ABL of a particular stimulus is given by the average of the intensity of the sound in dB SPL (sound level SL) across both speakers. $P_0 = 20\mu Pa$ is the reference pressure of the SPL scale. (The scheme was inspired in the scheme presented in Figure 3.2 **C** elaborated for the rat's experiments, Pardo-Vazquez et al., 2018).

4.1.2 Detailed description of the behavioral task

Sound Localization Task: Temporal and outcome contingencies

Every single trial was initiated by the subject pressing both shoulder buttons of the gamepad. This was always followed by variable fixation period (of duration equal to the sum of 700ms and an exponentially distributed (mean 400ms) duration truncated at 2.7 s), when a fixation cross would appear on the interface screen. Note that, by establishing a variable fixation time (FT) we aim that the subject is unable to predict exactly when the sound will start. After this short but variable FT, the sound was played binaurally, through the headphones, until the subject had made their decision by releasing the shoulder button on the side on which the sound was louder (Figure 4.1). The sound was played for a maximum of 4 seconds. As soon as the subjects were certain about their answer they should immediately communicate their choice, always with the understanding that accuracy was valued over speed. That is, the subjects were not compensated, point-wise, by a faster response, only for a correct one. As mentioned above, correct answers were rewarded with 10 points, and incorrect answers were penalized with -10 points and a time penalty of 1.5s. This feedback message was played at the end of each trial, being the compensation for a correct answer displayed in green on the screen for 1.5s, and the negative message for an incorrect response was shown in red, blinking for a total of 3s. The feedback message was always followed by a fix inter trial interval of 0.5s associated to a screen with no information. The score was always present at the left upper corner of the screen, adding or subtracting the punctuation of every single trial to keep the subject motivated.

It was always considered as an abort if the subject took longer than the maximum sound duration (4s) to communicate their answer, or if they would try to respond within a time interval shorter than the average smaller reaction time, 100ms. Note that, in the experiment reaction time was calculated as:

$$RT = Releasing\ trigger - Sound\ start \tag{4.1}$$

Each session had a set duration - approximately 20 minutes (variations occurred since different subjects have different performances) - and it was divided into 2 blocks of different ABLs - one block per ABL, each block constituted by 120 trials. As in the rats experiments ILDs varied pseudo-randomly across trials, that is, we guaranteed that each condition (ILD) was played the same number of times in a block, but always with a random position. The first block always corresponded to the ABL of 40 dB SPL and the second block to the ABL of 60 dB SPL, consequently by session a subject completed a total of 240 trials. A break was offered to the subjects within blocks, accompanied by a message of the subject score in the block in comparison to the maximum amount of points it was possible to obtain. A similar message was displayed at the end of each session.

Sound Localization Task: Training

The training session started with very easy discriminations (ILDs between -16 dB and 16 dB in 2 dB steps) and, as soon as the subject had reached a certain level of accuracy (80%), they would move on to a harder level of smaller ILDs. This procedure was followed until the subject had reached an accuracy degree of 80% in the level we would perform the actual experiment, ILD discriminations between 0.3 dB and 2.4 dB, always presented pseudo-randomly from trial to trial. In this way, the number of completed trials in the training session and, consequently, its duration was dependent on the subject performance. But on average, a subject needed 214 trials (from a range of 112 trials -best performance - and 288 trials - worst performance) to reach the desired performance level.

Chapter 5

Results and Discussion

The original experiment performed in rats uncovered a new temporal regularity - the TIED - describing how RT changes as a function of the stimulus' absolute level. The found relationship showed that the only effect of ABL in the RTD was a rescaling of the effective unit of time, with the RTDs, associated with different ABLs, overlapping almost perfectly if this time rescaling was applied. This result allowed to conclude that the stimulus' absolute level (ABL) determines, exclusively, the effective unit of time of the sensory process, narrowing the nature of the discriminative choices process to a bounded exact accumulation of evidence. By performing an adaptation of the rodents task to a human task, we aimed to analyze if the TIED was also a psychophysical regularity of the humans' sensory system, extending the validity of the proposed model and mechanistic explanation of Weber's Law.

A total of 13 human subjects performed our adapted version of the rodents task. After the initial training session, where an accuracy of 80% for the hardest conditions (same ILDs used in the final task) had to be achieved, subjects maintained this same accuracy level on most sessions. Only in these sessions where performance (accuracy above 80%) and RTs were stable, were used in the following analysis. On average, subjects completed 9 sessions in steady-state conditions, which will be analyzed in this chapter. Additionally, not all subjects suppressed the necessary criteria to pursue with the analysis of the results regarding the TIED. We start by presenting the criteria used to exclude these subjects.

To assess subjects accuracy, we used the SDT statistics for sensitivity (d') and criterion (c), defined as (Green and Swets, 1966):

$$d' = Z(Hits) - Z(False\ Alarms) \tag{5.1}$$

$$c = -0.5 \times \left[Z(Hits) + Z(False\ Alarms)\right] \tag{5.2}$$

Therefore, completed trials with $RTs < x_{Q(0.99)}$ (that is, excluding the trials were RTs were abnormally big) were divided into four categories:

- Hits: correct responses to the right.
- False alarms: incorrect responses to the left.
- Correct rejections: correct responses to the right.
- Misses: incorrect responses to the right.

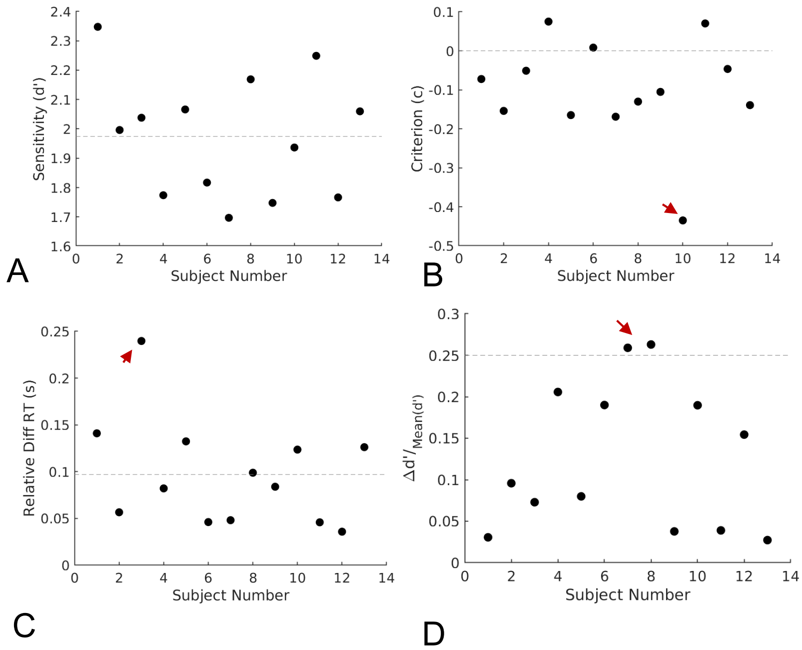

Figure 5.1: **(A)** Values of sensitivity (d') for each participant in the experiment. The dashed gray line represents the mean value of sensitivity across all subjects. **(B)** Values of criterion (c) for each one of the 13 participants. The dashed line represents $c = 0$ which means a perfect unbiased performance. The red arrow indicates the subject excluded by presenting a value of $c > 2s.t.d$ away from the mean across subjects value. **(C)** Values of the Relative Difference in RT between both side responses, defined as $Average\ \Delta\ RT\ sides/Average\ RT$. The dashed gray line represents the average of this value across subjects, and the red arrow, points out the subject who presented a strong bias in his RT pattern ($Relative\ Diff > 2s.t.d$ across subjects). **(D)** Value of $\Delta d'/Mean(d')$ for each participant in the experiment. Subjects who presented $\Delta d'/Mean(d') > 0.25$ (red arrows) were discarded, as we considered that this value was too large to be compatible with WL, making this a criterion of exclusion. All these quantities presented in all 4 plots were calculated for all subjects exclusively for trials where$RTs < x_{Q(0.99)}$.

The values of sensitivity and criterion were then estimated by applying the standard z-transform formulas. This first analysis of these values made us exclude one subject who was extremely biased, presenting a value of criterion (c) > 2 std away from the across-subject mean in criterion (Figure 5.1 B). Besides, the analysis of subjects RTs made us exclude another subject who showed biased values of RT, > 2 across-subject deviations away from the across-subject mean in the relative difference in average RT between left and right choices (Figure 5.1 C). Two other subjects were also excluded, as the first analysis of their results showed a clear violation of Weber's Law, $\Delta d'/Mean(d')$ (Figure 5.1 D). The remaining 9 subjects showed sufficiently unbiased results to pursue with further analysis.

WL states that the discrimination between two distinct stimuli (different intensities) does not depend on the overall stimulus magnitude but, exclusively, on their fraction (Weber, 1834). As ILD, by definition, corresponds to the difference between the sound level presented to both ears (Equation 3.1 and 3.3), in accordance to WL, the subject's ability to decide on the sound localization (louder side) should depend exclusively on the ILD (how different stimulus are) and not on the overall level of the sound, the ABL. This was exactly what was observed in the rodents experiment and it was what one expected to observe also in humans behavior.

As shown in Figure 5.2 A, in agreement with WL and the results obtained for the rodents experiment, on average, subjects' accuracy in correctly selecting "right" depended exclusively on the ILD and did not depend on the ABL - psychometric functions for both ABLs overlap. As expected, with higher accuracies associated with bigger ILDs and smaller accuracies associated with smaller ILDs. Additionally, differences in sensitivity (d') for ILD discriminations at both ABL levels did not reach statistical

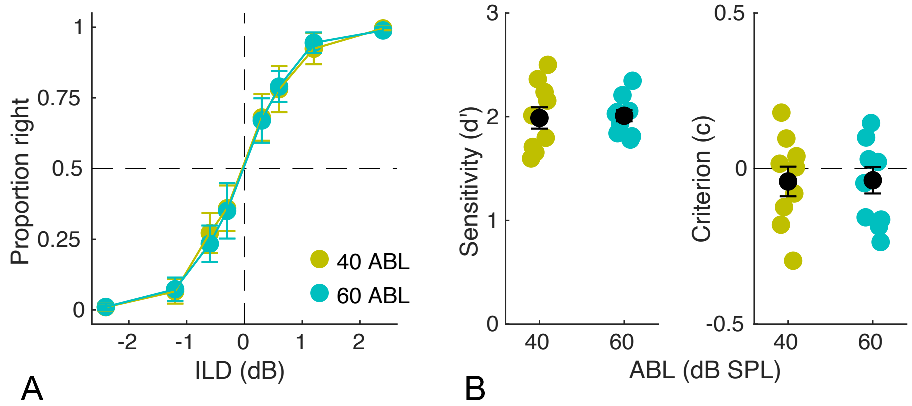

Figure 5.2: **(A)** Choose right probability (psychometric functions) for both ABL used in the human task (n = 9 human subjects; circles, across-subject mean; lines, across subjects standard deviation).**(B)** Sensitivity (d') and criterion (c) for each subject (black filled circles represent the mean; error bars are $\pm$ 1 s.d. across subjects).

difference either at the group level (Fisher's exact permutation test, evaluated at the confidence interval of 5%, $p = 0.984$), and neither for the majority of subjects (Fisher's exact permutation test, confidence interval of 5%, n=9, Table 1 of Appendix 1). Therefore, just like for rats (Pardo-Vazquez et al., 2018) and other species (Racanzone and Beckerman, 2004; Nodal et al., 2008) humans showed level-invariant ILD discrimination, corroborating previous results (Sabin et al., 2005). Thus, ILD discrimination in humans also obeys WL.

In resemblance to the rodents experiment design, this task also followed a RT configuration. To accomplish this, as explained, subjects had to continue pressing both shoulder buttons of the gamepad until they consider to know the correct localization of the auditory stimulus. As soon as they considered knowing the answer, they should release the shoulder button on the side corresponding to what they considered to be the stimulus's location. Thus, our RT measure consisted of the amount of time between the sound onset until the release of one of the gamepad shoulder buttons (Equation 4.1) - amount of time subjects needed to sample the stimulus to make the decision.

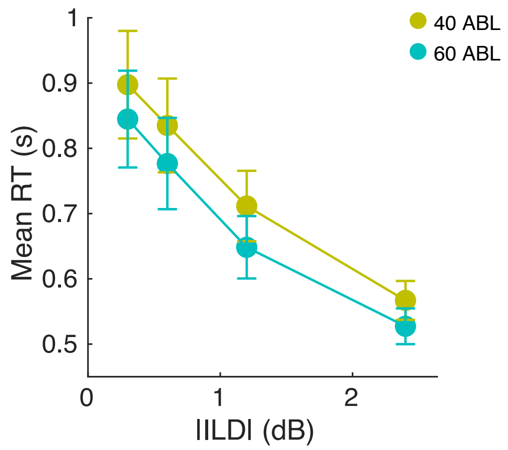

Figure 5.3: Average RT across subjects (mean $\pm$ s.e.m. across subjects) as a function of difficulty for each ABL separately.

The analysis of these RT measures showed that more difficult discriminations (smaller ILDs) are associated with longer RTs. This result exhibits that, as expected, for more difficult decisions, subjects are willing to collect sensory evidence for a longer period of time as an attempt to maintain their accuracy. This pattern, widely observed (Schouten and Bekker, 1967; Wickelgren, 1977; Chittka et al., 2009; Uchida and Mainen, 2003) is named speed-accuracy trade-off and a signature of the bounded accumulation of evidence process (Ratcliff, 2002).

In addition, as in the rodents task results, it was observed that discriminations involving overall quieter sounds (ABL = 40 dB SPL) took longer on average (Figure 5.3), showing that ABL also plays a role in the average RT presented by the subjects. To analyze the significance of both ILD and ABL effect on the average RT, we applied a repeated measure analysis of variance test (RM- ANOVA) to the RT data of each subject and to the averages across subjects, that is, for the group level. The results revealed ILD to play a significant effect on average RT for every subject and at the group level ($F(3, 24) = 16, 66$, $p < 0.0001$), and ABL to be significant at the group level ($F(1, 8) = 6, 20$, $p = 0.0375$) and for some subjects (Appendix 1. Table 2). These results lead us to conclude that, as in the rodents experiment, smaller ABLs are associated with longer RTs. However, the nature of this result was not as marked as it was for rats, with human subjects showing a stronger ILD effect than ABL on the average RT. It is noteworthy that this less accentuated effect of ABL, might be due to the fact we used in this version of the task only 2 different values of ABL (Pardo-Vazquez et al., 2018), while in the rodents version were used 3 different ones, naturally, more sparsed. Besides, both internal auditory mechanisms (human/rodents) might involve different readings of the sound pressure (SPL), but this was not tested and it is not being further developed. Resting the conclusion, that even weaker, there was an overall effect of ABL in the values of human subjects RTs, being quieter sounds associated with bigger values of RT.

To test if the TIED was also held for humans, we initially examined the dependence of the full RTDs on ABL. To do so, we started by excluding very small RTs impossible to account for the decision process. Therefore, we excluded $RTs < 200ms$ as it has been documented that RTs are around $100ms$ for the cells in sensory systems (Luce, 1986; Ditterich et al., 2003); and when associating a reaction time to the performance of a task, one must also account for the non-decision time included, when, for instance, motor information is transmitted. According to this, a RT below $200ms$ hardly accounts for the processes necessary for perception and motor response to occur, leading us to exclude these trials, which reflect anticipation and impulsivity.

The RTDs, originated by applying the kernel density estimation technique, were revealed to be right-skewed, as it is characteristic of integration to bound models, and were more narrow for louder sounds (Figure 5.5), as expected from the mean RT data. Being this result more accentuated for the harder conditions. Additionally, as expected from the previous analysis of the average RTs, the observed shift to the right of RTD associated with less intense stimuli was less accentuated in this experiment; however, it was still present.

To prove that the shape of the distributions was independent from stimulus intensity (ABL) we scaled the distributions of ABL = 40 dB SPL to maximize the overlap with the one for ABL = 60 dB SPL. By doing this, we aimed to prove that the distributions were the same, only differing in the time scale of the axis. Therefore, we applied a linear stretch of the time axis by considering the following relationship, $\rho(t) = \alpha\rho'(t)$, which implies:

$$\int_0^\tau \rho(t)dt = \alpha \int_0^\tau \rho'(\alpha t)dt = \int_0^{\alpha t} \rho'(t')dt' \tag{5.3}$$

for all θ. This relationship implies that the quantiles of the distributions are proportional. That is, if the value of previous integrals is $Q/100$, then τ is the Q^{th} percentile $PC_\rho(Q)$ of $\rho(t)$ and $\alpha\tau$ is the Q^{th} percentile $PC_{\rho'}(Q)$ of $\rho'(t)$. This fact allowed us to use the slope of a linear fit of the percentiles of the two distributions to identify the temporal rescaling factor, α, corresponding to a given change in ABL, and the R^2 of the fit to quantify its precision (Figure 5.4). As the ILD was shown in our results to be relevant in the determination of RT and it has been suggested by others to also lead to an approximate rescaling of the RTD, we performed this analysis for each pair of ILD used.

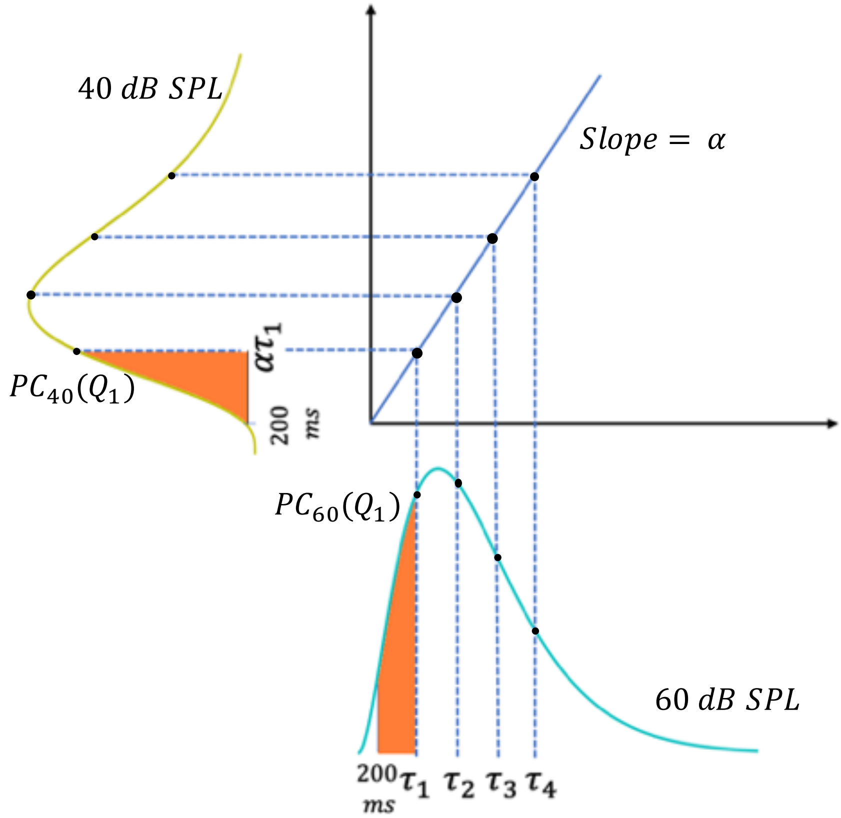

Figure 5.4: Illustration of the method followed to calculate α, the temporal rescaling factor.

The application of this temporal rescaling showed that the rescaled distributions are almost identical for each difficulty and for all difficulties combined (Figure 5.5). Even though, this process offered a smaller precision than for rats, on average, the RTDs at one ABL explained more than 99% of the variance in the shape of the RTDs for the other ABL (mean $R^2 = 0.992$). This result confirms that the sole effect of overall stimulus intensity in the RTDs is a temporal rescaling of the same, indicating the TIED to be a psychophysical regularity also present in humans perception. By other words, it is also observed in human subjects that changes in the absolute intensity of two stimuli being discriminated under a fixed intensity ratio are completely equivalent to a change in the effective unit of time, with which the discrimination duration is measured (Pardo-Vazquez et al., 2018).

Identifying the TIED in the rodents experiment allowed to constrain the characteristics of the model (from the broad class of accumulation to a bound models) able to explain WL. That is to say, the TIED allowed to build a robust explanation of how the discriminability process between 2 stimuli is independent of the absolute magnitude of the same. This because the TIED brings a stronger constraint to the discrimination process than the WL itself (Pardo-Vazquez et al., 2018). That is, while WL effectively

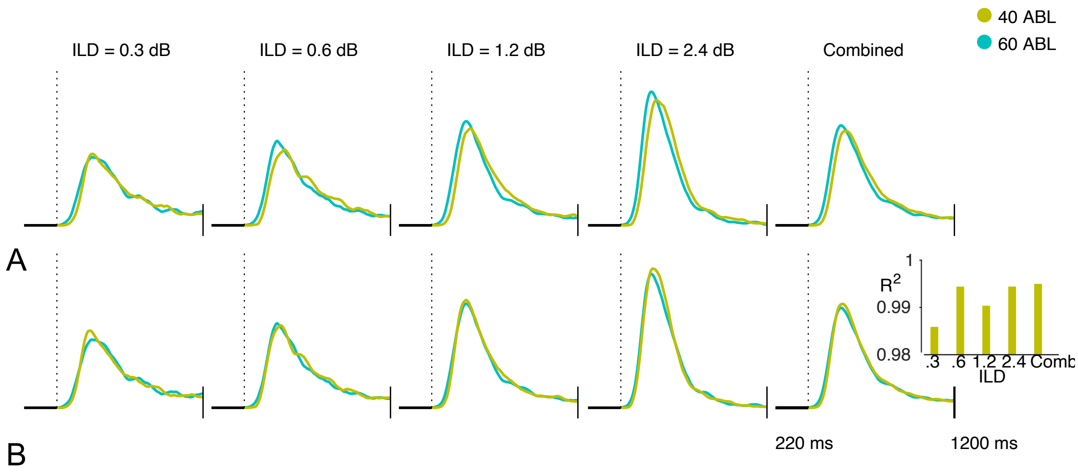

Figure 5.5: **(A)** RTDs for the two ABLs are shown separately for each difficulty (|ILD|), and combined across difficulties (far right). For all RTDs, the broken line signals the value of 200ms at which we started considering RT as valid RTs. Each RTD contains all data for that condition from all 9 human subjects. **(B)** For each difficulty, we uniformly rescaled the time to maximize the overlap of RTD correspondent to ABL = 40 dB SPL with that for ABL = 60 dB SPL. The inset shows the fraction of variance (R^2) that the RTD at ABL = 60 dB explains about the rescaled RTDs at 40 dB SPLs.

constrains the accuracy of an ILD discrimination at one ABL given its value at a different ABL, the TIED constrains not only the accuracy of this discrimination but also the full shape of its associate RTD.

Thus, it was possible to identify the essential characteristics of the computational mechanism underlying ILD discrimination: (1) Bounded exact temporal accumulation of evidence; (2) Poisson-like variability and (3) Power-law encoding of stimulus intensity, for the reasons pointed out in the previous chapter. The implementation of a model with these characteristics revealed not only to be necessary to account for the TIED, but it was additionally sufficient to account for almost all variance in the accuracy and RT of the rats (Pardo.Vazquez et al., 2018). Note that, the results found with this experiment, showed not only the TIED to be present in human perception, but importantly, revealed the same fundamental behavioral patterns, that is:

- Accuracy independent of the overall stimulus magnitude;
- Accuracy extremely dependent on ILD;
- Faster RTs for easier conditions (higher values of ILD) and longer RTs for harder conditions;
- On average, longer RTs for quiter sounds.

Therefore, if the TIED is the sufficient argument to conclude about the model's characteristics capable of explaining the process of sensory discrimination (Pardo-Vazquez et al., 2018), whose features are qualitatively shared by these results, then it can be conclude that the same model is able to describe the features of the discrimination process in humans.

Chapter 6

Conclusions

This work was developed within the scope of a major project which goal was to settle about the mechanistic foundation of Weber's Law. Inspired by the successful experiment performed with rodents, which identified a new psychophysical regularity - the TIED - able to bring consensus to the computational explanation of WL, we developed a human adaptation of the same task. By performing this adaptation, we aimed to answer if the TIED was a behavioral signature also present in other species, proving not only its generality but the generality of the suggested mechanistic foundation of WL, sustained by the own existence of the TIED.

The obtained results showed that human subjects, similarly to rodents, presented level-invariant ILD discriminations. In other words, subjects' accuracy in identifying the localization of the sound did not depend on the overall intensity of the stimuli (ABL), but, exclusively, on the intensity difference between right and left ears (ILD). To conclude, human subjects' behavior obeyed Weber's Law.

Along with the analysis of subjects' performance, we also analyzed the subjects' RTs. As found in the rodents' behavior (Pardo-Vazquez et al., 2018), human subjects presented, on average, longer reaction times for discriminations involving quieter sounds, showing RTs to be dependent on stimuli's intensity. However, when analyzing subjects RTDs for both ABLs, we found that contrarily to the average values of RT, the shape of the RTDs was the same regardless of the overall level of the sound. This leads us to conclude that the sole effect of changes in the overall level of the sounds was a change in the effective units of the sensory process, showing the TIED to be present also in human behavior.

The identification of the TIED in humans behavior shows that the computational mechanism of bounded accumulation of evidence with stochastic encoding of stimulus intensity, proposed to explain the discrimination choices of rodents (Pardo-Vazquez et al., 2018), is also the only possible model able to account for humans' behavior. As this mechanism is not only necessary for the TIED to hold but is also sufficient to provide a virtually complete quantitative description of the main features of the rats' behavior, which are shared by the human behavior: (1) Level-invariance but not condition-invariant accuracy; (2) RTs distributions with longer RTs associated with quieter stimuli.

Consequently, the obtained results lead us to conclude that the TIED seems to be a regularity intrinsic to the sensory discrimination process, sustaining the hypothesis of the bounded accumulation of evidence mechanism to be the most reliable explanation of WL (results published in the Nature Neurosciences paper Pardo-Vazquez et al., 2019). Nevertheless, it is important to highlight that in order to actually take the TIED as an inherent regularity of the discrimination process conducted by the brain in general, other experiments would have to be conducted - experiments in other species, other sensory modalities and, importantly, experiments using sequential, as opposed to simultaneous presentation of stimuli, since many of the studies on WL were done using sequential presentation.

Part II

Study on Confidence, Vigor and Cost of Time

Chapter 7

Introduction

It is observed that when faced with choices, subjects not only decide what action to perform but also, when and how fast one should perform it. Different models and experiments, as the one mentioned in the first part of this thesis, have been dedicated to the study of how and why an animal selects an option over another. However, why subjects perform these actions with different speeds - vigor - is not so well understood and, only recently, this important dimension of behavior has been receiving more attention in studies of decision making.

When performing the Weber's Law task, rats were free to select the speed with which they indicated their decision (how fast they move from the central port to the side ones) and we observed that they modulated their speed. More specifically, it seems that animals were modulating this speed according to how certain they were about the correctness of their decision, moving faster when they were more confident about their choice. Even though this relationship has been proposed a couple of times (Kiani et al., 2014; Seideman et al., 2018), establishing movement speed as an implicit measure of decision confidence is important for future work. Nevertheless, the recent studies approaching vigor suggest a causal relationship with cost of time, as the passage of time delays the acquisition of reward and that represents a cost. However, it is not clear how this value of cost of time is computed by the brain. Some suggest the brain computes a value of how much each unit of time is worth based on past experience, defining the cost of time, and of course, the vigor with which one moves. Others observe an important role of expected reward in determining cost of time. However, all these studies have been leaving confidence aside.

The aim of this study is then to address how decision confidence and reward are reflected on vigor, by proposing a vision where decision confidence is seen as the degree of certainty that reward will be received, engaging the theory where cost of time is determined by expected reward.

7.1 Reward, Reward Rate and Cost of Time

Decision-making is a cognitive process of selection of actions among several alternative options. If a certain reward is assigned to every option, the subject aims to select the option which offers him the biggest amount of reward or the reward with higher utility. However, it has been observed that a subject does not always perform the same action with the same speed. For example, a thirsty rat moves faster towards water than it would when not thirsty (Dickinson and Balleine, 2002). Moreover, when thirsty in a scenario where he can obtain water, the animal not only moves faster to get water, but generally performs all actions faster (Hull, 1943; Brown, 1961; Bolles, 1967). That is, reward and its utility (subjective value of the reward according to one's state) have an invigorating effect on behavior, which is often reflected in faster actions and smaller decision times (Niv et al, 2007) .

In the past 20 years, this invigorating effect has been connected to the concept of cost of time. The link was first proposed by Niv and her colleagues (Niv et al., 2006; Niv et al., 2007), based on the idea that the time it takes to perform a movement carries a cost, as it delays the acquisition of the expected future rewards. As a consequence of this idea, they suggested that the optimal policy followed by a subject is not only to maximize reward but reward per unit of time – the average reward rate - being the optimal vigor of an action a trade-off between the higher cost of time inherent in a slower action and the energetic costs associated with faster actions. Succinctly, the theory introduced the idea that the time itself also envisages a cost, higher the higher the value of what the subject can obtain from the environment. For example, if a subject is thirsty and can obtain water for a correct reward, he will move faster as the passage of time in this situation embodies a higher cost for the subject than in a scenario where the animal is not thirsty. When introducing this opportunity cost of time in their reinforcement learning model, Niv et al. were very successful in explaining the observed modulations of vigor in decision-making environments. Consequently, the model has been broadly accepted and has inspired future works.

An important and recent debate is how this value of cost of time is formalized by the brain. According to Niv et al. and others (Guitar-Masip et al., 2011; Beierholm et al., 2012; Constantino and Daw, 2015; Otto and Daw, 2017), the cost of time is computed by an average of the richness of the environment and is reported by tonic levels of dopamine in the nucleus accumbens (Niv et al., 2007). That is, subjects compute a long-run estimate of how much they should receive per unit of time - the expected average reward rate - based on the rewards received in the past. The rate computes the cost of time and then determines the optimal speed of an action, which is the speed that allows to maintain exactly the average reward rate. For example, the higher the average reward received in the past, the larger the value 'allocated' to each unit of time. By acting slower, the subject is losing more than in a scenario where little was received in the past and, consequently, a lower estimate of how much the unit of time costs is computed. Computational and empirical experiments have been giving strong evidence supporting this theory (Guitar-Masip et al., 2011; Beierholm et al., 2012; Constantino and Daw, 2015; Otto and Daw, 2017). Particularly, because this interpretation not only allows to predict the increase in vigor of actions that lead to higher rewarding states, but also the observed general increase of vigor in all actions, including ones that do not themselves lead directly to the rewards.

However, if this theory makes sense in a stationary environment, where the past can be a good prediction of the future, what about in a constantly changing environment? The model of Niv et al. does not mention 'local' fluctuations of the expected reward, as the agent only maximizes the long term reward in a stationary environment. Additionally, it neglects the fluctuations in motivation that can occur after collection of reward (i.e a hungry rat after being successful in collecting food a couple of times does not feel the same urge to work for it). Another relevant aspect lies on the fact that the theory predicts a subject to move with the same speed in order to obtain a reward of great and less value; while there is evidence that the subjects move faster and present a shorter reaction time towards rewards they value more (Kawagoe et al., 1998; Wilson et al., 2009; Sackaloo et al., 2015). Here is where the debate begins.

Recent experiments have been performing this study of vigor in a non-stationary environment, an environment where the reward offered for a correct decision is always changing, but where subjects are aware of the amount of reward being offered for every decision and consequent action. On the one hand, and in alignment with the model and results of Niv et al., some recent experiments have demonstrated that immediate reward plays no role in vigor modulation, being this mainly predicted by an average of the 'near past' reward (Guitar-Masip et al., 2011; Constantino and Daw, 2015;Constantino et al., 2017;

Otto and Daw, 2017). On the other hand, other experiments clearly show that subjects present shorter reaction times and bigger movement speeds when pursuing bigger rewards (Milstein and Dorris, 2007; Wilson et al., 2009; Opris et al., 2011; Yamamoto and Hikosaka, 2013; Reppert et al., 2015; Sackaloo et al., 2015; Summerside et al., 2018; Revol et al., 2019). That is, while knowing what is the payment for a correct action, subjects not only initiate movement faster but they also move faster. Supporting this view, studies of dopamine in the midbrain, indicate that the dopamine signal accounts for future reward and the motivation towards that reward (Tachibana and Hikosaka,,2012; Hamid et al., 2015; Walton et al., 2018).

Therefore, it is not clear what is the influence of expected reward and past reward on cost of time and, consequently, on vigor. In order to disentangle this question a new experiment needs to be performed. Additionally, it is important to highlight the fact that not all the mentioned experiments use the same approach to what is vigor. While some experiments refer to vigor as reaction time together with movement time (Guitar-Masip et al., 2011; Beierholm et al., 2013; Otto and Daw, 2017), not distinguishing both, and standing out the impact of reward on reaction time; others, analyze vigor as movement velocity or pressing rate (Niv et al., 2007; Summerside et al., 2017). Then, it is also fundamental to investigate the extent of the impact of reward in every single one of these behavioral dimensions.

7.2 Confidence and Vigor

In the Renart Lab, when animals are free to select not only what action to perform but the speed – vigor- with which they execute it, one observes that the time between exiting the central port (moment where the animal has made his decision) and entering the lateral port (left/right) – designated Movement Time (MT) – has the following three characteristics:

1. X pattern: MT decreases as a function of evidence strength (ILD) for correct trials but increases as a function of ILD for incorrect trials;

2. MT decreases with accuracy, when both measures are marginalized over difficulty.

3. For a given difficulty, accuracy increases the smaller the MT.

These characteristics of vigor (Figure 7.1), in this case corresponding to 1/MT (Movement Speed - MS), have striking resemblances with decision confidence (DC), which leads to the belief that both are related.

DC refers to the subjective belief, prior to feedback, that a decision is correct (Kiani et al., 2014). In a human experiment, where explicit reports of confidence – self- report of one's certainty in his answer - were collected, one observes the same behavior of MT with strength of evidence (Kepecs et al., 2016). That is, the apparent overconfidence in choices based on uninformative evidence, and the strange decreasing of confidence with increasing evidence strength for erroneous choices (Figure 7.2). A relationship between vigor and DC has been pointed out before (Kiani et al., 2014; Seideman et al., 2018) supporting our understanding and observation of a monotonic behavior of vigor with DC. One of the goals of this work is, therefore, to expose vigor as being an implicit measure of DC.

However, if vigor reflects the degree of certainty that a decision is correct, how is this in line with previous studies that identify vigor as being the reflex of cost of time? Our prediction departs from the understanding that DC corresponds to the degree of belief that a certain reward will be acquired. That is, that DC sets the value of expected future reward. In alignment with experiments which demonstrate that expected near future reward modulates vigor, in our view, being very confident corresponds to being certain that the reward will be received, making the subject move faster, as the passage of time is more costly than if one is sure that no reward will be received. More succinctly, our view proposes vigor to be influenced by the expected reward, $< Reward >= DC \times Reward$, that is, cost of time to be determined by expected reward. Testing this prediction, is therefore one of the goals of this work.

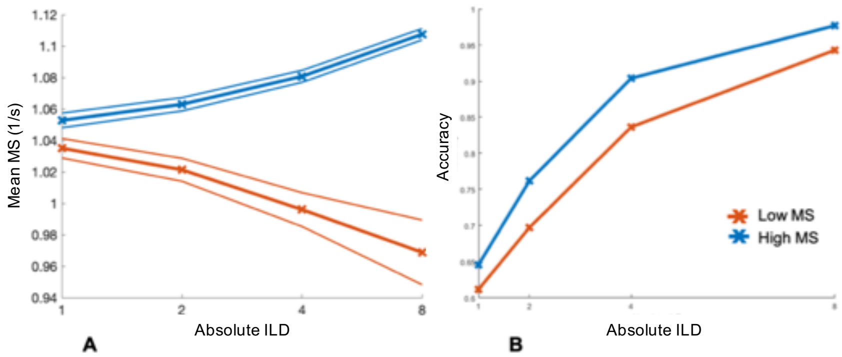

Figure 7.1: Results regarding the Movement Speed (1/MT) of the rats when performing the Weber's Law described in the first part of this thesis. **A** MS as function of ILD (evidence strength) for correct responses (in blue) and incorrect responses (in red). The results represent average results across subjects. **B** Accuracy as function of ILD (evidence strength) for faster responses (blue) and slower responses (red). Faster responses are associated to a higher accuracy.

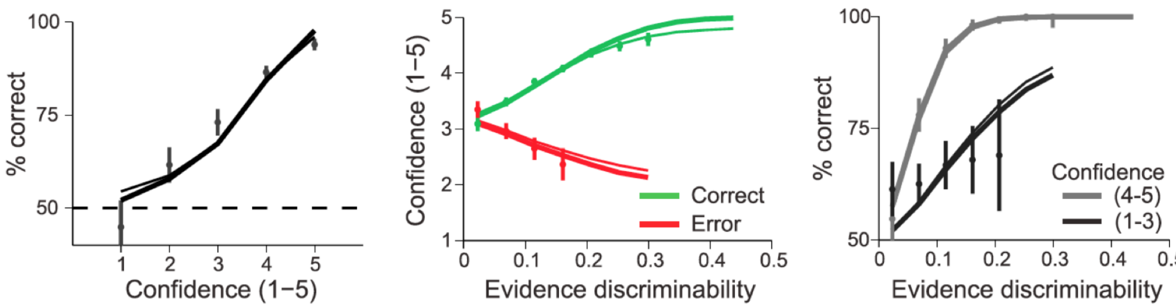

Figure 7.2: Results of Kepecs experiment when studying human DC. In their experiment, subjects indicated the level of confidence on their decision on a 5-division scale between a random guess (1) and high confidence (5). The plots represent the confidence patterns of a single subject. Thick lines show parameter-free normative statistical model sitmulations. Thin lines show one-paramenter model fits with a efficacy parameter. (Image adapted from (Kepecs et al., 2016)).

7.3 Temporal Discounting of the Reward

In humans the cost of time has been strongly associated with the discount of reward. For example, several experiments reveal that a lot of people prefer to receive a smaller amount of money now rather than a bigger one in the future (Myerson and Green, 1995). Naturally, this decision depends on the amounts being offered and the difference in time that separates them. However, if a smaller amount of money is selected it means that the value of the bigger one is devalued by the passage of time. That is, that time has a cost.

Some studies have been proposing different functions for how time devalues reward (Shadmehr et al., 2010; Berret and Jean, 2016). The most accepted one is the hyperbolic reward discount function:

$$V(t) = \frac{\alpha}{1 + \beta t} \tag{7.1}$$

Which translates how much will be the value of the reward α after the passage of time - t. Being β the rate at which one discounts the reward. That means, the cost of the passage of time is in terms of reward given by

$$C(t) = \alpha \left(1 - \frac{1}{1 + \beta t}\right) \tag{7.2}$$

This equation is important because it directly exposes the findings mentioned previously. First, it reveals the longer the movement duration, t, is, the larger the loss of reward is. Additionally, it explains why a bigger reward would involve a bigger loss of reward, justifying why, when there is more at stake, a subject moves faster.

Experiments where a subject has to pick between different amounts of reward in different moments in time are important, because they allow one to measure exactly how the passage of time devalues reward. That is, they allow one to compute the value of β, finding in this way a value for cost of time. It is, for example, the case of Shadmehr's study (Shadmehr et al., 2014) where it was measured how willing a subject is to wait for a reward. In this experiment, there is a second stimulus that has a chance to appear with a delay in every trial and signalizes the reception of reward. Therefore, if the subject wants to catch the reward he has to wait long enough for this second stimulus to appear, without never knowing if the current trial is a trial with this 'special' stimulus and when exactly, if present, it will appear. Strategically, every time the subject succeeds on waiting for this second stimulus, the delay increases in their experiment. When the subject fails to catch it, the delay decreases. In this way, it was possible to calculate for each subject the exact amount of time one was willing to wait for the reward, computing individual estimates of cost of time. Their results were successful in proving that the subjects who waited less in general - deeper reward discounters - were the ones who presented more vigorous saccades. This is extremely relevant, as it directly connects the cost of time with the vigor of the movement.

Based on Shadmer's experiment, it is our understanding it would be important to include a similar temporal discounting protocol in an experiment that aims to relate cost of time with vigor. By including a discounting protocol, one can evaluate how changes in reward and DC not only affect vigor (by measuring directly its variations), but how they directly affect cost of time, by measuring if they play a role in the time one is willing to wait for a reward. Note that according to Equation 7.2, and the theory of temporal discounting of reward, the offer of a bigger reward would make the subject wait less. The goal is to identify if this relationship is better predicted by reward, solely, or expected reward, that is, if also suffers an influence of the level of expectation the subject has about receiving the reward, the DC. As it is our previously mentioned prediction, which we aim to study.

7.4 Specific Aims

In this project we aim to create a human experiment able to approach all the questions mentioned above, that is:

1. To establish vigor as an implicit measure of DC;

2. To address how/if DC determines the vigor of a movement, by determining the expected reward. As mentioned above, our hypothesis predicts that expected reward, proposed to be defined by $<$

$reward >= DC \times\ reward\ on\ offer$. This study aims to respond to this proposed relationship.

3. To study if cost of time is computed by an average of the reward by unit of time received in the near past, or reflects the near future expected reward.

A summary of the findings that lead us to establish these goals were described above. The next chapter will explain how we designed an experiment able to respond to these challenges.

Chapter 8

A sound localization task in a changing environment with temporal discounting

To address how movement vigor is affected by - decision confidence, reward, past reward - evaluating how each of these variables changes the way one values the passage of time, we developed a human sensorial task. This chapter dedicates itself to explain each one of the features of the experimental design.

8.1 Methodology

A two-alternative forced choice (2AFC) human auditory task with a temporal discounting protocol was developed. To concatenate measurements of all the variables proposed to be studied, the task followed a particular design, as explained below.

In order to address the first goal - establish vigor as an implicit measure of DC -our task had to:

(1) Evoke uncertainty among subject responses;
(2) Assess self-reports of decision confidence;
(3) Measure the movement speed associated with the report of each one of the confidence levels.

(1) DC is commonly studied in the paradigm of perceptual decision-making, that is, through choices based on sensory evidence gathered through the senses. Naturally, a perceptual decision-making task has different sources of uncertainty (e.g., noise in the sensory system and noise in the sensorial evidence). However, the ability to directly manipulate the stimulus condition allows the control of the decision's difficulty, by generating different levels of uncertainty. As a result, we decided to implement an auditory two-alternative forced decision-making task in this study, using streams of sound clicks as sensory evidence, since this type of stimuli has been shown empirically and computationally to easily relate to responses' uncertainty (Radillo, et all., 2019).

(2) Humans are the only species able to give confidence ratings using qualitative semantic classification (e.g., "absolutely certain", "almost certain", "very uncertain", etc.) or quantitative scales (Green and Swets, 1966; Granziano and Sigman, 2009; Smith et al., 2012). Given this, for long, confidence judgments were considered to be a uniquely human cognitive ability (Flavell, 1979; Metcalf and Shimamura, 1994; Tulving, 2005). Nowadays, we know that is not the case, although humans are the only species able to explicitly report confidence, without using another variable as an approach. For that reason, our

experiment was naturally performed with humans.

(3) To measure the vigor associated to different levels of DC, we recurred of a Numpad to collect our measurements (Figure 8.1 A). We used a three-level rating of DC signalized by different levels of the Numpad (Figure 8.1 A):

- Level 1 - Not confident.

- Level 2 - In-between.

- Level 3 - Confident.

In this way, we were able to collect explicit values of DC by asking the subjects to report their choice - left or right (blue or yellow - Figure 8.1 A) - in the adequate level of the numpad. Simultaneously, this design allowed us to measure the velocity associated with the report of each value of DC, by measuring the time (MT) each subject took to move the finger from the green key until the selected response key in each one of the confidence levels.

Vigor has been defined as the speed of an action. By measuring the time one takes to indicate his/her response, one can assess the value of speed - MS. MS was our primary way of analyzing vigor. However, other studies (Niv et al., 2005, Niv et al., 2007) have reported that animals modulate the rate with which they press a lever that leads to food obtainment. This lever-press rate has been in these experiments their measure of vigor, additionally to the speed with which the animal moves towards the lever. Similarly, in this study, we decided to include a system where subjects had to press multiple times in the selected response key to lock their response. We measured the time between each one of the presses - response press rate - to test if varied according to the other task's parameters.

The second and third aim of our experiment involved the study of reward and its effect on vigor. Therefore, we introduced a reward system in our experiment, where subjects were on each trial informed about the reward available for a correct answer. In this way, we could associate the measurement of vigor to the reward value being offered at the trial. Also, the value of reward was made variable across trials. By varying its magnitude on each trial, one could decouple the effect of DC from reward on vigor. Another advantage of making the immediate available reward completely random on each trial (so as to maximize local fluctuations) and independent from trial to trial is creating an environment where the past does not indicate anything about the future. Allowing one to test if the policy followed by the subjects complies with the retrospective hypothesis, which suggests that subjects compute a running average of the past and that the current trial has less (or no) weight; or with our hypothesis, which implies subjects care about the reward available on the given trial.

In addition to measuring cost of time indirectly through response vigor, which relies on the assumption cost-benefit analysis to determine action kinematics, the standard way of measuring empirically the subject's cost of time is, as mentioned, through discounting experiments. Given the success of Shadmehr et al., we decided to include a similar discounting protocol in our experiment. We did it by adding the possibility of a bonus to appear with a certain delay after the sensory response lock in our task. The delay with which the bonus appeared was dependent on subject's performance, but subjects never knew the trial or the delay with which the bonus would appear. Therefore, on each trial they had to decide to wait or not and for how long. The amount of time each subject was willing to wait for the bonus constituted our measure of cost of time, allowing us to directly test if it was affected by DC or reward in our experiment.

In sum, our experimental design aimed the measurement of the following behavioral dimensions:

1. Self-reports of DC;

2. RT, as some other works indicate a reflection of reward in the amount of time one takes to decide

(Guitar-Masip et al., 2011; Beierholm et al., 2013; Otto and Daw, 2017) ;

3. MS;

4. Response Press Rate;

5. Waiting Time (WT) - Amount of time each subject was willing to wait for an extra reward. Our approximation of cost of time.

The final result was a 2AFC auditory task implemented in humans, where subjects had to indicate which side (left or right) they considered to present a higher probability of playing a sound click. Each trial started with the display of the amount of reward at stake for the trial, followed by a variable fixation time (FT) where subjects had to press the green key uninterruptedly. The sound clicks' display started after the fixation period, where the subject had to remain pressing the green key. As the task was RT-oriented, subjects should only release the green key after they had made their decision. The time between the beginning of the stimulus display and the subject release of the green key was our measure of each trial RT, Figure 8.1 B. After releasing the green key, subjects had to indicate their choice (left/right) simultaneously to their confidence level Figure 8.1 B). As mentioned, the time between releasing the green key and the first press on the selected response key was our measure of MT ($vigor = distance/MT$) (Figure 8.1 B). Besides, to lock their response, subjects had to press several times on the selected response key until the bar showed on the screen was filled. By measuring the time between each one of these presses, we accomplished an additional measure of vigor (Press Rate in Figure 8.1 B).

After the response had been indicated, subjects had to decide to wait or not for the bonus. If they decided to wait, subjects would go back to green key and press the key at a natural rate for as long they felt like waiting (Figure 8.1 B). As soon as subjects had decided it was no longer worth to wait for the bonus, they would cash their trial reward by pressing the red button. In case of being a trial where the bonus was present, its appearance was marked by a general beep sound and a bright green screen. In trials where the bonus was present, and subjects successfully waited enough, they did not need to cash the reward by pressing the red button. The feedback of their sensorial response (correct/incorrect) plus the payment for waiting would appear immediately after the beep sound. If not, and subjects decide to move to the next trial by pressing the red button, feedback would appear after the press. Feedback was given informing the subject about the correctness of their sensory decision and about their performance regarding the bonus, which can be:

- The value of the bonus if the subject succeeded in waiting for it;
- Information about type of bonus trial: "There was no bonus"/"You lost a bonus".

8.2 General Parameters of the Task

8.2.1 Subjects

All procedures were reviewed and approved by the Ethics Commission for human studies of the Champalimaud Centre for the Unknown, and all subjects signed a consent form also approved by the Ethics Committee.

A total of 20 healthy human subjects, 13 female and 7 male, aged between the 20 and 50 years old performed the task. All subjects had right handiness except one, and all performed the task with their dominant hand. No subject had knowledge of any hearing impairment, and no abnormality was found in the analysis of the data. All subjects were naive about the aims of the task. 19 of the subjects performed

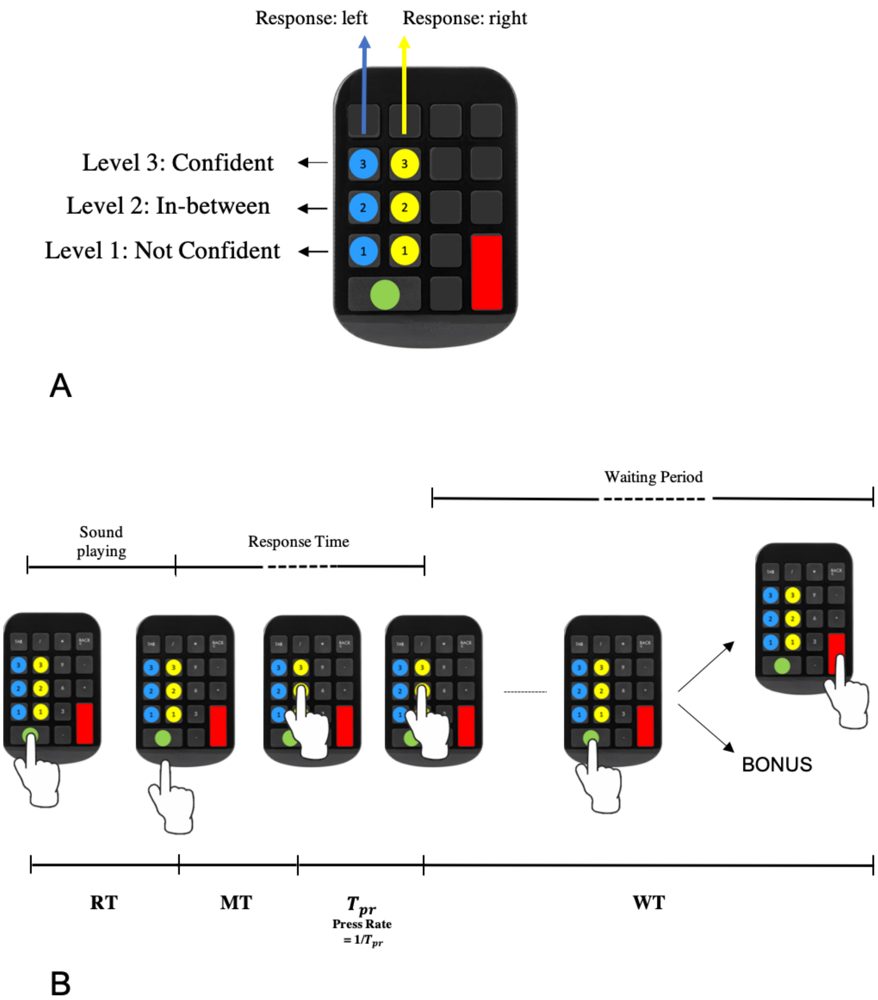

Figure 8.1: **(A)** Scheme of the Numpad used to collect data. Only the colored buttons were used in the experiment. To report choice subjects should either press left side (blue) or right side (yellow) reporting, simultaneously, their DC by reporting the response on the adequate key' level. Level 1 - meant to be unsure, level 2- in between and level 3 - confident. **(B)** Scheme of the different task's moments and how our temporal behavioral measures where collected. RT corresponds to the amount of time passed between the beginning of the stimuli' play and the moment where subject has decided marked by the release of green key. MT corresponds to the time between subject's releasing of the green key and the first press of the selected response key. The time between presses until the final lock of the sensory response corresponded to the average presses rate. At last the time between beginning waiting by pressing green at natural rate after the indication of the sensory response, and the final of the waiting period marked by the pressing of the red button or either the appearing of the bonus correspondent to our measure of WT on each trial.

a total of 20 sessions and a variable number (3-5) sessions of training. The training consisted in slowly introducing the subject to the several contingencies of the task and making sure the subject was able to perform them. All subjects passed the training sessions, although one of the subjects, given to schedule incompatibilities was unable to finish the 20 sessions, and therefore, was discarded in the following analysis. Two other subjects were excluded from the analysis given the bias in their performance. One subject was excluded under the criterion of presenting an average value of sensitivity (d') inferior to two standard deviations across subjects away from the mean value of sensitivity across subjects; and the other subject under the criterion of presenting a value of criterion (c) superior to two standard deviations across subjects away from the mean value of c across subjects. This analysis will be better depicted in the Results and Discussion's section.

The subjects were divided in 2 groups of 10. Each group performed a slightly different version of the task with different rewarding systems, but which translated in similar amounts of money by session, that was on average 5 Euros. However, the amount of compensation received per session was very dependent on the subject performance and on the amount of time one was willing to wait for the bonus (there were compensations between 0 Euros and 10 Euros). Each session was composed by 48 trials, but the amount of time spent on each made the session duration be extremely variable.

8.2.2 Material and Setup

As in the first experiment presented in this thesis, the task was developed using the software Matlab version R2018a recurring of the Psychtoolbox-3 (http://psychtoolbox.org) to produce the images on the screen and the auditory stimuli (in parallel with the sound card - Asus Xonar DX PCI Express 7.1 Audio Card). To deliver the auditory stimuli we used the over-the-ear headphones (Sennheiser HD 300 PRO) with a good sound isolation power. Subjects used a number pad (brand at home) to responde. In this numpad we marked the keys corresponding to the different confidence levels and different localization responses (right or left) with color stickers. The number pad was not fixed to the table, so the subjects could maximize their comfort while doing the task. However, we requested the use of the same hand and the same finger (index finger) to indicate the responses. The task was performed in a close, small room with reasonable sound isolation and the headphones were placed in the subject's head at the beginning of every session.

8.2.3 Auditory Stimuli

The auditory stimuli consisted of streams of sound clicks played to both ears. The clicks were presented with a frequency of 20 Hz, which means that by second 20 clicks were generated, but never simultaneously in both ears - if one click was being presented to the right, the left side was silent.

The decision-making task consisted in identifying which side had a bigger probability of presenting a click. For that, 3 difficulty conditions were used, that is, 3 proportions of clicks (0.52 - hard, 0.55 - medium, 0.60 - easy) towards both sides. If the selected proportion of clicks to the right was 0.55, to the left would be 0.45 and vice-versa. That is why, throughout the rest of this thesis, we will only refer to the proportion of clicks used to the right (0.40, 0.45, 0.48, 0.52, 0.55, 0.60), as they represent the 3 difficulty conditions for both sides. In the way the streams of clicks were generated, each one of these conditions represents a number of intended clicks to play on the right and on the left. For example, if the selected condition was 0.55 means that a total of 110 clicks ($clicks\ frequency \times condition \times max\ sound\ duration$) was predetermined to play on the right side, and 90 clicks to play on the left. However, the positions of each one of these clicks was randomly selected on each trial. Therefore, the total number of clicks or condition only set the probability of a click being played to each side, because the actual number of clicks heard on each side was extremely dependent on the amount of time spent listening to the stimulus. To exemplify, for the hard condition the number of clicks predetermined to play on both sides is very close. As their position is determined randomly, it can happen if the subject listens to a small portion of the clicks, to hear a bigger number of clicks to the incorrect side. Therefore, the subject will identify the incorrect side as having a higher probability of displaying a click.

Each click consisted in a sum of pure tones (at 2,4,6,8 and 12 kHz) convolved with a cosine envelope of 3 msec in width. The speakers calibration was made guaranteeing each one of the pure tones was correctly detected, within a time window of 3 msec. Once again, the calibration was performed in a soundproof box, with a styrofoam head with the Bruel $\&$ Kjaer Free-field 14 microphone placed inside of the ear of the styrofoam head, thus replicating as much as possible the distance to the eardrums and the aerodynamics of the sound waves in this path.

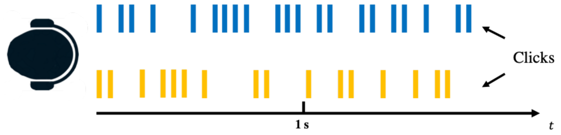

Figure 8.2: Schematic of the clicks' streams used in the task. In blue is represented the stream of clicks presented on the left and in yellow the stream presented to the right. The auditory stimulus was always initiated with a simultaneous click, but posteriorly, it consisted on a pseudo-random alternation of the clicks from the right to the left, but ensuring the total number of clicks presented on the right corresponded to - $condition \times 20Hz \times 10sec$. Being $20Hz$ the frequency of the clicks and 10 seconds the maximum amount of time the stimulus was displayed. The condition corresponded to one of the proportions of clicks (0.40, 0.45, 0.48, 0.52, 0.55, 0.60). Therefore, the total number of clicks to the right could be (80, 90, 96, 104, 110, 120). More importantly, as the position of the clicks was randomly selected on each trial, not always a proportion of the sound would reveal the correct answer.

8.2.4 Reward

The reward constituted a fundamental part of our task. The 20 subjects were divided in two groups and each group completed the task under 2 different reward systems. For the first group (Group 1) we developed a point-accumulating system. At the beginning of each trial it was announced how many points the subject could gain with a correct response. In case of an incorrect answer there was no penalty in terms of reward, only a small time penalty ($2s$). In the left upper corner on the screen we kept the amount of points accumulated so far, to keep the subject motivated. The amount of points on offer in each trial was generated from a uniform distribution (1-100 points). The subjects were not aware of the real conversion factor between points and euros, but they were informed that on average the amount of reward received per session was around 5 $Euro$. The real conversion factor was 100 points = 0.25 $Euro$ and at the end of each session a message showing how much the subjects had made in cash was displayed.

The second group of subjects (Group 2) performed the task with a discounting reward system. Each subject would start the task with a balance of 5 $Euro$. At the beginning of each trial they were informed of how much they could lose (in Euros) with an incorrect answer. A correct answer did not increase their balance. The amount of reward at stake on each trial was also generated from a uniform distribution, with values between (0.1 $Euro$ and 1 $Euro$).

For the discount protocol, the subjects were aware that after they had indicated their response of which side (left/right) presented the higher chance of presenting a click, there was a certain probability that a second stimulus could appear. If they waited long enough and indeed there was a second stimulus, they could increase their points balance (point accumulating system) or they could earn an extra amount of money (discounting reward system). The subjects were not informed a priori of how much they could make by waiting for the 'bonus', that is, what was the exact value of the bonus. The values were, once again, generated from a uniform distribution on both cases, 1-50 points or 0.01 $Euro$ - 0.50 $Euro$, respectively.

8.3 Detailed description of the behavioral task

Sound Localization Task in a changing environment with a temporal discounting protocol: Temporal and outcome contingencies

Every single trial was initiated by the subject pressing any key of the numpad. This prompted the appearance of a screen indicating the amount of reward (gain in points for group 1 and loss in Euros for group 2) for 1.5 seconds (Figure 8.3 A). Following this, a fixation cross appeared on the screen initiating a variable fixation period (of duration equal to the sum of 700ms and an exponentially distributed (mean 400ms) duration truncated at $2.7s$) during which the subject had to press continuously the green key, otherwise it was considered a too short fixation trial and fixation was initiated again (Figure 8.3 B). After the variable fixation period, the sound was played binaurally, through the headphones, for as long as the subject kept pressing the green key up to a maximum stimulus display of $10s$ (Figure 8.3 C). The amount of clicks played was always higher to one side and the subject had to choose the side he/she considered to present a higher probability of playing a sound click. To indicate the response, after releasing the green key the subject had to report both choice and confidence level by pressing multiple times on the selected response button (Figure 8.3 D), according to the confidence rate showed in Figure 8.1. Each one of these buttons presses had a probability of 20% of filling a level of the bar that was displayed on the screen at the moment of response (Figure 8.3 D). To record his/her response the subjects needed to fill the whole bar, which translated in a different number of presses in every trial.

Subjects were informed about the possibility of encounter a bonus after a certain period following their sensory response, offering the possibility of earning an extra amount of money. The presence of the bonus happened in 20% of the trials per session but subjects were not aware of this ratio as neither the amount of reward possible to earn with each bonus. The appearing of the bonus was signalized by a pure frequency beep sound and with the display of a green screen for $0.75s$. Therefore, in every trial, after the sensory response, the subject had to decide to:

a. Wait - by repeatedly press the green button of the numpad at a rate felt natural to the subject (Figure 8.3 E).

b. Not wait any longer - by pressing the red button of the numpad signalizing the decision of moving to the next trial (Figure 8.3 F).

Subjects were only informed of the outcome of the trial after they have encountered the bonus or after they have cashed in their sensory choice (Figure 8.3 F). It is important to note that, regardless of whether or not the trial had a bonus, and whether or not the subject got the bonus, if the sensory response was correct the amount of reward indicated at the beginning of the trial would sum to subject current balance (Group 1) or not discounted of the current subject balance (Group 2). The reason is that, otherwise, we would be implicitly creating an incentive for the subjects to wait longer if they were more confident about their response according to recent works (Kepecs and Mainen, 2018), making it impossible to analyze the effect of DC on cost of time, as we would be inducing a contrary effect. In the same way, in case of an incorrect response subjects of group 1 were only penalized with time ($2s$) and the subjects of group 2 would lose the amount of reward indicated at the beginning of the trial and would also receive a time penalty of $2s$. The feedback message was displayed in the screen always for $1.5s$ and in addition to inform about the sensory response, informed about the bonus, which could be:

a. The subjected waited enough and it was a "bonus" trial, they would hear the general beep sound and they would receive automatically the bonus reward (number from a uniform distribution between 1point and 100points (group1) and 0.01 and $0.5euro$.

b. If the trial had a bonus but the subject did not wait long enough, the feedback of the trial included a message saying "you missed a bonus!".

c. In the case where the trial had no bonus, the feedback included the message "There was no bonus!".

This feedback message was always followed by a fixed inter trial interval (ITI) of $0.5s$ associated to a black screen with no information. The score was for both groups always present at the left upper corner of the screen, adding or subtracting reward according to the subject performance on each trial. At the end of each session, subjects of group 1 received a message with information of how much money they have earned accordingly to the amount of points they had summed.

Each session had always the same number of trials (48) and this information was given to the subjects. Each one of the stimuli conditions was played a total of 8 times per session, and the duration of the session was naturally really dependent on subject performance and average amount of time spent waiting for the bonus.

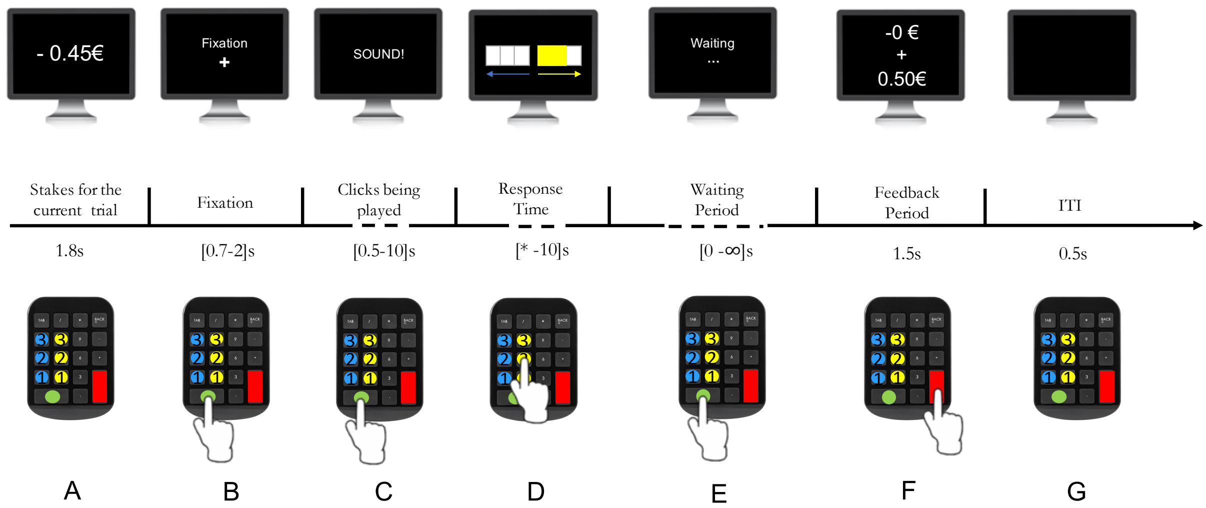

Figure 8.3: Task structure and stimulus set - Sound localization task in a changing environment with temporal discounting: On top is a schematic depiction of the different trials events on the screen. In the bottom is depicted what the subject has to do on every moment of the task, described in the timeline presented in the middle.

Training

The design of our task turned out to be rather complex, with each trial involving several moments as described above. To ensure each subject was well acquainted with all the task contingencies, a training was carried out before pursuing to the final version of the task. In this way, the training involved 2 different versions of the final task. The first, where the subject was introduced to the auditory stimuli, only having to report the side considered to present a higher probability of displaying a sound click, left or right. In this version, subjects had to exclusively report their response on the level 2 of the numpad (Figure 8.1 A) in order to not create a bias on MTs. In resemblance to the final version, subjects should indicate their response as soon as they felt they knew the correct side, that is, as soon as they considered to not be necessary to listen more of the auditory stimuli. After reaching an accuracy of at least 70% in this simple version of the task, subjects could advance to the next level of training. In this new stage, subjects had to simultaneously report the sensory response and level of DC. Therefore, subjects were instructed about the meaning of the different levels signalized on the numpad (Figure 8.1 A). The reports of DC were evaluated for each subject, and if they followed what was expected - reports of higher levels of confidence on easy trials and reports of low levels of confidence on more difficult trials - subjects were ready to advance to the final version of the task.

Note that the number of necessary training sessions was, naturally, linked to subject performance on the same. Besides, it is worth mentioning that even though subjects knew they were not being compensated for their time in these training sessions, the reward feedback was included from the beginning to facilitate adaptation.

Chapter 9

Results and Discussion

In this chapter, the results from the behavioral task are presented along with their discussion. Our experiment was ambitious, and it targeted the measure of several dimensions of behavior. Therefore, the following chapter divides itself into four main sections, each dedicated to the analysis of specific variables, with concrete goals:

(1) Accuracy, Reaction Time and Decision Confidence - Presents the analysis of accuracy, RT, and DC to approach our sensorial task's decision-making process. Aims to address what determined DC in our experiment and the reliability of the collected self-reported confidence levels.

(2) Movement Speed and Decision Confidence - After showing the sensitivity of the reports of DC we are in the conditions of analyzing the relationship between movement speed and decision confidence, exposing their correlation. Proposes to settle on one of the main questions of this experiment - Is MS and implicit measure of DC?

(3) Reward and Vigor - Reward and its potential impact in movements' speed were among the main study's objects of this thesis. Thus, this section will approach the results of vigor in an environment reward-wise always changing, but where subjects are aware of what is at 'stake' for each decision.

(4) Discounting protocol and Cost of Time - To directly link the modulations of vigor with cost of time we introduced in our experiment a temporal discounting protocol. It is also another way of assessing the relationship of reward and DC directly with cost of time. Therefore, this section will be entitled to approach the results of subjects' WTs.

9.1 Accuracy, Reaction Time and Decision Confidence

In our 2AFC task, subjects received auditory information simultaneously on both ears and had to decide which side (left or right) they considered to present a higher probability of playing a sound click. With this protocol, we sought to induce uncertainty among decision responses to collect reliable measurements of DC. This section is then dedicated to showing that subjects were able to learn the task and that our DC measures were trustworthy, understanding they were related with strength of the auditory evidence and, the time one took to decide, RT.

9.1.1 Training

Before subjects entered the final version of the task they had to complete training to guarantee they were well acquainted with all the task contingencies. The evaluation of each subject suitability to advance in the experiment followed a two-step process, which results and analysis are presented here.

Succinctly, subjects from both groups started by performing a very simple version of the task which goal was introducing the auditory stimuli to the subjects. As soon as subjects reached a performance of 70 % accuracy, they could proceed to the next stage, where in addition to the report of their sensory response, they had to report the confidence level on their responses' correctness. Thus, the progressing to the final level of the task was assessed by analyzing subjects' psychometrics, RTs and confidence reports. As a consequence of these criteria and analysis, together with the fact different people take naturally different amounts of time to accommodate the task' contingencies, not all subjects performed the same number of training sessions (Appendix 2. Table 3).

To learn and to be able to perform a sensory task means, typically, to perform "well" above chance (50% accuracy). The results presented in Figure 9.1 reveal that subjects from both groups performed above chance, and the majority improved their performance during training, showing adaptation. However, one also observes that not all subjects finished the training with 70% accuracy (criteria to move to the 2nd training stage). Even so, these subjects moved to the final level of the experiment with the understanding that the sensory task was more challenging for these subjects, but that they understood the task as the following analysis will show.

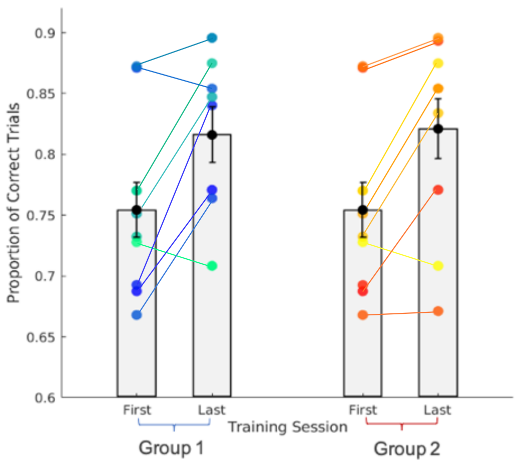

Figure 9.1: Results of accuracy (percentage of correct answers) on the first and last training session. On the left side we have the results for Group 1(cold colors) and on the right side for Group 2 (warm colors) of subjects. Each one of the colorful dots corresponds to the accuracy of one of the subjects in that specific session. The bar represents the mean value for accuracy across subjects and the errorbar corresponds to the standard error of the mean. Both groups of subjects increase their mean value of accuracy from the first training session to the last one.

The psychometric function is, by definition, an analytic function that relates the proportion of correct responses in a perceptual sensory task to some physical stimulus value (Treutwein B., Strasburger H., 1999). Therefore, we analyzed the proportion of correct "to the right" responses given as a function of the stimulus condition (pre-determined proportion of clicks presented to the right) for each subject. In Figure 9.2 A are presented the results of the logistic fits to the means across subjects of this analysis (psychometric) for the first and last training sessions. Generally, subjects of both groups presented an increase in the psychometric slopes and asymptotes (Figure 9.2 B and C), which translates an increase in discriminability for a given stimulus strength. This led us to conclude that subjects showed perceptual learning (Gold et al., 2010).

We defined the stimulus and the task to be performed under a RT paradigm. Therefore, we expected subjects to stop sampling the auditory stimuli and to indicate their decision as soon as they felt they knew the answer. To confirm subjects were learning and respecting this task' contingency, we analyzed

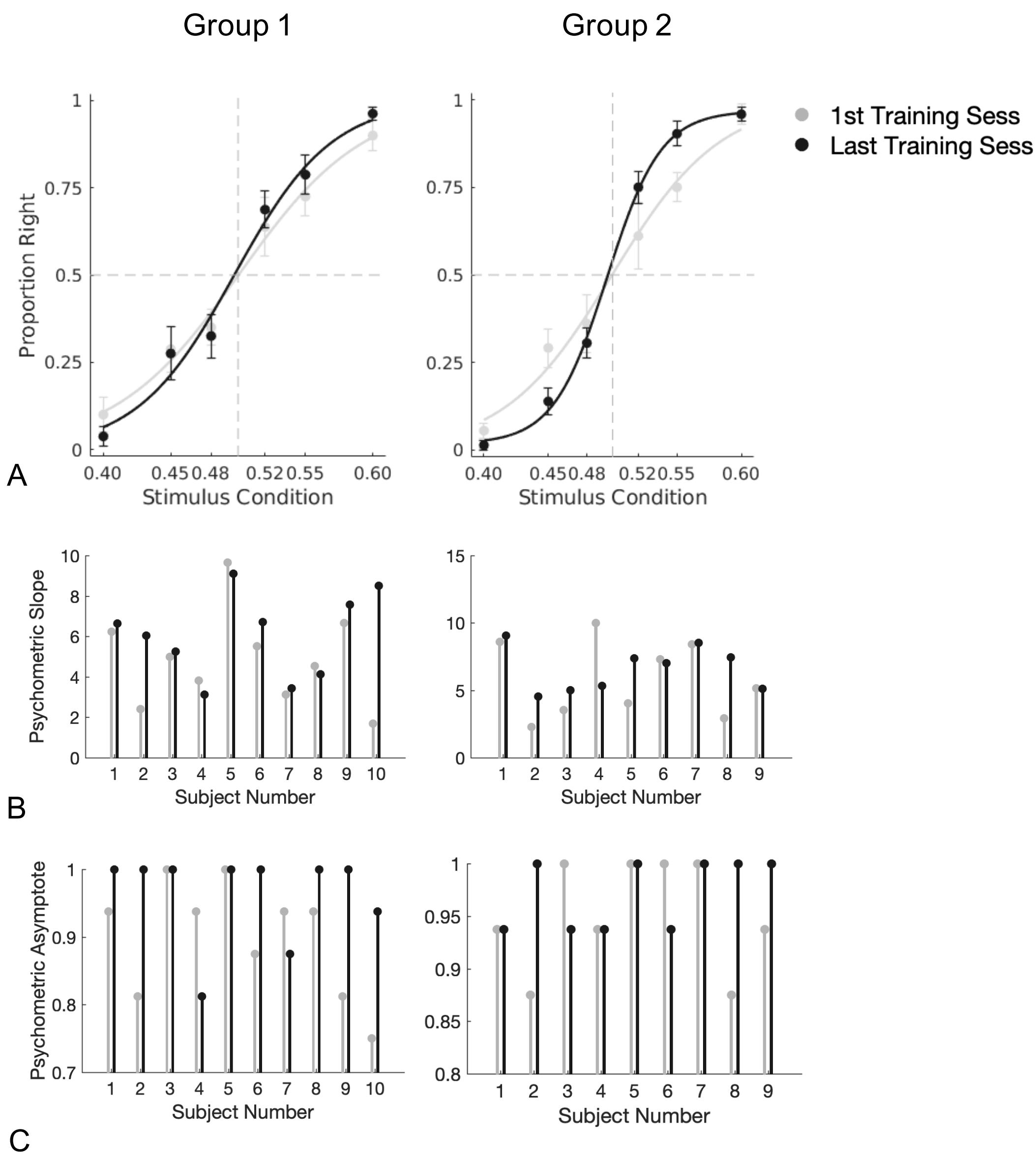

Figure 9.2: On the first row **(A)** are represent the fittings of psychometric functions for the mean values of both groups of subjects. In light gray we have the fit of the psychometric function of the mean across subjects in the first training session. In black we have the fit function on the last training session. On the second row are presented the results of the psychometric slopes in the first and last training session of all subjects from both groups. We observe that in general there is an increase of the slope value which represents perceptual learning. In the last row of plots of this figure one can find the values of the psychometric asymptotes in the first and last training session for all subjects of both groups. There is a tendency of this value to increase from the first to the last training session for almost all subjects, showing once again evidence that there was perceptual learning.

their RTs to verify they were not waiting for the automatic end of the stimulus presentation to give their response. The results in Figure 9.3, confirmed on average, subjects took less than 10s (maximum duration of stimulus presentation) to give their responses. Even though one can observe a slight increase in the average RT from the initial to the final training, we considered this increase as a natural adaptation to the stimuli and task, being a consequence of the subjects' increase of performance.

The last and crucial aspect of the training was to guarantee the presented stimuli evoked responses within the three levels of DC; and that subjects correctly understood the meaning of each one of these levels. As shown in Figure 9.4, subjects presented bigger values of mean DC for easier stimuli conditions

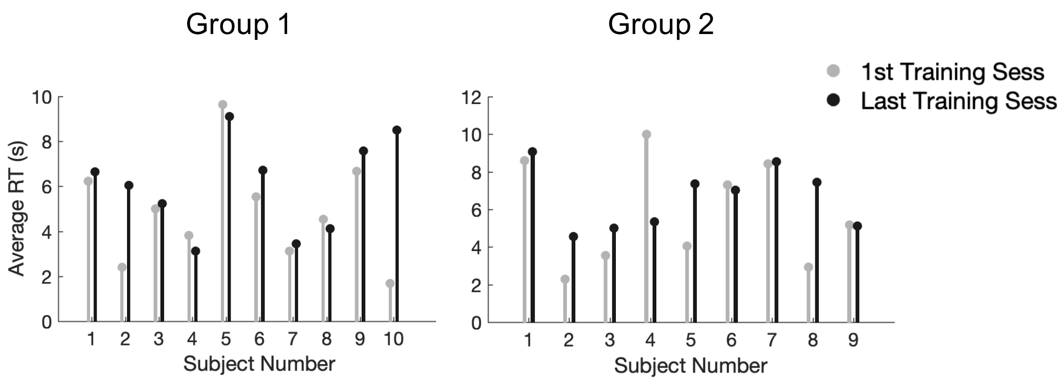

Figure 9.3: Average RT, in s, for the first and last training session of each subject of both groups of subjects.

and smaller DC values to the hardest conditions. Additionally, subjects pressed more often the level of confidence 3 when easy conditions were presented and more often confidence level 1 when hard conditions were presented (Figure 9.4). This behavior shows that the presented stimuli evoked uncertainty in subjects' responses and that subjects learned to classify their responses correctly with lower levels of confidence associated to the display of hardest stimuli conditions, and so on.

From this analysis, we understood that all subjects of both groups met the necessary conditions to move to final version of the task.

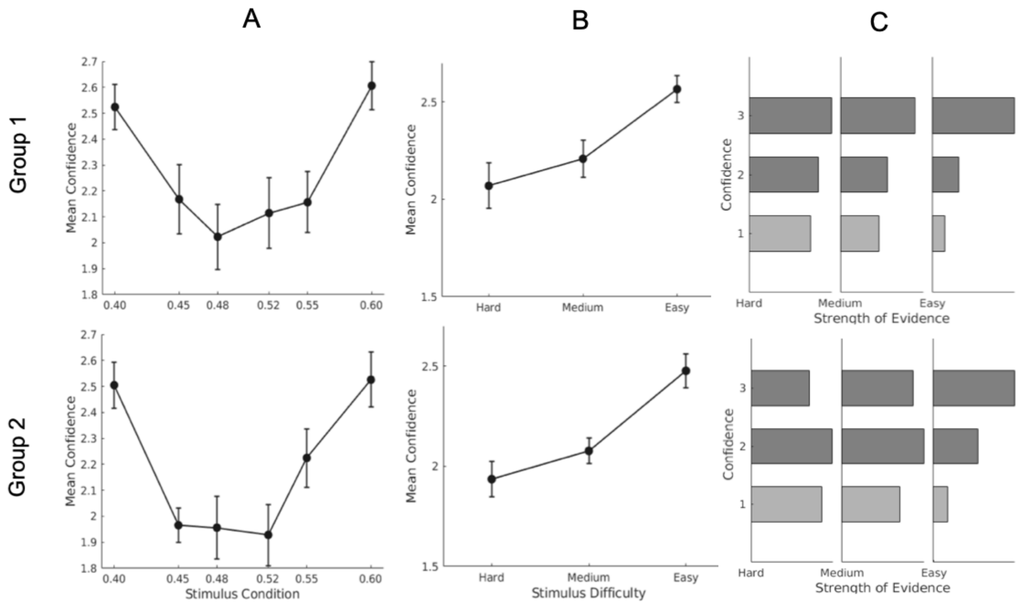

Figure 9.4: Mean values of confidence across subjects as function of stimulus condition **(A)** and as function of stimulus difficulty **(B)**. These results represent the mean of the mean values of reported DCs of each subject, computed by its as the average of the mean values for each training session performed with the confidence protocol. As not all subjects performed the same number of training sessions with the confidence protocol, with some subjects having exclusively completed one training session with confidence, we do not compare the evolution across training sessions. Also, the more relevant result is that subjects were on average well succeed in reporting higher values of confidence when easier stimuli conditions were presented, and smaller values of confidence to more difficult stimulus conditions. In plots **(C)** are presented the number of times each one of the confidence levels were presented by all subjects together in the training sessions. Recall that confidence 1 meant - "Not confident", 2 - "In between" and 3-"Confident". In the top row one finds the results for Group 1 of subjects and in the bottom row the results for Group 2.

9.1.2 Final version of the task

All subjects from both groups reached the final level of the task, and all subjects performed a total of 20 sessions of this final version, except one (group 2) due to schedule incompatibilities, who will be excluded from further analysis. This section will analyze the results for accuracy, RT and reported level of confidence to understand how the subject decides upon the two-alternative possible responses (left/right), how long it takes to decide, and how confident it is according to the sensory evidence he collects.

As previously mentioned, subjects were separated in two groups which performed the same 2AFC task with a protocol only differing in the type of reward system used. The first group followed an accumulating point system and the second group followed a discounting monetary system. Besides this, all other contingencies of the task were the same – auditory stimuli, maximum time of sound presentation, confidence levels, hardware (numpad) and buttons in question. Therefore, we evaluated if there was any significant difference between the averages across the groups for the variables accuracy, RT, and DC to verify if they could be analyzed together for both groups.

Variable		Mean	Std.Deviation	Std.Error Mean	t-test
Accuracy	Group 1	0.8877	0.0213	0.0071	p = 0.1178
	Group 2	0.8574	0.0608	0.0192	
RT	Group 1	7.3872	1.1544	0.3848	p = 0.6038
	Group 2	7.3717	2.2258	0.7039	
Confidence	Group 1	2.0187	0.2104	0.0701	p = 0.6038
	Group 2	2.0646	0.2146	0.0678	

Table 9.1: Summary of the statistical results when testing for differences between groups of subjects for the behavioral measurements - accuracy, RT, DC. The performed test was the Mann-Whitney as all variables were tested for normality and tested negative.

The results summarized in Table 9.1 reveal that the mean values of these variables are not significantly different between groups, consequently its analysis will be performed together for both groups until the variable reward is incorporated into the analysis.

To assess subjects performance, similarly to what was performed for the first experiment of this thesis we calculated the values of sensitivity (d') and criterion (c) for each subject. The analysis of these results made us exclude 2 subjects (Group 1). One for presenting a really biased accuracy with a value of criterion (c) higher than 2 standard deviations of the mean of criterion across subjects. The other subject as a result of a poor performance in a lot of the sessions, showing a value of sensitivity (d') 2 standard deviations below of the mean value of sensitivity (d') across subjects. It was our understanding that the bias and the performance significantly inferior to the mean of subjects could compromise the study variables of this experiment, in particular DC, in a way we could not account for. Therefore, both subjects were removed from further analysis.

The other subjects' psychometric functions, relating the condition of the played auditory stimulus to the proportion of times the response "right side" was correctly selected, were plotted in Figure 9.6 in light grey. From the analysis of Figure 9.6 one observes that performance varied as function of strength of evidence, being almost 100% to the easiest conditions and around 75% to the hardest conditions for each side, consistently across subjects. These results show that subjects learned the task and confirm that the auditory stimuli used in the experiment led to different levels of sensory uncertainty, even after the extensive exposition to the same.

A fundamental aspect of the task was its reaction time paradigm. That is, subjects were free to select the time they needed to make a decision and how long they were exposed to the stimulus. In tasks

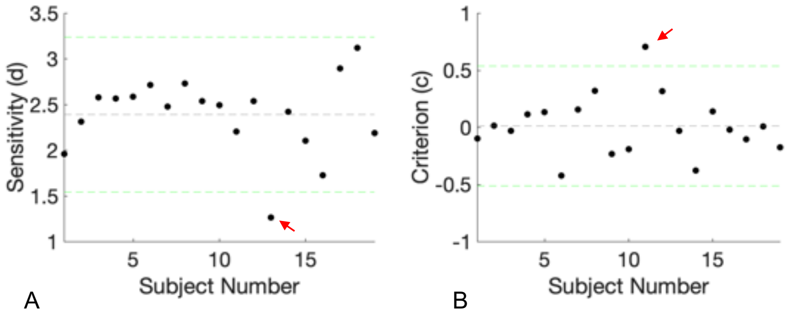

Figure 9.5: **(A)** Average values of d' (sensitivity) for each subject who performed the experience. The dashed gray line represents the mean value of d' across subjects, and the dashed lines in green the value of 2 standard deviations away from the mean value. There is only one subject which value of d' is located outside the green boundaries. This subject was excluded of further analysis. **(B)** Mean values of criterion (c) for each subject. The dashed gray line represents the mean values of c across subjects and the dashed green lines the values of 2 standard deviations across subjects away from this value. The subject who is located outside these boundaries was excluded for extreme biased performance.

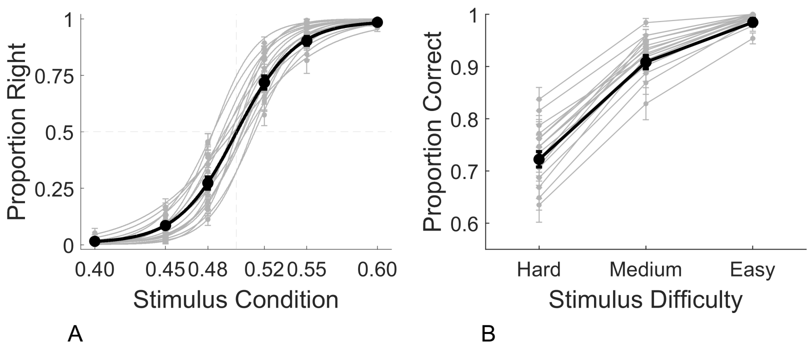

Figure 9.6: **(A)** Logisitc fits of psychometric functions for the 17 subjects (light grey) and the mean across subjects (black), showing that choice varies as function of stimulus condition. **(B)** Percentage of correct responses for easy (proportion of right clicks being equal to p = 0.40 or p = 0.60), medium p = [0.45, 0.55] and hard conditions p=[0.48, 0.52], showing performance increases with stimulus strength of evidence. Error bars on both plots correspond to standard error.

were this reaction time paradigm is applied are observed difficulty effects: higher average RTs for more difficult conditions (Uchida et al., 2014, Roitman and Shadlen, 2002). This effect happens because more difficult conditions are associated with uncertainty about the stimulus or noise in the sensory system responsible by the creation of the percept. Therefore, to reach a decision subjects need to collect a bigger amount of sensory evidence.

In this particular task, because of the way the auditory stimuli - streaming of clicks – were generated the correctness of a response was even more dependent on the amount of time a subject was exposed to the display of the stimulus. That is, difficulty was not only present in the difference of number of clicks played to both ears, but also in the own sequence of clicks which was generated pseudo-randomly on each trial . Therefore, it would be expected to observe a strong effect of stimulus condition on RT and a strong accuracy trade-off (SAT), which corresponds to the subject decision of compromising time over accuracy or vice-versa.

As shown in Figure 9.7 B, RTs increased as the stimulus difficulty increases, as expected. The effect is more pronounced for some subjects, but it was always significant (with the exception of one subject) (Appendix 2. Table 4) and at the group level (F(2,48) = 9.94, $p < 0.001$) for a confidence interval of 5%.

The speed accuracy trade-off (SAT) corresponds to the complex relationship between an individual's willingness to respond slowly and make relatively fewer errors compared to their willingness to respond

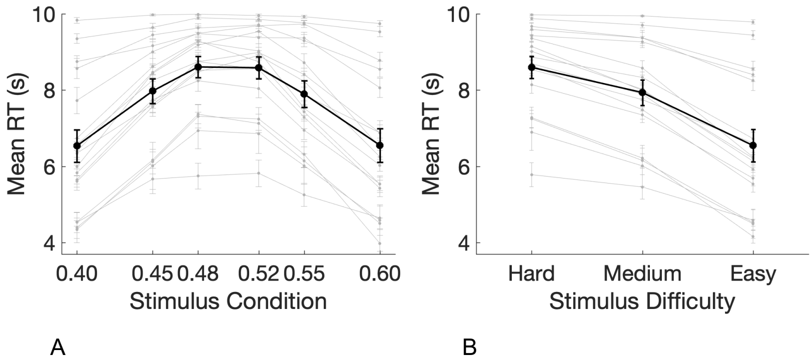

Figure 9.7: **(A)** Average reaction times for the 17 subjects (light grey) and the mean across subjects (black), as function of stimulus condition. **(B)** Mean reaction times for easy (proportion of right clicks being equal to $p = 0.40$ or $p = 0.60$), medium $p = [0.45, 0.55]$ and hard conditions $p = [0.48, 0.52]$ for every subject and mean across subjects (black). Error bars on both plots correspond to standard error.

quickly and make relatively more errors (Zimmerman, 2011). For some tasks this trade-off is stronger and the chronometric curves show a constant improvement as the time the animals are subject to the stimulus increases (Brunton et al. 2013). However for other tasks, the dependence of accuracy on sampling time is weaker, resulting in a saturation of the performance (Uchida and Mainen, 2003).

For this particular work, when plotting the average across-subjects chronometric curves (Figure 9.8), one observes that subjects seem to benefit from listening to the auditory stimulus for longer. Even when the easiest and medium condition's performance do not seem to change much from 4 and 6 seconds onward, respectively, for the hardest condition, subjects seem to improve their accuracy with the increase of the RT. Although these chronometric curves do not show such a strong dependence on the stimulus sampling time as one would predict, specially for the easiest conditions, subjects seem to prioritize accuracy over RT. The evidence in favor of this affirmation lies in the following two observations:

(1) Some subjects, even for the easiest conditions, wait, on average more than 9 seconds to make a decision - Figure 9.7.

(2) Even subjects who, in comparison to the mentioned in (1), presented on average smaller RTs, commonly presented a significant portion of trials with a RT up to the maximum duration of stimulus presentation - histograms of the Figure 9.9.

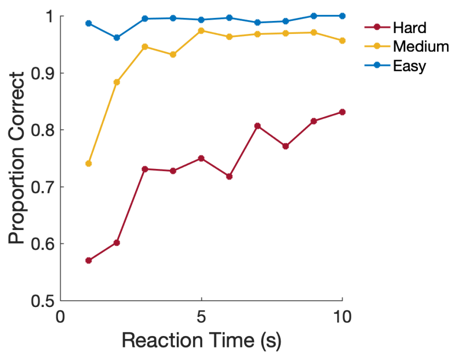

Figure 9.8: Average of chronometric curves of the 17 subjects by difficulty. For each difficulty, performance improves as a function of RT, except for the easiest condition where the accuracy is always high, even for small RTs.

Importantly, these results evidentiate that the designed 2AFC task successfully evoked different levels of difficulty in making a decision, with stronger evidence (easier conditions) leading to improved accuracy and faster RTs, as expected from previous works (Palmer et al., 2005; Roitman and Shadlen,

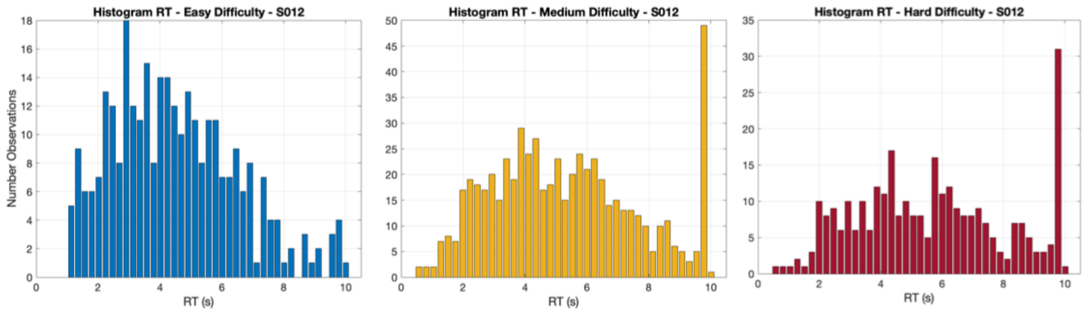

Figure 9.9: RT Histograms for different stimulus difficulty of an example subject, Subject 12. In blue it is plotted the histogram for stimulus of easy difficulty, and in yellow and in red, medium and hard difficulties respectively. When easier the condition of the stimulus presented a shift to the right of the pick of the histogram is observed. However, for both hard and medium stimulus difficulties, a big portion of the trials was performed with the maximum stimulus display time, that is, 10s.

2002). However, one still needs to understand if these levels of uncertainty were correctly translated into different levels of confidence, and what it is the dependence of these self-reported levels of confidence on each one of these factors - strength of evidence and RT.

To assess these self-reported levels of confidence (explicit values of DC), subjects had to report their choice and confidence about the localization (left/right) of a higher probability of playing a sound click simultaneously. In experiments performed with a similar protocol (Kiani et al., 2014), DC was shown to be directly correlated with strength of evidence and to be inversely correlated with RTs, suggesting that elapsed decision time informs certainty because it serves as proxy for task difficulty (Kiani et al., 2014). However, before moving to the analysis of the relationship between these variables in our experiment, we have to take into account that some of the collected measurements of RT, might not be trustworthy, which can influence the following analysis.

As mentioned before, there were trials where subjects waited until the end of the stimulus presentation to indicate their decision. In those trials, one cannot assume the value of 10s as being the real value of RT, since the decision was most likely not triggered internally but imposed by the time contingency. Subjects would probably have benefited from more time to decide. For subjects where these trials were a majority ($RT > 9s$ in over 60% of the trials), this might have compromised other dimensions of behavior such as DC, which has been shown to be conditioned by the time one takes to make a decision. Therefore, the analysis of DC was for these subjects (total of 5 subjects, 4 subjects of group 1 and 1 subject of group 2) carried out separately.

In Figure 9.10, we plotted the relationship between DC, stimulus condition, and stimulus difficulty, as well as the relationship between DC and RT for the not discarded subjects. The results show similar relationships between DC, strength of evidence, and RT as the ones found by Kiani et al. 's work. All subjects reported higher confidence levels when the stimulus condition was easy, and smaller values of DC when the condition was hard.

The Figure 9.10 C shows that for RTs bigger than 3 seconds, DC decreases with the increase of RT. Contrarily, we observe that for very small RTs, the mean of the confidence values reported are smaller, smaller the RTs. This is likely because for very small RTs the amount of evidence collected is very little, therefore, the uncertainty about the stimuli is naturally still very big. However, a very neat relationship between RT and DC is exposed in Figure 9.10 D, with bigger RTs associated with reports of uncertainty about the decision (Confidence=1) and smaller values of RT associated with reports of being confident about the decision (Confidence=3).

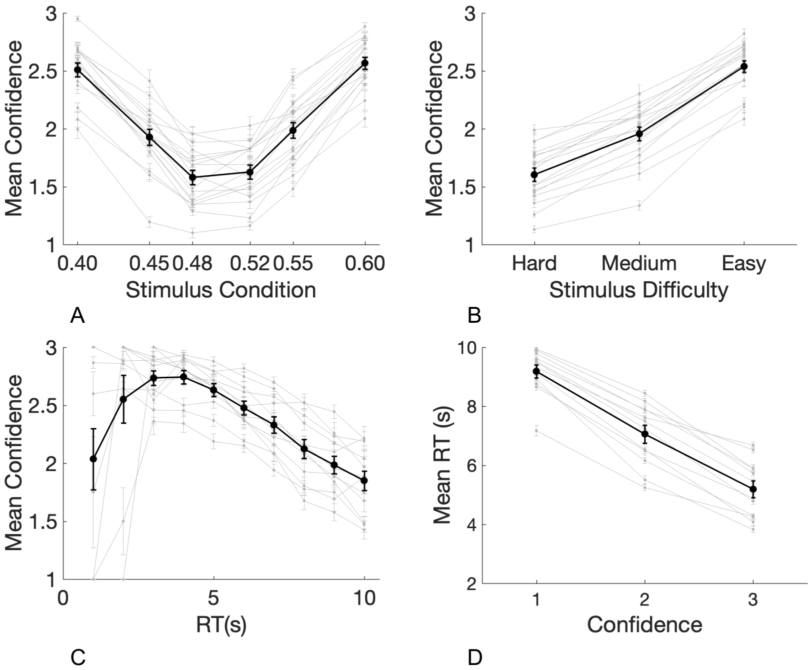

Figure 9.10: **(A)** Mean of the reported levels of confidence as function of stimulus condition. It is observed that for both sides confidence increases with strength of evidence **(B)** Mean of the reported levels of confidence as function of stimulus difficulty. **(C)** Mean values of the reported levels of confidence as function of the RT, evaluated for chunks of a second. **(D)** Mean values of RT by reported level of confidence, where confidence number 1 meant - " Not confident", number 2 - " In -between" and number 3 - " Confident". In light grey we have the means across sessions for each subject and in black the means across subjects. The errorbars represent standard errors of the means.

To test the real nature of the relationship between DC, strength of evidence and RT the following multiple regression analysis was performed:

$$DC = \beta_0 + \beta_1 S + \beta_2 T \tag{9.1}$$

Where DC corresponds to Decision Confidence, S to strength of evidence (hard, medium, easy), T is reaction time, and β_i are the regressor coefficients. For all subjects the model was significant (Appendix 2. Table 5), and the null hypothesis of lack of relationship between DC and RT ($H_0 : \beta_2 = 0$) and lack of relationship between DC and strength of evidence ($H_0 : \beta_1 = 0$), were rejected for all subjects at a level of confidence of 5% ($p < 0.001$ for all subjects). The results confirm the existence of a direct relationship between DC and strength of evidence and of an inverse relationship between DC and RT as it is observable in Figure 9.10 D to all 12 subjects (total number of subjects minus the subjects for now discarded of this analysis) in analysis. More importantly, it clarified that neither the effect of RTs nor the effects of strength of evidence could be described by the other one. That is, for a fixed RT, trials with lower stimulus strength had lower DC reports, and for a fixed stimulus strength, trials with longer RTs had lower DC reports as illustrated in Figure 9.11.

According to these results, for the subjects who in most of the trials waited until the end of stimulus presentation ($RT > 9s$), their confidence reports would presumably be less impacted by the RT values. As, in most trials, these subjects waited a long time to indicate their responses being their decisions imposed by the end of the stimulus and not by an internal mechanism. Thus, one would predict these RTs to not be very informative. To test this hypothesis and verify that independently of their reaction time distributions (RTD), confidence reports were still trustworthy to pursue with other analyzes, we

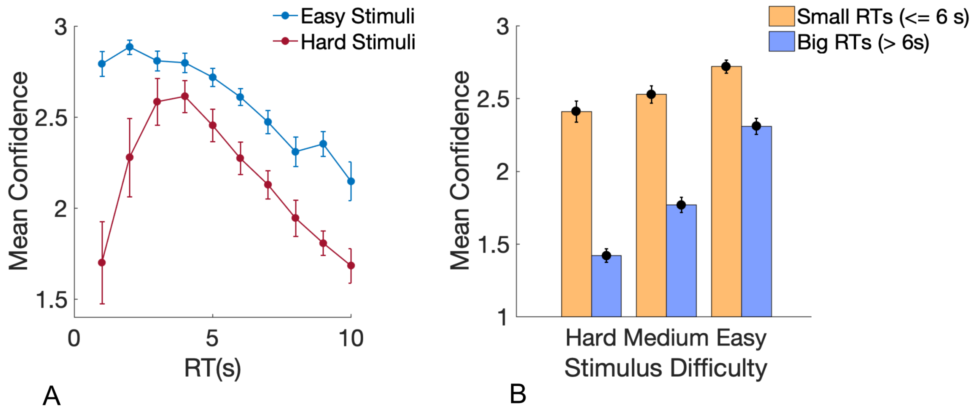

Figure 9.11: **(A)** Mean of the self-reported levels of confidence as function of RT (chunks of 1 s) for the easiest and hardest stimulus condition. It is observed that the mean values of the reported levels of confidence for the same RT value is always higher for easy stimulus than for hard stimulus. **(B)** Mean values of the self-reported levels of confidence for the 3 stimulus difficulties divided in RTs smaller than 6s and RTs bigger than 6s. The results of both plots were computed using the averages across subjects. Errorbars represent standard error of the mean

evaluated how their self-reports of confidence varied according to strength of evidence and RT.

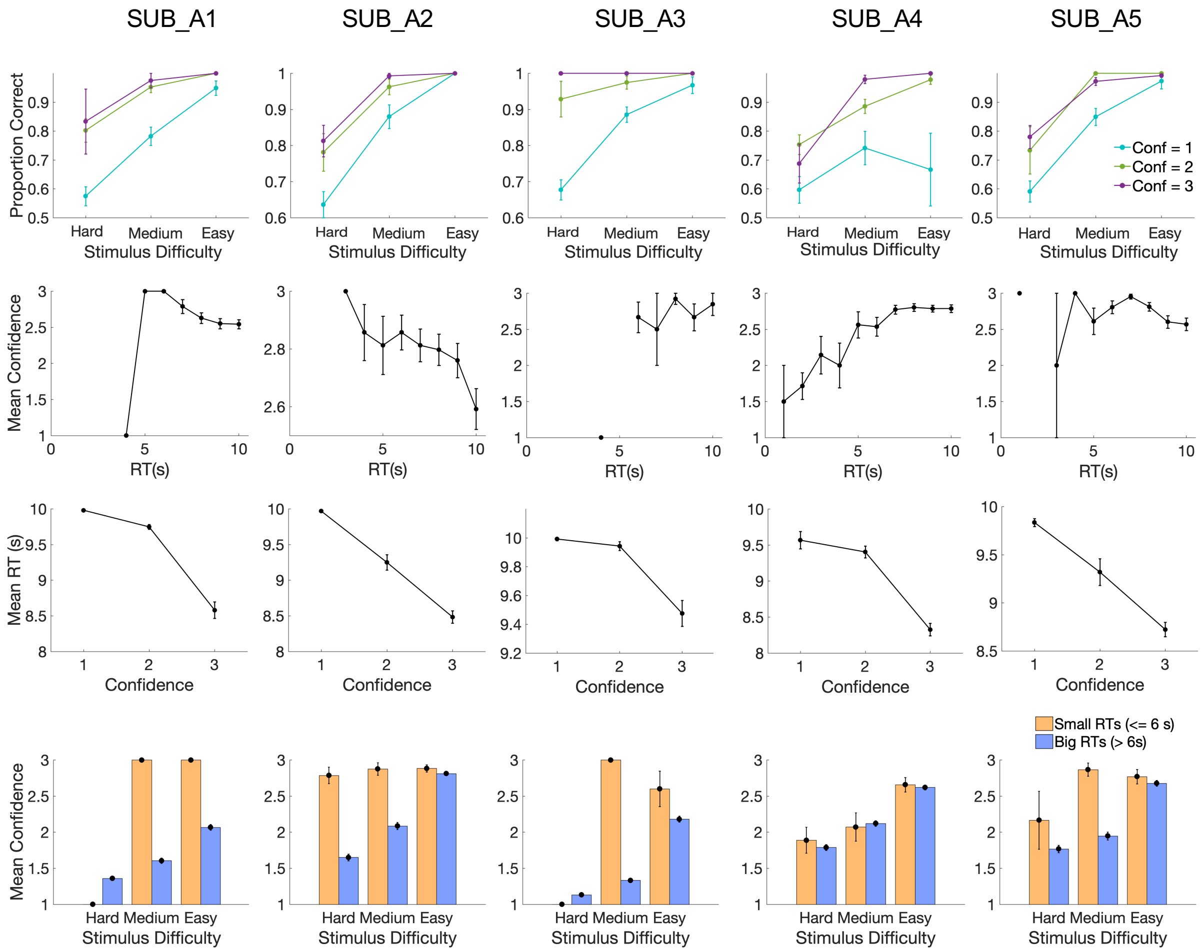

Figure 9.12: DC confidence results for the subjects who presented $RT > 9s$ in more then 60% of the trials. Figure first row of plots, shows accuracy marginalized over stimulus difficulty for the 3 different levels of DC. The second row presents the mean confidence as function of RT for the 5 subjects. The third row contains plots of the average RT, in s, as a function of the confidence level. In the final row we find the mean values of the reported values of DC as function of stimulus difficulty for two classes of RT: $RT <= 6s$ or $RT > 6s$.

Contrarily to our prediction, as shown in the second row of Figure 9.12 the inverse relationship between DC and RT was still observed for these subjects (Appendix 2. Table 6) except subject A4. Even though subject A4 also showed smaller average RTs associated with bigger values of DC, this relationship seemed to be, for this subject in particular, exclusively based on the fact that higher levels of DC are associated with easier stimuli' conditions. It is noteworthy that the found relationship between DC and RT were weaker than those of the other group of subjects, except subject A2 (Appendix 2. Table 6). These results suggest that these subjects (except A4) also used information of the elapsed time necessary to formalize their self-reports of confidence. Subjects who presented longer RTs on average, reported their values of DC faithfully - being higher values of DC associated with higher accuracies (first row 9.12).

Studies of perception are, traditionally, based on three behavioral measurements: accuracy, RT, and confidence ratings. These studies have identified that all three measures are primarily affected by stimulus strength/difficulty. Although accounting for the exact qualitative relationships is nontrivial, it seems natural that a low quality evidence would be associated with worse accuracy, slower response times, and lower confidence ratings. Indeed, if confidence reports are all meaningful, it ought to reflect accuracy, on average, even if imperfectly (Drugowitsch et al., 2014). This trend is apparent in our experiment, reassuring us that our subjects' reports were sensible. Additionally, our results identify a clear relationship between DC and the time it takes to make a decision, showing that DC is informed both by strength of evidence and RT, as suggested by Kiani et al.

In sum, our sensory 2AFC task successfully evoked uncertainty among subjects' responses, and the collected measures of DC were reliable and sensitive to the sensory evidence and amount of time to decide.

9.2 Movement Speed and Decision Confidence

One of the main aims of this project was to expose movement speed as an implicit measure of decision confidence. As shown in the previous section the designed task was successful in evoking different levels of DC. This section will be dedicated to demonstrate that movement speed follows a monotonic relationship with DC.

Recent work performed to address the human sense of confidence (Sanders et al., 2016) revealed a very particular pattern relating confidence to choice and strength of evidence: an increase of confidence with strength of evidence for correct choices and a decreasing confidence with increasing evidence strength for erroneous choices. In Renart Lab, when animals perform ILD discrimination experiments it is observed that movement time (MT) - defined as the time between exiting the central port where stimuli are delivered and entering the lateral ports where rewards are given -the exact same pattern was found. That is, the following three properties were found:

- "X pattern": MT decreases as a function of ILD for correct trials but increases as a function of ILD for incorrect trials;

- MT decreases with accuracy when both measures are marginalized over difficulty;

- For a given difficulty, accuracy increases the smaller MT.

These properties lead us to believe movement speed ($MS = 1/MT$) behaves monotonically with DC under Kepecs normative theory and others.

To test this prediction, in this work we measured the time subjects took from the moment where they have made their decision until the first press when they indicate their choice, the movement time (MT).

$$MT = t_{first\ key\ press} - t_{stop\ sampling\ auditory\ evidence} \quad (s) \tag{9.2}$$

Where $t_{stop\ sampling\ auditory\ evidence}$ corresponds to the moment where the subject releases the green key marking that has decided, Figure 8.1, and $t_{first\ key\ press}$ the moment where the subject presses the selected response key indicating, simultaneously, his decision and level of confidence. From the measure of the MT, we were able to calculate the MS on each trial:

$$MS = \frac{x_{response\ key} - x_{green\ button}}{MT} \tag{9.3}$$

As in our experiment, the selected key represented not only the subject choice but also the level of confidence, this allowed us to calculate the MS associated with the report of each one of the three levels of confidence.

Figure 9.13 A shows how DC varied with stimuli difficulty. In the previous section, we showed that our measures of DC were sensible to strength of evidence. These results evidence that similarly to the Kepecs normative theory, DC increased with strength of evidence for correct trials and decreased with strength of evidence for incorrect trials for most subjects (Appendix 2. Figure 2 and Figure 3). However, not all subjects presented the expected "X pattern", so we did not observe a perfect "X pattern" across subjects, with easy conditions for incorrect trials associated with a bigger average DC report, than for harder conditions. Nevertheless, overall we observe a pronounced increase of average DC with strength of evidence in correct trials across subjects and a small decrease of the average DC from the hard to the medium conditions in incorrect trials.

In addition, when analyzing the accuracy marginalized over difficulty for different values of DC (Figure 9.13 C), one observes that for the same difficulty accuracy grows with DC's value, meeting once again the results of Kepecs et al., and others (Peirce and Jastrow, 1884; Lichtenstein et al., 1981).

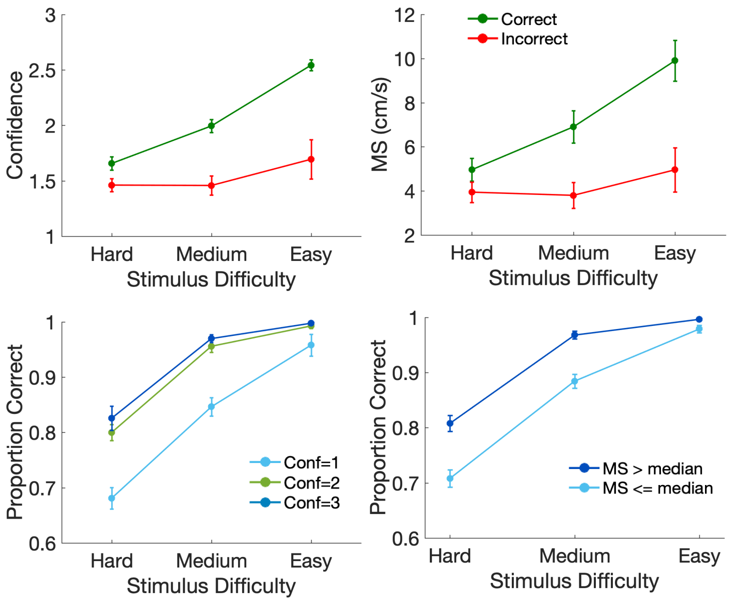

Figure 9.13: In the first row its plotted the mean values across subjects of the explicit values of confidence for correct (green line) and incorrect responses (red line) and movement speed (cm/s) on the right. On the bottom line we have accuracy as function of stimulus strengths for the 3 confidence levels (left) and as function of smaller or faster movement speeds on the right. The values were all computed from the means across subjects and the errorbars represent standard errors across subjects. The results of both plots were computed using the averages across subjects. Errorbars represent standard error of the mean across subjects.

As predicted when analyzing MS as function of strength of evidence for correct and incorrect responses the same pattern was found, that is, MS increased as function of strength of evidence for correct responses and decreased as function of strength of evidence for incorrect trials, for almost all subjects (Appendix 2. Figure 2 and Figure 3). Additionally, higher levels of accuracy were associated with faster movements' responses. This mirroring of the DC behavior by MS suggests that, indeed, MS behaves monotonically with DC. To test this prediction, we perform the following linear regression analysis:

$$MS = \beta_0 + \beta_1 * DC \tag{9.4}$$

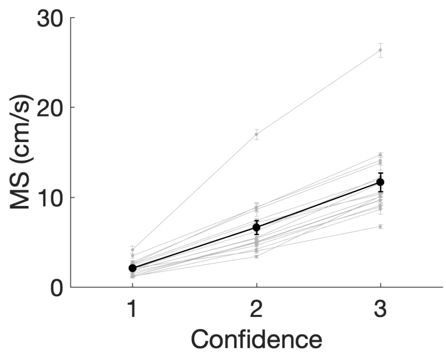

Figure 9.14: Calculated values of MS (cm/s) as function of the self-reported level of confidence. The lines in light grey represent the means across sessions for every one of the 17 subjects and in black is plotted the means across subjects. Errorbars show standard errors of the mean.

Where β_i stands for the regressor coefficients.

The results proved that the model was significant for all subjects (Appendix 2. Table 7 with a $p < 0.001$ for a significance level of 5%) and the null hypothesis ($H0 : \beta_1 = 0$) was rejected for all subjects (Appendix 2.Table 7 with a $p < 0.001$ for a significance level of 5%), confirming the existence of a monotonic relationship between the measures of MS and the self-reported levels of DC, Figure 9.14.

Observing again the scheme of the numpad used to collect subjects' responses (Figure 8.1 A), one sees that the buttons associated with different levels of DC are located at different distances of the central green button. Therefore, one cannot exclude the hypothesis that the observed results, showing bigger MSs associated with bigger DC' reports, are driven from natural different movement trajectories. That is, different MSs might not be the result of different uncertainty levels, but the result of different movement distances. To test this possibility, we performed a control experiment where subjects pressed the button indicated on the screen. With its results, one could compare the speeds of movements unlinked to any value of decision's confidence to the speeds of movements representative of a certain level of confidence. The experiment and its results are presented below.

Are different movement speeds caused by different movement distances?

In this control experiment, 4 subjects naive to the aim of the same, performed the simple task of pressing in the numpad the button correspondent to the one being displayed on the screen (Figure 9.15). On each trial, after a variable fixation period ($FT = [0.5 - 2.5]s$) where the subject had to press the central key on the numpad (the previously named green key), a scheme of the numpad would appear on the screen up to a maximum duration of 1.5s. As soon as the subject knew the key he had to press on, he should release the central key and press the correspondent key on the numpad, always using the index finger exclusively. Responses with a display of the scheme for less than 0.1s were considered aborts, as well as trials where subjects took more than 3s between releasing the green key and pressing the correct key on the numpad. After pressing the key on the numpad, feedback would appear on the screen in the form of a green check mark (correct button pressed) or a red cross (incorrect button pressed), mainly to avoid distractions or impulsive responses. All 4 subjects performed a total of 4 sessions, each one with 96 trials. Subjects were instructed to press each one of the 6 keys used in the main experiment the same number of times per session, which corresponded to a total of 16 times per key on every session. None of these 4 subjects participated in the main experiment, since we considered that any previous assigning value to any key might have influenced the results.

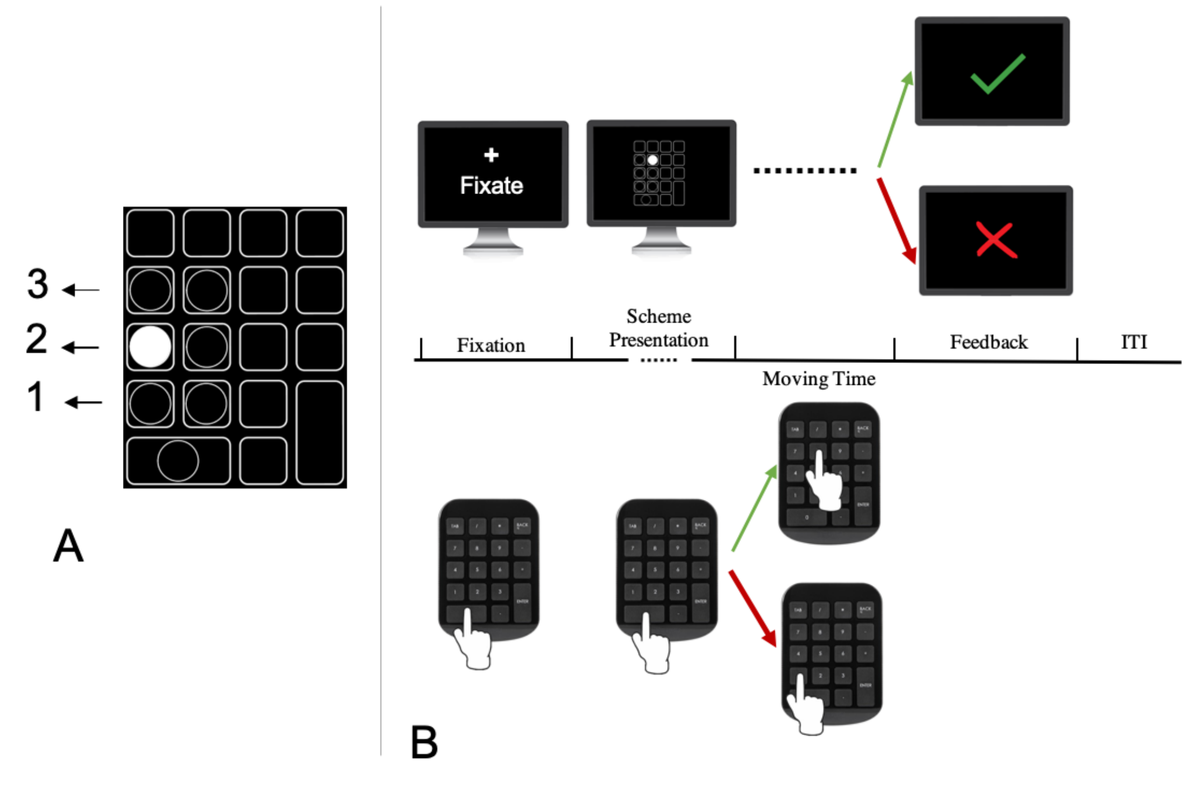

Figure 9.15: Task Structure - Control Experiment: **(A)** Example of the numpad scheme displayed on the screen in every trial. The keys with circles are the ones used in the auditory 2AFC experiment and therefore the ones that are being used in this control experiment. When the circle is filled subject should press the correspondent key on the numpad. In this pictured case, subject must press the key with the number 4 of the numpad. The numbers 1, 2, 3 next to the arrows represent the distance level to the central key where subject is pressing before moving towards any of the keys. Recall that level 1 was previously assigned to represent confidence 1 ('Not confidente'), level 2 the self-report of confidence 2 ('in between') and level 3 confidence 3 ('Confident'). The results will be presented as function of these 3 levels, independently if it was on the right or left side, as the distance to the central one (on the numpad number 0) is symmetric. **(B)** Schematic depiction of the different trial events in this control experiment. Subject starts by pressing the central button (0' key) for a variable fixation period ($FT = [0.5 - 2 - 5]$) and the numpad scheme is after displayed while the subject continues to press the central key. After the subject has acknowledge the button he must press, he should release the central key and move to the correspondent key on the numpad. After pressing feedback is given followed by an inter-trial interval (ITI) of $1.2s$.

In this way, it was possible to measure the MS associated with the pressing of the same buttons as the ones used in the auditory 2AFC task with a discouting protocol, and evaluate if different movement distances (the press of different buttons) could explain the monotonic relationship observed between MS and DC. Naturally, this control experiment should have been performed with the same original 17 subjects, before any value had been assigned to the keys in question. However, we considered that even without being possible to mimic exactly the natural movements of those subjects towards the different keys, this experiment to be sufficient to disentangle the role that different movement distances play in the speed of the executed movement.

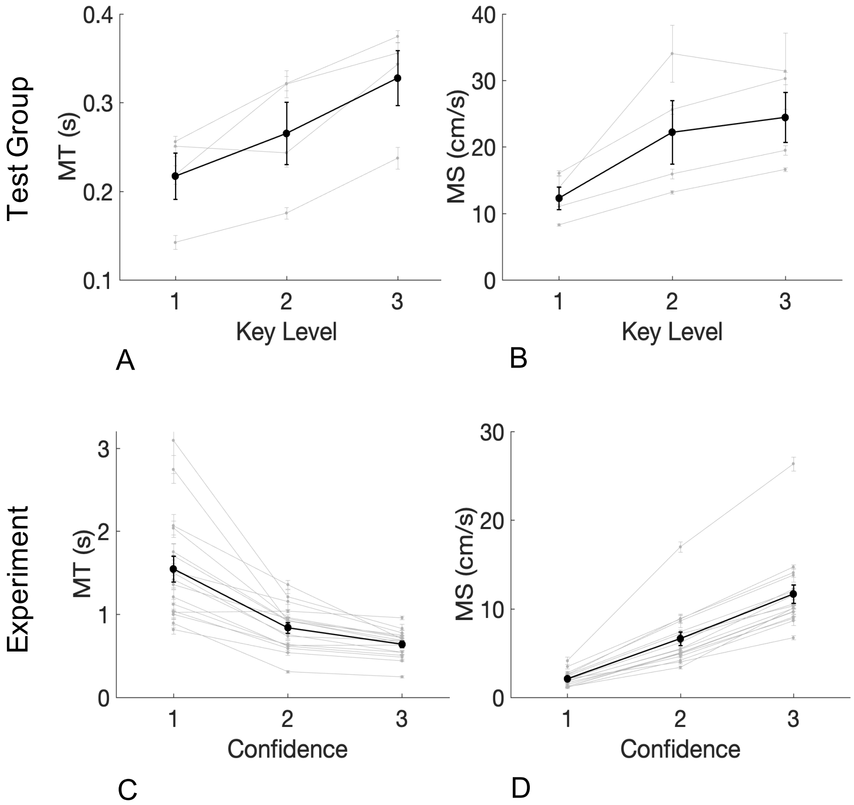

Figure 9.16: **(A)** Values of MT in seconds as function of the distance level to the central button. Level 1 - keys '1' and '2' of the numpad (closer keys); Level 2- keys '4' and '5' of the numpad (middle distance to the central key); Level 3 - keys '7' and '8' of the numpad (furthest keys) . All 4 subjects take on average a longer amount of time to reach the furthest keys on level 3 than the closer ones. **(B)** MS as function of the key level. Subjects even taking a longer time to reach the keys on level 3 of the numpad, they move towards them at a faster rate when comparing to the speed of the movements to reach keys on level 1 and 2. **(C)** Values of MT as function of confidence level in the auditory 2AFC task. Subjects contrarily to what was shown to the control experiment take less time to reach the keys on level 3 that represented reports of being confident about the decision (confidence 3) than to reach the keys closer to the green button representing the self-reports of 'Not confident' (confidence =1) and 'In between' (confidence = 2). **(D)** MS as function of the confidence level in the auditory 2AFC task. In all plots the light grey lines represent the mean across sessions for each one of the 4 control group subjects (top row) and the 17 subjects of the auditory task (bottom row). The errorbars represent standard errors of the mean.

As shown in Figure 9.16 all 4 control subjects presented faster movements when pressing more distant buttons than closer buttons (Appendix 2. Table 10). However, subjects also took on average longer to press the furthest buttons (level 3) than the closest ones (level 1 and 2), as represented in Figure 9.16. The nature of this relationship was significant for 3 of the 4 control subjects (Appendix 2. Table 10), even though it was not significant at the group level (F(2,9)=3.24, p= 0.087, one-way ANOVA) for a confidence interval of 5%. Remarkably, the same does not happen when confidence values are assigned to the different buttons (Figure 9.16 D). That is, contrarily to what is observed regarding MT on this control experiment, in the auditory 2AFC task was observed that subjects significantly took less time to reach the buttons in level 3, when they are confident, than to reach the buttons on level 1 if they are not confident about their decision, with this relationship being significant for all subjects, with the exception of one subject (Appendix 2. Table 8) and significant at the group level (F(2,48)=26.49, $p < 0.001$, one-way ANOVA). These results suggests that even though distance plays a role in the movement velocity, the observed differences in MS regarding different key levels of the numpad are steeper when values of

DC are assigned to the keys.

In order to test this result, we compared the slope of the linear fit to the values of MS in the control experiment with the linear fit to the values of MS to in the auditory 2AFC task. As expected, the slope of the fit is smaller in the test group (0.600) in comparison to the fit to the data of the 2AFC task (1.082)(Figure 9.16 C and D). We tested the significance of the difference between the slopes with a permutation test (Appendix 2. Figure 1), where we shuffled the z-scored MS as a function of confidence/key level values for all subjects together (main experiment plus control subjects) and calculated the values of the slopes dividing subjects in two groups with the same number of subjects (one with 4, and other with 17), repeating this process until a distribution of the difference was obtained. The results showed that the original difference between slopes of the fits to the MS data was significant ($p < 0.001$). This result implies that the monotonic relationship between MS and DC cannot be accounted just by the position of the keys, and there is indeed a large component intrinsic to the DC itself. However, it is noteworthy that the control experiment was only performed by 4 different subjects and the amount of collected data was relevantly smaller. Nevertheless, the striking difference between the pattern of MT in the control experiment and in the auditory 2AFC task, as the results of the linear fit showing significant differences of the slope values, exclude the hypothesis of the level of certainty about a decision not to be reflected in the response MS (Figure 9.13 D).

Does decision confidence affect other measurements of vigor?

In this experiment, we introduced different ways of evaluating the vigor in reporting a decision. The first, above studied, was MS, which consisted of measuring the speed with which a subject moved his index finger to the selected response key. The second consisted in measuring how fast the subject pressed the response button until the response was blocked - response press rate. By introducing different response vigor measures, one had better chances of understanding how vigor can be affected by other variables and evaluate these effects' extension. Therefore, after exposing such a clear relationship between DC and MS, we will now evaluate if the response press rate also varies according to the self-reported level of confidence.

In our experiment, to obtain a measure of the response press rate, the subject had to press an unpredictable number of times in the selected response key until the bar displayed on the screen was full, locking the response. Each key press had a chance of filling a rectangle of the bar by 20%. As to lock one's response was necessary to fill 3 of these rectangles, the number of presses necessary to mark the response was unpredictable and independent from trial to trial.

We started by analyzing the average press rate on each trial as function of the self-reported confidence level. The results, showed in Figure 9.17, revealed a significant relationship between the press rate and DC for some of the subjects (in green, Figure 9.17). When significant, most subjects decreased the time between presses, that is, press faster, higher the confidence level, except for two subjects that significantly decrease their response rate when they are more confident about their response at the confidence interval of 5% (Appendix 2. Table 11). 6 out of the 17 subjects present a non-significant difference in their average press rates for different confidence levels. Although the increase in the response press rate with the increase of confidence level was significant at the group level (F(2,48) = 3.25, $p = 0.047$, one-way ANOVA, confidence interval of 5%), we observed a less pronounce relationship between DC and the press rate than between DC and MS.

We also measured the time between all the presses necessary to fill the bar and block the response. In this way, we calculated how the press rate varied throughout each response (Figure 9.18 A). We looked

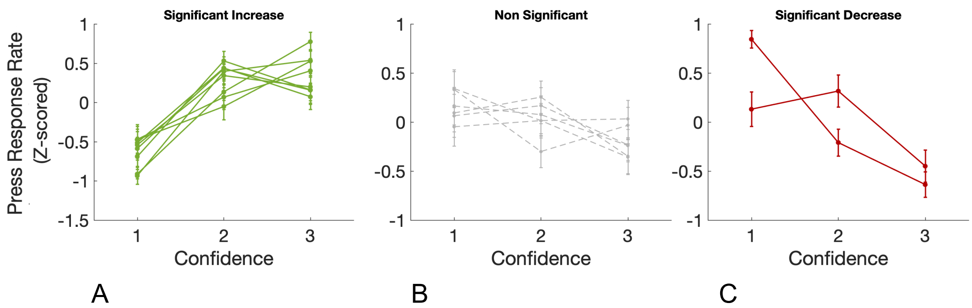

Figure 9.17: Z-scored mean values of response press rate across sessions for each subject of both groups in (A) showing a significant increase as function of DC, (B) no significant effect and (C) a significant decrease as function of DC.

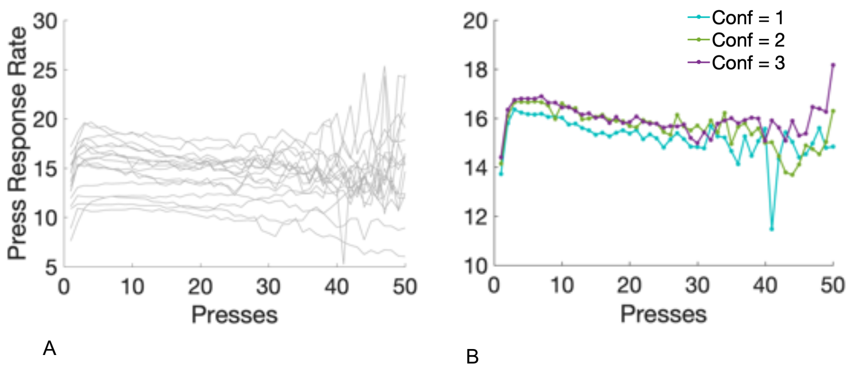

Figure 9.18: **(A)** Evolution of each subject response rate (1/time between presses) with the number of presses in the response key. Each line represents a subject, being the values plotted an average of the rate profile on each trial. **(B)** Average response rate across subjects as a function of the number of presses for the three different confidence levels (Conf 1- 'Uncertain', Conf 2 - 'In between', Conf 3- 'Confident').

into each subject press rate profile to see if any potential difference was present according to the reported level of confidence. Overall, there was a tendency to press the response button faster until the response was locked when the reported confidence level was higher (Figure 9.18 B). If the relationship was not present for all subjects, it was very notorious for the majority (Appendix 2. Figure 4 and Figure 5).

To conclude this point, similarly to what we observed for MS, there was an overall tendency of pressing faster the response key when subjects were more confident about their responses than when they were unsure. However, it is important to highlight that the relationship was not as strong as the one found between MS and DC.

9.3 Reward and Vigor

As demonstrated, response movement vigor, that is, the speed with which one performs the action marking a decision, undergoes modulations. In the previous section, we have shown that these vigor modulations are strongly linked with the level of certainty about the decision being made. However, one still needs to understand why being very sure about a decision leads to faster response movements, as faster movements involve a higher energy expenditure. It is necessary to comprehend the different factors that might play a role in selecting movement vigor, to better understand subjects' policies under perceptual decision-making paradigms.

Previous works approaching decision-making vigor suggest cost of time as the primary definer of MS (Niv et al., 2007). As reviewed, reinforcement learning works (Niv et al., 2007, Constantino and

Daw, 2015; Otto and Daw, 2019) suggest cost of time to be determined by a running average of the reward received in the near past - the average reward rate (Guitar-Masip et al., 2011; Contantino and Daw, 2015). According to this theory, higher the average of rewards received in the past, the higher the average reward rate, computing a higher 'opportunity' cost of time (cost of being sloth by saving energy), leading the subject to move faster. However, it is not clear, according to this view, how DC influences vigor, even though, our results clearly show vigor to depend on DC.

To include DC in this cost of time computation of vigor, we propose a view in which DC sets the local future expected reward of an action:

$$< Rew >= Rew \times DC \tag{9.5}$$

That is, if the subject just made a confident decision, that means by this definition that he will expect a higher reward when he gets to the reward site, compared to a trial with low confidence. Thus, on a time of the order of the action's duration, the expected reward will grow with confidence. If we are willing to believe in a theory, as proposed by others (Milstein and Dorris, 2007; Wilson et al., 2009; Opris et al., 2011; Yamamoto and Hikosaka, 2013; Reppert et al., 2015; Sackaloo et al., 2015; Summerside et al., 2018; Revol et al., 2019; Tachibana and Hikosaka,,2012; Hamid et al., 2015; Walton et al., 2018), where cost of time is influenced by the reward expected to be received; we would, by defining DC in this way, exposing how DC influences cost of time and, consequently, the vigor of an action.

In order to substantiate this hypothesis, we would have to show that the vigor of an action is affected by the expected reward. Thus, in our experiment we introduced a reward variable and independent from trial to trial, in order to analyze whether different amounts of immediate available reward had an effect on vigor (decoupled from the DC effect). Furthermore, by creating this environment where the past does not allow to predict the future, we could directly analyze if the average reward rate, even in a constantly changing environment, affects the current action vigor. Hence, we would be directly testing which vision - ours (prospective computation of cost of time) or Niv et al. (retrospective view) - best describe cost of time computation in an unpredictable environment.

Additionally, we set 2 different variable reward systems. An accumulating monetary system (Group 1) and a discounting monetary reward system (Group 2). By doing this, we aimed to test how subjects behave if what is at 'stake' on each trial is not an increase in their balance but a decrease. In other words, we tried to understand if what influences vigor, that is, what motivates the subject to move with different speeds, is exclusively the possibility of winning a reward or the value /the weight of what is at stake in the trial. Consequently, in this section, dedicated to address the influence of reward on movement vigor, the results will be presented separately for both groups of subjects, which performed the experiment under different reward systems.

To analyze the effect of these different variables - immediate reward, average past reward, expected reward (Equation 9.5) - we performed a multiple regression analysis using the mentioned variables as predictors, and others, we considered that might have influenced vigor. The model used was the following:

$$MS = Diff + Rew + Conf + Rew * Conf + Av_{Rew} + RT \tag{9.6}$$

where $Diff$ stands for stimulus difficulty (Hard: $p = [0.48, 0.52]$, Medium: $p = [0.45, 0.55]$, Easy: $p = [0.40, 0.60]$), Rew for the reward at 'stake' on each trial, either the amount displayed represented a loss or a gain, $Conf$ for the self-reported value of confidence, $Rew * Conf$ amount of reward at 'stake' in the trial times confidence, Av_{Rew} for average reward rate and RT for reaction times. The model was fitted to the data of each subject individually and we defined average reward rate, in units of rew using the following update rule (Constantino & Daw, 2015):

$$Av_{Rew}(t) = (1-\alpha) \times Av_{Rew}(t-1) + \alpha \times Rew \tag{9.7}$$

α was calculated by fitting a linear regression for every subject and selecting the $alpha$ which offered the best R^2 to the model.

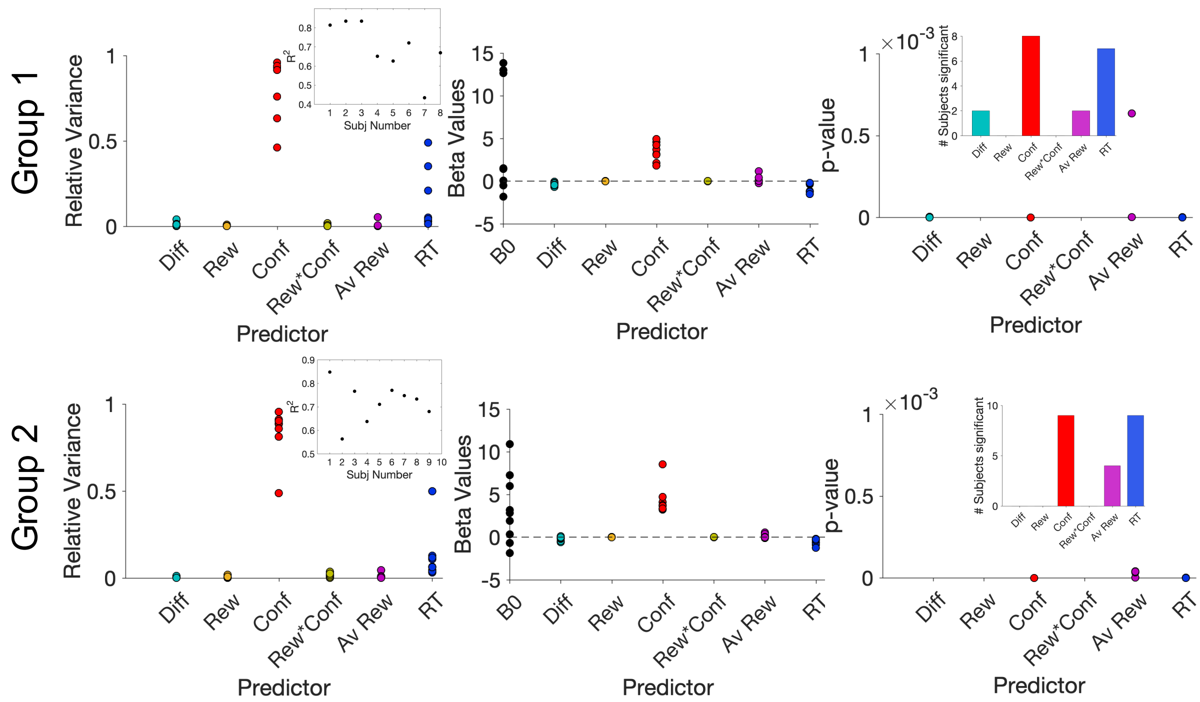

Figure 9.19: Results of multiple regression analysis for MS: **(A - Left Column)** Fraction of variance of the linear regression model explained by the stimulus difficulty (Diff), trial reward (Rew), confidence (Conf), trial reward times confidence (Rew*Conf), average reward rate felt in the past (Av Rew) and reaction time (RT). Each circle represents a subject of group 1 (top row) and of group 2 (bottom row). The inset shows R^2 of the model for each model. That is, the proportion of variance of MS explained by the model. For all subjects expect subject 7 (group 1) and 2 (group 2) the model explains over 60% of the variance explained in the data. **(B - Middle Column)** Predicted values of β for each one of the predictors. **(C - Right Column)** p-values of the different predictors, to schematize which predictors were actually significant for the model. The inset shows the number of times, that is the number of subjects, to who the predictors showed as significant were indeed significant. For example, only for a small fraction of subjects of Group 1, the predictor Av Rew was actually significant. Also on Group 2 only a small subjects models showed Av Rew to be a significant predictor.

As shown in Figure 9.19 and as expected from our previous analysis, DC has a close relationship with MS explaining a significant percentage of the variability in MS data for all subjects of both groups ($p < 0.001$ for all subjects Figure 9.19 plots on the right). However, the multiple regression analysis results showed expected reward (Equation 9.5) to be a non-significant predictor of MS at the confidence interval of 5%. This result contradicts our prediction where expected reward sets the cost associated to the duration of an action and, by consequence, should be a very significant predictor of MS.

Surprisingly, results (Figure 9.19) showed that both immediately available reward, Rew, and near past average reward, Av_{Rew} were not significant in predicting MS. That is, neither reward or average past reward explained a significant percentage of variability in MS data of each subject for both groups. Even with Av_{Rew} being, sometimes, a significant predictor able to explain very little MS variability, its associated values of beta were very low, revealing the weakness of the effect. These results contradict all theories previously mentioned (retrospective and prospective). To confirm these results we performed a simple t-test, for each subject individually, comparing the means of MS in trials where the 'stakes' (Group 1: points one can gain, Group 2: points one can lose) were smaller than the median with trials where the 'stakes' were bigger than the median. The results Figures (Appendix 2. 6, 7, 8, 9) did not make us reject the null hypothesis, showing that MS for both big and small reward categories were not significantly different at the confidence interval of 5%, for both groups of subjects (Group 1: t(14)=0.389, $p = 0.703$, Appendix 2. Table 12; Group 2: t(16)= -0.05, $p = 0.964$, Appendix 2. Table 13).

As the reward was randomly selected on each trial from a uniform distribution, with values ranging from 1 to 100 points (Group 1) and 1 to 100 cents (Group 2), we decided to test the possibility that only very 'striking' values of reward actually 'meant' something to the subject and, therefore, influenced the response's MS. Thus, we tested if the means of MS were significantly different in trials where the reward was very small ($<$ 10 points - Group 1 and $<$ 10 cents - Group 2) from trials where the reward was very large ($>$ 90 points - Group 1 and $>$ 90 cents Group 2). The results showed no significant difference at the confidence interval of 5% for both groups of subjects (Group 1: t(14) = 0.561, $p = 0.583$, Appendix 2. Table 14; Group 2: t(16)=0.461, $p = 0.931$, Appendix 2. Table 15). These results, confirm the output of the multiple regressions analysis, showing immediate reward to have no significant impact on response MS.

A more detailed analysis was performed to better understand the relationship between the average reward felt in the near past and MS. Since, the multiple regression analysis showed the average reward to be a significant predictor for some subjects, even though only explaining a very small percentage of MS data variability. To calculate the average reward, we use the update rule described in Equation 9.7 and the individual values of α calculated before by maximizing the R^2 of the model for each subject. However, if the multiple regression analysis showed the average reward to be a significant predictor of the model for some subjects, this analysis showed no significant difference between the MS of high and low average rewards (Group 1: Appendix 2. Figure 12 and Table 16; Group 2: Figure 13 and Table 17). Result which, contrasts with the theory of Niv et al. (Niv et al., 2007).

Also, RT revealed itself to be a significant predictor of MS for almost all subjects ($p < 0.001$ for 7 out of the 8 subjects in the first group study and the subjects' totality in the second group). This result is not surprising if we take into account that RT was, in this experiment, strongly linked to strength of evidence, influencing, as already shown, the level of confidence self-reported by the subject. Nevertheless, several of the previously mentioned works which have been dedicating themselves to understand the role of immediate reward and average reward rate in response vigor, analyze vigor in the form of RTs (Guitar-Masip et al., 2011 and Beierholm et al., 2013). In their interpretation and according to their findings, when the opportunity cost of time becomes bigger, that is, the felt average reward rate was higher, subjects decide faster, presenting smaller RTs even at the expenditure of accuracy. Their results showed that immediate reward did not impact RT as significantly; although, there were noticeable longer RTs when the available reward on offer was large. According to these results, we decided to test the hypothesis that in our experiment, RT was being influenced by the average reward rate or immediate reward.

In the first place we started by testing if the RT means were significantly different between classes of small ($<$ median reward) and big ($>$ median reward) immediate available reward. The results (Appendix 2. Figures 14 and 15) revealed that the means were not significantly different for the majority of subjects (Appendix 2. Table 18 and Appendix 2. Table 19), and at the group level (Group 1: t(14)=-0.293, $p = 0.774$ and Group 2: t(16)=-0.345, $p = 0.735$) for a confidence interval of 5%. Even though the RTs means were not significantly different, one observed a subtle tendency of subjects to wait longer to respond when they could gain more points (Group 1) or lose a large amount of money (Group 2)

(Appendix 2. Figures 14 and 15). As there was this observable tendency, we decided to test once again for very extreme amounts of reward. However, the results (Appendix 2. Figure 16 and 17) showed no significant difference difference between extreme categories of reward either (Group 1: t(14) = -0.02, $p = 0.986$, Appendix 2. Table 20; Group 2: t(16)=0.0163, $p = 0.987$, Appendix 2. Table 21, $p = 0.987$) at the confidence interval of 5%. These results, leads to the conclusion that even though it was an observed tendency of subjects to listen the stimulus for longer when the 'stakes' of the trial were higher, as other suggested (Guitar-Masip et al., 2011 and Beierholm et al., 2013) this relationship was not significant in our experiment. In addition, it is still noteworthy this result might be conditioned by the fact that in a significant portion of trials, subjects waited almost until the maximum sound presentation to give a response, leaving aside other effects either than stimulus condition, to play a role. As in these trials the decision was not triggered internally, but it was timely imposed. Therefore, it is possible that in a scenario where all decisions are triggered internally (actual RTs), this effect could have been more notorious.

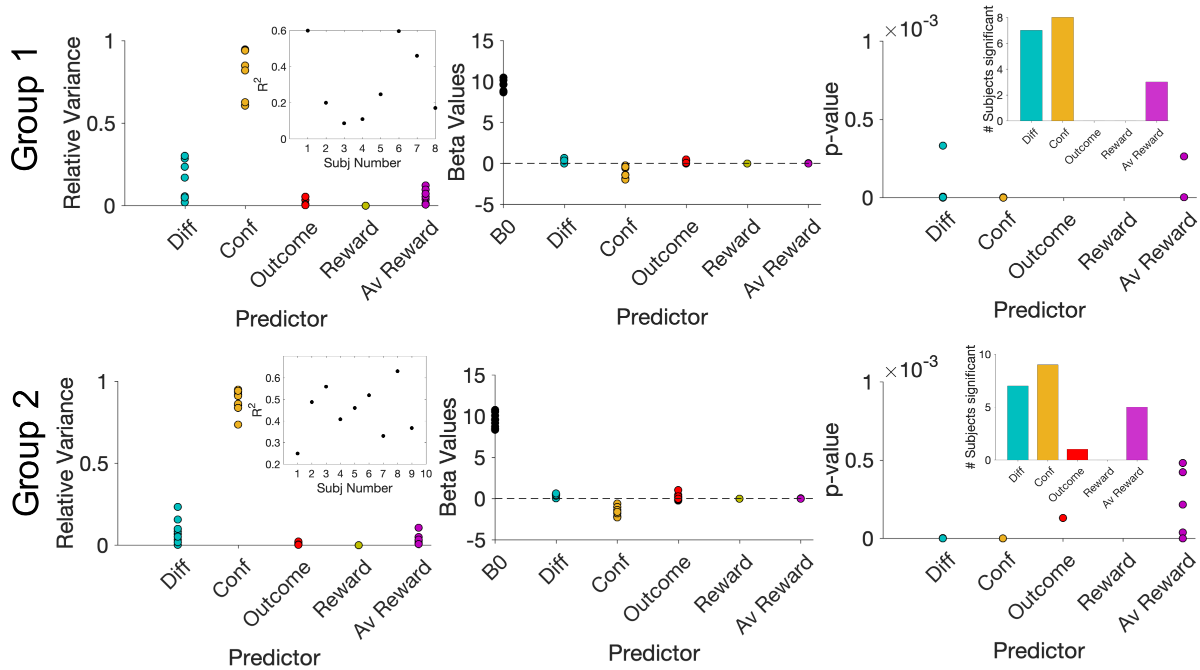

Figure 9.20: Results of multiple regression analysis for RT:**(A - Left Column)** Fraction of variance of the linear regression model explained by the stimulus difficulty (Diff), trial reward (Rew), confidence (Conf), trial reward times confidence (Rew*Conf), average reward rate felt in the past (Av Rew) and reaction time (RT). Each circle represents a subject of group 1 (top row) and of group 2 (bottom row). The inset shows R^2 of the model for each model. That is, the proportion of variance of MS explained by the model. **(B - Middle Column)** Predicted values of β for each one of the predictors. **(C - Left Column)** p-values of the different predictors, to schematize which predictors were actually significant for the model. The inset shows the number of times, that is the number of subjects, to who the predictors showed as significant were indeed significant. For example, only for a small fraction of subjects of Group 1, the predictor Av Rew was actually significant. Also on Group 2 only a small subjects models showed Av Rew to be a significant predictor.

As shown in the results of the multiple regression analysis performed to predict RT presented in Figure 9.20, average reward seems to be a significant predictor of RT to some of the subjects of both groups, being its relative power in explaining the variance of RT data superior to the immediate reward predictor. However, even if sometimes revealed as significant, the average reward predictor's beta values are always very small. The analysis of the nature of this result with a t test (Appendix 2. Figure 18 and Figure 19) show a not significant result at the confidence interval of 5% (Group 1: t(14)=-0.07, p=0.944, Appendix 2. Table 22; Group 2: t(16)=-0.162, p=0.873, Appendix 2. Table 23). This results evidences, that contrarily to many other works which many times have included RT as their measure of vigor (Guitar-Masip et al., 2011; Beierholm et al., 2013; Constantino and Daw, 2015), in our experiment average reward did not impact the velocity of making a decision.

Summarily, until now, we have shown that against all our predictions, expected reward, immediate available reward or average past reward do not significantly influence neither MS or RT. The predominant factor in determining MS and RT is, from all the analyses presented above, DC. Before discussing these results, it is still necessary to conclude about the role that immediate reward or average reward might have played in the response press rate, our other measure of vigor.

In resemblance to the analysis performed to MS and RT we used multiple regression analysis to approach the average press rate. The multiple regression analysis was performed for each subject individually. The model used to predict the average press rate on every trial, took into account: (1) the variables we considered that could be linked with response rate; (2) the ones we wanted to test -immediate reward, average reward rate and expected reward; (3) and DC solely, which we concluded above to have an impact on response rate. We also included as predictors the session number and the number of clicks necessary to block the response, once, one observes in Figure 9.18 a tendency of response press rate to decrease with the number of clicks necessary to block a response. Consequently, the model used to predict press rate was the following:

$$PR = Diff + Conf + Outcome + Reward + Av_{Rew} + Sess + Presses \tag{9.8}$$

Nevertheless, for a big portion of the subjects (80%) of both groups, we could not predict more than 10% of the variability present in the press rate data. Thus, all considerations removed from this analysis were considered insufficient and, consequently, not presented. It is noteworthy that this incapability of the model to predict press rate emphasizes its variable character. It shows that the response press rate is intrinsically linked with the subject in question, being influenced by factors we could not control in the experiment (e.g., tiredness, late for a post-experiment event, distracted while pressing the button).

When we analyzed the effect of immediate reward solely, we found no significant difference between the average press rate of different categories of reward: small ($< median$) and large ($> median$), at the confidence level of 5% (Appendix 2. Figure 20 and 21). No effect was found when performed the same analysis testing the impact of average reward rate (Appendix 2. Figure 22 and 23), as there was no significant difference between the press rate associated with the low and high categories of average reward for both groups' subjects. Also, here, the average reward rate was calculated with the update rule present in 9.7, with individual α being calculated by maximization of the R^2 of the model.

Now that we have demonstrated that not either MS, RT or response press rate are influenced by immediate available reward or average reward rate, we can conclude that against all our hypothesis, we showed in this experiment human subjects did not adjust, significantly, the vigor of their motor behavior based on an estimate of the near past average reward or as a function of the currently available reward. What one can conclude from these results in that in this experiment, subjects cared mainly about being correct or incorrect, prioritizing accuracy significantly over time and reward.

The first indication of this result is the observation of subjects' tendency to increase their RTs significantly when the stimulus condition is hard—showing a marked speed-accuracy trade-off. Especially if we take into account that many subjects waited until the end of stimulus presentation in a considerable amount of the trials, which reveals not only a possible difficulty associated with the task but also an intention of the subject to be certain about the decision at the cost of a time expenditure. Besides, if one observed an effect of reward in the RT of some subjects, the effect was small and not significant, showing that subjects prioritized being correct at the cost of spending more time listening to the sound regardless of the amount of money they could lose or gain.

Secondly, there was only one effect we could clearly identify as being truly determinant of response movement vigor - the decision confidence. More specifically, what one can conclude from this extensive analysis is that subjects move faster and press faster when they are sure they would be correct, regardless of the amount of reward in question. In this section, we aimed to find why a subject is willing to spend more energy to respond when he is confident about the decision made. Consequently, we tested the two hypotheses regarding how reward affects cost of time and, consequently, vigor - the retrospective (Niv et al., 2007; Guitar-Masip et al., 2011; Beierholm et al., 2013; Constantino and Daw, 2015; Otto and Daw, 2019); vs. prospective hypothesis (Kawagoe et al., 1998; Hamid et al., 2015; Sackaloo et al., 2015; Walton and Bouret, 2018; Summerside et al., 2018). In particular, our prediction extended the prospective hypothesis (the amount of future reward defines the cost of time, influencing how fast one moves to catch it) to include DC in the following way:

$$Cost\ of\ time \sim Expected\ reward \sim rew * conf$$

However, what one found was vigor to be exclusively influenced by DC, without reward increasing or decreasing that effect. It might have been the case that the values of reward used in the experiment or allocated in each trial were not appealing enough to out stand the role that DC has in determining the vigor of a response. In order to test this, another experiment had to be performed with a different type of reward. One can question why then, in other experiments with similar reward values (Guitar-Masip et al., 2011; Beierholm et al., 2013; Constantino and Daw, 2015;), the reward was identified as decisive for response movement vigor. It is noteworthy that the stimulus condition was kept constant in somel of these experiments, and exclusively reward was made vary across trials (Guitar-Masip et al., 2011). This, naturally, restricts significantly the levels of uncertainty associated with a decision, minimizing the potential effect that this variable may have on response vigor. What we observe in our experiment is that when different values of DC are induced, vigor is selected according to the level of confidence.

Another important conclusion of this analysis is that the past rewards did not significantly influence the current MS or average press rate in our experiment. If there were cases where average reward might have revealed itself as a significant predictor, the changes induced in the current response vigor were very small compared to DC. Once again, this result contrasts with the work of Guitar et al., and others, which have shown in a similar experimental environment, that subjects modulate the current trial MS according to the near past average reward. However, it is noteworthy that Guitar et al. 's experiment generated moments of high average reward and low average reward, being the reward currently on offer not completely independent from past reward. Our results show that in a situation where the current reward is entirely independent of the previous rewards, the past seems not to influence the time allocated to make a decision (RT), the time allocated to respond (MS), or the time between presses on the response button (press rate).

9.4 Discounting Protocol and Cost of Time

In addition to measuring cost of time indirectly through response vigor (MS and response press rate), the standard way of empirically measure cost of time is through discounting experiments. This type of experiment can test how a subject values time by measuring how long the subject is willing to wait for a certain amount of reward. Inspired by Shadmehr et al. work and results, which showed a correlation between vigor (measured in their experiment as the speed of saccades) and discounting, we introduced a very similar protocol in our experiment. The idea was to measure how long each of the subjects was willing to wait for a certain amount of reward and test if this value was modulated by the factors we aimed to test as impacting cost of time - DC, expected reward, and average reward rate.

Thus, in our experiment, we instructed the subjects about the possibility of receiving a bonus (unknown amount) that would appear with a delay after the report of their sensorial response. This delay was made variable according to subject's performance. It increased by half a second every time the subject succeeded in obtaining the bonus until the first miss; after that, the delay started to increase by 100ms only. Additionally, every time the subject failed to obtain the bonus, its latency decreased by 100ms. As a result, on each trial, subjects did not know if there was a bonus or when it would appear, even though we informed them that if present, the bonus would appear with a similar delay as the previously caught bonus. Therefore, subjects had to decide if they wanted to wait and for how long they were willing to wait. The amount of time each subject was willing to wait for their reward on each trial was our variable of interest, which we called waiting time (WT).

By imposition of this type of protocol, we expected to observe an average WT, naturally, related to the current bonus latency, but with, expectably, small variations which we aimed to test if were correlated with the variables we predicted to influence cost of time. This is, we predicted to find on each trial a WT given by:

$$WT = subjective\ bonus\ latency + \varepsilon \qquad (s)$$

Where $subjective\ bonus\ latency$ would initially correspond to an estimate of how long subjects considered necessary to wait for the bonus, that is, a prediction of when the bonus would appear, if present. However, imposed by our protocol, this value would grow, predictably, until an asymptotic value correspondent to the maximum amount of time each subject was willing to wait for the bonus - individual value of cost of time. According to Shadmehr results and other discounting experiments (Shadmehr et al., 2010; Berret and Jean, 2016; Shadmehr et al., 2018), we expected this asymptotic bonus latency to be different for each subject since not all subjects discount the value of reward equally over time. Some subjects are steeper discounters than others, that is, are less willing to wait for the same reward, as the passage of time carries more weight for those. Besides, as subjects never knew a priori the value being offered for the bonus, we predicted this asymptotic latency to be computed in function of the bonus's mean value, as subjects were aware of the maximum amount of reward possible to obtain with the bonus. In sum, we expected these asymptotic values of the bonus latency to reflect the amount of time each subject was willing to wait for the average reward offered with the bonus, reflecting these amounts of time an estimate of each subject's cost of time.

Besides, we predicted one's willingness to wait for the bonus, WT, to suffer variations from trial to trial - ε - since we introduced variables in our experiment, shown by others, to influence cost of time. For example, a notable result of discounting experiments is that the value of reward affects how long one is willing to wait for its reception, being greater values of reward associated with shorter waiting periods in comparison to smaller amounts of reward (Myerson and Green, 1995; Shadmehr et al., 2010; Berret and Jean, 2016). In our experiment, subjects only received the reward of their sensorial decision after their waiting period. By doing this, we aimed to test if the expected value of reward conditioned the time one was willing to wait for the bonus + the sensorial reward, by changing the value of ε. According to the results of the mentioned discounting experiments, we should expect smaller WTs when what is at stake is a large amount of reward compared to trials where smaller amounts of reward are at stake. In addition to understanding the role that reward plays in the WT (i.e., cost of time), this design would expectably allow us to know if one's level of belief in acquiring the same reward (DC) also influenced WT.

In summary, our discounting protocol aimed to:

- Measure individual values of cost of time - asymptote of the bonus latency. In this way, one could study if there was a relationship between how long subjects are willing to wait for the bonus (how steeply one discounts the value of reward) and their movement vigor. If so, one exposes a causal relationship between cost of time and the vigor subjects presented to indicate their response.

- Identify which factors (DC, reward) influence the steepness with which each subject discounts the reward over time, which variables interfer and how with WT's values - tackle changes in ε.

Our results of each subject WT (Appendix 2. Figure 24 and Figure 25) indeed showed that subjects on each trial present a variable WT, which initially grows with the bonus latency increase. However, from the analysis of the same figures, one can also identify several problems about the way we implemented the discount protocol in our experiment.

We observed that subjects were initially very patient about the bonus, which set its latency value very high. However, if subjects were willing to wait a lot for the first sessions' bonus, most subjects were unwilling to wait that same amount of time until the end of the experiment, that is, for the entire 20 sessions. Consequently, for most subjects we observe a rupture moment (in general, corresponds to a session) where they start to wait for the bonus only in a few trials or give up waiting, ultimately. On the other hand, few other subjects never stop increasing their WT. In either case, subjects never reach an asymptotic value for the bonus latency, as predicted, showing that the way we codded our bonus latency was inefficient in accompanying subjects' behavior. Either because it did not decrease fast enough when subjects decide the amount of time they were waiting was excessive, or because it did not increase enough in cases where subjects did not stop increasing their WT until the end of the experiment.

In addition, we also observed that some subjects waited, on average, considerably longer than the necessary to get the bonus, especially in the early sessions (Figure 9.21). This observation might be because it was not clear for all subjects that the bonus, when present, appeared around the same amount of time. The misunderstood of this information naturally carries problems for the experiment. Since it is the belief that a bonus can occur at any moment in time, utterly unrelated to when the last bonus appeared, that allowed the bonus latency to grow initially so much. The bonus latency does not represent the time a subject is willing to wait for it, but the amount of time he considers mandatory to wait for it. Naturally, one can argue if subjects wait that long it is because they are willing to, and that is true. On the other hand, as already mentioned, we observe that many subjects enter in rupture, giving up of waiting at a certain moment of the experiment, which evidences that the value they were waiting felt excessive. The point necessary to make is that the possible failure in transmitting this contingency of the bonus latency to subjects, together with the weak capacity of the way we coded the bonus latency to follow subjects' behavior, made it very hard to track each subject cost of time. As a result, except Subject 15, non of the subjects reached an asymptotic value for the WT.

Even though subjects' WT behavior was not as expected, revealing an inability of our protocol to lead the subject until its natural cost of time, we aimed to study if there was a relationship between each subject's vigor and average availability to wait for the bonus.

We started by approaching the correlation between the average MS of each subject and average waiting period. To establish this relationship, we first asked if there were significant differences between each subject MS, so that we could use MS as an individual characteristic. We performed a one-way ANOVA with subject identity as the group effect. The results showed a significant effect of subject identity ($p < 0.001$) at the confidence interval of 5%, indicating that some subjects moved their fingers

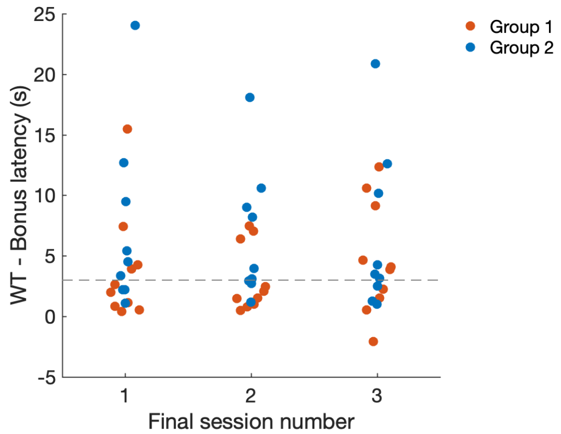

Figure 9.21: Excess time (WT - bonus latency) that each subject waited on average for the bonus in the first 3 sessions. At least half of the subjects (both groups together) wait more than 3 seconds than the necessary to get the bonus.

with reliably higher velocities than others.

As shown in Figure 9.22 A, we observed a tendency of the subjects who presented a higher average MS to wait less for the bonus to arrive. However, the correlation between the variables was not significant enough to affirm a direct relationship between MS and the time one is willing to wait for the bonus. We considered the hypothesis of average WT not be a good measure of one's cost of time for the reasons pointed above, especially after subjects start to wait for the bonus only in some trials and not in others, ruling their behavior by luck. To respond to this issue, we tried to look for other ways of measuring each subject's real willingness to wait for the bonus. The ones that showed a higher correlation with MS were the maximum bonus latency and the first missed bonus's latency.

The results revealed a subtle tendency of subjects who presented higher MSs to show lower values for the maximum bonus latency and to miss the bonus earlier than subjects with smaller MSs (Figure 9.22 D and G). Although not significant, the character of this result was present between MS and the three approaches we considered to measure subjects' availability to wait for the bonus, that is, cost of time. To look for some robustness for this relation between vigor and willingness to wait, despite all the mentioned problems and lack of rigor in measuring cost of time, we analyzed the relationship between other measures of vigor and these three measurements of one's willingness to wait.

The most solid relationship we have found so far with this experiment is that subjects vary the velocity with which they indicate their response according to the reported confidence level - being the slope of this relationship significantly more pronounced for some subjects than others (Appendix 2. Table 7). Therefore, we considered the hypothesis that the way subjects alter their velocity as function of the confidence level to be related with the time one is willing to wait for the bonus. The results showed a small correlation between the average slope of each subject MS and the average WT, maximum bonus latency and latency of the first missed bonus (Figure 9.22 B, E and H), corroborating the character of the relationship found between MS and these measurements. Naturally, we also analyzed the relationship between the average response press rate used to mark a response and the average WT (Figure 9.22 C,F and I). Once again, the results showed a small negative correlation between the average press rate and our measures of one's willingness to wait for the bonus.

The fact that one finds the same type of relationship between different measures of vigor and willingness to wait, points in the direction of existing a correlation between vigor and cost of time in our experiment. As in the work of Shadmehr et al., subjects who moved faster seem to discount time more

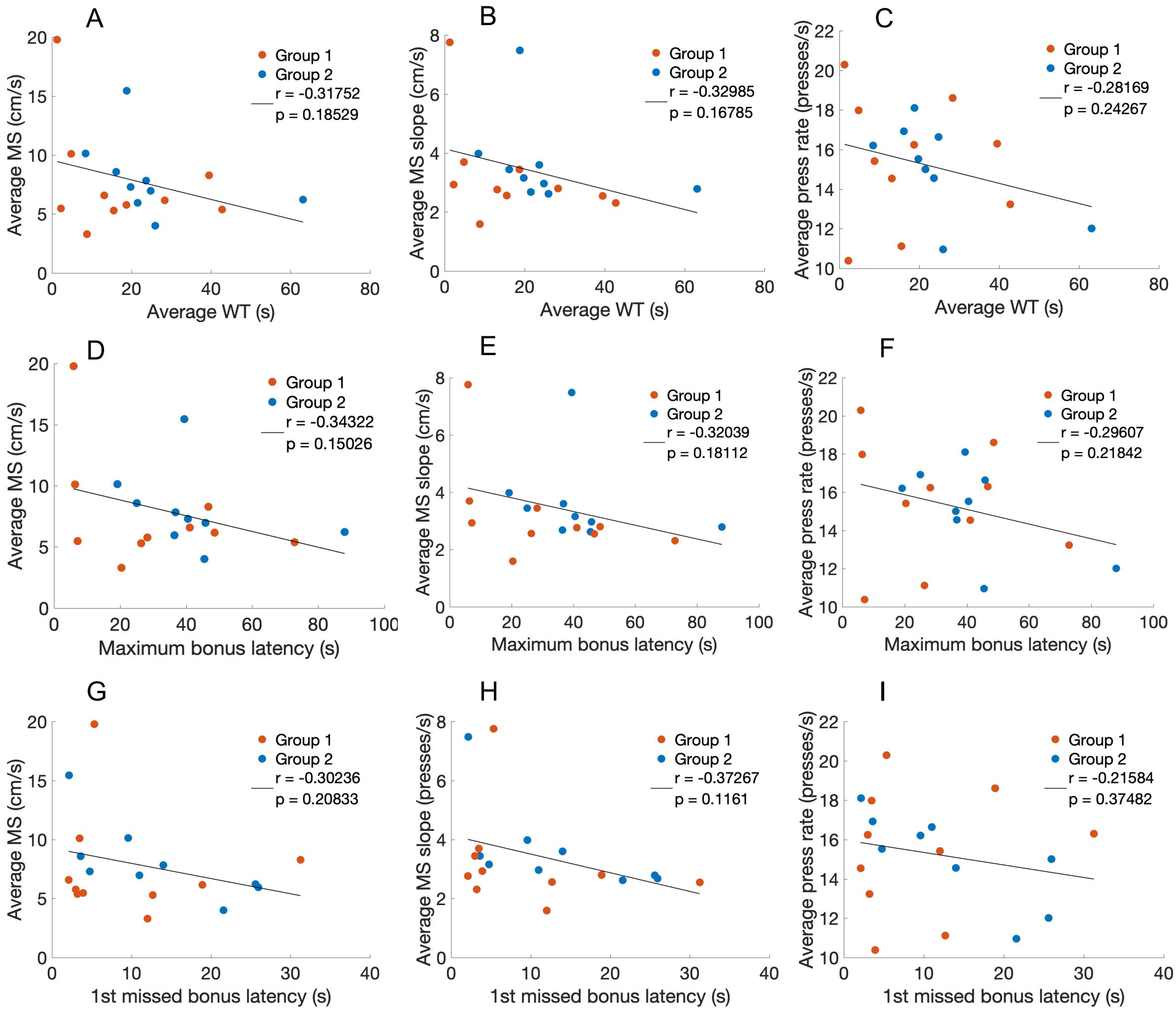

Figure 9.22: **(A)** Average MS as function of the average WT. Each blue dot represents a subject of Group 1 and in orange a subject of Group 2. There is a small negative correlation between both variables, r= -0.32 which was found non significant $p > 0.05$. **(B)** Slope of MS growing with confidence for each subject as function of the average WT. Each blue dot represents a subject of Group 1 and in orange a subject of Group 2. There is a small negative correlation between both variables, r= -0.33 which was found non significant $p > 0.05$. **(C)** Average response press rate as function of the average waiting time. There is a small negative correlation between both variables, r= -028 which was found non significant $p > 0.05$. **(D)** Average MS as function of the maximum reached latency for the bonus. There is a small negative correlation between both variables, r= -0.30 which was found non significant $p > 0.05$. **(E)** Average slope of the MS as function of the maximum reached latency for the bonus. There is a small negative correlation between both variables, r= -0.32 which was found non significant $p > 0.05$. **(F)** Average response press rate as function of the maximum reached latency for the bonus. There is a small negative correlation between both variables, r= -0.30 which was found non significant $p > 0.05$. **(G)** Average MS as function of the latency of the first missed bonus. There is a small negative correlation between both variables, r= -0.30 which was found non significant $p > 0.05$. **(H)** Average slope of the MS as function of the latency of the first missed bonus. There is a small negative correlation between both variables, r= -0.37 which was found non significant $p > 0.05$. **(I)** Average response press rate as function of the latency of the first missed bonus. There is a small negative correlation between both variables, r= -0.21 which was found non significant $p > 0.05$.

steeply, and by consequence, to wait less for the bonus on average, even though in our case, the correlations found between our variables were not significant. However, it is noteworthy that our discounting protocol limited our approaches to quantify each subject cost of time.

To follow our second aim, we analyzed how WTs varied according to the level of confidence and reward on offer. Ideally it would be possible to discern both $subjective\ bonus\ latency$ from ϵ. However, as most subjects never reached an asymptotic value for their WT this separation becomes even harder.

As shown in Figure 9.23, our results showed a very tenuous tendency of subjects to wait more for the bonus when they were more confident about their choice (Appendix 2. Table 24). This result, even only significant for very few subjects (Group 1: $F(2,21) = 2.53$, $p=0.099$; ; Group 2: $F(2,24) = 0.002$, $p=0.002$, Appendix 2. Table 24) and showing a very small effect, contradicts our predictions where expected reward computes the value of cost of time. It implies that a more confident response should lead to an increase in cost of time and, therefore, a smaller availability to wait for the bonus. However, given the difficulties we found with the implementation of our discounting protocol, not much can be concluded from this small tendency, which contradicts our results. If anything can be concluded is that our discount protocol failed in measuring each subject availability to wait for the bonus, which makes the analysis of the motivations behind small fluctuations on average WT impossible to account.

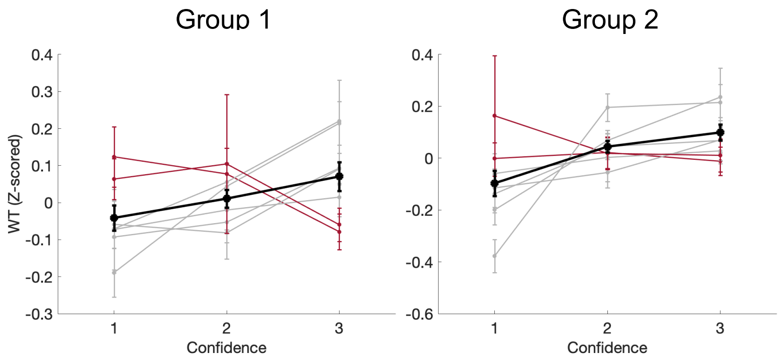

Figure 9.23: Average waiting time as a function of the self-reported confidence level for both groups of subjects. In red are signalized the relationships which contradict the tendency of the group of waiting more where confidence choices were reported. Presented in black is the relationship of the average of the group. Errorbars represent standard error across sessions (light gray and red) and across subjects (black).

The average WT analysis as a function of the immediate available reward and average reward revealed no significant effect of reward on the average WTs (Appendix 2. Figure 26, Figure 27, Figure 28, Figure 29 and Tables 25, 26), as expected from our previous results, which have shown reward to have no impact on our subjects' behavior.

9.5 Problems encountered with the experiment and Future Work

The developed experiment aimed to concretely study the impact of reward and DC on movements' vigor and how these effects are related to variations of the cost associated to the passage of time. For this, we developed a 2AFC auditory task with confidence measurements and a time discount protocol. The experimental design and the whole proposal of the experiment turned out to be quite complicated, resulting in flaws and limitations impossible to anticipate. This section dedicates itself to the enumeration of some of these flaws, followed by a reflection of how they can be transported to future works.

One of the main observations of this work is that neither reward or average reward rate seem to evoke a significant change in subjects' vigor, contradicting many previous findings (Niv et al., 2007; Guitar-Masip et al., 2011; Beierholm et al., 2013; Choi et al., 2014; Constantino and Daw, 2015; Hamid et al. 2015; Ottto and Daw, 2019). Indeed, the results show that subjects care mainly about being correct or incorrect about their response, moving faster when they consider being correct and slower when they think is likely to be incorrect. The opposing nature of this result to other findings leads us to question our reward system.

The first critique one may consider falls directly on the monetary value of each trial reward. In the sense the amounts of money offered may have been potentially too low to be relevant for the subjects.

However, we used two reward systems: one used points and the other one used monetary compensations. The results of both groups showed the amount of reward at stake to not influence movements' vigor. This fact makes us question if the nature of these results are indeed related to the amount of reward at stake or about what people value the most. In other words, do subjects move faster when they are more confident because they know they will receive the reward (regardless of its value) or because to be correct is rewarding for the subject? A possible way of disentangling this question might be to perform the same experiment where subjects are only monetarily compensated in some trials. In this way, it would be possible to directly address if the introduction of reward affects subjects' response speed and if it does, how the value of the reward being offered affects this change.

Additionally, to approach the question of what defines the cost of time - an average of the reward received in the near past or the expected future reward - we made the reward variable and independent from trial to trial. In our experiment, we observed that the average reward rate seems not to influence vigor. This result might be conditioned by the values of reward used (not significant for the subjects), or the fact that in a scenario like this, where the reward fluctuates a lot, it is not possible to discern the role of the average reward rate as this value itself does not vary much. Meaning that even though we did not observe any effect, the nature of this result might have been limited by the way we computed the environment of rewards. Therefore, this analysis would also benefit from an experimental design where subjects are only compensated for their responses in some trials, since this introduces a very localized variation in the average reward rate.

The way we computed the bonus latency also showed insufficiencies, as noted above. In order to get around this difficulty, in addition to the clear change in the computation of the bonus latency, it would be potentially preferable in a first version of the experience to let the subjects reach the maximum time they are willing to wait for the bonus, that is, their subjective time cost value. Then, only after the subjects reached this value, the subjects would be introduced to the final version where factors such as reward and confidence in the decision would be integrated. In this way, it would be easier to analyze the variations that these variables introduce in the value of the temporal cost found for each subject.

However, in addition to the inability to adapt the bonus latency to each subject's waiting disposition, there are other relevant observations to make. One is the way the instructions are given. When carrying out experiments on humans, instructions are very important. One does not want to influence the behavior of the subjects, but it is essential to inform about the different aspects of the experiment. The instructions given in our experiment probably conditioned the subjects to wait long for the bonus. First, by not making it clear enough that the bonus latency was similar to the previous bonus latency. Second, because it was not clear to the subjects that there was a patent relationship between money and time in our experiment. Naturally, behavior cannot be influenced by directly mentioning the objective of the experiment, but somehow this should be a decision that the subjects should be aware of. However, qualitatively, the subjects' behavior, especially in the first attempts of the experiment, was mainly to obtain the bonus as if it were our experimental objective.

It is important not to forget that even though a way was found to overcome the problem behind confidence levels and their uneven distance to the central button, in future work the buttons associated with different confidence levels should imply the same movement distance. If not possible, the significance of the buttons should be shifted during the experiment.

Apart from this, our experiment was trial-fixed and not duration-fixed. We made this decision because we considered that having a duration-fixed experiment was necessarily influencing subjects to wait less for the bonus, as they could compensate by performing more trials. However, I consider that having a bonus with an appealing value and duration-fixed experiment would have brought awareness to the subjects about the time-money paradigm. That is why, in a lot of the experiments with the goal of studying vigor, they perform experiments with a fixed duration and not with a fixed number of trials.

The other problem encountered in our experiment and more challenging to get around is the extremely variable subjects' behavior. The experiment had a duration of 20 sessions, and subjects given to schedule incompatibilities and other factors did not perform a session per day. Not only sessions sometimes became very sparse in time, but in addition to that, factors as being late for the session, have other appointments after the session, stress, and others may have contributed to the variations in behavior observed within sessions and contaminated the results of particular sessions. Besides, all subjects who performed the experiment were scientists at the Champalimaud Foundation. It is important not to exclude the possibility that pre-conceived ideas about the experimental goal may have contaminated some results as well.

Chapter 10

Conclusions

In this work, a perceptual decision-making task with collection of confidence reports and a discounting protocol was developed in order to investigate vigor and the nature of its relationship with decision-confidence and reward.

The sensory task involved the presentation of sound clicks with a different total number of clicks to both ears. After collecting evidence for as long as subjects felt necessary, they had to indicate which side they considered to have a higher probability of displaying a click. The experiment results showed that subjects were able to learn this sensory task and that the task itself successfully evoked uncertainty among subjects' responses. More importantly, our analysis showed that these different levels of uncertainty were reliably translated into different decision confidence reports, with reports of response's certainty associated with easier stimuli conditions and smaller RTs, and reports of uncertainty associated with more difficult conditions and bigger RTs. Pattern also found in previous studies of DC (Kiani and Shadlen, 2009; Fetsch et al., 2014).

The experiment design and the way it included the measurement of one's movement duration while reporting choice and decision confidence allowed to directly analyze the movement speed associated with each level of confidence. The analysis of these results exposed the existence of a strong correlation between one's decision confidence and the speed used to report this response, with reports of more confident responses associated with higher movement speeds. However, as the report of different levels of DC implied different finger displacements, given the distribution of the buttons on the numeric keypad, we consider the hypothesis of this relationship to be exclusively a consequence of different movement trajectories. In order to test this hypothesis, we performed a simple version of our main task without sensory stimulation, where subjects had to perform the same finger trajectories but now without any sensorial value attributed to any of the buttons. The results showed that even though subjects presented higher velocities when performing longer movements (same trajectories as the ones necessary to report more confident choices) than shorter movements (trajectories to report uncertainty), the difference between these trajectories velocities was significantly smaller than in the case where these movements were associated with reports of confidence. The experiment was only performed by 4 human subjects, and ideally should have been performed by the same subjects who performed the main experiment before any confidence value had been attributed to the buttons. Nevertheless, the results showed that the difference of velocities associated to the report of different confidence levels could not be entirely explained by different trajectories, confirming that more confident responses are associated with more vigorous movements.

In addition, the found relationship between MS and DC of the main experiment showed to be robust, as both variables were not only linked by the significance of a direct relationship but when marginalized over strength of evidence for correct and incorrect responses, the pattern presented by both variables (MS and DC) was exactly the same. All these results show evidence that movement speed is an implicit measure of DC, as suggested by others (Seideman et al., 2018). It is noteworthy that the nature of this result is not only relevant to understand better the mechanism behind the selection of one action's movement speed, but it is also of extreme importance for works where the collection of explicit measures of DC is not possible.

Our experiment involved the collection of another vigor' measurement - the response press rate. The analysis of its relationship with DC revealed a tendency of pressing faster on the response press button when reporting higher levels of DC. Even though this relationship was not evident in all subjects' behavior and the correlation found was less pronounced when compared to the relationship between MS and DC, it corroborates with the observation that reports of a higher level of confidence are associated with more vigorous movements (Seideman et al., 2018).

Notably, this result corroborates with the proposed prospective hypothesis, where the value of the expected reward defined as $< Rew >= rew * DC$, determines the cost associated to the passage of time and sustained by other works (Kawagoe et al., 1998; Hamid et al., 2015; Sackaloo et al., 2015; Walton and Bouret, 2018; Summerside et al., 2018). However, our results showed no significant reward effect in the collected vigor measurements - MS and response press rate. In opposition to our proposal, the first works that have suggested cost of time to be a brain's computation fundamental to decide the vigor of an action, state that the cost of time should be influenced by the average reward received in the near past, and not by the expected future reward (Niv et al., 2007). To test this theory, a variable reward system was also introduced in this experiment, which allowed us to investigate whether in a changing environment, where the past does not predict the future, the cost of time can also be predicted by an average of the reward received in the near past. However, our experiment results revealed that no variability of subjects' vigor (MS and response press rate) could be accounted by the average of the reward received in the near past, contradicting many previous findings (Niv et al., 2007; Guitar-Masip et al., 2011; Beierholm et al., 2013; Choi et al., 2014; Constantino and Daw, 2015; Hamid et al. 2015; Ottto and Daw, 2019).

Therefore, from the results of this experiment, one can conclude that neither reward nor average reward rate seem to influence subjects' action vigor. Indeed, what was observed is that confidence is able to predict most of the variability in vigor data, being the only variable with a significant and clear role in vigor modulation across subjects from both groups of reward system.

In addition to measuring vigor directly by accessing movements' speed, in our experiment, we aimed to measure vigor indirectly by measuring the time one is willing to wait for a reward. The inclusion of this discounting protocol, shown by others to be successful in linking vigor with cost of time (Shadmehr et al. 2010) showed some limitations in our experiment. In the first place, the way we codded the time each subject had to wait for the reward was unsuccessful in finding each subject's real availability to wait for the average reward received with the bonus. As a result, similarly to what Shadmehr et al. found, our results showed a tendency of subjects who presented a bigger average vigor in their movements to wait on average less for the bonus. However, this relationship between average vigor and average WT shows a very weak correlation, most likely due to the fact that almost no subject reached a value truly representative of the actual amount of time they were willing to wait for the bonus. The fact that our discount protocol was unable to guide the subjects to a plateau value of the WT, leading in several cases subjects to a rupture point where the behavior stops to be related with the factors of the experiment, made it impossible to find the real relationships between reward or even DC and the average WTs in our experiment.

Consequently, this experiment failed to respond to the way reward affects the vigor of our movements. Once, it was not possible to find a significant effect of the available reward or the average reward rate both on our direct measures of vigor (MS and response pres rate) and on its indirect measure through the time someone is willing to spend waiting for a reward. Therefore, this experiment was not able to reach to a conclusion about the prospective and retrospective hypothesis of the cost of time computation, and more specifically how can DC be included in these normative theories of movement's vigor. Even though, by concluding about a monotonic relationship between vigor and DC in our experiment, we are showing evidence in favor of our proposed prospective view. As the value of the reward did not show any influence on vigor no further conclusion can be reached.

With all the works that have highlighted the role of reward both in the direct definition of cost of time (Shadmehr et al., 2010; Berret and Jean, 2016) or vigor (Niv et al. 2007; Guitar-Masip et al. 2011; Constantino and Daw, 2015; Sackaloo et al., 2015; Summerside et al., 2017, Ottto and Daw, 2019), this result can only indicate that there were flaws and limitations in our experience, which we were unable to predict and that were impossible to correct within the scope of this thesis. However, their identification is important to promote the success of future studies:

- Contrary to expectations, human beings seem to attach more value to being correct than to the amount of reward itself. One possible way to unravel this observation would be to make the reward bigger (most likely the rewards were too small to have real value for the subjects) and available only in some attempts. In other words, it could be tested whether the introduction of a reward directly affects the speed of a movement, or if for the same level of confidence the vigor does not vary from a situation in which the reward is available to a situation in which it is not.

- With all the difficulties encountered in measuring each subject's value for cost of time and consequent analysis of any effect of DC or reward, the experiment discount protocol should naturally be reviewed. Besides reviewing the bonus latency's computation, the experiment would most likely benefit from a separation in two phases. An initial phase dedicated in making subjects reach a plateau value of cost of time and a second phase where other factors would come into play in the experiment - reward and DC - in order to analyze their effect in each subjects' value of cost of time.

Chapter 11

Conclusion

Within the scope of this thesis, two perceptual decision-making experiences were developed. The first, based on the observation of a new psychophysical regularity in rodents - the TIED - specifically sought to verify the existence of this regularity in human perception. The developed experiment and analysis of subjects' accuracy and RT confirmed this regularity as inherent to humans' perception. This result is a significant step in the empirical support of the bounded accumulation of evidence mechanism, which TIED specifications allow to determine, as the mechanistic foundation of the primordial Weber's Law (work included in the article Pardo-Vazquez et al., 2019).

The observation of the rats' behavior while performing the original version of this first experiment showed that the animals indicated their responses using different speeds. The remarkable similarity of the pattern followed by their MS as a function of strength of evidence with the behavior observed for DC across stimulus difficulty in other works, led us to carry out another experiment that sought to study the vigor and its relationship with DC. However, a work that aimed to study vigor could not exclude the recent findings that have exposed vigor as being related to reward by the computation of cost of time. Thus, the developed experiment sought to address vigor and its relationship with DC and include DC in these normative theories that expose cost of time as decisive in determining response vigor. This experiment's results allowed us to expose the existence of a monotonic relationship between the response vigor and the level of certainty about the correctness of that response. However, the experiment failed to respond as proposed to how this result falls into the normative theories that expose the vigor of a reaction as being related to what the environment can offer, both in the immediate (prospective hypothesis) and based on the past (retrospective hypothesis). Even though we were unable to answer this second question, it is important to note that this experience has taken steps in understanding:

(1) DC should not be excluded in a study of perceptual vigor;

(2) the reward system used when we study vigor in humans has to be carefully designed, as well as all the time contingencies of the task, offering ideas for an upcoming experiment with the same objective.

Chapter 12

Bibliography

Banks, W. P. (1970). Signal detection theory and human memory. Psychological Bulletin, 74(2), 81–99.

Beierholm, U., Guitar-Masip, M., Economides, M., Chowdhury, R., Duzel, E., Dolan, R., & Dayan, P. (2013). Dopamine Modulates Reward-Related Vigor. Neuropsychopharmacology, 38, 1495-1503.
Berret, B. & Jean, F. (2016). Why Don't We Move Slower? The value of Time in the Neural Control of

Action. The Journal of Neuroscience, 36(4):1056-1070.

Bolles, R. C. (1967). *Theory of Motivation*. New York: Harper & Row.

Brown, J. S. (1961). *The Motivation of Behavior*. McGraw-Hill.

Brunton, B. W., Botvinick, M. M. & Brody, C. D. (2013). Rats and Humans Can Optically Accumulate Evidence for Decision-Making. Science, 340(6128), 95-98.

Celesia, G. & Hickok, G. (2015). Sound Localization in: The Human Auditory System. (1st Edition). Volume 129, Elsevier.

Chittka, L., Skorupski, P., & Raine, N. (2009). Speed-accuracy tradeoffs in animal decision-making. Tends in Ecology & Evolution, 24(7), 400-407.

Choi, J. E. S., Vaswani, P. A. & Shadmehr, R. (2014). Vigor of Movements and the Cost of Time in Decision Making. The Journal of Neuroscience, 34(4), 1212-1223.

Clark III, R. D. (1988). On predicting minor influence. European Journal of Social Psychology, 18(6), 515-526.

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. Cognitive, Affective, & Behavioral Neuroscience, 15, 837-853.

Cornsweet, T. N & Pinsker, H. M. (1965). Luminance discrimination of brief flashes under various conditions of adaptation. The Journal of Physiology, 176(2), 294-310.

Cools, R., Nakaruma, K., & Daw, N. D. (2011). Serotonin and Dopamine: Unifying Affective, Activational, and Decision Functions. Neuropsychopharmacology Reviews, 36, 98-113.

Cutnell, J. D. & Johnson, K. W. (1997). *Physics*. Wiley, Volume 1.

Dayan, P. (2012). Instrumental vigour in punishment and reward. European Journal of Neuroscience, 35, 1152-1168.

Deco, G., & Rolls, E. T. (2006) Decision-making and Weber's Law: a neurophysiological model. European Journal of Neuroscience, 24(3), 901-916.

Dezfouli, A., Balleine, B. W. & Nock, R. (2019). Optimal response vigor and choice under non-stationary outcome values. Psychonomic Bulletin & Review, 26, 182-204.

Dickinson, A. & Balleine (2002). The Role of Learning in the Operation of Motivational Systems. Wiley Online Library, https://doi.org/10.1002/0471214426.pas0312

Ditterich, J., Mazurek, M. E., & Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. Nature neuroscience, 6(8), 891-898.

Drugowitsch, J., DeAngelis G. C., Klier E. M., Angelaki, D. E. & Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. eLife, 3:e03005. Edwards, W., 1965. Optimal strategies for seeking information: models for statistics, choice reaction times, and human information processing. Journal of Mathematical Psychology 2, 312–329. Fastl, H., & Zwicker, E. (2007) Psychoacoustics: facts and models. Springer, Berlin, Heidelberg.

Fechner, G. T. (1860). *Element der psychophysik.* Breitkopf and Harterl.

Fetsch, C. R., Kiani, R., Newsome, W. T., and Shadlen, M. N. (2014). Effects of cortical microstimulation on confidence in a perceptual decision. Neuron 83, 797-804.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American Psychologist, 34(10), 906–911.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J.(2016). Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. Annual Review of Psychology, 67, 641-666.

Gaydos, H. F. (1958). Sensitivity in the Judgment of Size by Finger-Span. The American Journal of Psychology, 37(3), 557-562.

Gescheider, G. A. (2013). *Psychophysics: the fundamentals.* (3rd Edition). Psychology Press.

Gold, J. I., Law, C. T., Connolly, P., & Bennur, S. (2010). Relationships between the threshold and slope of psychometric and neurometric functions during perceptual learning: implications for neuronal pooling. Journal of neurophysiology, 103(1), 140-154.

Granziano, M., & Sigman, M. (2009). The Spatial and Temporal Construction of Confidence in the Visual Scene. PLoS ONE 4(3): e4909.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. John Wiley.

Guitar-Masip, M., Beierholm, U. R., Dolan, R., Duzel, E., & Dayan, P. (2011). Vigor in the Face of Fluctuating Rates of Reward: An Experimental Examination. Journal of Cognitive Neuroscience, 23(12), p.3933-3938.

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Weele, C. M. V., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. Nature Neuroscience 19, 117-126.

Hartline, E. R. & Graham, C. H. (1932). Nerve impulses from single receptors in the eye. Journal of Cellular and Comparative Physiology, Vol 1, No 2, 227-295.

Hauser, C. K. & Salinas, E. (2014). Perceptual Decision Making. Encyclopedia of Computational Neuroscience, pp 1-21.

Hull, C. L. (1943). Principles of behavior: an introduction to behavior theory. Appleton-Century.

Indow, T. & Stevens, S. S. (1966). Scaling of saturation and hue. Perception & Psychophysics 1, 253-271.

Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998) Expectation of reward modulates cognitive signals in the basal ganglia. Nature Neuroscience 1, 411-416.

Kiani, R., & Shadlen M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. Science, 324(5928), 759-764.

Kiani, R., Corthell L., & Shadlen M. N. (2014). Choice Certainty is Informed by Both Evidence and Decision Time. Neuron, 84(6), 1329-1342.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. New York: Academic Press.

Laming, D. (1985). Some principles of sensory analysis. Psychological Review, 92(4), 462–485.

Laming, D. (1986). *Sensory Analysis*. London: Academic.

Laming, D. (2009). *Inside psychology: A science over 50 years*. Oxford University Press. Chapter: Weber's Law, pp 179-191.

Lauer, A. M. , & May, B. J.(2011) The medial olivocochlear system attenuates the developmental impact of early noise exposure. Journal of the Association for Research in Otolaryngology,12(3), 329–43.

Leshowitz, B., Tavb, H. B., Raab, D. H. & College, B. (1968) Visual detection of signals in the presence of continuous and pulsed backqrounds. Perception & Psychophysics 4, 207-213.

Lichtenstein, S., Fischhoff, B., & Phillips, L.D. (1981). Calibration of Probabilities: The State of the Art to 1980 (DTIC Document).

Link, S. W. (1992). *The Wave Theory of Difference and Similarity.* (1st Edition). Psychology Press.

Luce, R. D. (1986). *Response times* (No. 8). Oxford University Press.

Mazzoni, P., Hristova, A. & Krakauer J. W. (2007). Why Don't We Move Faster? Parkinson's Disease, Movement Vigor and Implicit Motivation. The Journal of Neuroscience. 27(27): 7105-7116.

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). Metacognition: Knowing about knowing. The MIT Press.

Meter, D. V., & Middleton, D. (1954). Modern Statistical approaches to reception in communication theory. Transactions of the IRE Professional Group on Information Theory. 4(4), 119-145.

Miller, G. A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. Journal of the Acoustical Society of America, 19, 609–619.

Milstein, D. M., & Dorris, M. C. (2007). The Influence of Expected Value on Saccadic Preparation. Journal of Neuroscience, 27(18), 4810-4818.

Mulder, M. J., Wagenmakers, E-J., Ratcliff, R., Boekel, E., & Forstmann B. U. (2012). Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. Journal of Neuroscience, 32(7), 2335-2343.

Myerson, J., & Green, L. (1995). Discounting of Delayed Rewards: Models of Individual Choice. Journal of Experimental Analysis of Bahevior, 64(3): 263-276.

Nachmias, J., & Steinman, R. M. (1965). Brightness and discriminability of light flashes. Vision Research, 5(9): 545-557.

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. Trends in Cognitive Sciences. Vol 10, No 8.

Niv, Y., Daw, N. & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology 191, 507-520.

Nodal, F. R., Bajo, V. M., Parsons, C. H., Schnupp, J. W., & King, A. J. (2008). Sound localization behavior in ferrets: Comparison of acoustic orientation and approach-to-target responses. Neuroscience, 154(1), 397-408.

Oberlin, K. W. (1936). Variation in intensive sensitivity to lifted weights. Journal of Experimental Psychology, 19(4), 438–455.

Otto, A. R. & Daw, N. D. (2019). The opportunity cost of time modulates cognitive effort. Neuropsychologia 123, 92-105.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. Journal of Vision, 5(1), 376-404.

Panek, W. & Stevens, S. S. (1966). Saturation of red: A prothetic continuum. Perception & Psychophysics 1, 59-66.

Pardo-Vazquez, J. L., Castiñeiras J., Valente, M., Costa, T., & Renart, A. (2018). Weber's law is the result of exact temporal accumulation of evidence. bioRxiv, https://doi.org/10.1101/333559

Pardo-Vazquez ,J. L., Castiñeiras- de Saa, J., Valente, M., Damião, I., Costa, T., Vicente, M. I., Mendonça, A., Mainen, Z. F., & Renart, A. (2019). The mechanistic foundation of Weber's law. Nature Neuroscience 22, 1493-1502.

Peirce, C. S. & Jastrow, J. (1884). On small differences in sensation. Memoirs of the National Academy of Sciences, 3:75-83.

Peterson, W., Birdsall, W; & Fox, W. (1954). The theory of signal detectability. Transactions of the IRE Professional Group on Information Theory. 4(4), 172-212.

Radillo, A. E., Veliz-Cuba, A, Josic, K., & Kilpatrick, Z. P. (2019). Performance of normative and approximate evidence accumulation on the dynamic clicks task. Neurons and Cognition.

Ratcliff, R. (1978). A theory of memory retrieval. Psychological Review, 85(2), 59–108.

Ratcliff, R. (2001). Diffusion and random walk processes. InInternational encyclopedia of the social and behavioral sciences (Vol. 6, pp. 3668–3673). Oxford: Elsevier.

Ratcliff, R., & Smith, P. L. (2004). A Comparison of Sequential Sampling Models for Two-Choice Reaction Time. Psychological Review, 111(2), 333–367.

Ratcliff, R., Smith, P. L., Brown, S. D., & Mckoon, G. (2016) Diffusion Decision Model: Current Issues and History. Trends in Cognitive Sciences, 20(4), 260-281.

Recanzone, G. H. & Beckerman, N. S. (2004). Effects of intensity and location on sound location discrimination in macaque monkeys. Hearing Research, 198(1-2),116-124.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task.The Journal of neuroscience, 22(21), 9475- 9489.

Sabin, A. T., Macpherson, E. A., & Middlebrooks, J. C. (2005). Human sound localization at near-threshold levels. Hearing Research, 199(1-2),124-134.

Sackaloo, K., Strouse, E., & Rice, M. S. (2015). Degree of Preference and Its Influence on Motor Control When Reaching for Most preferred, Neutrally Preferred, and Least Preferred Candy. SAGE Journals, 35(2).

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a Statistical Computation in the Human Sense of Confidence. Neuron 90, 499-506.

Schnupp, J., Nelken, I., & King, A. (2011). *Auditory neuroscience: Making Sense of sound*. Cambridge, MA: MIT press.

Schouten, J. F., & Bekker, J. A. M. (1967). Reaction time and accuracy. Acta Psychologica, 27,143-153.

Seideman, J. A., & Stanford T. R. & Salinas, E. (2018). Saccade metrics reflect decision-making dynamics during urgent choices. Nature Communications, 9, 2907.

Shadmehr, R., Xivry, J. J. O., Xu-Wilson, M., & Shih, T. (2010). Temporal Discounting of Reward and the Cost of Time in Motor Control. The Journal of Neuroscience, 30(31), 10507-10516.

Simen, P., Vlasov, K., & Papadakis, S. (2016). Scale (in)variance in a unified diffusion model of decision making and timing. Psychological Review, 123(2), 151–181.

Smith, J. D., Couchman, J. J., & Beran, M. J. (2012). The highs and lows of theoretical interpretation in animal-metacognition research.

Stellmack, M. A., Viemeister, N. F., & Byrne, A. J. (2004) Monaural and interaural intensity discrimination: Level effects and the “binaural advantage”. The Journal of the Acoustical Society of America 116, 1149.

Stevens, S. S. (1957). On the psychophysical law. Psychological Review, 64(3), 153–181.

Stone, M. (1960). Models for choice-reaction time. Psychometrika 25, 251-260.

Sun, L., & Zhong, X. (2015). Dynamic binaural sound source localization with interaural time difference cues: Artificial listeners. The Journal of the Acoustical Society of America, 137(4), 2226-2226.

Summerside, E. M., Shadmehr, R. & Ahmed, A. A. (2018). Vigor of reaching movements: reward discounts the cost of effort. Journal of Neurophysiology 119: 2347-2357.

Tachibana, Y., & Hikosaka, O. (2012). The Primate Ventral Pallidum Encodes Expected Reward Value and Regulates Motor Action. Neuron, 76(4), 826-837.

Tanner, W. P., Jr., & Swets, J. A. (1954). A decision-making theory of visual detection. Psychological Review, 61(6), 401–409.

Teodorescu, A., Moran, R., & Usher, M. (2016). Absolutely Relative or Relatively Absolute – Violations of Value Invariance in Human Decision Making. Psychonomic Bulletin & Review 32(1):22-38.

Thurstone, L. L. (1927). A law of comparative judgment. Psychological Review, 34(4), 273–286.

Treutwein, B. & Strasburger, H. (1999). Fitting the psychometric function. Perception & Psychophysics 61, 87-106.

Tulving, E. (2005). Episodic Memory and Autonoesis: Uniquely Human? In H. S. Terrace & J. Metcalfe (Eds.), The missing link in cognition: Origins of self-reflective consciousness (p. 3–56). Oxford University Press.

Uchida, N. & Mainen Z. F. (2003). Speed and accuracy of olfactory discrimination in the rat. Nature Neuroscience, 6, 1224-1229.

Uchida, N., Poo, C., & Haddad, R. (2014). Coding and transformations in the olfactory system. Annual review of neuroscience, 37, 363-385.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. Neuron 31(4), 523-535.

Wald, A. (1947). *Sequential Analysis*. New York: Wiley.

Walton, M. E. & Bouret, S. (2018). What is the Relationship between Dopamine and Effort? Tends in Neurosciences, Volume 42, Issue 2, P79-91.

Weber, E. H. (1834). *De pulsu, resorptione, auditu et tactu : Annotationes anotomicae et physiologicae, auctore.* Lipsiae : prostat apud C.F Kochler.

Wesolek, C. M., Koay, G., Heffner, R. S., & Heffner, H. E.(2010). Laboratory rats (Rattus norvegicus) do not use binaural phase differences to localize sound. Hearing Research, 265(1-2), 54-62.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. Acta Psychologica, 41(1), 67–85.

Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. Experimental Brain Research, 196, 475-481.

Yamamoto, S., Kim, H. F., & Hikosaka, O. (2013). Reward Value-Contingent Changes of Visual Responses in the Primate Caudate Tail Associated with a Visuomotor Skill. Journal of Neuroscience, 33(27), 11227-11238.

Yost, W. A. & Schlauch, R. (2001). Fundamentals of Hearing: An Introduction (4th edition). The Journal of the Acoustical Society of America, 110, 1713.

Zimmerman M.E. (2011) Speed–Accuracy Tradeoff. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) Encyclopedia of Clinical Neuropsychology. Springer, New York, NY.

Appendices

.1 Human Task - Looking for the TIED

	Sensitivity (d')			
	$Mean \pm SD$		$p\ value$	$\#\ Blocks$
	ABL 40	ABL 60		
Subject 1	2.430±0.122	2.374±0.365	0.589	7
Subject 2	1.953±0.299	2.263±0.493	0.140	10
Subject 4	1.656±0.244	2.317±0.354	**0.003**	10
Subject 5	2.249±0.406	2.034±0.337	0.291	9
Subject 6	1.730± 0.340	2.027±0.187	**0.049**	12
Subject 9	1.772± 0.162	1.813± 0.232	0.734	7
Subject 11	2.338± 0.235	2.268± 0.411	0.870	12
Subject 12	1.689 ± 0.163	2.000± 0.176	**0.003**	10
Subject 13	2.203 ± 0.350	2.121 ± 0.447	0.450	11
Sample	2.002±0.305	2.135±0.183	0.984	

Table 1: Accuracy results for each subject not discarded from previous analysis. In the first 2 columns are presented the values of sensitivity (d') for each ABL, 40 and 60 dB SPL. Next, we report the p-values of a sensitivity comparison between both values of ABL used, statistical analysis performed with a Fisher's exact permutation test. At last number of blocks in stable conditions used to perform this analysis. At the group level, there are no significant differences in sensitivity between ABLs.

	Reaction Times ANOVA					
	ABL			ILD		
	F	$d.f$	$p\ value$	F	$d.f$	$p\ value$
Subject 1	5.47	1	0.02	44.82	3	0.00
Subject 2	1.16	1	0.29	37.26	3	0.00
Subject 4	).83	1	0.00	20.45	3	0.00
Subject 5	.36	1	0.55	11.06	3	0.00
Subject 6	1.33	1	0.00	30.05	3	0.00
Subject 9	.36	1	0.55	7.40	3	0.00
Subject 11	.29	1	0.59	20.85	3	0.00
Subject 12	.33	1	0.25	19.05	3	0.00
Subject 13	.15	1	0.70	56.89	3	0.00
Sample	6.20	1	0.04	16,66	3	0.00

Table 2: Results for two-way ANOVA testing differences in mean reaction times across ABL (left three columns) and ILD (right three columns). The effect of ABL is not significant for all subjects but is significant at the group level. The effect of ILD is significant for all subjects data and at the group level.

.2 Sound localization task in a changing environment with temporal discounting

# Training Sessions				
Subject Number	Group 1		Group 2	
	Version 1	Version 2 (with DC)	Version 1	Version 2 (with DC)
0	2	1	1	3
1	1	1	2	1
2	2	1	2	1
3	2	1	1	2
4	2	2	1	2
5	2	3	1	2
6	2	2	1	2
7	2	1	1	2
8	2	2	1	2
9	2	2	1	4

Table 3: Description of the number of training sessions (both versions) performed by all subjects of both groups of subjects.

	Reaction Time ANOVA		
	Difficulty		
	F	*d.f*	*p value*
Subject 11	122.04	2	<0.001
Subject 12	12.69	2	<0.001
Subject 13	2.62	2	0.081
Subject 14	13.62	2	< 0.001
Subject 15	3.42	2	0.039
Subject 16	49.53	2	< 0.001
Subject 17	86.66	2	< 0.001
Subject 18	14.86	2	< 0.001
Subject 21	7.32	2	0.001
Subject 22	43.70	2	<0.001
Subject 23	54.34	2	<0.001
Subject 24	49.74	2	0.092
Subject 25	7.00	2	0.002
Subject 26	119.29	2	<0.001
Subject 27	13.97	2	<0.001
Subject 28	185.32	2	<0.001
Subject 29	37.04	2	<0.001
Sample	9.94	2	< 0.001

Table 4: Results for one-way ANOVA testing differences in reaction times across different levels of stimulus difficulty (Easy, Medium, Hard). The effect of difficulty is significant for all subjects (exception subject 13) and also significant at the group level.

	Decision Confidence **Linear Regression Analysis** **DC = $\beta_0 + \beta_1 S + \beta_2 T$**					
	Difficulty		*RT*		*Model*	
	β_1	*p value (H0: β_1=0)*	β_2	*p value (H0: β_2=0)*	*F*	*p value*
Subject 10	-0.23	<0.001	-0.23	<0.001	770	<0.001
Subject 16	-0.13	<0.001	-0.13	<0.001	129	<0.001
Subject 17	-0.16	<0.001	-0.23	<0.001	698	<0.001
Subject 18	-0.41	<0.001	-0.15	<0.001	577	<0.001
Subject 22	-0.18	<0.001	-0.16	<0.001	476	<0.001
Subject 23	-0.25	<0.001	-0.19	<0.001	682	<0.001
Subject 24	-0.27	<0.001	-0.20	<0.001	343	<0.001
Subject 25	-0.25	<0.001	-0.14	<0.001	411	<0.001
Subject 26	-0.26	<0.001	-0.19	<0.001	581	<0.001
Subject 27	-0.22	<0.001	-0.16	<0.001	255	<0.001
Subject 28	-0.16	<0.001	-0.27	<0.001	789	<0.001
Subject 29	-0.23	<0.001	-0.14	<0.001	303	<0.001

Table 5: Results of the linear regression analysis performed for the subjects not excluded under the criteria ($RT > 9s$ in 60% of the trials) with the model $DC = \beta_0 + \beta_1 S + \beta_2 T$ where S corresponds to stimulus strength or difficulty (Easy, Hard, Medium) and T to RT. The results indicate that both predictors, S and T (Difficulty, RT) were significant in predicting decision confidence, for each subject with a confidence interval of 5%. In the table we can also verify the values of β_1 were negative for every subject and at the group level, showing that an increase in stimulus difficulty lead to a significant decrease of the self-reported level of confidence for every subject. The values of β_2 were also negative for all subjects showing an increase in RT represent a significant tendency to report smaller values of DC. The last 2 columns of the table, shows us as well that the model was significant for every subject reassuring the quality of the results.

	Decision Confidence **Linear Regression Analysis** **DC = $\beta_0 + \beta_1 S + \beta_2 T$**					
	Difficulty		*RT*		*Model*	
	β_1	*p value (H0: β_1=0)*	β_2	*p value (H0: β_2=0)*	*F*	*p value*
Subject A1	-0.29	<0.001	-0.12	<0.001	255	<0.001
Subject A2	-0.47	<0.001	-0.34	<0.001	234	<0.001
Subject A3	-0.50	<0.001	-0.14	<0.001	298	<0.001
Subject A4	-0.37	<0.001	-0.03	0.056	158	<0.001
Subject A5	-0.36	<0.001	-0.16	<0.001	133	<0.001

Table 6: Results of the regression analysis performed for the subjects excluded under the criteria ($RT > 9s$ in 60% of the trials) testing the model $DC = \beta_0 + \beta_1 S + \beta_2 T$ where S corresponds to stimulus strength or difficulty (Easy, Hard, Medium) and T to RT. The results indicate that for these subjects the predictor S (Difficulty) was still significant for every subject for a confidence interval of 5%. Surprisingly, the same was verified for the predictor T (RT) with the exception of subject A4. In addition, by comparing the values of β_2 with the ones found for the subjects with smaller average RTs, we notice that in general these values are smaller in magnitude, which reveals a weaker relationship between DC and RT for these subjects, however, still significant. The last 2 columns of the table, shows us as well that the model was significant for every subject reassuring the quality of the results.

	Movement Speed **Linear Regression Analysis** **MS = $\boldsymbol{\beta_0 + \beta_1 DC}$**			
	Decision Confidence		*Model*	
	β_1	*p value (H0: β_1=0)*	*F*	*p value*
Subject 11	4.42	<0.001	1.11E3	<0.001
Subject 12	11.23	<0.001	2.81E3	<0.001
Subject 13	5.41	<0.001	1.50E3	<0.001
Subject 14	5.25	<0.001	2.13E3	<0.001
Subject 15	4.25	<0.001	2.96E3	<0.001
Subject 16	5.85	<0.001	1.35E3	<0.001
Subject 17	4.73	<0.001	2.46E3	<0.001
Subject 18	4.16	<0.001	1.85E3	<0.001
Subject 22	3.85	<0.001	3.62E3	<0.001
Subject 23	3.47	<0.001	1.43E3	<0.001
Subject 24	2.29	<0.001	2.49E3	<0.001
Subject 25	4.72	<0.001	1.57E3	<0.001
Subject 26	5.57	<0.001	1.43E3	<0.001
Subject 27	4.22	<0.001	2.23E3	<0.001
Subject 28	4.15	<0.001	1.02E3	<0.001
Subject 29	5.05	<0.001	672	<0.001

Table 7: Results of the regression analysis performed for all subjects testing the model $MS = \beta_0 + \beta_1 DC + \beta_2 T$. The results indicate that for all subjects DC was a significant predictor of MS with an increase of DC leading to an increase of MS.The last 2 columns of the table, shows us as well that the model was significant for every subject reassuring the quality of the results for a confidence interval of 5%.

	Movement Time ANOVA		
	Confidence		
	F	*d.f*	*p value*
Subject 11	70.18	2	<0.001
Subject 12	40.63	2	<0.001
Subject 13	37.04	2	< 0.001
Subject 14	21.46	2	< 0.001
Subject 15	47.56	2	< 0.001
Subject 16	64.35	2	< 0.001
Subject 17	30.89	2	< 0.001
Subject 18	66.09	2	< 0.001
Subject 21	56.50	2	< 0.001
Subject 22	90.22	2	<0.001
Subject 23	21.46	2	<0.001
Subject 24	2.48	2	0.092
Subject 25	43.11	2	<0.001
Subject 26	10.51	2	<0.001
Subject 27	79.67	2	<0.001
Subject 28	25.80	2	<0.001
Subject 29	25.77	2	<0.001
Sample	26.49	2	< 0.001

Table 8: Results for one-way ANOVA testing differences in mean movement times across self-reports of DC. The effect of DC is significant for all subjects (exception subject 24) and also significant at the group level for a confidence interval of 5%.

	Movement Time ANOVA		
	Key Position		
	F	*d.f*	*p value*
Subject T1	6.57	2	0.017
Subject T2	11.59	2	0.003
Subject T3	0.84	2	0.461
Subject T4	4.79	2	0.038
Sample	3.24	2	0.087

Table 9: Results for one-way ANOVA testing differences in mean movement time across different levels of key position. The effect of key distance is significant for all subjects (exception subject T3) and it is not significant at the group level for a confidence interval of 5%.

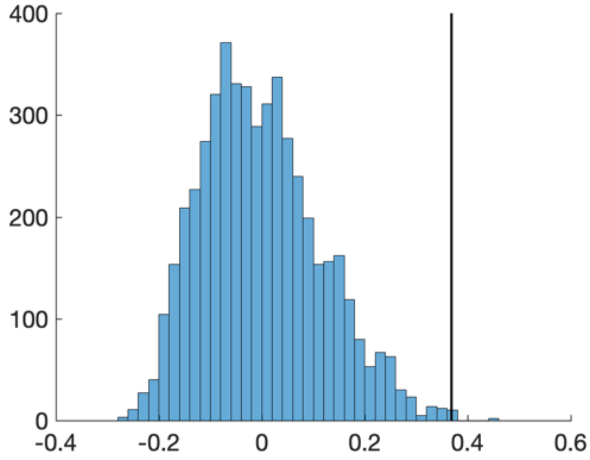

Figure 1: Distribution of the difference between slopes generated by permuting the z-scored MS as a function of confidence/button values. The black line corresponds to the observed difference between the slopes fitted to the across-subject means of MS as a function of confidence/key levels to the subjects of the main experiment and control experiment.

	Movement Speed ANOVA		
	Key Position		
	F	*d.f*	*p value*
Subject T1	23.39	2	0.001
Subject T2	10.89	2	0.004
Subject T3	1.23	2	0.337
Subject T4	4.36	2	0.048
Sample	3.13	2	0.093

Table 10: Results for one-way ANOVA testing differences in mean movement speed across different levels of key position. The effect of key distance is significant for all subjects (exception subject T3) and it is not significant at the group level for a confidence interval of 5%.

	Average Press Rate ANOVA			
	Confidence			
	F	*d.f*	*p value*	*β*
Subject 11	7.76	2	0.001	0.44
Subject 12	8.52	2	0.001	0.50
Subject 13	8.18	2	0.001	0.37
Subject 14	2.87	2	0.065	-0.18
Subject 15	0.58	2	0.565	-0.14
Subject 16	4.38	2	0.017	0.00
Subject 17	2.91	2	0.062	-0.35
Subject 18	0.71	2	0.494	-0.20
Subject 21	5.68	2	0.006	-0.29
Subject 22	9	2	0.001	0.29
Subject 23	26.86	2	<0.001	-0.29
Subject 24	11.54	2	<0.001	0.37
Subject 25	1.99	2	0.146	-0.22
Subject 26	74.42	2	<0.001	0.84
Subject 27	51.33	2	<0.001	0.74
Subject 28	0.18	2	0.837	0.04
Subject 29	16.88	2	< 0.001	0.43
Group	**3.25**	**2**	**0.047**	**0.14**

Table 11: Results for one-way ANOVA testing differences in mean movement speed across different levels of self-reported DC. The effect is significant for 15 subjects, with 9 subjects showing a significant increase $\beta > 0$ and 2 subject showing a significant decrease $\beta < 0$. There is one particular subject, with a significant difference of pressing rate, when reporting confidence level 2, exclusively (gray row). However, this subject presents a similar press rate for confidence level 1 and 3, that is why the associated β value for the relationship between average press rate and confidence is 0. For this reason this particular subject was discarded from any conclusions regarding the relationship between these 2 variables.

DC and MS as a function of strength of evidence - G1

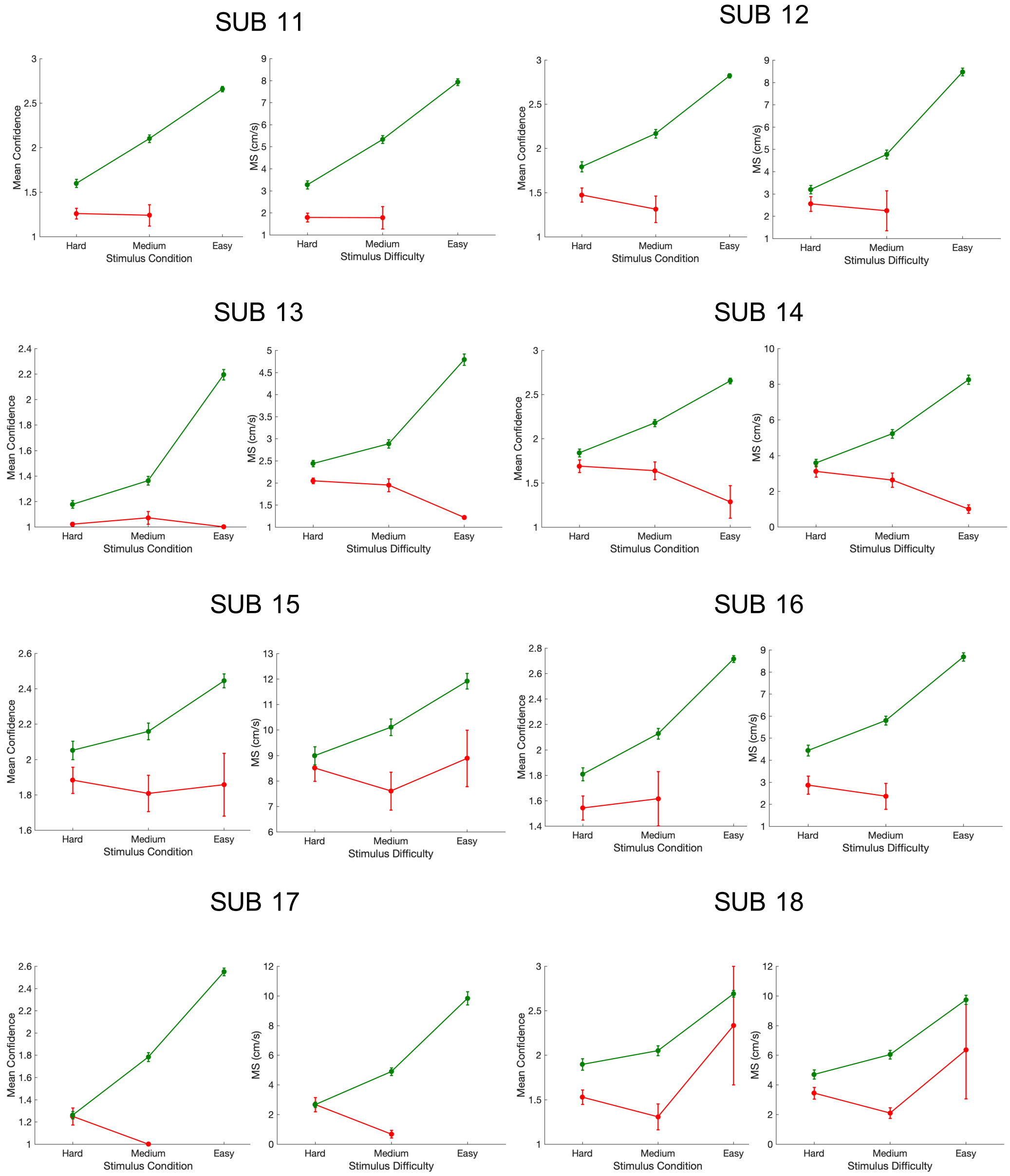

Figure 2: DC and MS as a function of stimulus difficulty in incorrect and correct trials for every subject of Group 1. The patterns of DC and MS are identical for every subject, showing that MS mirrors the behavior of DC, even though not all subjects present the expected X patter (Sanders et al., 2016). The error bars represent standard error across sessions.

DC and MS as a function of strength of evidence - G2

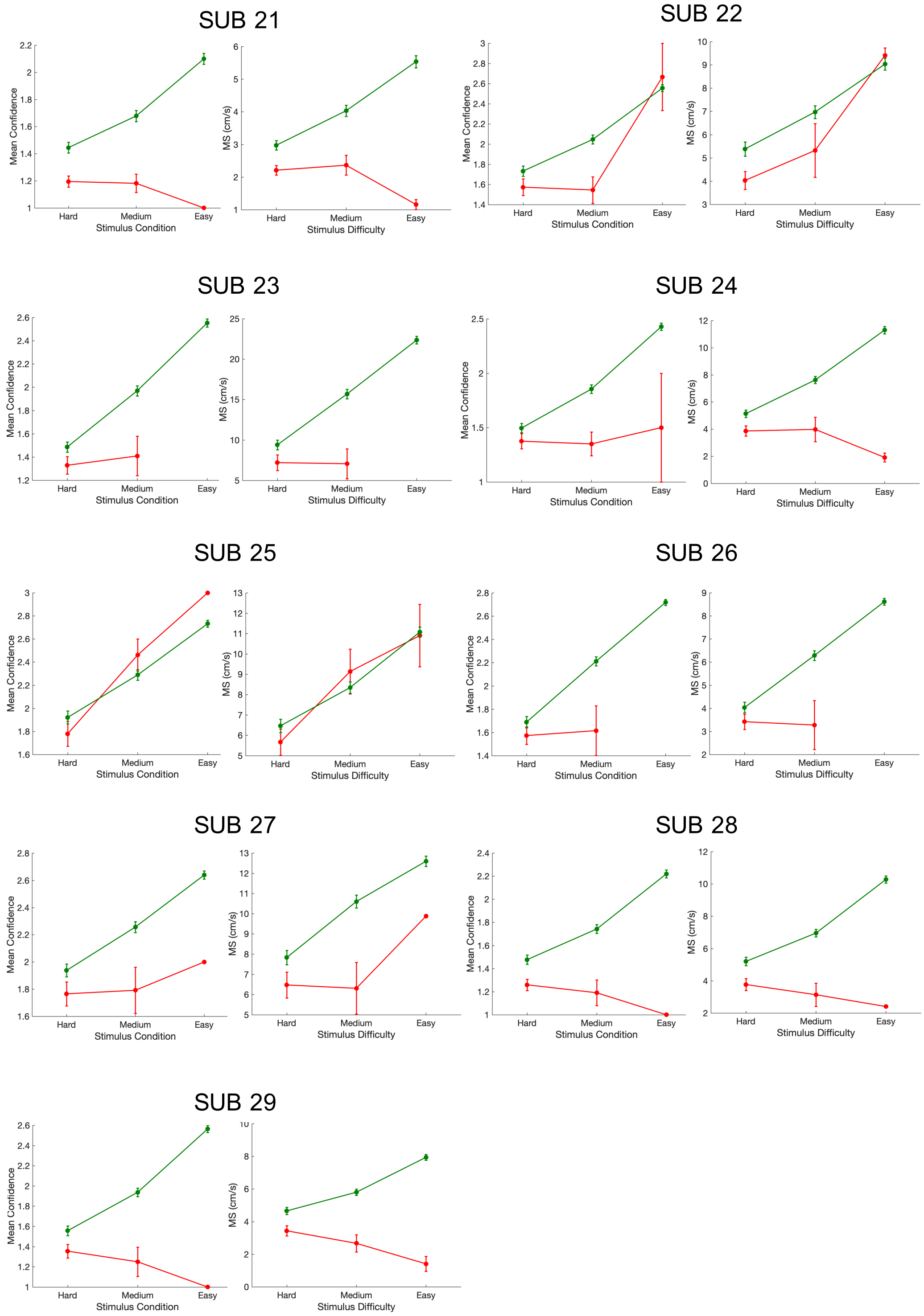

Figure 3: DC and MS as a function of stimulus difficulty in correct and incorrect trials for every subject of Group 2. The patterns of DC and MS are identical for every subject, showing that MS mirrors the behavior of DC, even though not all subjects present the expected X patter (Sanders et al., 2016). The error bars represent standard error across sessions.

Response Press Rate - G1

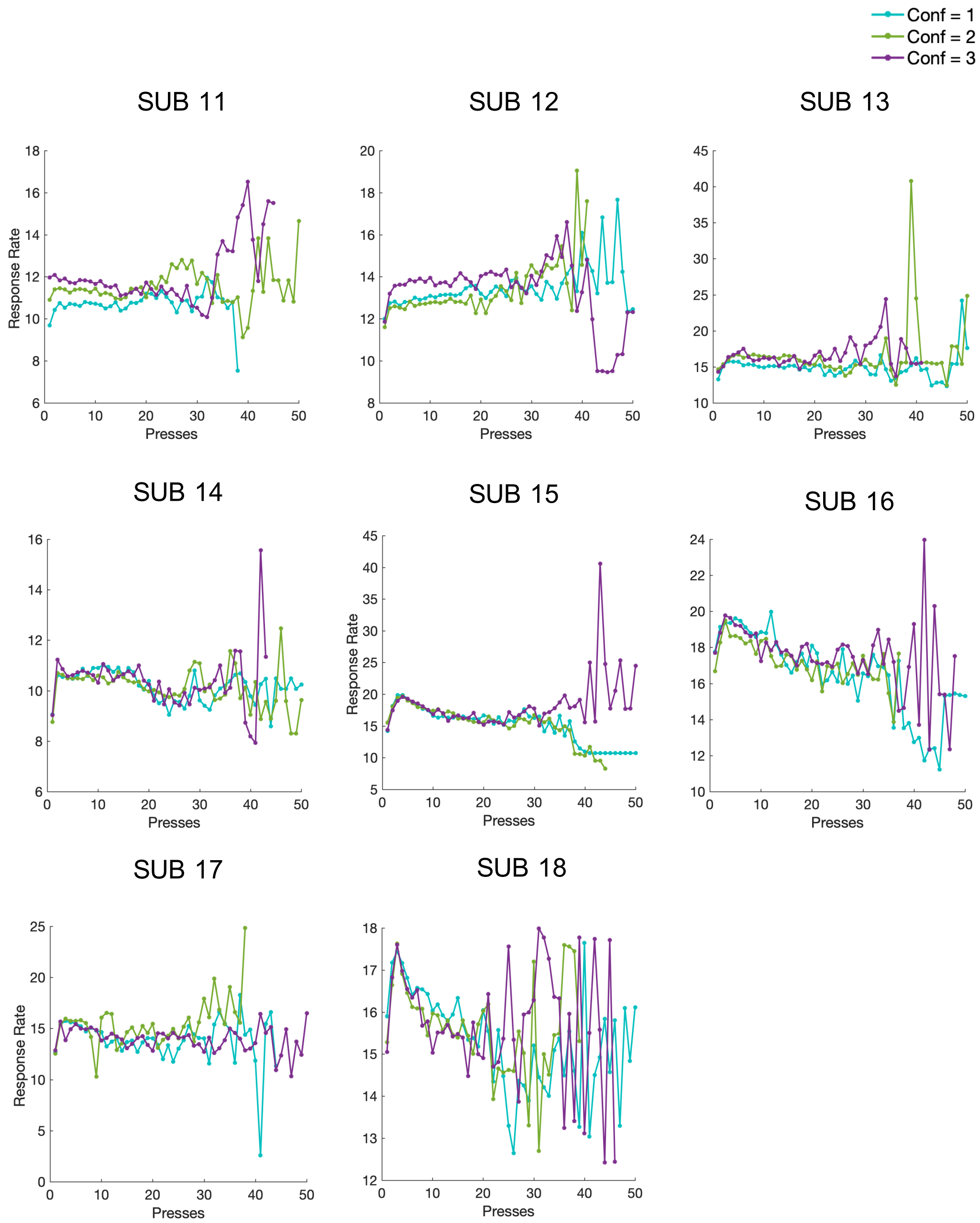

Figure 4: Evolution of response rate with the number of presses in the response key according to the self-reported level of confidence, for each subject of group 1.

Response Press Rate - G2

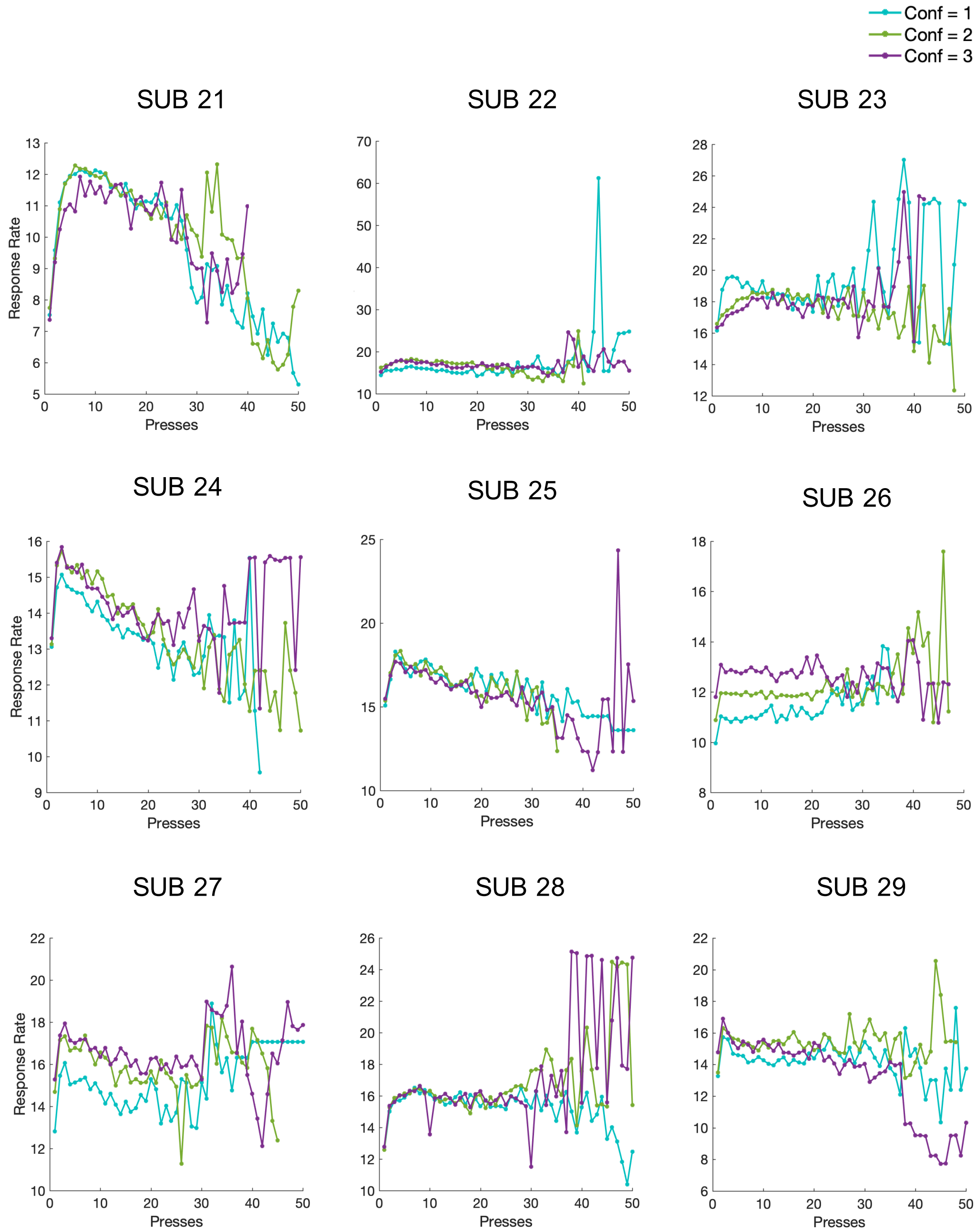

Figure 5: Evolution of response rate with the number of presses in the response key, according to the self-reported level of confidence,for each subject of group 2

MS as function of reward category - G1

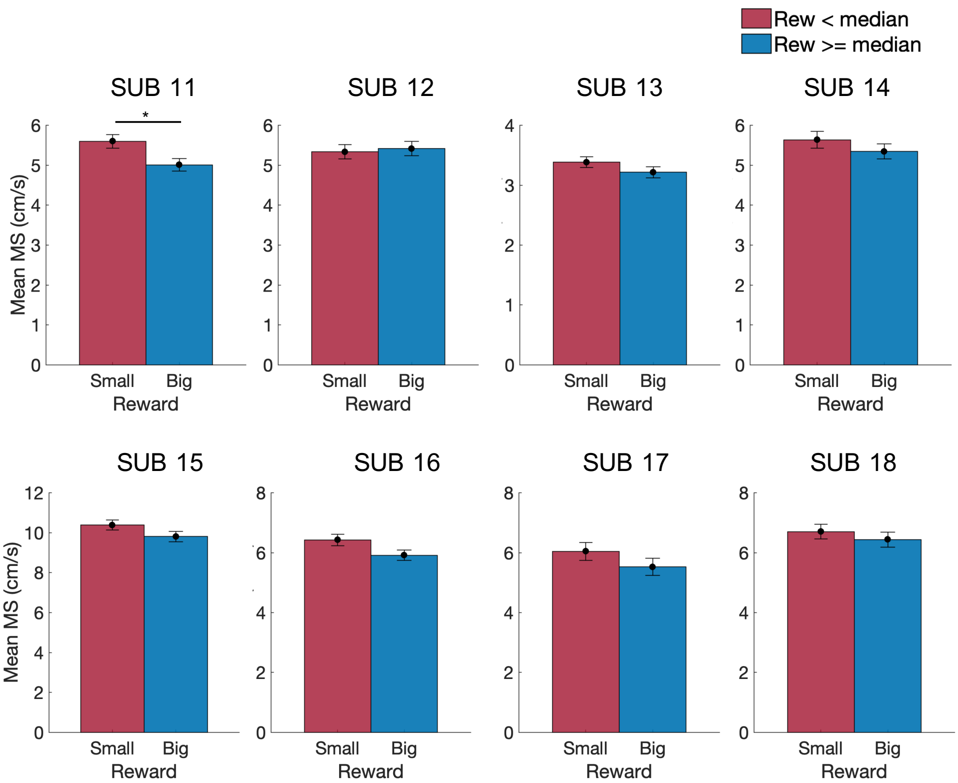

Figure 6: Mean movement speed in cm/s as function of the two categories of reward - $Reward < median$ (red) and $Reward >= median$ (blue) - for each subject of Group 1. Being Group 1 the group of subjects who performed the task under a point accumulating system. Only subject 11 revealed a significant difference between the means of MS for the 2 categories of reward, for a level of confidence of 5% (signalized with an $*$), Table 12. The errorbars represent standard error across trials.

	Movement Speed			
	Reward			
	Mean $\pm$ *SD*		*p value*	*df*
	Small	*High*		
Subject 11	**5.59 $\pm$ 0.78**	**5.01 $\pm$ 0.93**	**0.012**	38
Subject 12	5.33 $\pm$ 1.51	5.41 $\pm$ 1.32	0.706	38
Subject 13	3.39 $\pm$1.01	3.22 $\pm$ 1.10	0.636	N = 20
Subject 14	5.64 $\pm$ 1.58	5.35 $\pm$ 1.34	0.438	38
Subject 15	10.39 $\pm$1.49	9.80 $\pm$ 1.67	0.156	38
Subject 16	642 $\pm$ 1.62	5.91 $\pm$ 1.12	0.343	38
Subject 17	6.04 $\pm$ 2.84	5.53 $\pm$ 1.88	0.344	N = 20
Subject 18	6.70 $\pm$ 1.61	6.44 $\pm$ 1.86	0.997	38
Sample	6.22 $\pm$ 1.86	5.89 $\pm$ 1.69	0.703	14

Table 12: Results of the parametric test t-test testing differences in mean MS (cm/s) across the two categories of reward- Small: $Reward < median$ and High: $Reward >= median$, for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of MS across subjects, that is, the results for the group level.

MS as function of DC for the two reward categories - G1

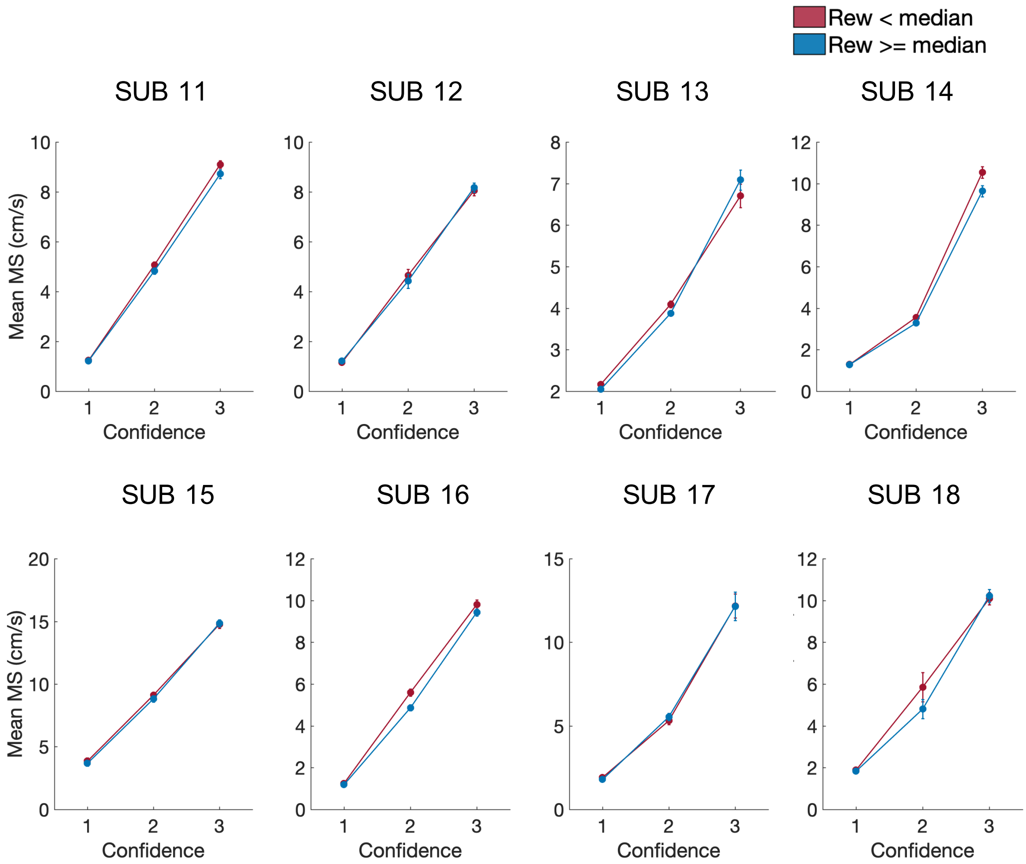

Figure 7: Mean movement speed in cm/s as function of the self-reported confidence level for the two categories of reward - $Reward < median$ (red) and $Reward >= median$ (blue) - for each subject of Group 1. The errorbars represent standard error across sessions. The statistical analysis performed with 2-way ANOVA for every subject showed ($p < 0.001$) for the columns effect (DC) and $p > 0.05$ and $p > 0.05$ for the row effect end for every interaction term, showing the observed results are likely, given that there is no interaction. This results highlights that the DC effect on MS is very significant, contrarily to the effect of reward.

MS as function of reward category - G2

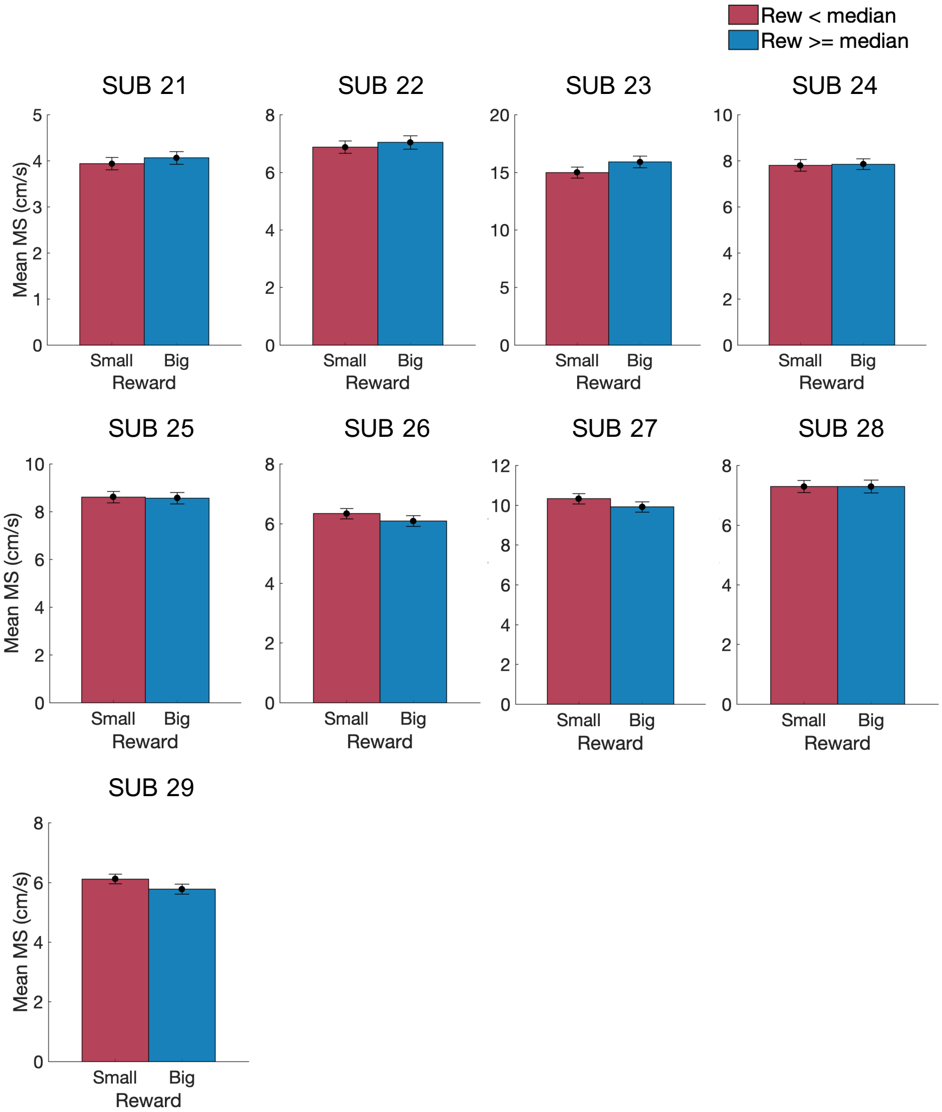

Figure 8: Mean movement speed, in cm/s as function of the two categories of reward - $Reward < median$ (red) and $Reward >= median$ (blue) - for each subject of Group 2. Being Group 2 the group of subjects who performed the task under a monetary discounting reward system. No subject revealed a significant difference between the means of MS for the 2 categories of reward, for a level of confidence of 5%, Table 13. The errorbars represent standard error across sessions.

	Movement Speed			
	Reward			
	Mean ± *SD*		*p value*	*df*
	Small	*High*		
Subject 21	3.94 ± 1.18	4.06 ± 0.74	0.798	38
Subject 22	6.88 ± 2.29	7.04 ± 1.87	0.801	38
Subject 23	14.52 ±3.91	15.52 ± 3.86	0.205	38
Subject 24	7.85 ± 2.34	7.92 ± 1.93	0.848	38
Subject 25	8.58 ±1.67	8.44 ± 2.12	0.642	38
Subject 26	6.45 ±1.44	6.14 ± 1.24	0.155	38
Subject 27	10.16 ± 2.92	9.90 ± 2.47	0.387	38
Subject 28	7.33 ± 1.21	7.31 ± 1.21	0.970	38
Subject 29	6.13 ± 1.25	5.88 ± 1.29	0.245	38
Sample	7.77 ± 3.25	7.95 ± 4.05	0.964	16

Table 13: Results of the parametric test t-test testing differences in mean MS (cm/s) across the two categories of reward- Small: $Reward < median$ and High: $Reward >= median$, for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The sample translates the results of the t-test applied to the mean values of MS across subjects, that is, the results for the group level.

MS as function of DC for the two reward categories - G2

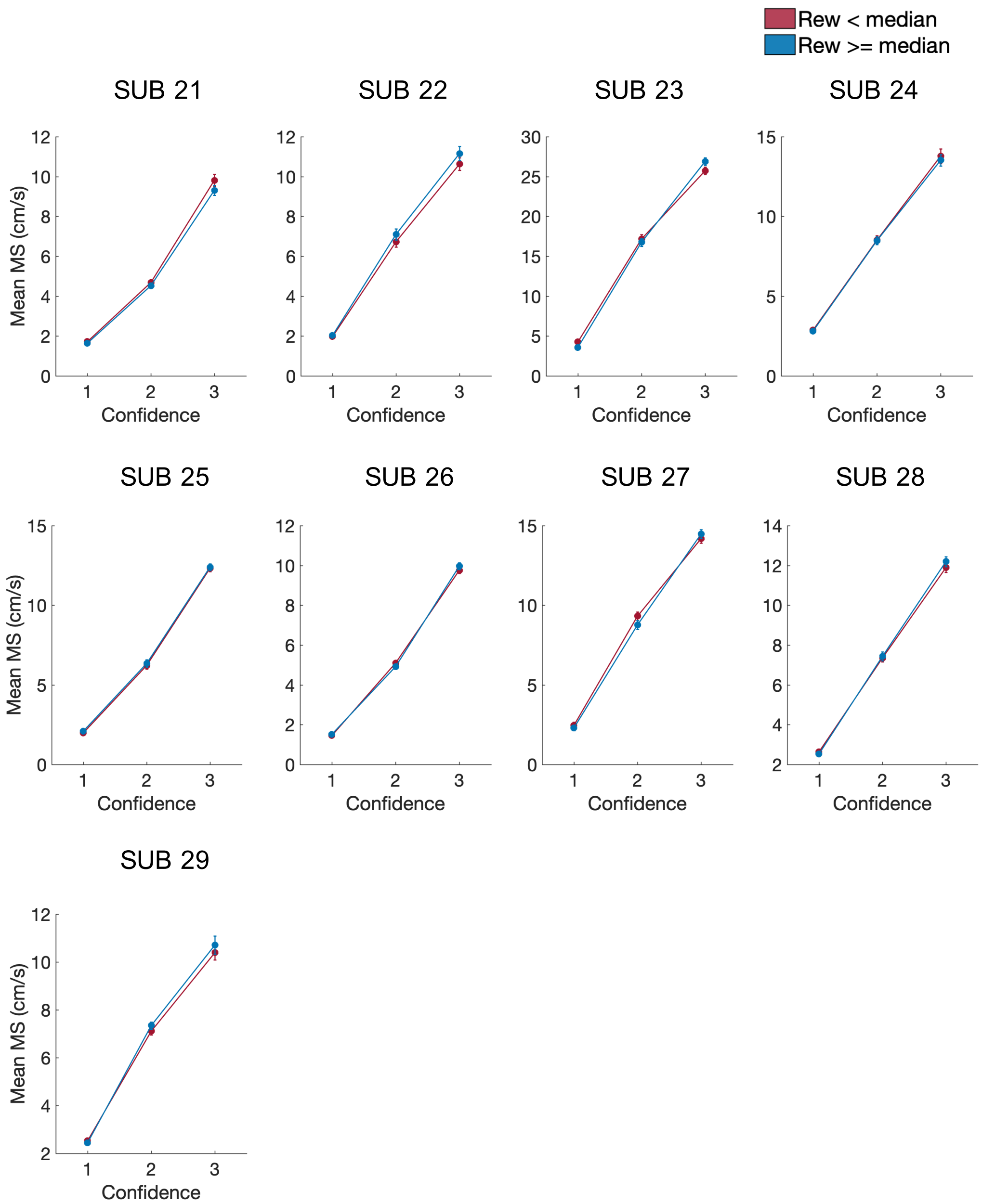

Figure 9: Mean movement speed, in cm/s, as function of the self-reported confidence level for the two categories of reward - $Reward < median$ (red) and $Reward > median$ (blue) - for each subject of Group 2. The errorbars represent standard error across sessions. The statistical analysis performed with 2-way ANOVA for every subject showed ($p < 0.001$) for the columns effect (DC) and $p > 0.05$ and $p > 0.05$ for the row effect end for every interaction term, showing the observed results are likely, given that there is no interaction. This results highlights that the DC effect on MS is very significant, contrarily to the effect of reward.

MS as function of 'extreme' reward category - G1

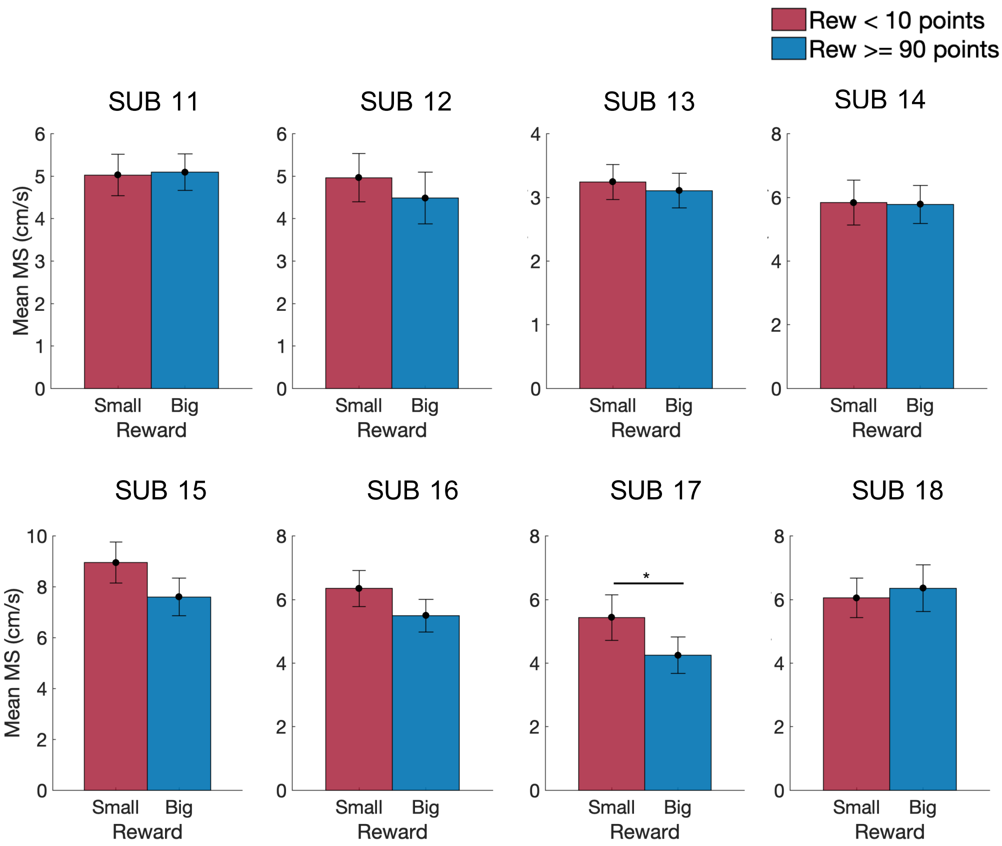

Figure 10: Mean movement speed, in cm/s as function of the two categories of reward - $Reward < 10points$ (red) and $Reward >= 90points$ (blue)- for each subject of Group 1. Quite surprisingly the subject who presented a significant difference between means of MS for the non-extreme categories of reward, did not present a significant difference for these categories of reward. With only subject 17 showing a significant difference on their MS data for the two categories of reward at a significant level of 5%, Table 12 . Once again, this analysis revealed no significant effect of reward for the majority of the subjects. The errorbars represent standard error across sessions.

	Movement Speed			
	Extreme Reward			
	$Mean \pm SD$		*p value*	*d.f*
	Small (< 10 points)	*High (>= 90 points)*		
Subject 11	5.28 $\pm$ 1.98	5.35 $\pm$ 1.63	0.891	38
Subject 12	5.21 $\pm$2.42	4.71 $\pm$ 2.65	0.543	38
Subject 13	3.41 $\pm$1.04	3.26 $\pm$ 1.05	0.563	38
Subject 14	6.13 $\pm$3.00	6.06 $\pm$ 2.44	0.939	38
Subject 15	9.40 $\pm$2.31	7.98 $\pm$ 2.39	0.068	38
Subject 16	6.67 $\pm$ 2.21	5.77 $\pm$ 2.04	0.171	38
Subject 17	**5.71 $\pm$ 3.09**	**4.46 $\pm$ 2.50**	**0.037**	**38**
Subject 18	6.35 $\pm$ 2.54	6.67 $\pm$ 3.11	0.703	38
Sample	6.09 $\pm$ 1.86	5.60 $\pm$ 1.59	0.583	14

Table 14: Results of the parametric test t-test testing differences in mean MS (cm/s) across the two categories of reward- Small: $Reward < 10points$ and High: $Reward >= 90points$, for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The sample translates the results of the t-test applied to the mean values of MS across subjects, that is, the results for the group level.

MS as function of 'extreme' reward category - G2

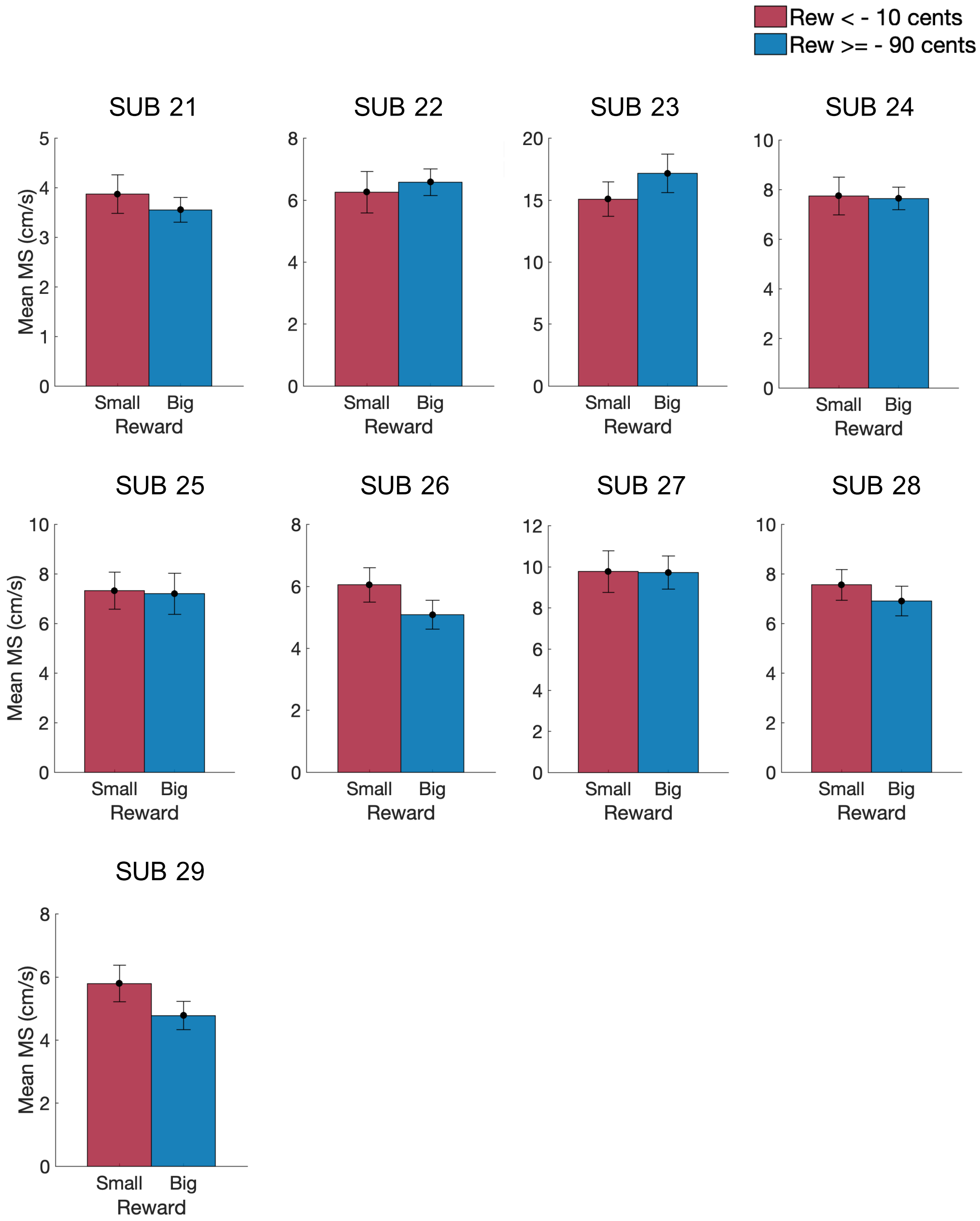

Figure 11: Mean movement speed, in cm/s, as function of the two categories of reward, divided in $Reward < -10cents$ (red) and $Reward > -90cents$ (blue) for each subject of Group 2. No subject revealed a significant difference of the average MS associated to the two categories of reward, Table 15. Errobars represent standard errors across sessions.

	Movement Speed			
	Extreme Reward			
	$Mean \pm SD$		p value	$d.f$
	Small (< - 10 cents)	*High (>= -90 cents)*		
Subject 21	3.87 ± 1.77	3.55 ± 1.14	0.445	38
Subject 22	6.26 ± 3.06	6.58 ± 1.95	0.620	38
Subject 23	15.43 ± 5.66	17.59 ± 6.27	0.239	38
Subject 24	7.75 ± 3.47	7.65 ± 2.09	0.776	38
Subject 25	7.45 ± 3.19	7.65 ± 3.44	0.857	38
Subject 26	6.17 ± 2.29	5.53 ± 1.99	0.331	38
Subject 27	10.17 ± 4.42	9.89 ± 91	0.756	38
Subject 28	7.68 ± 2.57	7.34 ± 2.28	0.285	38
Subject 29	5.91 ± 2.44	7.81 ± 1.95	0.199	38
Sample	8.00 ± 2.98	8.00 ± 3.27	0.963	16

Table 15: Results of the parametric test t-test testing differences in mean MS (cm/s) across the two categories of reward - Small: $Reward < 10points$ and High: $Reward >= 90points$ - for subjects of group 2. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The sample translates the results of the t-test applied to the mean values of MS across subjects, that is, the results for the group level.

MS as function of average reward (low and high categories) - G1

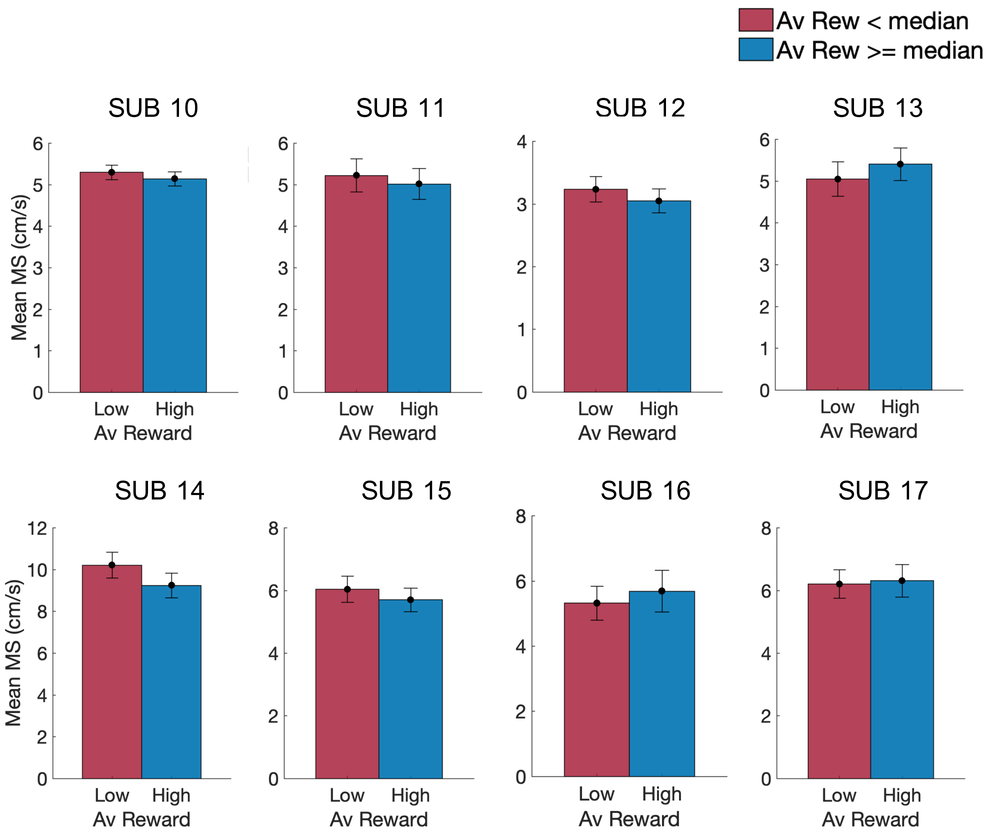

Figure 12: Mean movement speed, in cm/s, as function of low ($AvRew < median$) - red - and high ($AvRew > median$) average reward rate. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t - 1) + alpha \times Rew$, where $alpha$ was selected individually for each subject as the value which offered the biggest R^2 for the prediction of MS performed for each subject individually. The means of MS for high and low average reward were not significantly different for any of the subjects of Group 1 at the confidence level of 5%, Table 16. Errobars represent standard error across sessions.

	Movement Speed			
	Average Reward			
	Mean $\pm$ *SD*		*p value*	*d.f*
	Low	*High*		
Subject 21	5.38 ± 0.82	5.21 ± 0.65	0.792	38
Subject 22	5.48 ± 1.35	5.27 ± 1.37	0.989	n=20
Subject 23	3.40 ± 0.60	3.20 ± 0.51	0.336	38
Subject 24	5.30 ± 1.50	5.67 ± 1.51	0.595	38
Subject 25	10.10 ± 1.41	9.07 ± 1.46	0.864	N=20
Subject 26	6.34 ± 1.05	5.99 ± 1.15	0.817	38
Subject 27	5.59 ± 3.07	5.97 ± 1.80	0.091	N=20
Subject 28	6.52 ± 1.55	6.63 ± 1.73	0.198	N=20
Sample	6.01 ± 1.89	5.87 ± 1.64	0.816	14

Table 16: Results of the parametric test t-test testing differences in mean MS (cm/s) across the two categories of average reward - Low: Av Reward < median and High: Av Reward >= median - for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of MS across subjects, that is, the results for the group level.

MS as function of average reward (low and high categories) - G2

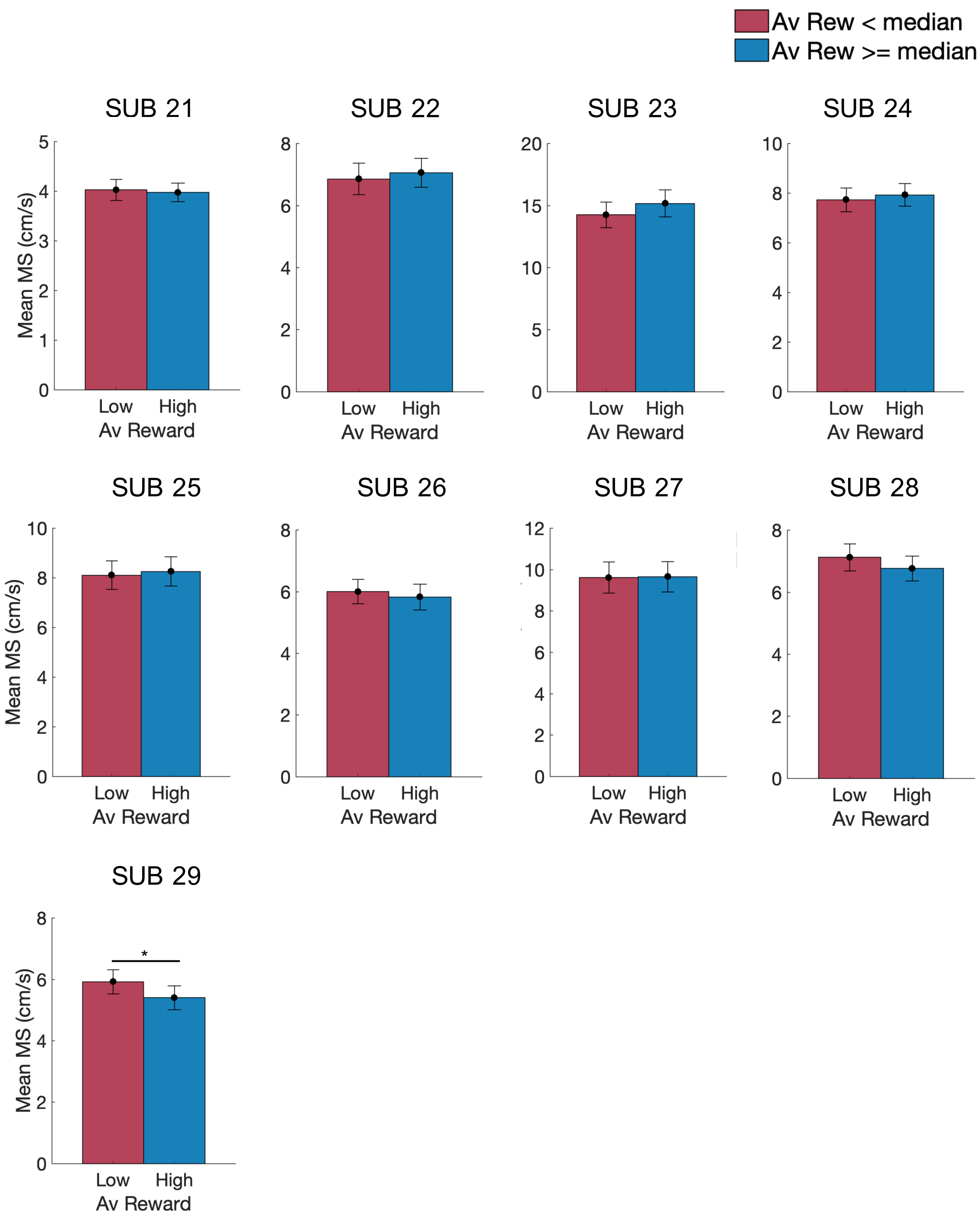

Figure 13: Mean movement speed, in cm/s as function of low ($AvRew < median$) - red - and high ($AvRew > median$) average reward rate. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t - 1) + alpha \times Rew$, where $alpha$ was selected individually for each subject as the value which offered the biggest R^2 for the prediction of MS performed for each subject individually. The means of MS for high and low average reward were only significantly different for subject 29 at the confidence level of 5%. Errorbars represent standard error across sessions.

	Movement Speed			
	Average Reward			
	$Mean \pm SD$		p value	$d.f$
	Low	*High*		
Subject 21	4.02 ± 1.26	3.98 ± 0.90	0.893	38
Subject 22	6.86 ± 2.61	7.06 ± 2.00	0.214	38
Subject 23	14.26 ± 4.16	15.18 ± 2.67	0.180	38
Subject 24	7.73 ± 2.65	7.93 ± 1.17	0.551	38
Subject 25	8.11 ± 1.42	8.26 ± 1.91	0.230	38
Subject 26	6.01 ± 1.70	5.83 ± 1.73	0.927	N = 20
Subject 27	9.62 ± 2.92	9.65 ± 2.68	0.279	N = 20
Subject 28	7.13 ± 1.16	6.76 ± 0.59	0.099	38
Subject 29	**5.92 ± 1.45**	**5.40 ± 1.29**	**0.025**	**38**
Sample	7.74 ± 3.17	7.78 ± 3.09	0.455	16

Table 17: Results of the parametric test t-test testing differences in mean MS (cm/s) across the two categories of average reward - Low: Av Reward $<$ median and High: Av Reward $>=$ median - for subjects of group 2. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of MS across subjects, that is, the results for the group level.

RT as function of reward category - G1

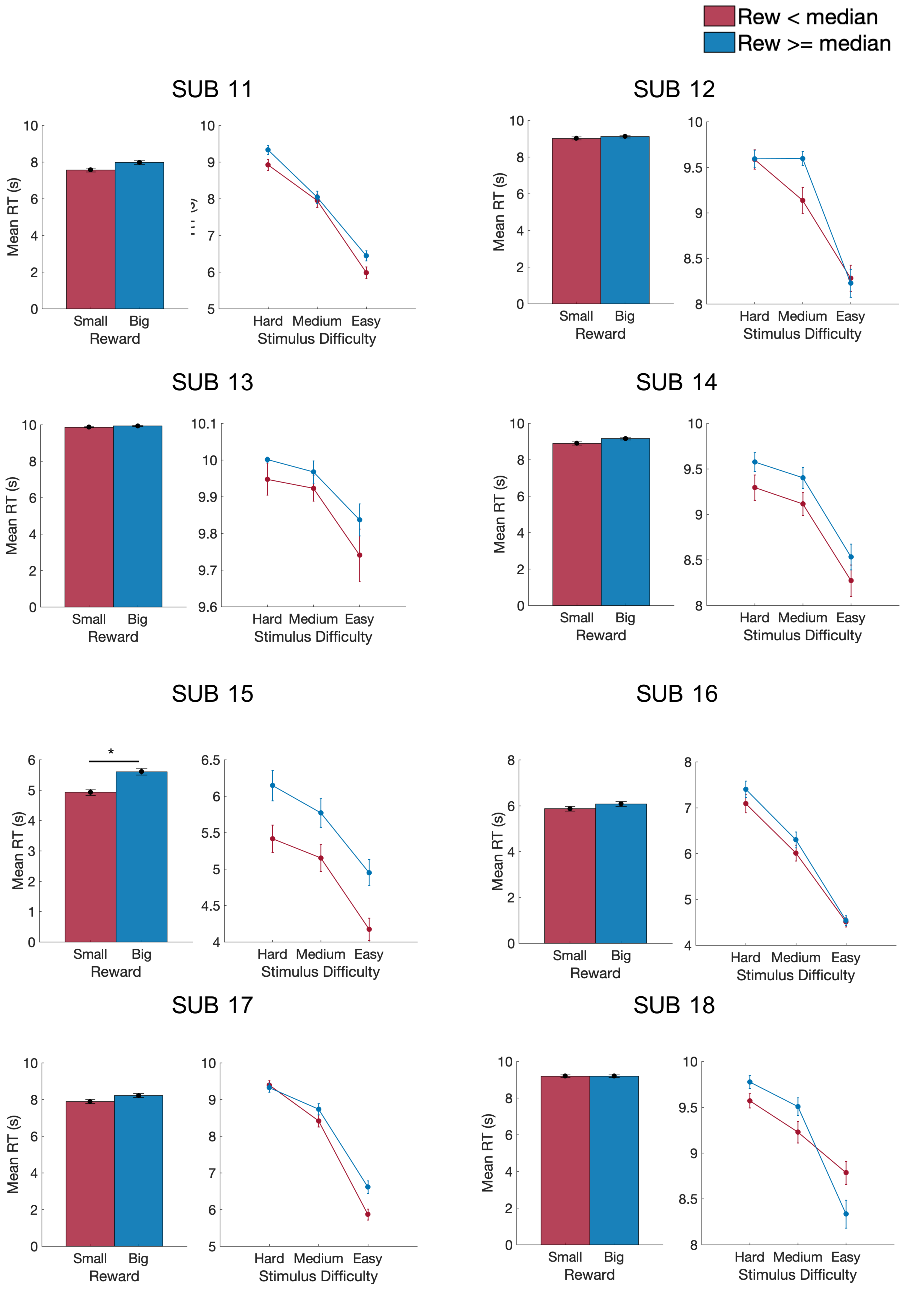

Figure 14: Mean RT, in s as function of the two categories of reward - $Reward < median$ (red) and $Reward > median$ (blue) - for each subject of Group 1. Being Group 1 the group of subjects who performed the task under a point accumulating system. Only one subject presented a significant difference between the means of both groups of rewards (subject 15), although it is possible to observe a general tendency of listen the stimulus for longer when the stakes are higher. No subject presented a significant difference between mean of reward groups. The errorbars represent standard error across sessions.

RT as function of reward category - G2

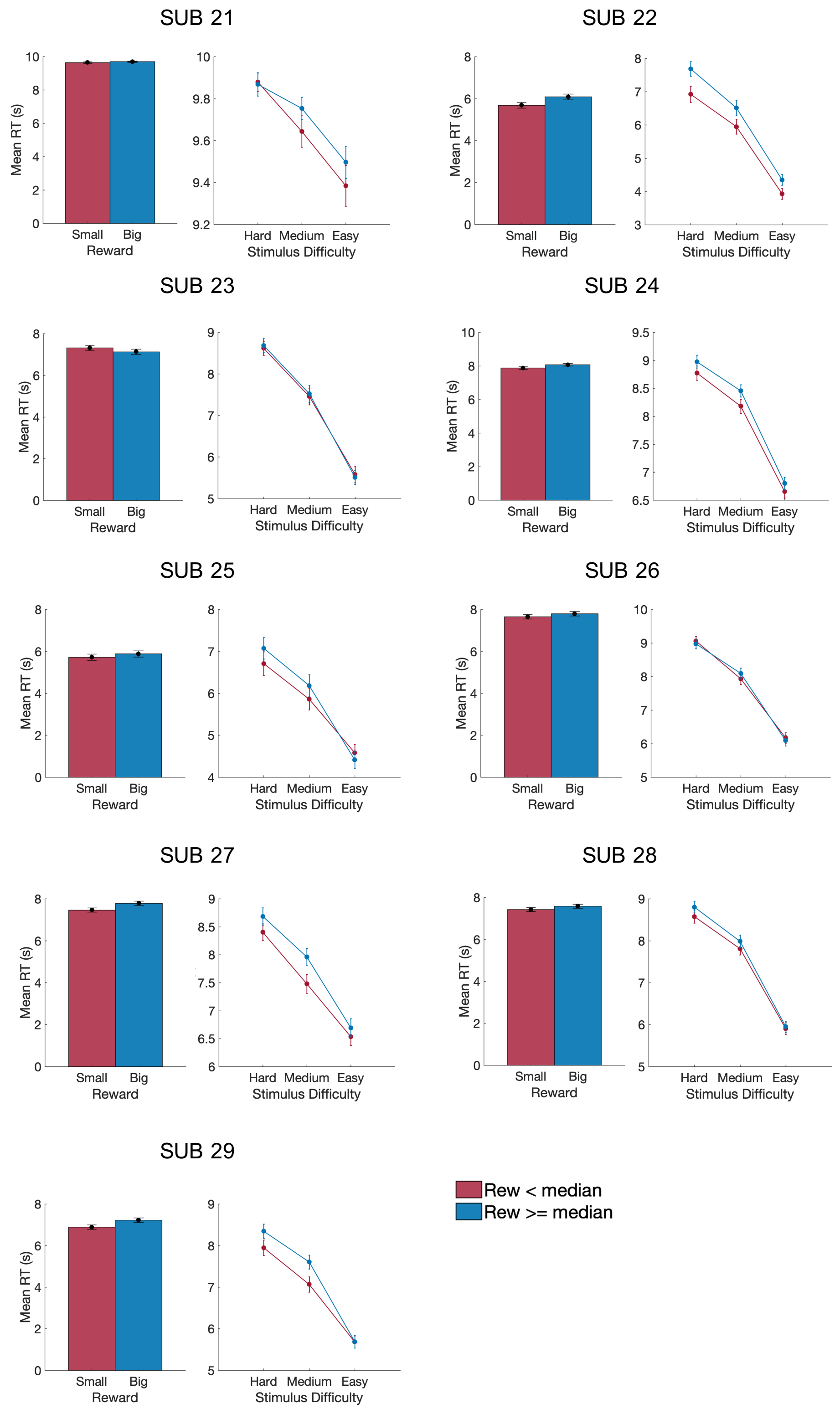

Figure 15: Mean RT, in s, as function of the two categories of reward, divided in $Reward < median$ (red) and $Reward > median$ (blue)- for each subject of Group 2. Being Group 2 the group of subjects who performed the task under a monetary discounting reward system. Even though no subject showed a significant difference between the means of RT for both groups of reward, it is observed a tendency of waiting longer when there is more at 'stake'. The errorbars represent standard error across sessions.

	Reaction Time			
	Reward			
	Mean $\pm$ *SD*		*p value*	*df*
	Small	*High*		
Subject 11	7.57 ± 0.79	7.99 ± 0.48	0.076	N = 20
Subject 12	9.02 ± 0.97	9.12 ± 0.86	0.967	N = 20
Subject 13	9.87 ±0.22	9.94 ± 0.09	0.560	N = 20
Subject 14	8.90 ±0.74	9.16 ± 0.64	0.260	38
Subject 15	**4.93 ±1.50**	**5.60 ± 1.53**	**0.032**	38
Subject 16	5.87 ±0.74	6.08 ± 0.82	0.467	38
Subject 17	7.89 ± 0.80	8.22 ± 0.72	0.364	38
Subject 18	9.20 ± 0.68	9.19 ± 0.64	0.480	38
Sample	7.94 ± 1.68	8.17 ± 1.50	0.773	14

Table 18: Results of the parametric test t-test testing differences in mean RT across the two categories of reward - Small: $Reward < median$ and $Reward >= median$ - for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of RT across subjects, that is, the results for the group level.

	Reaction Time			
	Reward			
	Mean $\pm$ *SD*		*p value*	*df*
	Small	*High*		
Subject 21	9.64 ± 0.41	9.70 ± 0.24	0.627	38
Subject 22	5.65 ± 1.15	6.11 ± 0.87	0.149	38
Subject 23	7.27 ±1.05	7.13 ± 0.86	0.631	38
Subject 24	7.85 ±0.74	8.08 ± 0.60	0.281	38
Subject 25	5.87 ±2.05	5.98 ± 2.10	0.853	38
Subject 26	7.64 ±0.54	7.81 ± 0.62	0.323	38
Subject 27	7.51 ± 1.23	7.79 ± 1.14	0.453	38
Subject 28	7.46 ± 0.52	7.61 ± 0.38	0.262	38
Subject 29	6.96 ± 1.00	7.29 ± 0.83	0.251	38
Sample	7.31 ± 1.17	7.50 ± 1.10	0.735	16

Table 19: Results of the parametric test t-test testing differences in mean RT across the two categories of reward -$Reward < median$ and $Reward >= median$ - for subjects of group 2. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The sample translates the results of the t-test applied to the mean values of RT across subjects, that is, the results for the group level.

RT as function of 'extreme' reward category - G1

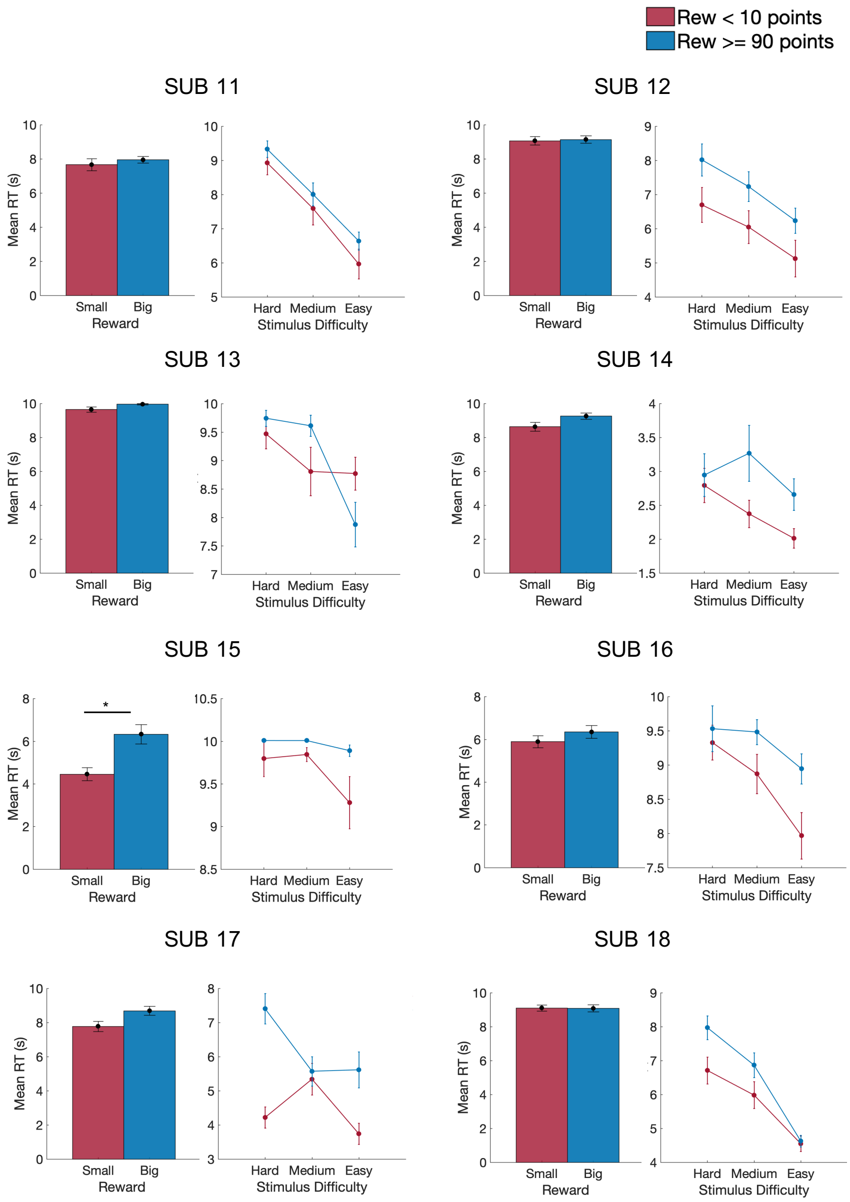

Figure 16: Mean RT, in s, as function of the two categories of reward, divided in $Reward < 10points$ (red) and $Reward >= 90points$ (blue) for each subject of Group 1. The only subject who showed significant differences between RTs for different categories of reward at the significance level of 5%, was the subject who had showed a significant difference for normal categories of reward. The tendency of longer RTs for higher amounts of reward on offer is, however, here more evident, even though not significant for more subjects. The errorbars represent standard error across sessions

RT as function of 'extreme' reward category - G2

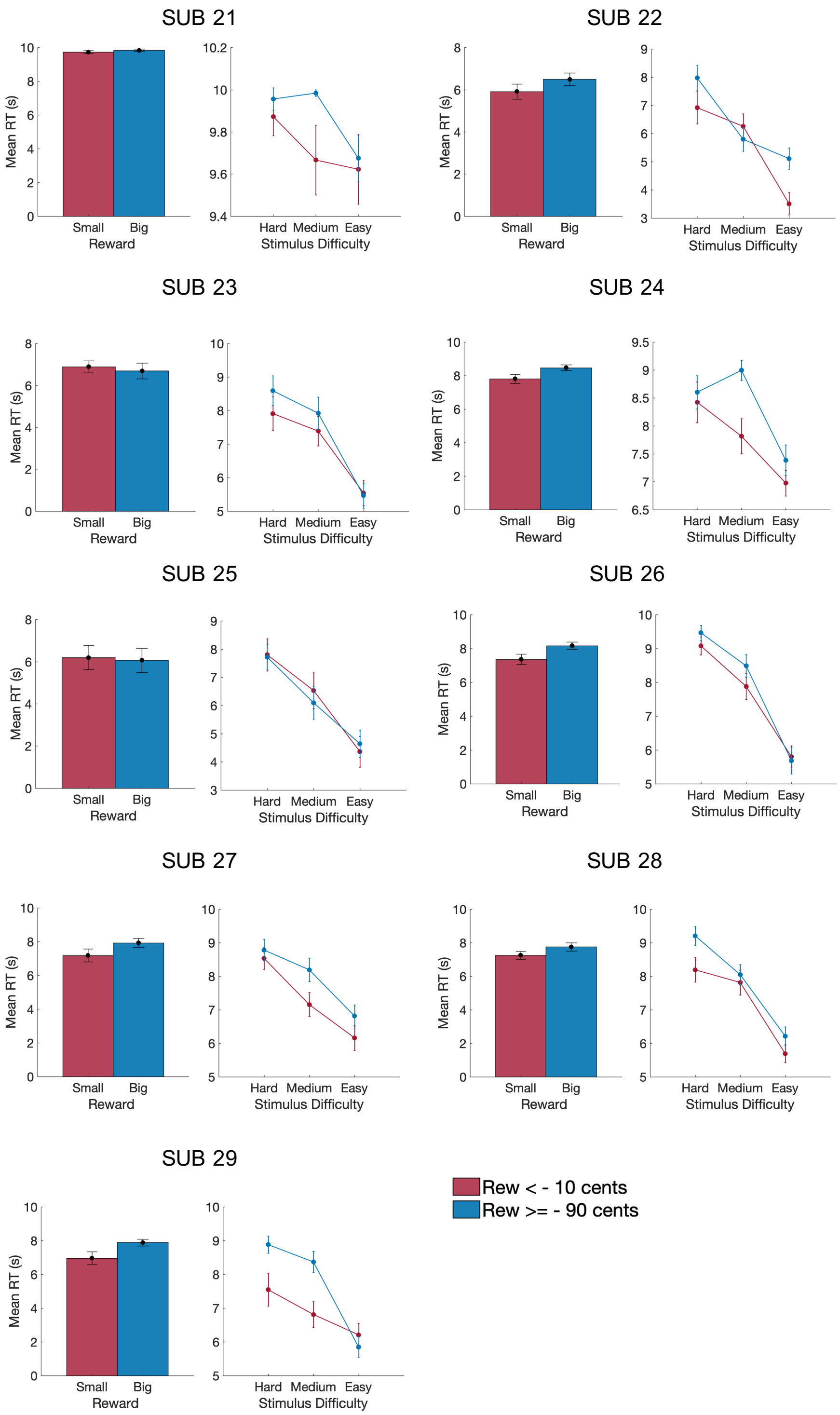

Figure 17: Mean RT, in s, as function of the two categories of reward, divided in $Reward < -10cents$ (red) and $Reward >= -90cents$ (blue) for each subject of Group 2. Surprisingly, no subject showed a significant effect of reward on RT at the significance level of 5%. However, it is also in this group observed a tendency of waiting more when the 'stakes' are higher. Once again, the errorbars represent standard error across sessions.

	Reaction Time			
	Extreme Reward			
	Mean $\pm SD$		*p value*	*df*
	Small (<10 points)	*High(>= 90 points)*		
Subject 11	7.67 $\pm$ 0.75	7.95 $\pm$ 1.06	0.073	38
Subject 12	9.65 $\pm$ 0.00	9.96 $\pm$ 0.00	0.076	N = 20
Subject 13	8.64 $\pm$0.73	9.26 $\pm$ 0.90	0.675	38
Subject 14	**4.46 $\pm$1.10**	**6.33 $\pm$ 0.97**	**0.046**	**N=20**
Subject 15	5.87 $\pm$2.05	5.98 $\pm$ 2.10	0.418	38
Subject 16	5.89 $\pm$0.47	6.35 $\pm$ 0.41	0.682	38
Subject 17	7.78 $\pm$ 0.94	8.69 $\pm$ 0.70	0.895	38
Subject 18	9.11 $\pm$ 1.27	9.09 $\pm$ 0.97	0.315	38
Sample	6.89 $\pm$ 2.36	6.89 $\pm$ 2.27	0.986	14

Table 20: Results of the parametric test t-test testing differences in mean RT across the two categories of extreme reward - Small: $Reward < 10points$ and High: $Reward >= 90points$ - for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of RT across subjects, that is, the results for the group level.

	Reaction Time			
	Extreme Reward			
	Mean $\pm SD$		*p value*	*df*
	Small (<-10 cents)	*High (>=-90 cents)*		
Subject 21	9.72 $\pm$ 0.51	9.84 $\pm$ 0.54	0.188	38
Subject 22	5.92 $\pm$1.31	6.150 $\pm$ 1.16	0.309	38
Subject 23	6.89 $\pm$0.83	6.69 $\pm$ 0.96	0.683	38
Subject 24	7.80 $\pm$0.73	8.47 $\pm$ 0.81	0.317	38
Subject 25	6.19 $\pm$1.56	6.07 $\pm$ 1.00	0.921	38
Subject 26	7.36 $\pm$ 1.08	8.16 $\pm$ 0.80	0.190	38
Subject 27	7.18 $\pm$ 1.23	7.92 $\pm$ 0.79	0.302	38
Subject 28	7.23 $\pm$ 1.20	7.76 $\pm$ 1.09	0.193	38
Subject 29	6.96 $\pm$ 0.94	7.89 $\pm$ 1.00	0.724	38
Sample	7.31 $\pm$ 1.17	7.50 $\pm$ 1.10	0.987	16

Table 21: Results of the parametric test t-test testing differences in mean RT across the two categories of extreme reward - Small: $Reward < -10cents$ and High: $Reward >= -90cents$ - for subjects of group 2. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of RT across subjects, that is, the results for the group level.

RT as function of average reward (low and high categories) - G1

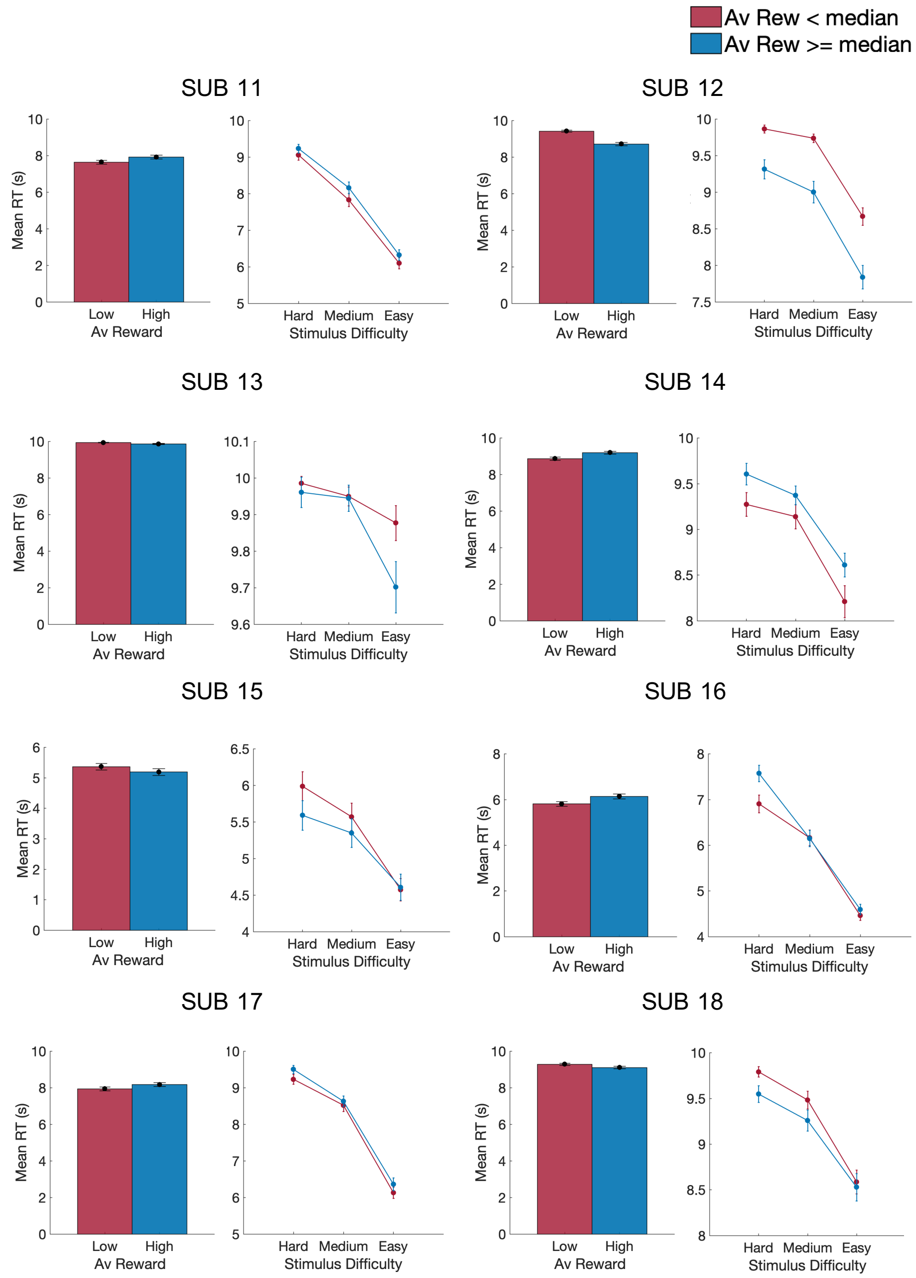

Figure 18: Mean RT, in s, as function of low ($AvRew < median$) - red - and high ($AvRew > median$) average reward rate for group 1. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t - 1) + alpha \times Rew$, where $alpha$ was selected individually for each subject as the value which offered the biggest R^2 for the prediction of RT performed for each subject individually. No subject showed a significant difference between the means of RT for high and low average reward at the confidence level of 5%. Errorbars represent standard error across sessions.

RT as function of average reward (low and high categories) - G2

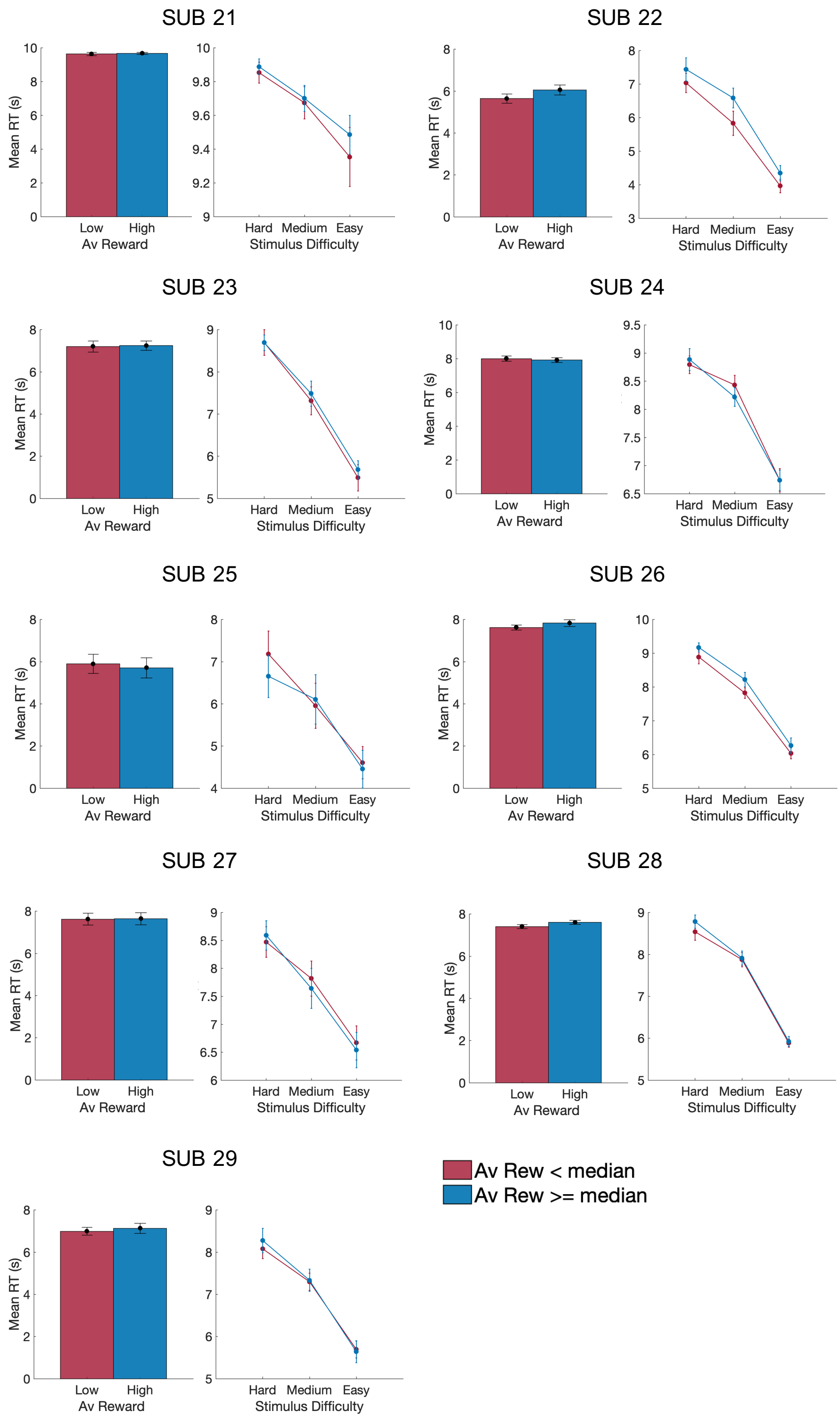

Figure 19: Mean RT, in s, as function of low ($AvRew < median$) - red - and high ($AvRew > median$) average reward rate for group 2. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t-1) + alpha \times Rew$, where $alpha$ was selected individually for each subject as the value which offered the biggest R^2 for the prediction of RT performed for each subject individually. The means of RT for high and low average reward were not significantly different for any subject of group 2 at the significance level of 5%. Errorbars represent standard error across sessions.

	Reaction Time			
	Average Reward			
	Mean $\pm$ *SD*		*p value*	*d.f*
	Low	*High*		
Subject 11	7.65 $\pm$ 0.66	7.91 $\pm$ 0.52	0.163	38
Subject 12	9.15 $\pm$ 0.82	8.98 $\pm$ 1.04	0.585	38
Subject 13	9.90 $\pm$ 0.13	9.91 $\pm$ 2.67	0.130	38
Subject 14	8.87 $\pm$ 0.70	9.20 $\pm$ 0.68	0.156	38
Subject 15	5.37 $\pm$ 1.58	5.19 $\pm$ 1.54	0.637	38
Subject 16	5.81 $\pm$ 0.85	6.14 $\pm$ 0.68	0.320	38
Subject 17	7.95 $\pm$ 0.83	8.18 $\pm$ 0.74	0.814	38
Subject 18	9.30 $\pm$ 0.63	9.11 $\pm$ 0.68	0.864	38
Sample	8.03 $\pm$ 1.59	8.08 $\pm$ 1.61	0.944	14

Table 22: Results of the parametric test t-test testing differences in mean RT across the two categories of average reward - Low ($AvRew < median$) and High ($AvRew > median$) - for subjects of group 1. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of RT across subjects, that is, the results for the group level.

	Reaction Time			
	Average Reward			
	Mean $\pm$ *SD*		*p value*	*d.f*
	Low	*High*		
Subject 21	9.64 $\pm$ 0.42	9.68 $\pm$ 0.27	0.739	38
Subject 22	5.64 $\pm$ 0.99	6.06 $\pm$ 1.07	0.213	38
Subject 23	7.20 $\pm$ 1.15	7.24 $\pm$ 0.99	0.894	38
Subject 24	8.00 $\pm$ 0.70	7.92 $\pm$ 0.65	0.704	38
Subject 25	5.90 $\pm$ 2.03	5.71 $\pm$ 2.16	0.776	38
Subject 26	7.62 $\pm$ 0.53	7.83 $\pm$ 0.74	0.316	38
Subject 27	7.62 $\pm$ 1.26	7.64 $\pm$ 1.27	0.960	38
Subject 28	7.41 $\pm$ 0.43	7.60 $\pm$ 0.44	0.160	38
Subject 29	6.99 $\pm$ 0.82	7.13 $\pm$ 1.06	0.632	38
Sample	7.33 $\pm$ 1.17	7.42 $\pm$ 1.15	0.873	16

Table 23: Results of the parametric test t-test testing differences in mean RT across the two categories of average reward - Low ($AvRew < median$) and High ($AvRew > median$) - for subjects of group 2. For each subject, the test was applied to the mean values across sessions and its normality was tested with a Lilliefors test. The rows with a light grey color are the ones where the normality test was negative, and therefore, it was applied the non-parametric version of the t test, the Wilcoxon rank sum test. The sample translates the results of the t-test applied to the mean values of RT across subjects, that is, the results for the group level.

Press Rate as function of reward category - G1

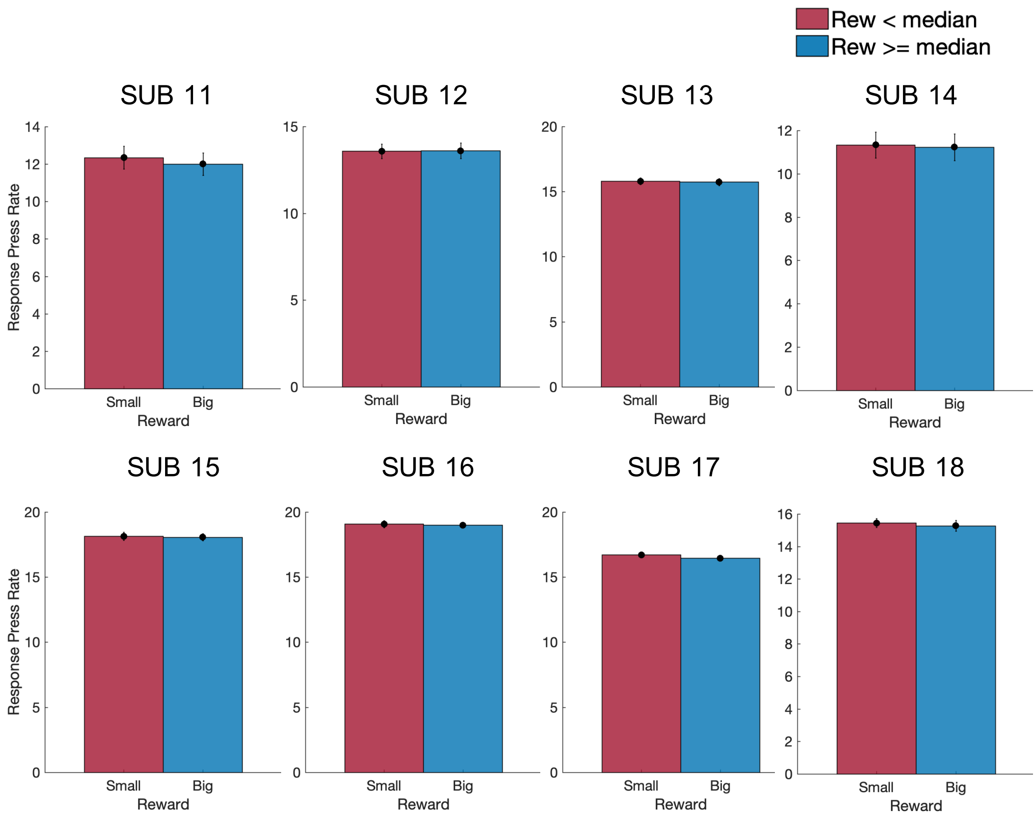

Figure 20: Mean response press rate, in $presses/s$, as function of the two categories of reward, divided in $Reward < median$ (red) and $Reward > median$ (blue) for each subject of Group 1. Being Group 1 the group of subjects who performed the task under a point accumulating reward system. No subject showed a significant difference between the means of press rate associated to each group of reward for a confidence level of 5%. The errorbars represent standard error across sessions.

Press Rate as function of reward category - G2

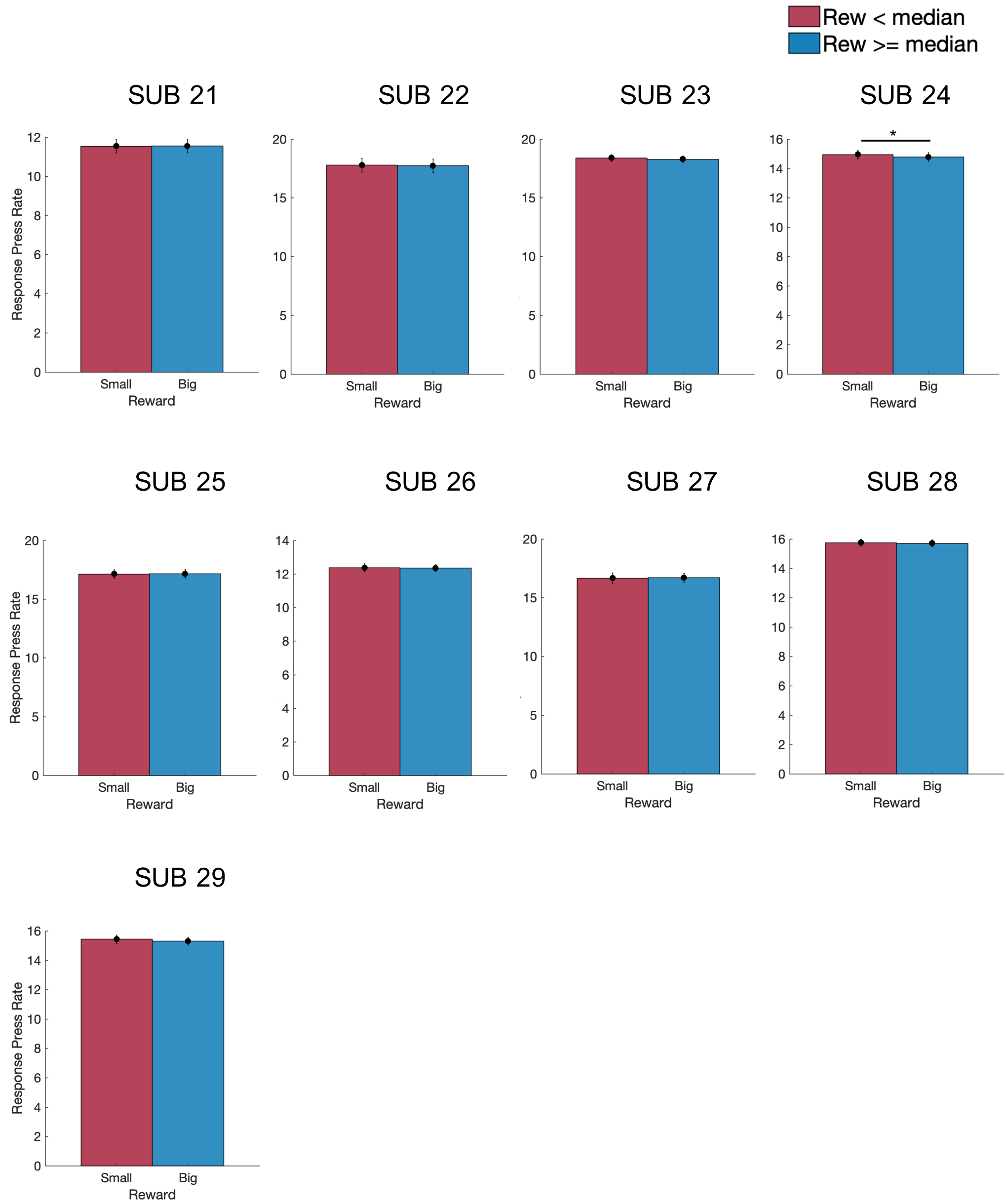

Figure 21: Mean response press rate, in $presses/s$, as function of the two categories of reward, divided in $Reward < median$ (red) and $Reward > median$ (blue) for each subject of Group 2. No subject showed a significant difference between the means of press rate associated to each group of reward for a significant level of 5%. The errorbars represent standard error across sessions.

Press Rate as function of average reward (low and high categories) - G1

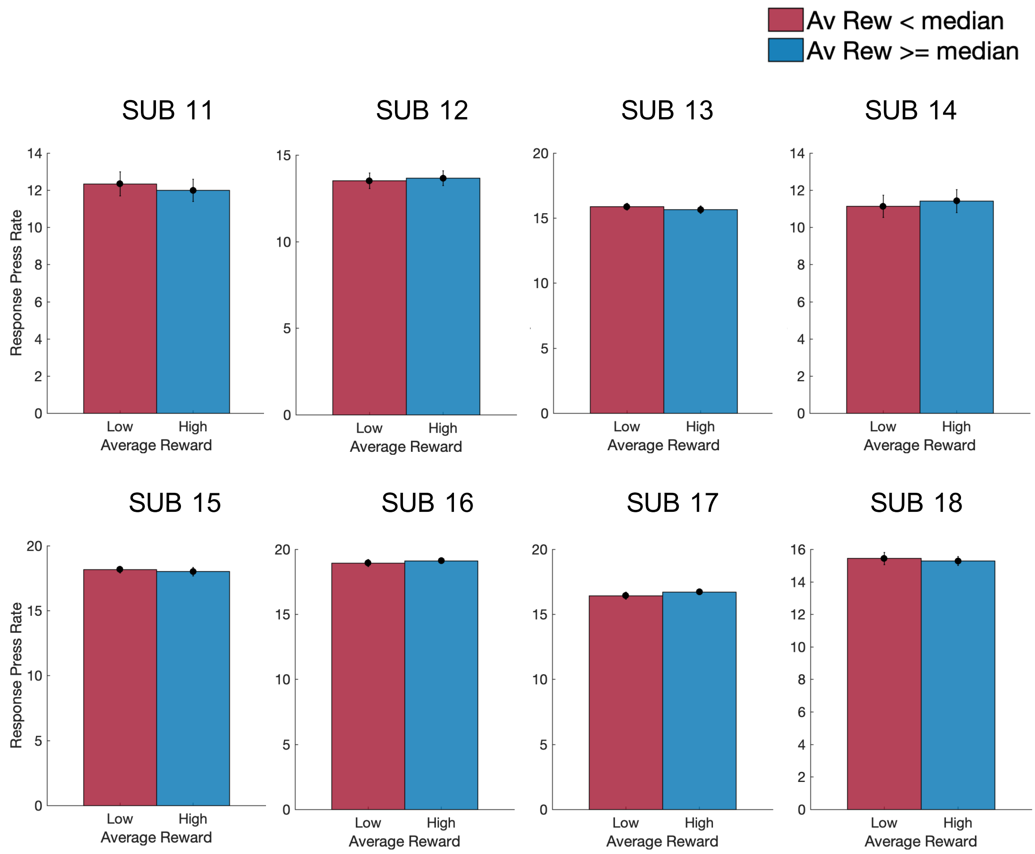

Figure 22: Mean response press rate, in $presses/s$, as function of low ($AvRew < median$) - red - and high ($AvRew > median$) average reward rate. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t-1) + alpha \times Rew$, where $alpha$ was selected individually for each subject as the value which offered the biggest R^2 for the prediction of press rate performed for each subject individually. The means of press rate for high and low average reward were not significantly different for any subject at the significance level of 5%. Errorbars represent standard error across tsessions.

Press Rate as function of average reward (low and high categories) - G2

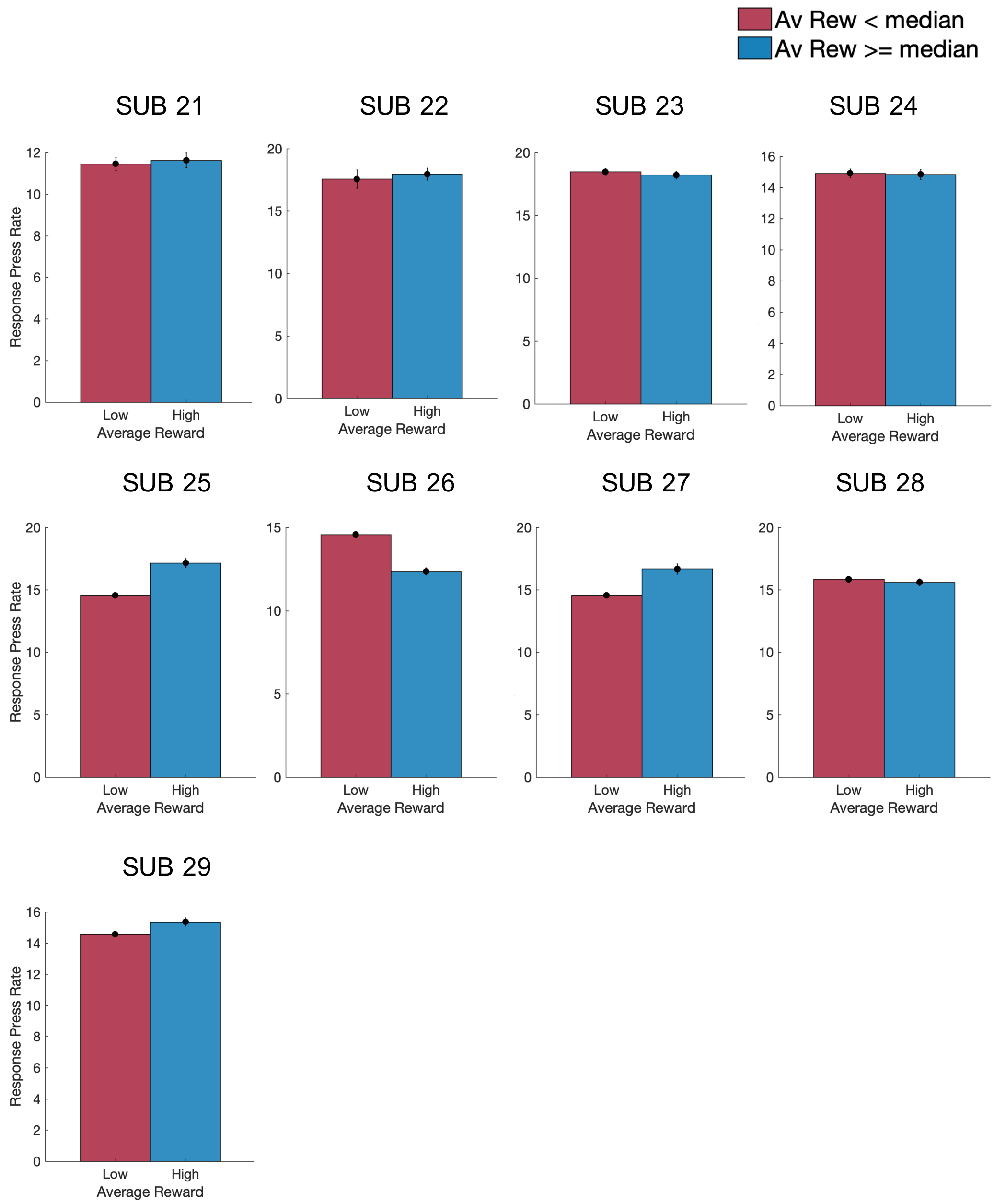

Figure 23: Mean response press rate, in $presses/s$, as function of low ($AvRew < median$) - red - and high ($AvRew > median$) average reward rate. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t-1) + alpha \times Rew$, where $alpha$ was selected individually for each subject as the value which offered the biggest R^2 for the prediction of press rate performed for each subject individually. The means of press rate for high and low average reward were not significantly different for any subject at the significance level of 5%. Errorbars represent standard error across sessions.

Waiting time behavior for each subject - G1

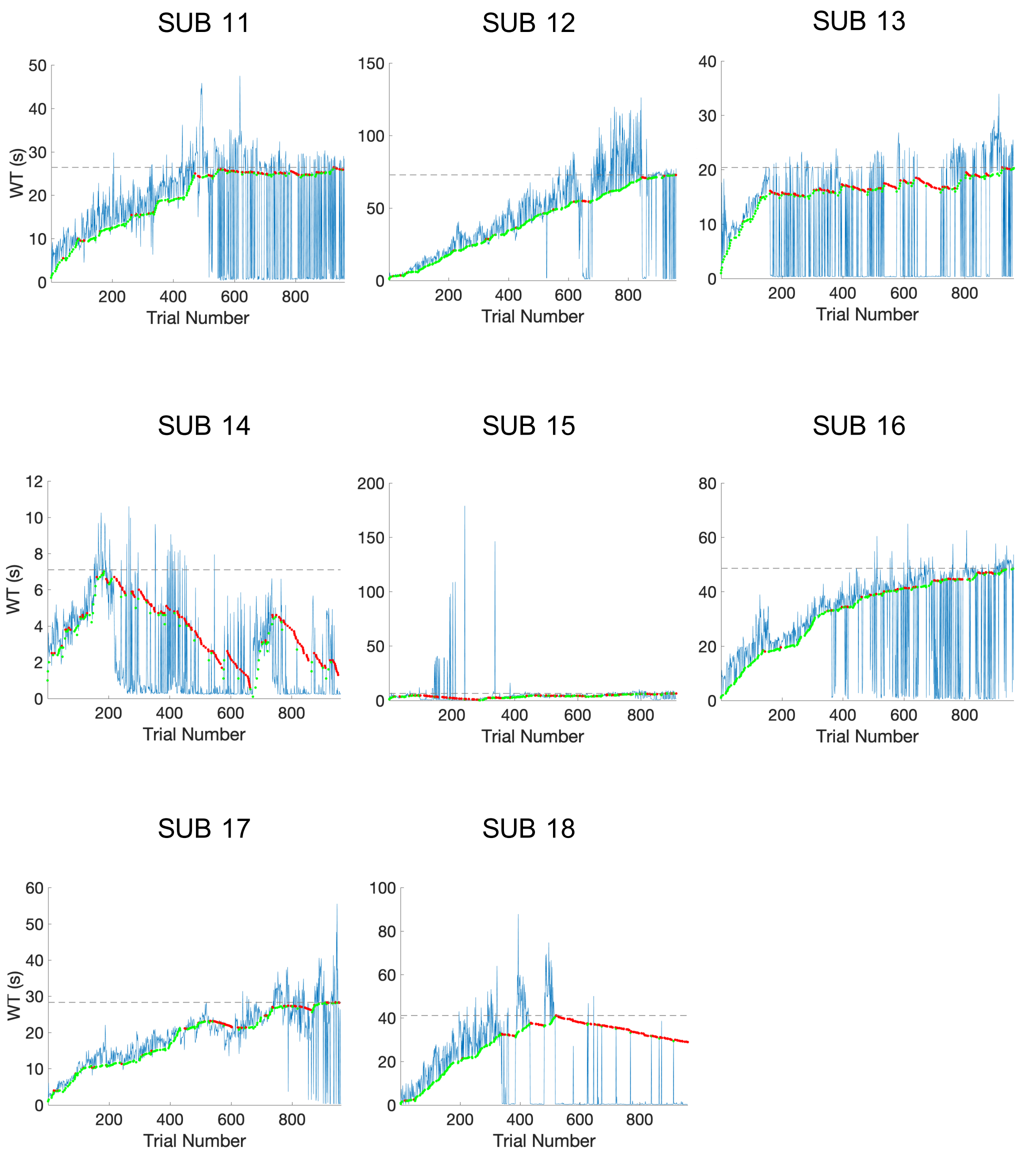

Figure 24: Behavior of each subject waiting time (WT) along the experiment, in blue. Each green dot represents a trial where the bonus was present and the subject succeeded in waiting for it. Each red dot represents a missed bonus by the subject.

Waiting time behavior for each subject - G2

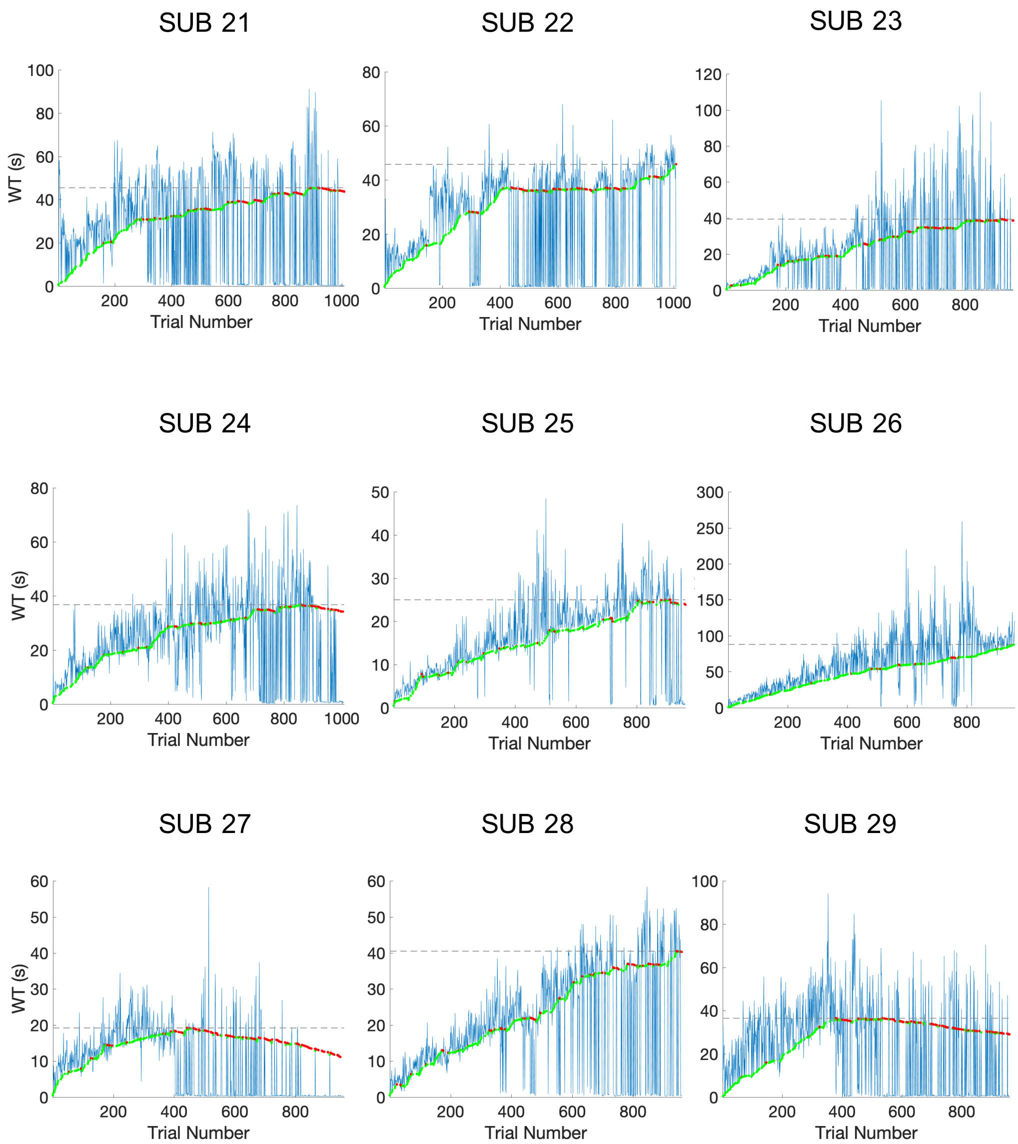

Figure 25: Behavior of each subject waiting time (WT) along the experiment, in blue. Each green dot represents a trial where the bonus was present and the subject succeeded in waiting for it. Each red dot represents a missed bonus by the subject.

	Waiting Time ANOVA		
	Conf		
	F	*d.f*	*p value*
Subject 11	0.75	2	0.473
Subject 12	2.85	2	0.059
Subject 13	3.47	2	**0.032**
Subject 14	0.62	2	0.536
Subject 15	3.24	2	**0.040**
Subject 16	2.06	2	0.129
Subject 17	10.66	2	**< 0.0001**
Subject 18	1.23	2	0.292
Sample	**2.53**	**2**	**0.099**
Subject 21	1.32	2	0.268
Subject 22	2.35	2	0.096
Subject 23	0.09	2	0.912
Subject 24	2.33	2	0.098
Subject 25	1.02	2	0.360
Subject 26	3.18	2	**0.042**
Subject 27	0.23	2	0.793
Subject 28	35.84	2	**<0.0001**
Subject 29	9.22	2	**<0.0001**
Sample	**7.94**	**2**	**0.002**

Table 24: Results for two-way ANOVA testing differences in mean waiting time across self-reported confidence level. The effect is significant for few subjects of both groups of subjects and significant at the group level for Group 2 at the confidence interval of 5%.

Waiting Time as function of reward category - G1

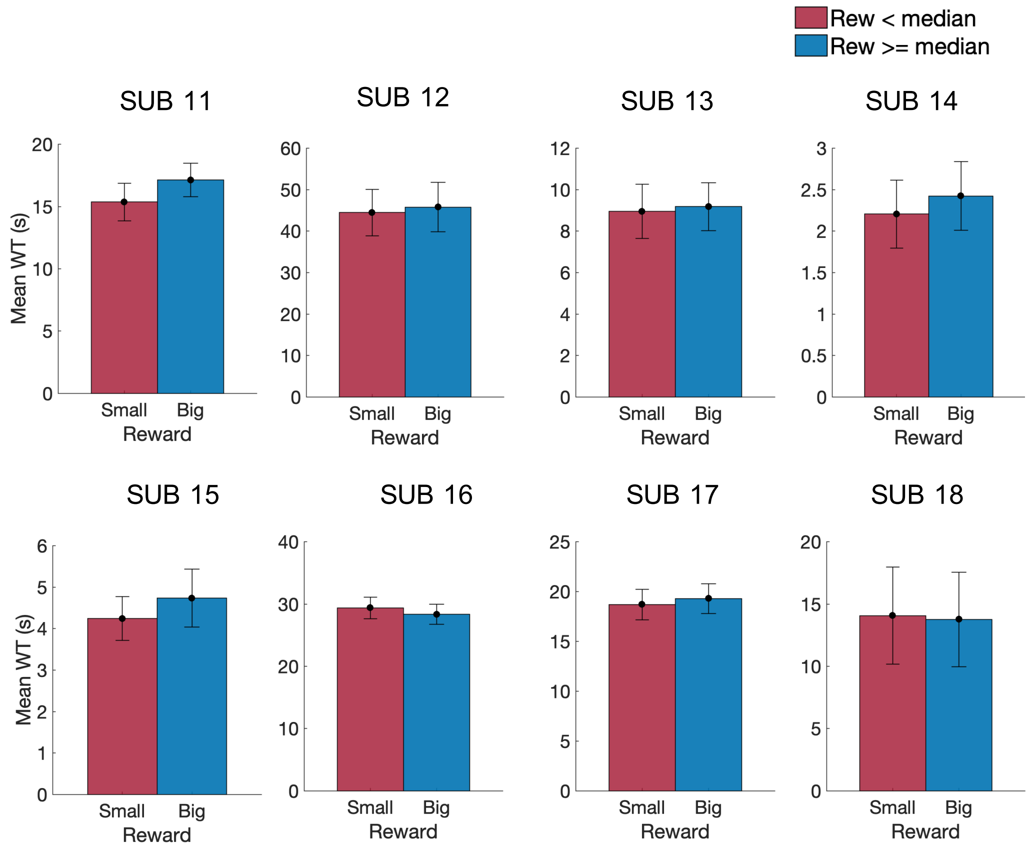

Figure 26: Average waiting time as a function of the two reward's categories - $Reward < median$ (red) and $Reward > median$ (blue) for each subject of Group 1. The errorbars represent standard error across sessions. No subject presented a significant difference between categories of reward.

Waiting Time as function of reward category - G2

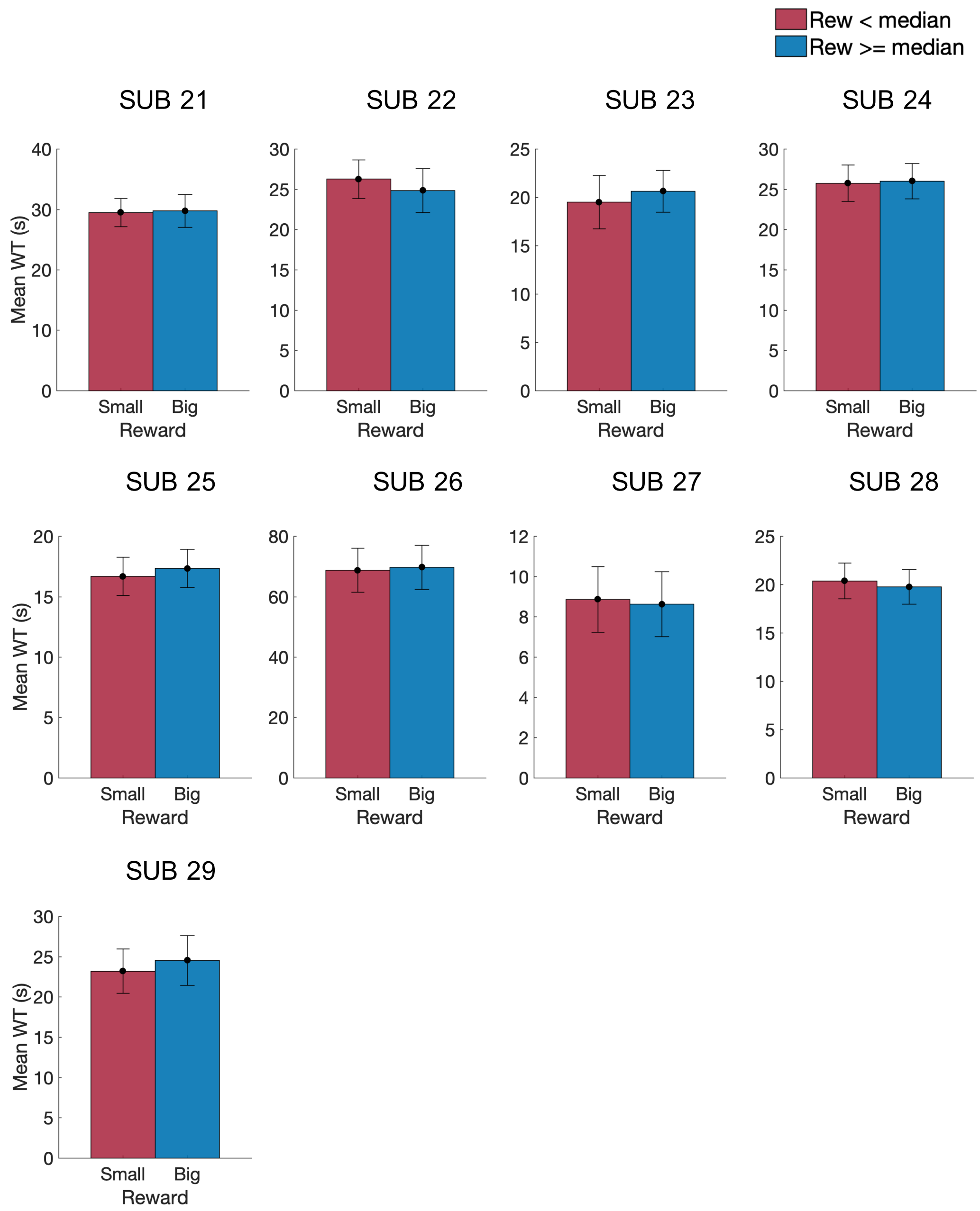

Figure 27: Average waiting time as a function of the two reward's categories - $Reward < median$ (red) and $Reward > median$ (blue) for each subject of Group 2. The errorbars represent standard error across sessions. No subject presented a significant difference between categories of reward.

	Waiting Time			
	Reward			
	Mean $\pm$ *SD*		*p value*	*df*
	Small	*High*		
Subject 11	15.36 $\pm$ 11.37	17.12 $\pm$ 11.34	0.472	38
Subject 12	44.47 $\pm$ 28.73	45.80$\pm$ 30.78	0.893	38
Subject 13	8.95 $\pm$ 10.10	9.18 $\pm$ 9.97	0.991	38
Subject 14	2.21 $\pm$2.48	2.42 $\pm$ 2.61	0.995	38
Subject 15	4.24 $\pm$8.76	4.74 $\pm$ 7.44	0.924	38
Subject 16	29.40 $\pm$ 17.15	28.35 $\pm$ 16.89	0.642	38
Subject 17	18.70 $\pm$ 8.00	19.29 $\pm$ 8.74	0.890	38
Subject 18	14.07 $\pm$ 18.56	13.77 $\pm$ 18.20	0.932	38
Sample	18.25 $\pm$ 16.02	18.84$\pm$ 16.58	0.936	14
Subject 21	29.51$\pm$ 23.64	29.77$\pm$ 23.08	0.871	38
Subject 22	26.27 $\pm$ 18.08	24.86$\pm$ 18.63	0.745	38
Subject 23	19.50 $\pm$22.43	20.64 $\pm$ 23.46	0.750	38
Subject 24	25.77 $\pm$16.33	26.01$\pm$ 16.81	0.981	38
Subject 25	16.68 $\pm$9.25	17.33 $\pm$ 9.45	0.841	38
Subject 26	68.77 $\pm$39.86	69.75 $\pm$ 40.00	0.943	38
Subject 27	8.87 $\pm$ 9.55	8.63 $\pm$ 9.80	0.862	38
Subject 28	20.39 $\pm$ 15.29	19.77 $\pm$ 14.55	0.780	38
Subject 29	23.21 $\pm$ 21.09	24.54 $\pm$ 21.65	0.790	38
Sample	26.48 $\pm$ 16.91	26.99 $\pm$ 17.81	0.923	16

Table 25: Results for t-test testing differences in mean waiting time across different groups of reward. The effect is not significant for any subject and neither at the group level for both groups of subjects.

Waiting Time as function of average reward (low and high categories) - G1

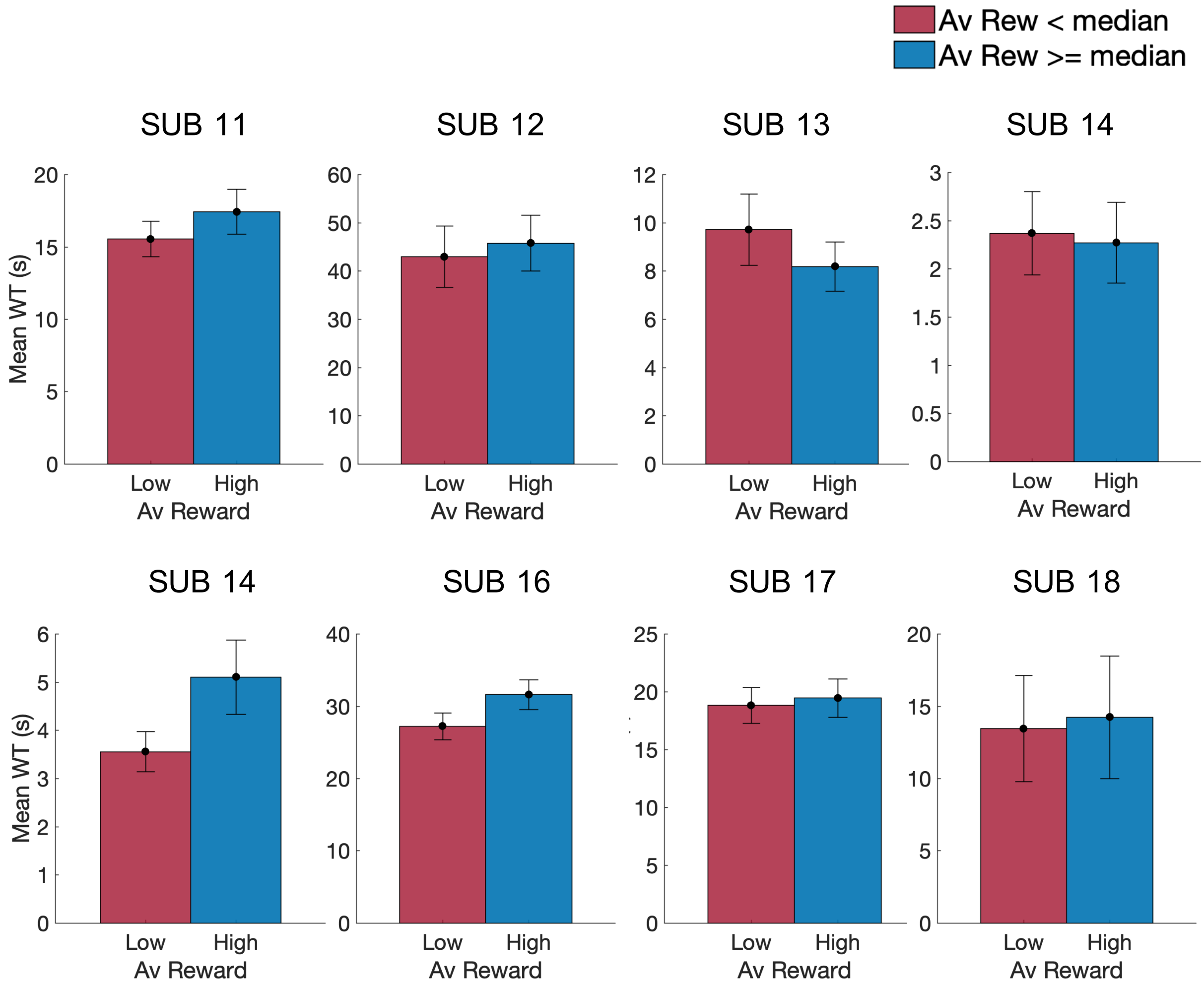

Figure 28: Waiting time in seconds as a function of average reward category - low ($AvRew < median$) in red and high ($AvRew > median$) for subjects of Group 1. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t-1) + alpha \times Rew$, where $alpha$ was the values of alpha obtained to maximize the prediction of each subject movement time. The errorbars represent standard error of the mean across sessions. No subject presented a significant difference between the means of WT for both groups of average reward

Waiting Time as function of average reward (low and high categories) - G2

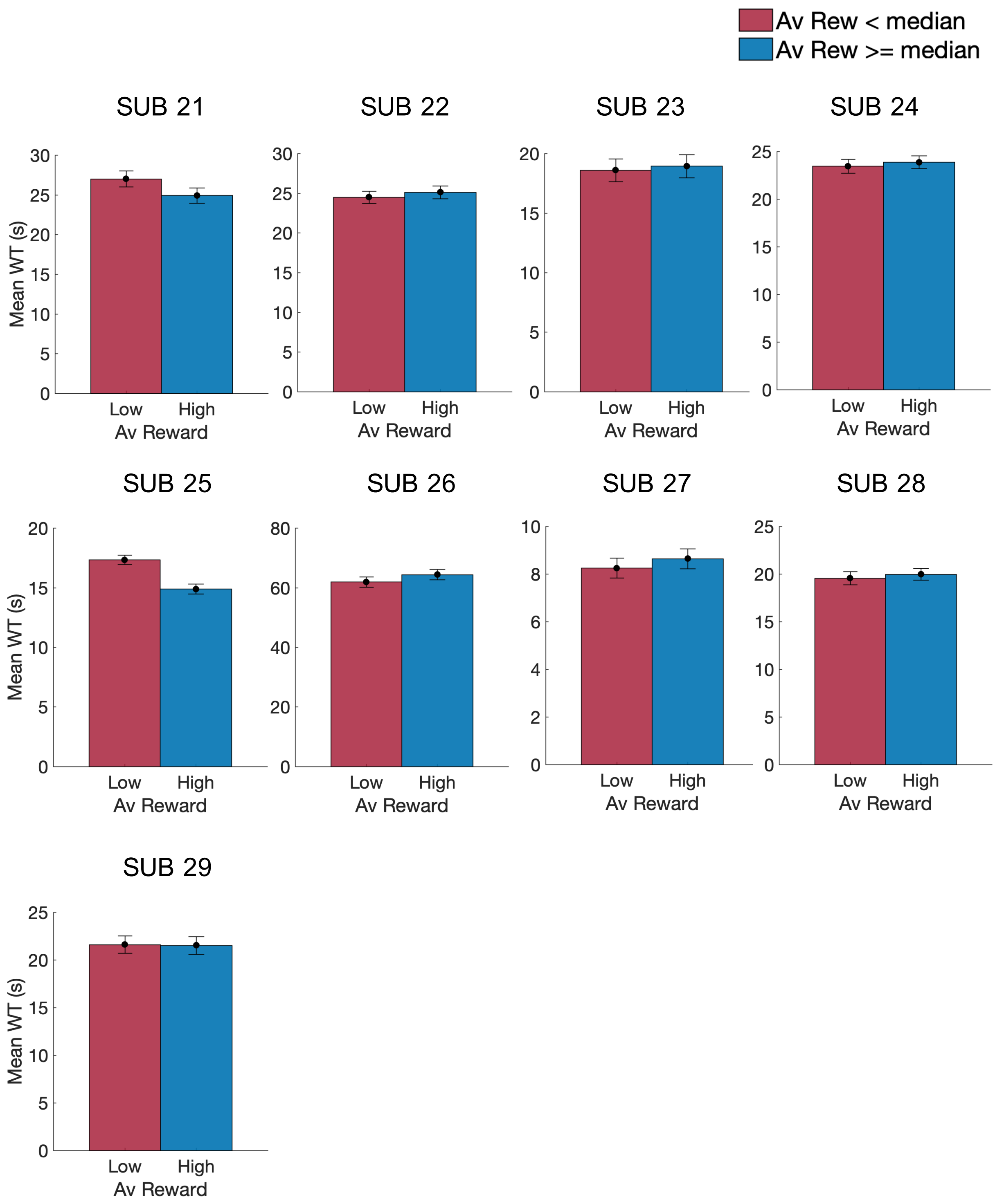

Figure 29: Waiting time in seconds as a function of average reward category - low ($AvRew < median$) in red and high ($AvRew > median$) for subjects of Group 2. The average reward was calculated with the following update rule $Av_{Rew}(t) = (1 - alpha) \times Av_{Rew}(t-1) + alpha \times Rew$, where $alpha$ was the values of alpha obtained to maximize the prediction of each subject movement time. The errorbars represent standard error of the mean across sessions. No subject presented a significant difference between the means of WT for both groups of average reward

	Waiting Time			
	Average Reward			
	Mean ± *SD*		*p value*	*df*
	Small	*High*		
Subject 11	16.00 ± 11.37	16.74 ± 11.33	0.869	38
Subject 12	45.61 ± 29.61	45.53± 29.83	0.842	38
Subject 13	9.48 ± 10.09	8.53 ± 9.97	0.386	38
Subject 14	2.28 ± 2.59	2.25 ± 2.50	0.552	38
Subject 15	3.87 ± 3.98	4.96 ± 10.69	0.800	38
Subject 16	28.63 ± 17.40	29.89 ± 16.65	0.897	38
Subject 17	18.85 ± 8.16	19.21 ± 8.59	0.649	38
Subject 18	13.71 ± 19.89	13.66 ± 16.55	0.772	38
Sample	17.75 ± 15.34	18.85± 16.29	0.877	14
Subject 21	28.75± 23.72	30.46± 22.98	0.670	38
Subject 22	24.97 ± 18.11	25.74± 18.65	0.815	38
Subject 23	20.40 ±23.97	19.31 ± 21.70	0.679	38
Subject 24	25.76 ±17.19	25.74± 15.95	0.913	38
Subject 25	16.89 ±9.05	17.16 ± 9.59	0.968	38
Subject 26	69.04 ±41.12	70.90 ± 38.68	0.897	38
Subject 27	8.35 ± 9.80	8.79 ± 9.56	0.976	38
Subject 28	19.83 ± 15.19	20.31 ± 14.60	0.967	38
Subject 29	25.37 ± 21.98	22.24 ± 20.77	0.887	38
Sample	26.48 ± 16.91	26.99 ± 17.81	0.951	16

Table 26: Results for t-test testing differences in mean waiting time across different groups of average reward. The effect is not significant for any subject and neither at the group level for both groups of subjects at the confidence interval of 5%.