# Sensorimotor decision-making with moving objects

David Aguilar Lleyda

**ADVERTIMENT**. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA**. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING**. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

UNIVERSITAT DE
BARCELONA

# Sensorimotor decision-making with moving objects

by

**David Aguilar Lleyda**

a thesis submitted in partial fulfillment of the requirements for the degree of

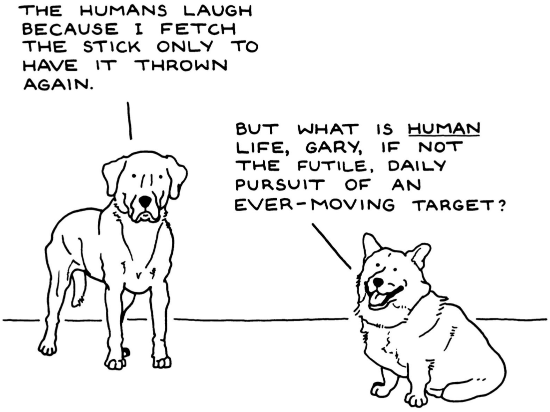
**Doctor in Psychology**

under the supervision of

**Dr. Joan López-Moliner & Dr. Elisabet Tubau Sala**

Doctoral program in Brain, Cognition & Behaviour
Departament de Cognició, Desenvolupament i Psicologia de l'Educació
Facultat de Psicologia
Universitat de Barcelona

October 2017



@DRAWINGSofDOGS

*A mons pares, Maria Margarida i Antonio,*
*i a mon germà, Víctor.*

# Acknowledgments

Tinc la immensa sort de comptar amb la meva família i els meus amics. El seu suport ha estat essencial per la finalització d'aquesta tesi. Espero poc a poc anar tornant-los el molt que m'han ajudat. També espero, però, que em perdonin per aquesta impersonal menció, ja que intentaré circumscriure els agraïments explícits a persones de l'àmbit acadèmic.

No puc sinó començar esmentant els meus supervisors. Crec que durant aquests anys hem sabut trobar un bon equilibri entre les tasques de cadascú. En cap moment m'he sentit abandonat, i les meves peticions d'ajuda sempre han sigut ràpidament ateses, fins i tot en caps de setmana i festius.

Elisabet, moltes gràcies per supervisar-me la tesina del màster, i per després emprendre una tesi que no sabíem si podria tenir finançament. També per buscar la manera perquè finalment em concedissin una beca per poder realitzar el doctorat. Si no fos per la teva perseverança no hauríem arribat fins aquí. Agraeixo també els teus consells durant la tesi, especialment sobre com encarar els articles que hem enviat. Has aportat sempre una perspectiva diferent, has aconseguit clarificar el nostre missatge, ii si algú no familiaritzat amb els temes de la tesi arriba a entendre alguna cosa, això serà gràcies a tu.

Joan, tu ets amb qui he acabat passant més temps. En tu ha recaigut la major part de la tasca de dotar-me d'eines per fer-me un investigador autònom. Això ha comportat hores i hores de dubtes, però cada cop crec que estic més preparat per enfrontar-me al món tot sol. Espero que un cop això hagi acabat encara et pugui molestar amb algun correu. També m'has donat l'oportunitat d'assistir a congressos i relacionar-me amb la comunitat científica, i això m'ha obert portes per poder seguir al món de la investigació després del meu doctorat.

Un cop fets aquests agraïments, voldria tirar enrere i reconèixer a la persona de la mà de la qual vaig entrar a la recerca. Miquel, d'aquí poc farà 10 anys que vaig posar els peus al departament per primer cop. Descriuria el teu mèrit científic, però de bons científics en mencionaré molts. Prefereixo destacar l'insòlit cas de l'acadèmic que, lluny d'encapsular-se en el seu tema d'estudi, mostra interès per tot. M'emporto amb mi no només coneixement sobre llenguatge, sinó també grans lliçons sobre literatura i música barroca.

Continuo amb les persones del grup de recerca en Visió i Control de l'Acció. Quan jo vaig arribar-hi ja hi havia un grapat de predocs i postdocs. El Borja, amb qui tan bé ens ho vam passar a Leeds, ara un pare de família feliç. La teva gran tesi m'ha servit d'inspiració. Alexis, antiguo ranger de Texas venido a más, y una de las personas con las ideas más claras de lo que debería ser la ciencia. Conversar contigo vale más que horas leyendo papers. Cristina, perquè sempre hi ha d'haver algú que et mostri com es fan les coses a la perfecció. Moltes gràcies per la preocupació que sempre has mostrat. Eric, hem parlat dels temes més diversos i vam passar una nit molt estranya a Lieja. Treballar amb tu m'ha ensenyat que no és temps d'enviar un article fins que cadascun dels autors no l'ha revisat 50 cops. Elena, de sonrisa contagiosa. Brendan, your words were not many but were always wise. Life takes many turns and some are weird, but I'm sure you'll always do well.

Mentre jo hi era tota aquesta gent ha anat marxant i ha passat a millor vida. Però n'han anat arribant d'altres. Els integrants del VISCA s'han anat succeïnt a un ritme vertiginós, i em fa una mica de vergonya admetre que ara mateix sóc l'estudiant més sènior. La primera va ser la Clara. Vas revolucionar el meu despatx solitari i ja res ha tornat a ser el mateix. La disciplina militar que ara s'hi respira es pot aguantar per tot el que hem rigut i pels teus petons al coll. Lisa, you were next. Probably the only monarchist I have ever liked. You have tolerated my bullshit better than anybody else. Your healthy orange skin shines bright and lights the way. Per si fóssim pocs, després va arribar el Björn. Una persona sorprenent. L'esportista suat que fa la metamorfosi i es converteix en un dandy amb corbata. El que ens demostra que es poden parlar 6 llengües a nivell nadiu, i també que mai no hi ha pantalons prou curts. Finalmente llegaste tú, Borja, ese extraño ser para quien comer no es un placer sino una necesidad vital. Estaré gustoso de dejar que seas tú quién usurpe mi lugar en el despacho.

Uns quants estudiants han fet estades al grup. Tots han deixat la seva petjada d'una manera o una altra. Paula, te mereces el mayor de mis respetos por ser la única en aceptar el reto de comerse un pijama. Agustín, el maestro de los loops, en tu honor aún conservamos tu bote de allioli en la nevera. Ljubica, I expect great things from you: mainly, to win a Nobel prize and to be my future boss. Lena and Valentina, thank you for being there during my final days.

El grup també consta d'uns quants investigadors amb qui he compartit tasques docents de l'assignatura Percepció i Atenció. José Antonio, has mostrat una gran flexibilitat per tal que adaptés les pràctiques al meu gust. Matthias, tú optaste por asumir que sabíamos lo que nos traíamos entre manos y dedicar las conversaciones a otros temas. Gracias a eso creo que ahora sé todo lo que hay que saber sobre bicis. Tot això sempre autoritzat pel nostre coordinador Ferran, de qui no puc dir res dolent per por a les represàlies. Un dels altres companys ha docència ha estat el Manel. Les mans més hàbils de tot Mundet, i una persona que té la capacitat de fer-te sentir bé només de dir "què passa senyor?".

Estenc els agraïments a persones amb relació diversa amb el grup. Alejandro, per les teves classes de mates. Juanma, per tot el que has fet durant aquests anys per ajudar-me a preparar els experiments. Cristina Pulido, per la teva diligència i consell. Mari Carmen i Graciela, per l'ajuda en els tràmits.

I would also like to acknowledge the people from the University of Leeds, who were extremely welcoming and helpful during my stay in Yorkshire. Faisal Mushtaq, for those great meetings where we discussed the project for half an hour and then football for the next two hours. You made it all easy, even cricket rules. Mark Mon-Williams, who first invited me up there, and someone whose Scottish complicity has always made me feel at ease. Pablo, el vato loco de la electroencefalografía, a quien ni siquiera el frío de Newcastle desalienta de llevar la playera de los Pumas. Zeynep, always cheerful, even when not drunk, the pound-stretcher who taught me how to survive in the UK by scavenging at dusk to get sandwiches at reduced price. And all the people from PaCLAB.

A tota la gent amb qui he compartit taula. Àngels, tan bona encaixant les meves bromes de mal gust com eficient solucionant problemes del departament. Javier, autoritat màxima en tots els àmbits de la vida que de veritat importen: música, videojocs indies, cinema i sèries, croissants i orxata. L'Òscar, un gran pastisser que intenta amagar la seva sensibilitat darrere d'un comportament fanfarró i milhomes.

Voldria esmentar també un grup variat de persones. Amb uns quants hem compartit despatx. L'Asia, la companya de laboratori que tothom voldria. No només per la seva bondat, sinó sobretot perquè és l'única amb qui he pogut parlar de heavy metal. També amb l'Adrià hem estat sota el mateix sostre. Inevitable perdre una hora cada dia discutint de política. Després hi ha uns quants individus amb qui hem coincidit pel passadís. Cal fer una menció al Toni, un dels pocs tribuneros decents. I la Maria. Definitivament el millor d'haver cursat el CCiL. El temps s'ha emportat la majoria dels coneixements que vaig aprendre durant aquell any, però per sort la nostra amistat perdura. También mencionar a Solène, a quién aún no he visto sin gafas de sol. Y a Ignacio, quién entre risas suelta las preguntas más dolorosas, y que hacen temblar los cimientos de cualquier estudio. And thanks to Angelica and Arnau for helping with this thesis' cover.

L'últim dels meus agraïments el vull dedicar a algú que pertany a moltes de les categories anteriors. Va ser membre del grup. Va tenir la generositat d'acceptar-me com a col·laborador en la seva investigació. També ha estat company de docència, deixant unes diapositives dignes de les millors universitats. Però, si hagués de classificar-lo en un rol, seria com el tercer dels meus mentors. Dani, ets la persona que més m'ha ensenyat sobre el món de la ciència en general. Un dels científics més rigorosos que he conegut. I, malgrat tot, abans que les nostres converses científiques, salvaria les que hem tingut sobre altres temes. També hauria d'agraïr-te haver marxat, ja que d'haver-te quedat no hagués tingut al·licient per acabar la tesi.

# List of abbreviations

ANOVA: analysis of variance

DV: decision variable

LMM: linear mixed model

SDT: statistical decision theory

TTC: time to contact

# Table of contents

# Summary

Moving is essential for us to survive, and in countless occasions we move in response to visual information. However, this process is characterized as uncertain, given the variability present both at the sensory and motor stages. A crucial question, then, is how to deal with this uncertainty in order for our actions to lead to the best possible outcomes.

Statistical decision theory (SDT) is a normative framework that establishes how people should make decisions in the presence of uncertainty. This theory identifies the optimal action as that which maximizes the expected reward (outcome) of the situation. Movement planning can be reformulated in terms of SDT, so that the focus is placed on the decisional component. Some experimental work making use of this theoretical approach has concluded that humans are optimal movement planners, while other has identified situations where suboptimality arises. However, sensorimotor decision-making within SDT has commonly eluded scenarios of interaction with moving objects. At the same time, the work devoted to moving objects has not focused on the decisional aspect. The present thesis aims at bridging both fields, with each of our three studies trying to answer different questions.

Given the spatiotemporal nature of situations with moving objects, we can plan our actions by relying on both temporal and spatial cues provided by the object. In Study I we investigated whether exploiting

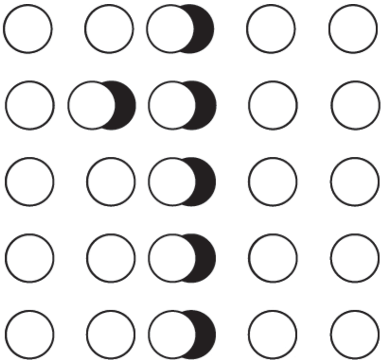
more one type of these visual cues led to a better performance, as defined by the reward given after each action. In our task we presented a target, which could vary in speed and motion time, approaching a line. Participants responded to stop the target and were rewarded according to its proximity to the line. Responding after the target crossed the line was penalized. We discovered that those participants planning their responses based on time-based motion cues had a better performance than those monitoring the target's changing spatial position. This was due to the former approach circumventing a limitation imposed by the resolution of the visual system. We also found that viewing the object for longer favored time-based responses, as mediated by longer integration time. Finally, we used existing SDT models to obtain a reference of optimality, but we defend that these models are limited to interpret our data.

Study II built on our previous findings to explore whether the use of temporal cues could be learnt. We took our previous paradigm and adapted it so that reward was manipulated after each task in order to foster exploiting temporal information. There was no evidence for learning taking place, since participants using temporal cues did so from the start of the experiment. Whether other methods reward can shape the use of certain cues, and why some people naturally tend to make more use of temporal information, still remain elusive.

Study III deepened our knowledge on which variability people consider when planning their responses. We hypothesized that the reason why people are suboptimal (as defined by SDT) in many situations is because they represent only their measurement variability, roughly equivalent to the execution noise, while excluding the variability created by sudden changes in their planning. We took previous data and used a

Kalman filter to extract each participant’s measurement variability. We then used it to compute SDT-derived optimal responses, and discovered that they explained well our data, giving support to our hypothesis. We also found evidence for participants using the information provided by reward both to avoid being penalized and to choose the point at which to stabilize their responses.

Taken together, our experimental work presents interaction with moving objects as a complex set of situations where different information guides our response planning. Firstly, visual cues of different origin. Secondly, our variability, coming from many sources, some of which may not be considered. Finally, the outcomes related to each action.



# Introduction

This introductory chapter intends to provide the reader with the necessary context to understand the motivation behind the present doctoral thesis. It also tries to give a quick glance at some of the models used in the experimental part. With this goal, the background is divided in two main parts.

In the first part, we review previous work on sensorimotor decision-making. We start by evidencing its biological relevance and characterizing it as a noisy process that generates uncertainty. We continue introducing statistical decision theory, one of the most general normative proposals to study decision-making under risk or uncertainty. We then present one model stemming from that theory that has been extensively used to test whether human movement planning can be considered optimal. After reviewing evidence in favor of optimal movement planning, we do the same with evidence for suboptimality, grouping it by the hypothesized source of this phenomenon.

The second part of the introduction is devoted to how people interact with moving objects. It starts by trying to sketch the foundations of how we can perceive motion. This section does not seek exhaustiveness, given the vast amount of topics that scientists of this area address. Alternatively, it is mainly focused on highlighting some mechanisms and phenomena, anticipating its future importance for the contents of the thesis later on. Once this has been explained, the next section dwells on the interaction between motion and position. The intention is to make clear how, although both are highly correlated, there are situations where motion can affect the perceived position of a moving object, and that this depends on the object's speed. We conclude with an overview of how people estimate the time to contact of an object relative to some reference, and the decisional variables they use to trigger an action to this object.

## 1.1. Making sensorimotor decisions under uncertainty

### 1.1.1. Perceiving and moving: essential, yet uncertain

Human beings owe the current state of their nervous systems to a long evolutionary history. Evolution is produced so that organisms can adapt to their environment, with the ultimate goal of surviving both as individuals and as a species. To survive, though, it is necessary to interact with the environment and modify it. Basic behavior like feeding, hiding or mating shows how the only way to achieve that is by moving our body. Moreover, since movement is framed as part of our interaction with the world, many times we move based on some sensory stimulus: seeing a ripe fruit hanging from a tree, hearing a possible predator, smelling the pheromones that indicate us that our partner is in heat. And while some movements are just a reflex reaction to a stimulus, most actions require some kind of decision: how much to extend our arm to pick the fruit, in which direction to start running from the predator, what is the best way to approach our mate. Finally, it has to be considered that all decisions and actions have a goal. That is, we decide how to move according to a desired outcome: being able to feed, not being eaten, reproducing.

The last paragraph gives us a gist of the elements we need to basically define our field of study: sensorimotor decision-making. This can be defined as the decisional processes related to using sensory information to produce a motor output, with the goal of achieving a determinate outcome.

Among all the modalities that can be the source of our sensory information, we will focus on vision. Constraining our field of study is justified if we consider the richness and complexity of visual information. In fact, conservative estimates point to around 30% of the human cortex being exclusively dedicated to processing vision (Van Essen, 2004; Van Essen & Drury, 1997), concentrated in the occipital cortex. The massive amount of space reserved in the brain for vision reveals the evolutive importance this sense has acquired. However, the storage capacity is still limited. This makes our visual system imperfect when trying to represent the outside world, especially when also considering the noise and system fluctuations generated at the different stages of visual processing (Faisal, Selen, & Wolpert, 2008).

At the other end of the perception-action cycle, we also face several problems when moving, even if the movements humans can perform are much more sophisticated than those available to our phylogenetic ancestors some millions of years ago. Noise is present both at the planning stage, when we select the motor commands, and at the execution stage, when these commands are sent to the muscles (van Beers, Haggard, & Wolpert, 2004). Whichever the stage noise is generated at, the result is some motor variability being produced. This means that the movement we initially wanted to do and the one we will end up doing will differ by some extent.

From the previous paragraphs it appears that both our perception and our movements have some uncertainty inherent to them. In the face of this problem, then, a crucial aspect is how we cope with this uncertainty when planning a movement to achieve an outcome. This will be discussed in brief, but first we will introduce a theoretical background that has commonly been used to study decisions under uncertainty.

### 1.1.2. Statistical decision theory as a normative decisional framework

Statistical decision theory (SDT, Blackwell & Girshick, 1954; Ferguson, 1967) has been developed to give an account of how people should make decisions in the face of incomplete information (uncertainty). It is a very general framework with many adaptations. However, models derived from it share one characteristic: they are normative, because they tell us how the best possible decisions should be reached. Following the review by Maloney & Zhang (2010), we will briefly outline the basic elements of SDT, which consist of three sets and three functions. The connections among them can be appreciated in Figure 1.1.

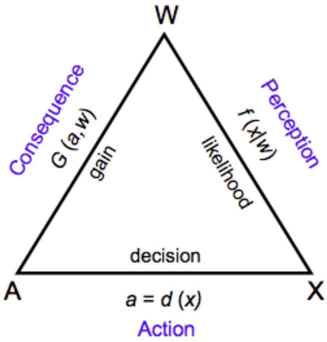


Figure 1.1. The basic elements of SDT. The vertices correspond to the sets of possible states of the world (W), possible sensory states (X) and possible actions (A). The sides correspond to the likelihood, decision and gain functions. Reproduced from Maloney & Zhang (2010).

W refers to the set of possible states of the world, while w refers to the particular state of the world (unknown to the decision-maker) at a particular moment. X is the set of possible sensory states, while x is the sensory state the decision-maker has access to at a particular moment. Given that instantaneous sensory state, the decision-maker has to select, among the set A of possible actions, a particular action a.

The functions that connect the different elements of the sets we just described are the following. The first is the likelihood function f(x|w), the sensory state given the state of the world. The second is the decision function a=d(x), that links the sensory state to a particular action. Finally, the gain function G(a,w) determines the outcome for a particular action.

Such a minimal carcass can be used as a starting point to build tools to model perceptual, sensorimotor, or even non-perceptual decisions. Bayesian decision theory (Berger, 1985) and signal detection theory (Green & Swets, 1988), for instance, are two well-known examples of what can be considered as adaptations of SDT. Most models derived from SDT revolve around the idea that the optimal decision is that which maximizes the expected gain or expected utility. Expected utility maximization has been long studied, and it is common to refer to the core theoretical work around this topic as expected utility theory (Bernoulli, 1954; von Neumann & Morgenstern, 1944). With its concepts and ideas preceding the formalization of SDT, expected utility theory has been central for it. Since some concepts will be very relevant for the present work, it is convenient to introduce them.

Choices have an associated outcome. This outcome can be measured in terms of gain, as already said before. A specific gain, though, will have a subjective value for a decision-maker. The subjective value of an outcome is known as utility, a concept first introduced by Daniel Bernoulli in 1738. However, assuming a context where the value is not distorted, we could use utility and gain interchangeably. From now on, and for the sake of simplicity, we will mainly use the term reward. This word, which is used in fields like reinforcement learning to denote an outcome, does not imply value distortion like utility, and avoids confusion with the other meanings of the term gain (like neural gain, or the Kalman gain we will use later on in this thesis).

It is commonly assumed that the larger the reward, the better an outcome. But uncertain situations, where our choices are related to an outcome via a probabilistic process, are not so straightforward. In those cases, we refer to the expected value of a choice as the average outcome if we repeated that choice multiple times. This is simply calculated by multiplying the probability of an event by the value of the outcome when that event is produced. For instance, imagine we win 1€ every time we get tails after tossing a coin. Assuming the coin is not rigged, the probability of obtaining tails is 0.5. Thus, the expected value associated to obtaining tails is 0.5 x 1€ = 0.5€. To unify the vocabulary, here we will also use the term expected reward. Finally, all the different rewards could be organized in a reward function, which would map how reward evolves as a function of some relevant variable related to the outcome of the choice. Some of the terms introduced so far are changed when framing reward in a negative way. In this case, we would refer to it as loss. More desirable outcomes would have a smaller loss, and losses would be arranged in a loss function. Then, we could summarize what many instantiations of SDT propose in two ways: we could say that, among all the possible choices, people should make the choice that maximizes expected reward, or that which minimizes expected loss.

For some decades, economic literature depicted humans as deficient decision-makers that constantly failed to maximize expected reward. These claims were mainly based on results derived from a particular kind of tasks. In them, participants chose between two monetary outcomes with different associated payoff probabilities, similar to lotteries. For example, they could be forced to choose between winning 5€ with a probability of 0.5, or 3€ with a probability of 0.8. Of course, optimal behavior would be revealed by consistently choosing the option with the largest expected reward (in our case, the first one). However, this was not what was commonly observed. Prospect theory (Kahneman & Tversky, 1979) and its update cumulative prospect theory (Tversky & Kahneman, 1992) are the most famous examples of those descriptive theories that tried to characterize the nature of observed behavior in this kind of situations. According to these accounts, one of the particularities of human decisions is that people undergo value distortion depending on the accumulated wealth, something already implied by the notion of utility. Another observed phenomenon is that risk seeking or risk aversion is shown depending on the gain or loss framing. One of the most famous claims is probability distortion, where people would overestimate events with low probability of occurring and underestimate those with high probability. Of note is that, strictly speaking, the lottery tasks where these behavioral phenomena were observed should not labeled as decisions under uncertainty, since probabilities were explicitly given to decision-makers. Instead, we should be talking about decisions under risk. In any case, they involve an uncertain outcome, and are similar to what we could experience in many sensorimotor situations. Actually, some researchers had the idea that movement planning could be formulated in terms of SDT. This is explained next.

### 1.1.3. Optimal movement planning

Imagine you are running late for an important meeting. You suddenly stop at a red traffic light and see that a car is approaching you, but it is still at a considerable distance. You really want to arrive at the meeting as fast as possible, so you consider crossing the street by running without waiting for the traffic light to turn green. However, it is preferable to be late than to be run over by the car, so you estimate your chances to make it across before the car reaches you. This estimation is uncertain. Will you be able to run as fast as you think? Did you estimate the car’s speed right? In these kind of situations, deciding a plan of action in order to achieve a situation-defined positive outcome (or avoid a negative one) requires estimating our own uncertainty. Are we humans able to take into account our uncertainty, so that we can perform as good as we possibly can?

In 2003, Julia Trommershäuser and colleagues (Trommershäuser, Maloney, & Landy, 2003b) published an article that gave start to a fruitful line of research. In it, they formulated movement planning in terms of SDT. They also conducted some experiments and, based on their results, they presented humans as efficient movement planners that were able to choose the motor trajectory that minimized the expected loss of the task. This means that participants performed optimally.

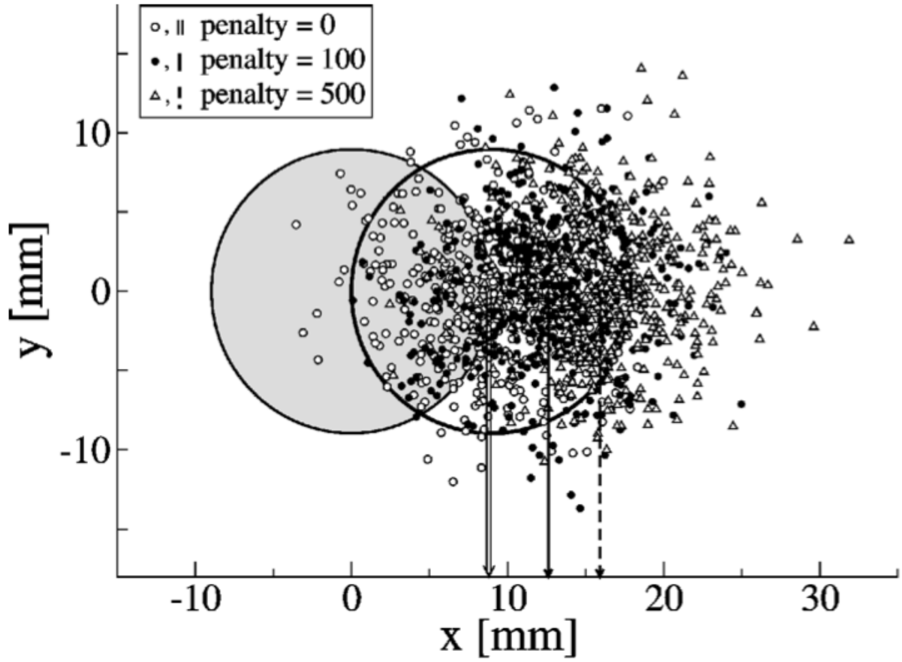


Figure 1.2. Figure taken from Trommershäuser et al. (2003b) that shows the target configuration of their experiments, with two partially overlapping circles. 500 simulated trials are plotted for each condition, where each condition differed in the penalty associated to hitting the filled circle. Arrows indicate optimal aimpoints. Response distributions were generated from a common standard deviation of 4.2mm.

In the experiments they conducted, participants had to reach a static target within a short time limit. The target was composed of two circles which partially overlapped, as shown in Figure 1.2. Depending on the region the finger finished in, a different number of points was received, and participants were instructed to maximize the points across trials. Landing into the region of the hollow circle (on the right in the figure) was rewarded with 100 points, landing outside any circle was neither rewarded nor penalized, and failing to reach any region within the time limit was penalized with -500 points. The outcome associated to landing inside the filled circle varied across conditions. Finally, landing inside the region where the two circles overlapped brought the outcomes of both circles. In one condition, finishing the movement within the filled circle had a consequence of 0 points. In that case, given that the hitting variability was normally distributed around the aimpoint, the best strategy was to always aim at the center of the hollow circle. Adopting this strategy would make that the final accumulated points depended only on the participant's variability, with less variable participants winning more. In another condition, hitting the filled circle was penalized with -100 points, so that landing at the overlapping region made the outcomes of both circles cancel each other out, so 0 points were given. In this case, the best strategy was to shift the aim point away from the filled circle (rightwards in the figure). Although doing this increases the probability of landing outside the hollow circle, it minimizes the chances of falling into the penalty area. This shift should be more pronounced for another condition where landing in the filled circle was penalized with -500 points. In that case, landing outside the circles was less undesirable than hitting the filled circle, so a considerable shift in the aim point should be seeked to minimize this. But how big should the shift be in each case? This depends on the hitting variability of each participant, given the time limit movements had to respect. The more variable a participant is, the bigger the shift to avoid being penalized. With the observed variability of each participant, experimenters calculated their optimal aimpoint for each condition, which is that minimizing the expected loss. Surprisingly, this optimal did not differ from the observed center of the distribution for the participants' responses. In other words, participants were optimal in their movement planning. According to the authors, participants could only achieve this through a knowledge of their hitting uncertainty, which was combined with the knowledge about the task's reward and penalty structure.

Of course, the idea of assessing optimal performance in perception and/or movement was not new. For instance, optimal control theory (see Todorov, 2006) had long been trying to explain movements in terms of minimizing cost functions, and research in multisensory integration showed how we combine visual and haptic information in a maximum-likelihood fashion when estimating some properties of an object (Ernst & Banks, 2002). However, the formalization by Trommershäuser and colleagues offered a simple and elegant way to account for how we integrate our intrinsic perceptuo-motor uncertainty with the extrinsic task constraints, in order to plan for a trajectory or endpoint that maximizes expected reward. This is especially useful when the study of movements one desires to pursue can overlook biomechanical issues, like those related to torque, muscles and joints.

Shortly after the first results were published, these were supported by a series of follow-up studies (for a review, see Trommershäuser, Maloney, & Landy, 2008). Evidence was found for humans being optimal in sensorimotor tasks even when there was a change in target overlap or number of penalty regions (Trommershäuser, Maloney, & Landy, 2003a), or when aiming at targets situated at different eccentricities (Gepshtein, Seydell, & Trommershäuser, 2007). Moreover, participants could optimally compensate for externally added task-relevant variability (Trommershäuser, Gepshtein, Maloney, Landy, & Banks, 2005). If needed, they could also take into account their speed-accuracy trade-offs (Dean, Wu, & Maloney, 2007). Plus, when having to choose between two targets, they picked the one with the largest expected value (Trommershäuser, Landy, & Maloney, 2006). In addition, Hudson, Maloney, & Landy (2008) showed how people also could optimally compensate for their temporal uncertainty when adapting movement times to different reward functions. However, all this evidence was mixed with the one produced by other studies, where results point to suboptimal movement planning. The next section will review that literature, grouping it by the argued source of suboptimality. It must be noted, though, that we will cover experimental work spanning over different types of movement, from simple key press to reaching movements. In all cases, observed performance can be compared with instantiations of SDT that, although adapted to the task, have very much in common with that used by Trommershäuser and colleagues.

## 1.1.4. Suboptimal movement planning

### 1.1.4.1. Not suboptimal, just inexperienced

Some studies defend that, although participants initially display suboptimal behavior in some sensorimotor, with experience they may learn how to be optimal. For instance, Neyedli & Welsh (2013a) found how, through feedback, their participants could learn the relationship between probability and task-defined reward, and progressively shift their aimpoint until it matched the optimal one. Somewhat similarly, Seydell, McCann, Trommershäuser, & Knill (2008) claimed that people can build a generative model to account for arbitrary randomness present in the task, and then act optimally. However, evidence has shown how, in many other cases, participants are never optimal.

### 1.1.4.2. Erroneous probability representations: distortions and overconfidence

An essential aspect of an optimal sensorimotor decision-maker is the fact that an agent makes decisions after successfully integrating two elements. The first is the agent's own uncertainty, which after all is information about probabilities (i.e. the probability of hitting a target given her motor variability). The second is the extrinsically-defined reward function of the task, which is a mapping between each possible action and its outcome (i.e. positive or negative reward). Although it could be that the failure to represent either of the two elements could still lead to an apparently optimal behavior (see section 1.1.4.4.), very easily that could be a source of suboptimality. Let us see evidence for suboptimal behavior when probabilities are not well represented.

Part of the impact of the previous studies was due to the fact that optimal behavior there contrasted with clearly suboptimal behavior shown in traditional decisional tasks, such as those prospect theory was based on. One of the main explanations for that difference revolved around the different way information about probabilities was acquired: in traditional tasks, it was explicit (i.e. prob. 0.8), while in motor tasks it was implicit in the participant's variability, which the participant was assumed to have some knowledge of. To test this, some attempts were made to design traditional and motor tasks that were completely equivalent. Wu, Delgado, & Maloney (2009) found that, while results of traditional tasks were replicated in terms of probability distortion, motor tasks also showed a distortion, but in the opposite pattern: underestimation of small probabilities and overestimation of large probabilities. Moreover, participants were more risk-seeking in motor tasks. The fact that probabilities were distorted in the motor task means that they did not act in an optimal way. The two task formats were also compared by Jarvstad, Hahn, Warren, & Rushton (2013). They found that, although their participants were quite good in both classic and sensorimotor tasks, they were not optimal, and that individual differences explained performance better than the type of task.

In some cases it could also be that, instead of sometimes being overestimated and sometimes being underestimated, probability information was distorted always the same way. One of the phenomena leading to such a situation is overconfidence. This can be defined as the overestimation of one's performance, as measured when comparing someone's actual performance with her beliefs about it (Olsson, 2014). In our daily lives we face multiple situations where an excess of confidence may take its toll. For instance, a baseball batter may underestimate his reaction time while waiting for the ball to be closer in order to hit it harder. Unfortunately, the batter may not swing the bat fast enough and the ball may end up at the catcher’s mitt.

Mamassian (2008) found overconfidence in a synchronization task where participants had to press a key when the third visual cue of a sequence appeared. A reward was provided if the response was within a narrow time interval of the visual event. Nevertheless, depending on the trial, a penalty was applied if they responded immediately before or after the reward interval. Before each trial, participants were informed about the location of the penalty interval. Considering that motor variability is present, when the penalty interval was after the reward interval, the optimal response time was earlier than the time of the visual event. That would maximize the expected reward, since it minimized the probability of responding during the penalty region. The larger the motor variability of a participant, the earlier this optimal time shift should be. Alternatively, when the penalty region was before the visual event, expected reward maximization should have required shifting responses forth in time. It can be easily seen how this timing task is very similar to Trommershäuser and colleagues’ original pointing task. However, here observed time shifts were mild compared to optimal ones, which were calculated by convoluting motor variability and temporal reward function, as seen in Figure 1.3. To disentangle whether this was due to participants not understanding the payoff distribution or to them having a misrepresentation of their variability, two more experiments were run. When the reward function was altered, participants shifted their responses in a coherent direction (although the shift was still suboptimal). On the other hand, participants did not update their time shifts at all when their variability was artificially increased through the addition of external noise to their responses. Thus, suboptimal performance was claimed to be mainly a product of participants being unaware of their actual motor variability, and especifically underestimating it.

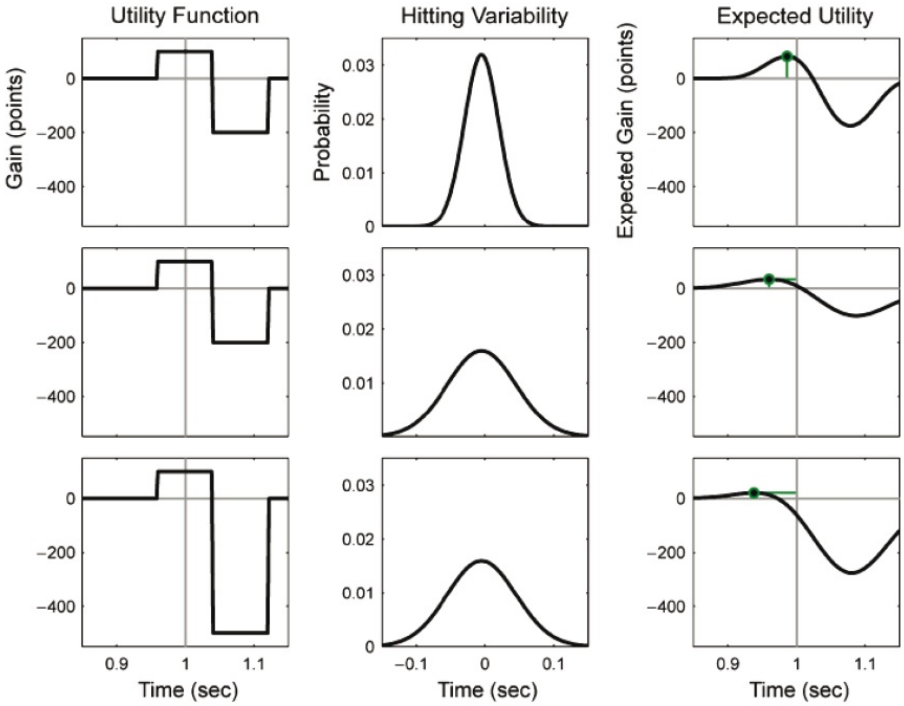


Figure 1.3. Calculating the optimal response time in Mamassian (2008). The top-left plot shows the task's reward (here utility) function, with the third visual cue appearing when

time = 1s, a reward (here gain) region awarding 100 points if the key was pressed during a period spanning ± 0.025s of the cue presentation time, and a penalty region subtracting 200 points if the key was pressed 0.025s to 0.075s after the cue. The top-central plot shows how the gaussian hitting variability would be distributed if the participant had a standard deviation of 0.025s. The top-right plot shows the result of convoluting the reward function and the hitting variability, with the optimal response time (in green) being the time that maximizes the expected reward. The middle row shows how the optimal response time would change if the participant had a hitting standard deviation of 0.05. The third row changes the penalty region with respect to the middle row, increasing the penalty up to 500 points.

Overconfidence has also been observed in similar paradigms. Ota, Shinya, & Kudo (2015) found that optimality in a simple timing task depended on the reward function. When it was symmetric, participants behaved in an optimal way. That happened both when the function was continuous, with reward progressively increasing with time until a maximum and then decreasing at the same rate, and also with a step function where a no reward area became a stable reward area, then a no reward area again. However, suboptimality arose when the reward function was asymmetric, with a temporal sequence where a continuously increasing reward was suddenly followed by a no reward area (or the opposite). This suboptimality acquired the form of overconfidence (or risk-seeking, as the authors called it), for observed responses were too close to the no reward area when compared to an optimal prediction. In another study, the authors saw that overconfidence persisted even after several days of practice (Ota, Shinya, & Kudo, 2016). Considering the difference among conditions, it could be argued that participants were simply having a problem with the reward function, not with their own uncertainty. However, the direction of the suboptimality matched overconfidence. Moreover, in those conditions of the experiment where optimality was found, the task simply required synchronizing the response time with the point of maximum reward of the reward function, without the need to calculate any response shift with respect to that point. In those cases, aiming at the maximum reward time was optimal because variability in that task was distributed around the aim point in a symmetric, gaussian way, and the reward function was symmetric as well. Nevertheless, other empirical work does show how suboptimality could be due to people having trouble with the task's reward function, as shown in the next section.

As a final piece of evidence, results coherent with overconfidence have also been found in movements more complex than key presses, like arm reaching and whole-body movements. O'Brien & Ahmed (2013) set participants in a situation where they had to move (either their arm or lean their body) forward, and the further away they stopped their movement in a trial, the more points they won. However, moving beyond a certain point (a "virtual cliff") was penalized. In a series of conditions where the externally-imposed penalty and cliff were manipulated, participants' final movement positions were too close to the cliff when compared to an optimal agent that maximized expected reward. This was interpreted as risk-seeking.

#### 1.1.4.3. Difficulties representing the reward function

Even if the decision-maker had a perfect representation of her uncertainty, she could struggle with the task's reward function. This could be rapidly changing, or simply too complex.

Following the logic about the difference between implicit and explicit information, some work points to suboptimality arising when rapid changes are produced in the structure of the explicit part of the information that has to be used to make the decision. Neyedli & Welsh (2013b) made participants point at one of the already familiar overlapping circle configurations. In one condition, the overlap (and thus the probability of hitting the penalty area) changed from trial to trial. Participants immediately shifted their aimpoints to be optimal. However, they failed to do so in another condition where the penalty value changed across trials. According to the authors, while some time might be needed to adapt to the new payoff structure, a change in the visual display is straightforward to interpret by the motor system.

Another source of suboptimality could be that the optimal strategy may be too difficult to learn. Wu, Trommershäuser, Maloney, & Landy (2006) found suboptimal movement planning in a task with two penalty regions where the value of one penalty was different from the other. In this asymmetric configuration, participants could not adopt the best aimpoint, which was located within the milder penalty area.

#### 1.1.4.4. From optimal to suboptimal: model limitations

A normative model should be able to establish the limit of optimal performance independently of task difficulty. However, when comparing the results of a pointing task with optimal simulations generated using Trommershäuser and colleagues' model, Jarvstad, Hahn, Warren, & Rushton (2014) saw that an increase in task difficulty turned optimal participants into suboptimal. This was done by decreasing target size and increasing target distance from the position where movements began. The authors argue that the model does not capture the fact that people may be failing to maximize precision. In their experiments, participants moved faster than necessary for close targets and relaxed their precision for larger targets. Since the data used to simulate optimal behavior comes from the participants' observed variability in the experiment, this decrease in precision for easier stimuli influenced optimal predictions, what made participants look more optimal than they really were.

Even some of the authors of the original model express their concern about how participants with a wrong representation of their motor uncertainty could still be classified as optimal by the model: for instance, a participant that assumed that her movement endpoints were uniformly distributed around the aim point, while the true distribution was gaussian (Zhang, Daw, & Maloney, 2013).

## 1.2. Perception and action with moving objects

### 1.2.1. Relevant topics in motion perception

As said in the very beginning, this section will only focus on topics of special relevance for the thesis. For a complete review of the advancements in the study of motion perception from the 60's until the 2010's, the reader is recommended the reviews by Nakayama (1985), Burr & Thompson (2011) and Nishida (2011). These have been the main secondary sources used to compose this section.

#### 1.2.1.1. Motion detectors

Changes in luminance are the main physical attribute our visual system uses to detect motion. The most popular models originate from Reichardt's model (Reichardt, 1957). This is based on a simple mechanism. Imagine two retinal receptors with spatially-contiguous receptive fields: for instance, one on the left and one on the right. These would send information to the same mediating neuron, but their timing would depend on the motion direction. If the rightmost receptor's output arrives later, the motion is interpreted as rightwards, and vice versa. Later proposals evidence the need for motion detectors to also act as spatio-temporal filters (Adelson & Bergen, 1985; Morgan, 1992; van Santen & Sperling, 1985; Watson & Ahumada, 1985) that perform a Fourier-like analysis, with the spatial filtering happening before the direction analysis (Morgan, 1992). In essence, motion detectors are temporally and spatially tuned, and direction-selective. The way these detectors would obtain information about speed is debated: either through a normalisation process by comparing detectors tuned to leftward, rightward and static energy (Adelson & Bergen, 1986), or by creating vector motion sensors after grouping linear filters with similar spatial frequency and location but different detection tuning, and then analyzing the temporal frequency of their responses (Watson & Ahumada, 1985).

#### 1.2.1.2. Second-order motion

So far we have described luminance-based motion, also called first-order motion. However, different work with lab-created stimuli (Badcock & Derrington, 1985; Chubb & Sperling, 1988; Derrington & Badcock, 1985; Derrington & Henning, 1987) showed how, even in absence of Fourier energy, participants (but not Reichardt-based models) could still detect motion. This second-order motion would mainly be based on changes in contrast, rather than in luminance, and would have a higher detection threshold (A. T. Smith, Snowden, & Milne, 1994), which would be based on a minimum displacement rather than a minimum speed (Seiffert & Cavanagh, 1998; A. T. Smith et al., 1994). Although evidence exists supporting both the same and different mechanisms subserving the processing of both kinds of motion, the latest evidence leans towards the latter hypothesis, at least for fronto-parallel motion. More complex motion like expansion might be based on displacement thresholds (de la Malla & López-Moliner, 2010).

### 1.2.1.3. Temporal resolution of the visual system

How fast can our visual system process changes in visual stimuli? The figures that have been reported enormously depend on the way stimuli are presented. For flashes of light presented at the same location, the flicker fusion rate ranges between 30 Hz and 60 Hz (Kelly, 1972), corresponding to periods of 33ms and 17ms, respectively. However, this limit in the temporal resolution of photoreceptors is lower with moving stimuli. When motion cues can be used, temporal order discrimination thresholds can be as low as 17 ms (Exner, 1875), and the precision can reach 2-5 ms when there is also adaptation (Sweet, 1953). For apparent motion experiments, there is evidence of thresholds of 2-4 ms (Westheimer & McKee, 1977). The fact is, then, that our temporal resolution seems to get better when there is motion. In these cases, an interesting question is whether the motion system may be responsible for this precision enhancement.

### 1.2.1.4. Integration time for speed perception

In order to obtain reliable estimates of an object's speed, signals are integrated over space (Watamaniuk & Duchon, 1992), but also over time. As seen in Figure 1.4., Snowden & Braddick (1991) showed how the difficulty in discriminating two speeds decreased as the stimulus' presentation time increased, eventually reaching an asymptote (as predicted by Bloch's law). This asymptote was reached earlier for faster speeds. However, the authors defend that this does not mean that temporal integration uses the total motion path length, but that faster speeds are dealt with by motion detectors with shorter temporal spans (see also McKee & Welch, 1985). Using Reichardt's motion detector model, the temporal span could be understood as the difference in arrival time between the non-delayed and the delayed signal when being stimulated at once (Snowden & Braddick, 1989). Thus, since temporal delays in their motion detectors are shorter, faster speeds require shorter integration times to obtain a reliable speed estimate.

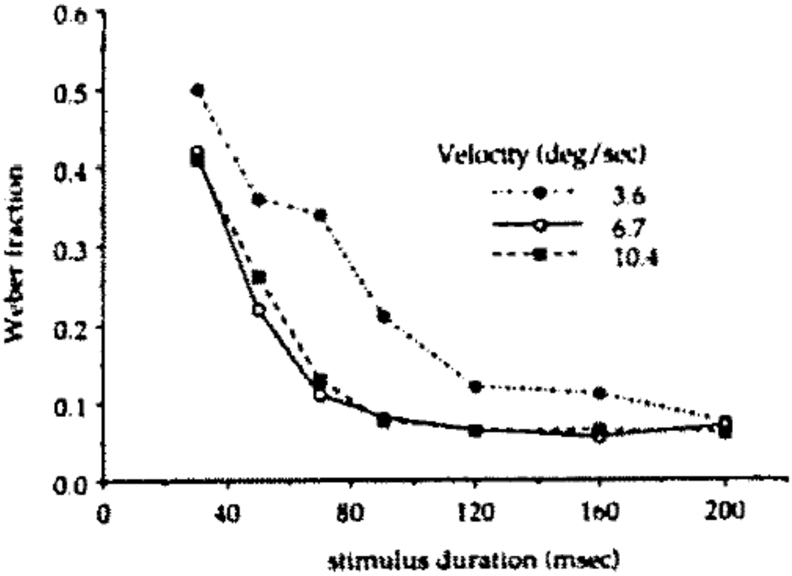


Figure 1.4. Weber fraction of velocity discrimination as a function of stimulus duration, reproduced from Snowden & Braddick (1991). Data is shown for one participant, and split by target velocity. As can be appreciated, the longer the stimuli are presented, the smaller the difference between target velocities that is needed to discriminate them. Stabilized Weber fractions can be interpreted as integration times having reached the asymptote.

Integration times depend on the stimulus type, with longer times for more complex motion. For contrast summation, integration time would be about 100 ms (Burr, 1981). On the other hand, according to some accounts (Burr & Santoro, 2001; Melcher, Crespi, Bruno, & Morrone, 2004; Melcher & Morrone, 2003; Neri, Luu, & Levi, 2006), two or three stages would be produced when analyzing complex stimuli, such as optic flow and biological motion: a local motion analysis (time constant of 200 - 300 ms), an intermediate stage (integration time of 1000 ms) and a global motion stage (time constant of 3000 ms). In the case of biological motion, integration time could exceed 8 s (Neri, Morrone, & Burr, 1998).

### 1.2.2. Interactions between position and motion

Our experience tells us that objects hardly undergo a sudden jump in their position. Slower or faster, a position change is normally mediated through motion. This strong correlation between motion and position, though, is sometimes dissociated. For instance, motion after-effects have been known for centuries (Addams, 1834) and are used to exemplify how motion can be experienced without a change in position (see Anstis, Verstraten, & Mather, 1998). But because motion and changing position normally happen together, sometimes the presence of one (motion) induces perceived changes in the other (position). The De Valois illusion (De Valois & De Valois, 1991) shows how a position shift is created by pattern motion (see below).

The dependencies and dissociations between motion and position perception made some investigate the extent to which the visual processing of visual movement is subserved by spatial or temporal mechanisms (de la Malla & López-Moliner, 2010; Nakayama & Tyler, 1981; Seiffert & Cavanagh, 1998). Close to the aim of the previous section, the present part will not try to review all the literature connected to this discussion. The scope here will rather be presenting issues on the interaction between motion and position that will later be relevant for the thesis' findings. We will start by two examples of how biases and variability in acknowledging the position of a moving object can be influenced by motion signals. First we will introduce motion extrapolation, where the perceived position of a moving object is biased in the direction of motion. However, it must be noted that we refer to experiments where the moving object was visible at all times. Considering the design and implications of the work of this thesis, motion extrapolation after target occlusion will not be discussed here. In case the reader is interested in that topic, the introductory part in Khoei, Masson, & Perrinet (2013) acts as a helpful review. The next part will present a hypothesis about how we report the position of a moving object after a cue. This has implications in terms of spatial variability, which depends on the object's speed. We will finish by presenting a Bayesian model that tries to integrate motion and position perception, and which accounts for the previous phenomena described in the section.

#### 1.2.2.1. Motion extrapolation

When perceiving an object, information takes some time to be transmitted along the visual pathways. If the object is moving, this delay means that when a percept about the object's position is formed, this is actually informing us of where the object was some instants ago, not of its current position. To compensate for these delays in information transmission, it has been suggested that some mechanism may bias an object’s perceived position in the direction of motion (Berry, Brivanlou, Jordan, & Meister, 1999; De Valois & De Valois, 1991; Jancke, Erlhagen, Schöner, & Dinse, 2004; Nijhawan, 1994; Sundberg, Fallah, & Reynolds, 2006).

Motion extrapolation has been shown even with static envelopes (De Valois & De Valois, 1991) with pattern motion embedded within the carrier. When fixating at these stimuli with the periphery, their position is seen as displaced towards the motion direction. However, most experiments trying to assess the existence of motion extrapolation come from studies based on what is known as the “flash-lag”. In this paradigm, an object moves and at some point a flash is suddenly presented, and the participant has to report the position of the object at the moment of that event. When the flash is presented aligned with the moving object, the latter is seen ahead of the former. Initial claims of motion extrapolation in this paradigm (Nijhawan, 1994) were later called into doubt by alternative explanations. According to Whitney & Murakami (1998), the neural delays for moving stimuli would be shorter than for flashed stimuli. According to Eagleman & Sejnowski (2000), it is the trajectory of the moving object after the flash, not before it, that would displace the perceived position. Other contributions have tried to defend motion extrapolation in the flash-lag effect describing it as the result of a predictive mechanism (Khoei, Masson, & Perrinet, 2017).

Beyond the discussion of whether motion extrapolation causes the flash-lag effect, other studies have focused on different setups. A particularly relevant situation is that where smooth pursuit is allowed: that is, the moving object can be followed with the gaze. This is common in most real life situations involving interaction with a moving object, in contrast with the aforementioned work, which made participants report the perceived position in tasks where they had to fixate a point. Studies addressing the localization errors of a flash while pursuing a moving target have shown that perceptual localization errors are reduced with respect to fixation (Kerzel & Gegenfurtner, 2003; van Beers, Wolpert, & Haggard, 2001).

### 1.2.2.2. Sampling the position of a moving object

In section 1.2.1.3. we have mentioned the temporal resolution of the visual system as a lower bound to detect changes in stimuli. When reporting the position of a moving object, this resolution constrains our spatial precision. This affirmation has been defended by some (Brenner & Smeets, 2000; Brenner, van Beers, Rotman, & Smeets, 2006) with the following reasoning. Reporting the position of the moving object can only be done with respect to another event. For instance, a cue that indicates that it is time to identify the position of the object. After the cue is presented, the position of the object will be sampled, but since this process takes time, a bias will be found in the reported position, which will be in the direction of the movement. The bias is explained by the temporal resolution of the visual system imposing a minimum sampling time. From this follows that the spatial variability in a localization task should increase for faster speeds. After all, for faster moving objects, the fixed temporal delay corresponds to a greater travelled distance. This sampling hypothesis was mainly used by the authors as a criticism of the motion extrapolation hypothesis in the flash-lag effect, since they found that the magnitude of this effect depended on the speed of the moving object after the flash, not before it. However, implications for spatial variability have been found in different occasions. Brenner et al. (2006) found that faster target speeds were related to increased spatial variability when having to report the position of a moving target after a flash or a tone had been presented. Linares, Holcombe, & White (2009) found similar results, with the distribution of responses matching a 70ms delay respective to the presentation of the cue. This held with many types of cue and different target trajectories.

### 1.2.2.3. An integrative framework of motion and position

In the preceding sections we have mentioned that motion and position are highly correlated. Had we evolved to optimize our resources, it would be sensible if we were able to exploit that redundancy by integrating motion and position in our perceptual estimates. Let us take a problem we sometimes experience: disambiguating the source of motion within an object. Imagine a moving ball, which is characterized by translation motion, but also by rotation. When our visual system processes local motion signals, it has to attribute that motion to one or the other source. A solution to disambiguate the motion source could be using the position information that is integrated over time. This, though, would need the aforementioned coupling between position and motion, where information about one helps inferring the cause of the other. Taking advantage of the correlation of moving stimuli over time is also useful for other purposes. Given the noisiness of our sensory signals, we could weigh incoming visual input according to its reliability and combine it with prior experience (information about the stimulus gathered earlier in time) in order to generate estimates regarding the current state of a moving object.

Could there be a common computational explanation to account for how we track objects over time and infer their generative causes? Kwon, Tadin, & Knill (2015) proposed one such model, where motion and position perception are integrated into an object tracking system. The model is based on a Bayesian process that tracks objects while relying more on past signals if the present sensory information is uncertain. It also tries to assign the origin of motion from our sensory measurements. To do this, this tracking system would also rely on an internal model of the world that incorporates certain priors arising from the phenomena that are most likely to happen: for instance, a prior towards objects slowing down, and another one where, in the face of uncertain position signals, motion is attributed as belonging more to global object motion rather than local pattern motion. Because the model tries to infer the generative causes of motion, it would be classified as a causal inference model. And because the noise corrupting our sensory estimates is assumed to be Gaussian, this Bayesian approach can be defined as a Kalman filter (Kalman, 1960).

This model provides the mechanistic explanation to a series of visual phenomena, such as the curveball illusion (Shapiro, Lu, Huang, Knight, & Ennis, 2010), where an oval moving downwards with embedded leftward carrier motion is perceived as drifting along a diagonal down and leftwards, or the aforementioned DeValois illusion (De Valois & De Valois, 1991). But the model can also predict motion extrapolation even with stimuli with no pattern motion. This happens when the relative reliability of velocity is higher than that from position signals, leading to extrapolation due to velocity states being weighted more heavily.

### 1.2.3. Acting to affect the course of a moving object

In those daily life situations where we have to interact with moving objects, we are seldom given a signal instructing us to start our movement. Instead, we have to decide how to set an internal criterion for action. The placement of this criterion is normally made with respect to some variables relative to the course of the moving object. A straightforward way to do so would be by tracking the position of the moving object until it reaches a point in space, and then start the action. In most catching and interception tasks, where the object approaches and thus there is radial motion (motion in depth), this strategy would be translated into acting when the object reaches certain retinal size. However, a lot of the literature has tried to explain our actions as based on an estimation of the object's time to contact (TTC). This refers to the time that the object will take to reach the point of observation. Of course, in sports situations involving balls, this point of observation is the person's position. But in other situations with lateral displacement, like clay pigeon shooting, the TTC can be calculated taking another reference. The following lines introduce the TTC literature, discussing some of the variables that could be used to estimate it. Then we discuss some ways people can use to signal the start of their actions.

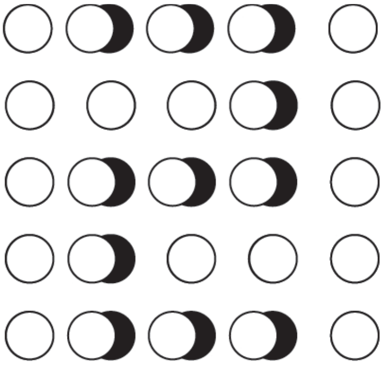
### 1.2.3.1. Estimating time to contact

Many proposals of TTC estimation have been formulated up to date. Some of the most successful ones are those derived from the gibsonian, or ecological, approach. According to this, rather than relying on complex computations, human beings employ variables immediately available from the environment. In line with this claim, some evidence (Lee, 1976; Regan & Hamstra, 1992; Regan & Vincent, 1995) supported TTC estimation to be based on tau ($\tau$). This is simply the ratio between the object's size ($\theta$) and its rate of expansion ($\dot{\theta}$). These are monocular variables that are available from the retinal image. However, estimation is not perfect when angular size is large, or when speed is not constant. Further studies have shown that, when other cues are available, these may be used as well: for instance, size knowledge would be important in grasping (Savelsbergh, Whiting, & Bootsma, 1991). Some of these cues can also be binocular, and they are more important when monocular information is not reliable (Gray & Regan, 1998; López-Moliner, Supèr, & Keil, 2013; Regan, 2002). Other work has tried to demonstrate that $\theta$ and $\dot{\theta}$ need not be combined as in $\tau$ (Keil & López-Moliner, 2012; López-Moliner & Bonnet, 2002; M. R. Smith, Flach, Dittman, & Stanard, 2001). The predominance we give to one or another cue could also change along the trajectory of the object (López-Moliner et al., 2013). In any case, some of the latest work assumes we can use many cues, and focuses instead on how we integrate them (de la Malla & López-Moliner, 2015).

TTC estimation variables have also been studied for lateral displacement. Tresilian (1990) suggested three variables that could be used: angular distance between the moving object and the reference, angular velocity of the object, and the local $\tau$. Tresilian (1994) affirmed that angular position would be used to calculate TTC. Bootsma & Oudejans (1993) claimed that the main cue for TTC estimation in lateral displacement is the relative rate of constriction of the optical gap that separates a moving object from the reference position. However, Smeets, Brenner, Trébuchet, & Mestre (1996) deny the use of gibsonian cues and defend that $\tau$ is not used with lateral displacement. Alternatively, they point towards the ratio between perceived distance and perceived velocity. This means that TTC is estimated from the ratio of two perceived variables, where both have to be transformed to spatial coordinates. Thus, contrary to the $\tau$ hypothesis, errors in perceived position or velocity will lead to errors in TTC estimation, which is what this study finds.

#### 1.2.3.2. Decisional variables with moving objects

One of the critiques addressed to $\tau$ accounts is the fact that they are purely phenomenological. Even at the neural level, they are just limited to fit a model that describes the firing pattern of the neurons, but fail to explain why neurons are activated. At the behavioral level, this would mean that they cannot account for the moment that would trigger an action in response to the visual information: for instance, hitting a ball or dodging something that is approaching. Hatsopoulos, Gabbiani, & Laurent (1995) found some neurons in the locusts' nervous system that fired as a function of what they called eta ($\eta$): the object's velocity multiplied with an exponential function of the object's retinal size. As opposed to $\tau$ not reaching a minimum while decreasing, the value of $\eta$ peaks before the retinal image is the maximum, thus allowing for collision-avoidance mechanisms. For equal speeds, bigger objects would make the peak be reached earlier. Sun & Frost (1998) also found these type of neurons in the nucleus rotundus of pigeons, along with others that fired as a function of $\tau$, and others that fired as a function of rho ($\rho$), the absolute rate of expansion. Keil & López-Moliner (2012) implemented a neurophysiologically plausible model to estimate TTC and implement collision-avoidance reactions. The model, that could account for both $\tau$ and $\eta$ sensitive neurons, allowed for collision-avoidance reactions to be initiated if some parameter reached a certain value. Thus, it seems that we could have evolved to use some sort of internal threshold in order to initiate some actions relative to moving objects.



# Aims of the thesis

Imagine a competition racer trying to decide when to steer the wheel to save time at a curve, a clay pigeon shooter waiting to pull the trigger, or a football goalkeeper trying to capture an approaching ball. All these situations share some common features. Firstly, there is an action to be made relative to the course of a moving object. Secondly, acting at one moment will not lead to the same outcome as acting at another moment. Finally, the action will previously require a decision on when to act. This decision is assumed to be based on different information: our visual sensory stimulation, a knowledge of how each action is related to an outcome, and a knowledge of our uncertainty when performing the planned action.

The study of human motor interaction with moving objects has so far not focused on the decisional process we have just sketched. From the opposite point of view, it can also be said that the field of sensorimotor decision-making has eluded situations involving interaction with moving objects. This can be exemplified by the introductory chapter of this thesis, composed by two big sections. One presented the study of sensorimotor decision-making, especially within a SDT framework. The other reviewed how we interact with moving objects. Despite both sections referring to literature about perception and action, there are very few obvious connections between the two. The present thesis aims at establishing a bridge between them. But why should we be interested in that? By formulating the main objectives of this thesis we will try to answer this question, and make clear how these two fields can benefit from each other.

The field of sensorimotor decision-making tries to understand how we combine different sources of information to make a decision. One of these sources, as already mentioned, is our sensory information. Yet sensorimotor situations involving moving objects are interesting in the sense that, given the spatiotemporal nature of the object's trajectory, our decision can be based on the use of different types of visual information. On one side, spatial information, that would derive from the retinal position the moving object has at each moment. On the other side, temporal information, coming from motion and velocity cues. Maybe exploiting more one type of information can lead to a better performance. Thus, **objective #1 is finding whether basing our decisions more on temporal or spatial information can improve performance in sensorimotor situations with moving objects**.

The study of our interactions with moving objects can use the formulations of SDT that the field of sensorimotor decision-making has been using for years. By framing performance in terms of the reward related to the outcome of our actions, we can use SDT to provide an assessment of optimal behavior. At the same time, extending previous SDT models to the study of moving objects will let us see its generalizability and formulate caveats regarding their limitations. **We can formulate objective #2 as testing how current SDT models used to study sensorimotor decision-making can provide a valid normative account of interaction with moving objects.**

The first objective dwells on how exploiting certain information can improve performance, where improvement is understood as achieving actions with a better outcome. The second objective has the definition of optimality at its core. Optimality can be loosely defined as obtaining the best possible outcome given some existing constraints. Deciding based on the outcome, then, plays a central role in both previous objectives. In our experiments, we implemented the outcome by giving explicit reward after each trial. **Objective #3 is to study how people use reward**

**information to plan their future actions. This is investigated both on a short timescale scale, where a certain reward can influence the decision on the next trial, and on a longer timescale, where reward shapes decisional criteria and the visual cues that are exploited.**

In the following lines we will articulate the specific questions that each of the studies of the thesis seeked to answer, and the way they relate to the global objectives. Since some studies were designed as follow-ups or to answer some questions previous studies had left unanswered, a brief outline of the results is also given.

**Study I**

- Does exploiting more temporal or spatial information improve performance when interacting with moving objects? This, that we tested in a coincidence anticipation sensorimotor task, directly tackles objective #1. We found improved performance, as measured through average reward per trial, for an increased use of temporal information.
- Is the type of information used influenced by the total time the object moves? We found that targets with longer motion times facilitated the use of temporal cues.
- Do SDT-derived optimal motor planning models offer a good measure of optimality in our task? This is exactly what objective #2 aimed at. Although current models determined general suboptimal performance, some inconsistencies revealed their limitations to interpret our data.

**Study II**

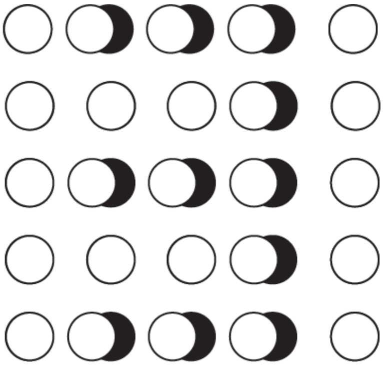
- Given the advantage of exploiting temporal information when interacting with moving objects, can people learn how to use temporal cues to plan their responses? This is an extension of objective #1.

- Can this learning take place through the manipulation of each trial’s outcome, so that exploiting the wrong cues leads to less reward? This implementation concerns objective #3, since we expect manipulated reward to influence future behavior, and specifically the decision variable. We found that our manipulation did not produce evidence of learning taking place, and that those people exploiting temporal cues did it from the very beginning, probably caused by individual differences.

**Study III**

- Can suboptimal behavior arise from participants planning their responses by not taking into account their whole observed variability? We found evidence for some participants responding at the point that maximized expected reward if only their measurement variability was considered. Since optimal responses were calculated from SDT models, this relates to objective #2.
- How does reward affect the next responses? We investigate the way corrections and aimpoint selection are driven by reward. This makes objective #3 more specific and applies it to some uninvestigated yet important phenomena from our previous studies. We find that participants use reward information to trigger corrections in order to avoid penalizations. We also hypothesize that they stabilize their aimpoint around specific reward regions.

An important element of our tasks must be justified. While we consider our work as studying sensorimotor decision-making, our experiments simply require that participants execute a motor response by pressing a key. Recent evidence has found how decisions can be influenced by the cost to act (Hagura, Haggard, & Diedrichsen, 2017; Marcos, Cos, Girard, & Verschure, 2015). To avoid a bias arising from the motor cost of the action, we keep this constant and control for it by implementing button press as the motor response. By doing this we also make non-applicable some of the limitations of SDT models denounced in section 1.1.4.4., that would be a concern if our experiments required more complex movements, such as reaching.



# Study I - Exploiting temporal information enhances performance in sensorimotor decision-making with moving objects

## 3.1. Summary

When interacting with moving objects, people can base their actions on temporal (motion), spatial (position tracking) information or their combination. We aimed at knowing, in such situations, to what extent performance depends on the type of information used to respond, and whether the information used changes in response to the time the stimulus can be seen. In our task, a target which could vary in speed and motion time approached a line. Participants stopped the target and were rewarded according to its proximity to the line. Responding after the target crossed the line was penalized. We identified whether responses were based on monitoring the target's changing spatial position or on time-based motion cues, and found that performance grew with more time-based responses. We can explain the results with two compatible mechanisms. Firstly, the temporal resolution of the visual system increasing spatial variability with faster speeds, bringing more penalizations when using spatial information. Secondly, increased position extrapolation across speeds, consistent with longer motion times allowing more motion integration and facilitating the use of temporal information. Our study evidences how exploiting motion information can help overcome some limitations of position tracking, and points out the need to consider the different sources of variability when providing a normative measure in situations with moving objects.

## 3.2. Introduction

In many daily life situations, such as those involving actions with moving objects (e.g. catching a flying ball or dodging an approaching bike), successful interaction requires updating the changing position of these objects efficiently. Motion generally concurs with changes in the object's spatial position, and these two types of information are interdependent (De Valois & De Valois, 1991; Linares, López-Moliner, & Johnston, 2007; Maus, Fischer, & Whitney, 2013; Whitney, 2002), although they can be processed independently in goal-directed actions (Smeets & Brenner, 1995). Given the noise present in sensory signals, it has been proposed that visual motion (e.g. through accurate velocity estimates) can be used to update the position more efficiently when this source is less reliable (Kwon et al., 2015), so humans would have the possibility to use either or both kinds of information. Sensorimotor actions could benefit from these proposed interactions between motion and position. For instance, a baseball batter may decide when to swing her bat based on some remaining time margin, be it obtained from optical variables (Lee, 1976), or based on an internal speed estimate (Smeets et al., 1996). Alternatively, she may base her decision to act on a spatial tracking of the ball's position that informs her on how close the ball is. We know that space or time-based information can be used as decision variables when timing actions with approaching objects (Bootsma & Oudejans, 1993; Keil & López-Moliner, 2012; Lee, 1976; López-Moliner & Keil, 2012; Smeets et al., 1996; Tresilian, 1990, 1994). An added difficulty in many scenarios is, though, that the decision-maker will commonly face constant changes in the environment: the speed of the moving object will not always be the same, just like the time available to accumulate evidence and make the decision.

Up to date, it has not been examined how, in changing situations where people have both temporal and spatial information available, the cues used to plan the response determine performance. This is what we aimed at in the present study. Specifically we wanted to investigate whether responses based on one type of information enhance performance when compared to responses based on the other. When planning an action involving a moving target, using spatial information to track the changing position of the target would compromise performance. This would be due to increased spatial variability for faster target speeds (schematised in Figure 3.1.A), caused by the limited visual temporal resolution (Linares et al., 2009). Furthermore, given that extracting motion information from a stimulus requires a minimum integration time (Burr & Santoro, 2001; Neri et al., 1998), responses could be limited to the use of spatial information in situations in which the moving object is seen during a short time, and thus reliable velocity estimates are hard to obtain.

In line with a common practice in sensorimotor decision-making studies, we defined performance as the reward that participants accumulated in a series of trials where they received positive or negative feedback as a function of their responses. In some previous studies, responses could only be planned by using temporal information (Mamassian, 2008; Ota et al., 2015). In others, performance depended on correctly hitting a series of spatially-defined reward or penalty areas (Gepshtein et al., 2007; Hudson, Wolfe, & Maloney, 2012; Neyedli & Welsh, 2013b; O’Brien & Ahmed, 2013; Trommershäuser et al., 2005, 2006, 2003a, 2003b, 2008). In contrast, in our experiment participants could plan their responses based on temporal or spatial information.

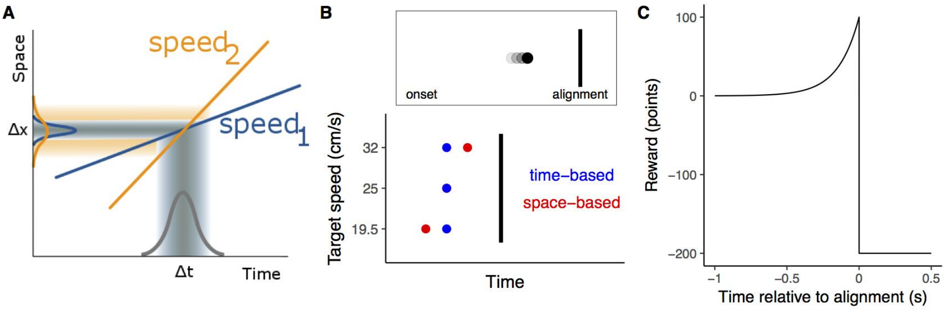


Figure 3.1. A. Spatial variability as a function of the temporal resolution for two different speeds (denoted by the oriented line in the space-time plot). Speed 2 is faster than speed 1 (shallower slope). When sampling the position at a given temporal resolution (illustrated by the Gaussian width on the x axis), the spatial variability of the sampled position will depend on the speed of the moving target. As a consequence, variability increases for faster speeds. B. The top inset depicts a representation of the stimuli. A circle appears at a starting position and moves rightward toward a vertical line. The circle stops when the participant presses a button. The bottom part shows the difference between time-based and space-based responses across target speeds: time-based responses are made at a similar time to alignment, while space-based responses are made later in time for faster speeds. The depicted location of the blue dots relative to the red ones does not obey any prediction, since it depends on the threshold used by every participant C. Reward given as a function of the time remaining to alignment in the experimental task.

The task we used was similar to those commonly referred to as coincidence anticipation (López-Moliner, Field, & Wann, 2007; Tresilian, 1995). As seen in Figure 3.1.B, at every trial a target underwent constant rightward motion toward a vertical line, and participants had to press a button to stop the target. A numerical reward was given according to how close the target was to the line when the button was pressed. Reward increased exponentially as a function of the time left for the target to align with the line, with the maximum reward given if the participant pressed the button at the moment of alignment. However, any response after alignment was penalized. Participants were instructed to accumulate as many points as possible across trials. From trial to trial, the target could change its speed. We identify time-based responses as those where mean response patterns are at a similar remaining time to alignment across all speeds. Alternatively, space-based responses are those where responses across speeds are at a similar spatial distance from the line. This way, we can distinguish the type of information participants were using when planning their responses, and whether the use of one type of information yielded a higher reward than the other. Introducing trial-to-trial changes in initial time to alignment enabled us to determine whether the information used changed across the different motion times.

Performance was analyzed at the level both of the responses and of the reward these responses were translated into. To anticipate, results confirmed our hypotheses: we found that performance was enhanced through the use of temporal cues. This resulted in more stable performance across target speeds. We also confirmed that, within participants, the type of information used was modulated by the total motion time participants were exposed to. While space-based responses were dominant for shorter motion times, longer motion times facilitated the use of temporal information, probably by refining speed estimates, and led to improved performance.

The present study had another objective. Note that, as just described, in our task participants received a numerical reward after each response. This way, participants had univocal feedback informing them of their response time relative to alignment. Participants had to decide how to respond based on the points that they would win given the asymmetric reward function, rather than always trying to synchronize their responses with an objective event, such as alignment. In our experiment, the maximum reward was produced when responding exactly at alignment. Nevertheless, considering that participants inevitably exhibit variability in their responses, they should have tried to respond before alignment, so that they did not incur into penalizations for responding after the line. Is there any response time for which reward in the long term is maximized? Previous work (Mamassian, 2008; Ota et al., 2015) has already assessed whether responses in coincidence anticipation tasks can be optimal. In these studies, observed responses are compared to simulations that maximize expected reward, in accordance with some frameworks derived from SDT. Choosing the action that maximizes the expected reward is achieved through integrating knowledge about the reward function (the mapping between each possible state of the world and its outcome) with knowledge about the agent's own variability (uncertainty) (Körding & Wolpert, 2006; Wolpert, 2007). We used these models to determine whether, when interacting with moving objects, performance could reach optimality in some situations. At the same time, though, we wanted to assess how idoneous these models were in the situation we were studying, and identify their limitations as a normative reference for our data.

## 3.3. Methods

### 3.3.1. Participants

21 participants took part in the experiment after giving informed consent (11 women, all right-handed, age range 18-32). All had normal or corrected-to-normal vision and were naïve about the aim of the experiment. The study complied with the local ethics guidelines, in

accordance with the declaration of Helsinki, and was approved by the University of Barcelona’s Bioethics Commission.

### 3.3.2. Apparatus and stimuli

Participants sat in a dimly lit room, approximately 50 cm in front of a Samsung SyncMaster 1100MB CRT monitor (21 inches, 1024x768 resolution, 120Hz refresh rate). They responded by pressing the button of an ancillary input device sampled at 120Hz refresh rate with their dominant hand. The experiment was run on a Mac Pro 4.1 Quad-Core Intel Xeon at 2.66 GHz.

A trial started with both a white vertical line and a white circular target appearing on a black background. The vertical line (10cm tall, 1px width) was positioned 15 cm right of the center of the screen and remained stable for the whole trial. The target had a radius of 0.3 cm and travelled from left to right. In each trial, the target constantly moved at one of three possible speeds (19.5 cm/s, 25 cm/s and 32 cm/s). The chosen speeds followed a geometric progression in order to compensate for a constant Weber fraction, so that the discriminability between the slowest and mid speeds was presumably similar to the discriminability between the mid and fastest speeds. At every trial, the target also had one of three possible motion times (0.8 s, 1 s, 1.2 s), where motion time is defined as the time from motion onset until the x-coordinate of the center of the target matched that of the vertical line. The moment when the center of the target and the line are at the same x position will henceforth be referred to as alignment. Motion times were clearly discriminable from each other (Regan & Hamstra, 1992), while at the same time all were above the 500 ms over which basic temporal recruitment is built up (Krekelberg & Lappe, 1999; Lappe & Krekelberg, 1998). The combination of speeds and motion times, which was random for each trial, resulted in nine different initial distances from the line (15.6, 19.5, 23.4, 20, 25, 30, 25.6, 32 and 38.4 cm, in ascending order for speeds and motion times). The target disappeared when the participant pressed the button or at a random point between 1.3 s and 1.4 s after movement onset, always when the target had already completely crossed the line. Visual feedback was given when required (see below). Note that interleaving different initial times to alignment prevented participants from basing a time-based response on a specific stimulus duration (e.g. 0.85 s) from motion onset rather than estimating some temporal (or other type of) threshold with respect to the time of alignment. The 9 different initial target positions would additionally require participants to link each initial position to its motion time. In addition, in comparison to traditional temporal interval reproduction tasks with multiple interleaved times (Block & Grondin, 2014; Grondin, 2010, 2012), participants in our experiment showed a much lower variable error (see results).

### 3.3.3. Procedure

Participants completed two blocks of practice trials (90 trials per block including all speeds and motion times). In these practice trials, if their response time was within a temporal window of 200 ms centered on the line, visual feedback was provided. The aim of the baseline session was to familiarize participants with the experimental paradigm, especially with the timing until alignment.

After the practice trials, the main part of the experiment started. Everything remained the same except for the feedback after each trial. Reward was introduced, so that each trial was rewarded according to the response. This reward increased exponentially as a function of time, so that the closer the remaining time to alignment when the response was made, the higher the reward was. Any response time beyond alignment was penalized with -200 points. After a non-penalized trial, the reward won in that trial appeared on the screen, as well as the total reward won in that block until the moment. After a penalized trial, “-200” appeared next to the total reward. As before, blocks had 90 trials. Participants completed 12 blocks. They started each block with 5000 points, and they were instructed to finish each block with as many points as possible. They were told that their final payoff would depend on the total points they accumulated throughout the experiment. When receiving the initial instructions, participants were shown a graphical depiction of the reward function, similar to Figure 3.1.C. However, no reference was made to reward being related to time or space.

Computing reward

The reward given after each response was a function of the response time *t.* Reward increased exponentially and the reward function *u(t)* was such that pressing the button at alignment, centered at $t$ = 1, was rewarded with 100 points. Responding after that was penalized with the subtraction of 200 points (Figure 3.1.C). Thus, for time-defined rewards:

$$u(t) = \beta \cdot exp(\alpha \cdot t) \quad \text{if} \quad t \leq 1 \qquad (3.1)$$

$$u(t) = -200 \quad \text{if} \quad t > 1$$

Where the linear parameter $\beta = 0.1008$ acted as a scaling parameter and the non-linear parameter $\alpha = 6.90$ defined the slope or rate of change of the reward as a function of time. Note that, for the motion times of 0.8 s and 1.2 s, to make alignment be always at *t* = 1 we simply shifted the reward function so that the same reward was given across motion times when the response was made at the same time to alignment.

Assessing response optimality

In our task, the reward function is the mapping between response time and reward, which follows an asymmetric progression: ascending exponentially until alignment, becoming a constant penalization afterwards. Given this reward function, the optimal response that maximizes the reward in the long run is dependent on the participant's variability: the more variable a participant is, the earlier her optimal response time will be, so that she can avoid being penalized too often. Thus, to act optimally, a participant is required to know not only about the task's reward function, but also about her own variability across the different conditions of the stimuli (i.e. motion times and target speeds).

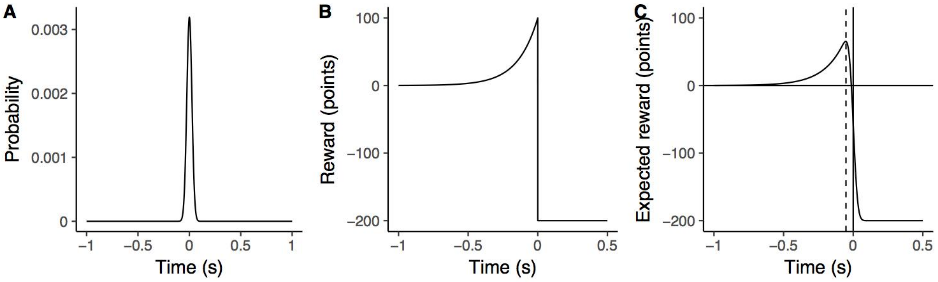


Figure 3.2. Example of the calculation of the optimal response. A. Response time distribution for an imaginary participant with a mean centered at 0 s and a temporal variability of 25 ms. B. Reward function of the task, as shown as in Figure 3.1.B, displayed here for direct comparison with the other plots of this figure. C. Convolution of

the response time variability and the reward function. Each response time is plotted against its expected reward. The vertical dashed line shows the optimal response time, which in this case would correspond to pressing the button 0.052 s before alignment. This mean response time would lead to the maximum expected reward for a variability of 25ms, which in this case would be 65.177 points.

We wanted to discover whether, in our task, participants could integrate the task's reward function and their own motor variability in order to respond at the point that maximized the expected reward. Figure 3.2. illustrates the process of calculating optimal responses. In this task, the temporal variability is normally distributed with mean $\mu_t$ and standard deviation $\sigma_t$ . Figure 3.2.A shows an arbitrary example of this normally-distributed variability. Response normality in our experiment was verified by comparing the quantiles of the observed and theoretical distributions through Q-Q plots (Evans, Hastings, & Peacock, 2000).

Given the reward function described in Equation (3.1), and displayed in Figure 3.2.B, the expected reward is the result of the convolution of the motor variability density with the task's reward function (e.g. Mamassian, 2008). Thus , the expected reward $g(t)$,

$$g(t) = v(t) * u(t) \tag{3.2}$$

where $v(t)$ is simply the standard deviation of the response times which are not considered outliers (see data analysis). The time value corresponding to a maximum reward $g^*$ is considered the optimal time $t^*$ (Figure 1C):

$$t^* = \arg\max_t (g(t)) \tag{3.3}$$

$$g^* = g(t^*) \tag{3.4}$$

We calculated $t^*$ for each participant, split by motion time and target speed. To see whether the optimal response was indistinguishable from the observed mean response, we checked whether $t^*$ fell within the 95% confidence intervals of the response times for each block, calculated with non-parametric bootstrap with 1000 resamples. Note that, besides of the constant reward function, $t^*$ only depends on the standard deviation of the responses. Note also that the standard deviation became smaller as the experiment progressed. To reduce the complexity of the analyses and be able to collapse across blocks, the data used to calculate $t^*$ was that corresponding only to the final 5 blocks of the experiment, where the standard deviation had already become stable. This way, we would also minimize any differences caused by initial learning.

## 3.3.4. Data analysis

For every participant, block, motion time and target speed we defined as outliers those trials whose response times and/or response locations were above or below 3 standard deviations. These outliers were removed before proceeding with the analyses.

Disclosing the decision variable

When adopting a response criterion, participants may rely more or less on temporal or spatial information, which translates to different extents to which either type of information is used as a decision variable (DV). A temporal DV would consist in, for all target speeds, aiming at responding when the time left to alignment is very similar. On the contrary, a participant using a pure spatial DV would aim at responding, for all speeds, when the target is at similar spatial positions. The idea of responding when the stimulus reaches a threshold based on a remaining time is a long-standing proposal that started with the tau hypothesis (Lee, 1976; Tyldesley & Whiting, 1975). Note, however, that we want to capture this feature in the overt responses and, therefore, in order to define a temporally or spatially consistent DV we will not depend on an invariant threshold for action, be it temporal or spatial, but rather on the difference across velocities in the response pattern. For example, suppose that one participant has a spatial DV by responding at a similar spatial distance to alignment across different speeds. We do not require this distance to be stable in absolute terms. For instance, this same participant may respond when the target is 1 cm away from the line, and eventually shift to when it is 2 cm away, without her DV being affected as long as this difference remains the same across speeds. Our experimental design allowed to determine the extent to which participants exploited more temporal or spatial cues. This is described below.

A pure temporal DV would require to estimate the time remaining to alignment, and set a moment in time acting as a response threshold, which would mainly be invariant across speeds. Note that we are less concerned about how subjects estimate this time: either by using optical variables (Bootsma & Oudejans, 1993) or relying on speed estimates (Smeets et al., 1996). When the target reached that temporal threshold $t_{th}$ participants would initiate their response (López-Moliner & Bonnet, 2002). This is exemplified in Figure 3.1.B. The observed response time $t_{resp}$ can be defined as

$$t_{resp} = t_{th} + t_m \qquad (3.5)$$

Where the threshold $t_{th}$ is a stimulus-dependent component and $t_m$ is an additive component independent of the stimulus or target that would include motor delays. We can further substitute the threshold $t_{th}$ to consider the speed of the target *t*:

$$t_{resp} = a\, vel^{-1} + t_m \tag{3.6}$$

Where $a$ would correspond to the remaining spatial distance to alignment when the the evidence reaches $t_{th}$, and $vel^{-1}$ is the speed. When the remaining space ($a$) is divided by the corresponding speed ($vel$), we obtain the corresponding temporal threshold. If we applied Equation (3.6) to a pure temporal data pattern with $a$ and $t_m$ as free parameters, the fit would estimate a value of $a$ very close to zero reflecting the invariance of the responses across speeds, but rendering the spatial meaning of $a$ void. This can be avoided by setting the $t_m$ (intercept) to a fixed value which would require further assumptions to separate $t_{th}$ from $t_m$.

Therefore, in order to keep our analysis as simple as possible while capturing the underlying response strategy, we can use the same linear model but interpreting the resulting slope (denoted as $s$ rather than $a$) as an indicator of the use of a temporal threshold. Thus, an $s$ value close to 0 would reflect time-based responses, as this would capture the invariance across speeds. On the other hand, a pure spatial DV would use a location in space as a threshold. Therefore, the $t_{resp}$ (i.e. remaining time) as a function of $vel^{-1}$ would have a linear relationship with a slope $s$ significantly different from zero. A negative $s$ would correspond to earlier response times for slower speeds and later response times for faster speeds, with bigger differences the more negative $s$ is. This would be congruent with a space-based DV, because aiming at a same spatial location corresponds to smaller remaining times for faster speeds (Figure 3.1.B). Of note is that, according to Equation (3.6), a constant location in space across speeds would correspond to tracking the changing position (Brenner et al., 2006) and aiming at responding when the target is at some distance before the line. Due to the motor delay to produce the action, one could argue that the target would move for a longer time in faster speeds. Synchronization mechanisms (Aschersleben, 2002; Aschersleben & Prinz, 1998) have been proposed to account for these delays so that participants with a spatial criterion would initiate the action slightly earlier in time for faster speeds, so as to make up for this motor delay. However, this is not captured by Equation (3.6), as we wanted it to include only the essential elements for our framework to capture the pattern of overt responses.

We calculated $s$ for each participant and motion time. And although values of $s$ closer to 0 informed us of a more time-based DV, we did not know which range of $s$values could be related to mainly temporal information being used to plan responses – and the same for spatial information. That is why, for each participant and motion time, and using her own data, we simulated the slope corresponding to a pure spatial DV, where a spatial threshold was used. For this purpose, we took the mean of the responses in space for the whole experiment (which would act as a unique spatial threshold), and dividing it by each speed, thus obtaining three response times. Those were fitted to Equation (3.6) as a function of its associated $vel^{-1}$in order to estimate a slope $s_{spat}$ that would correspond to an entirely spatial DV. This slope $s_{spat}$ was compared to $s$, the slope obtained from the observed data. To decide whether the participant exploited more temporal or spatial cues, we had to use a cutoff point. Our compromise choice was the value halfway between a purely spatial DV ($s_{spat}$) and a purely temporal DV (which, given our model, was always a slope with value 0). The midpoint between any value and 0 is half that value. Thus, if $s$ was closer to 0 to half the value of $s_{spat}$, the participant was considered to favour a time-based DV. Otherwise, her DV was categorized as space-based. In addition, by calculating a ratio between $s$ and half $s_{spat}$, we also obtained a value that was used as an indicator of how time or space-based responses were along a continuum. This value ("DV ratio") was calculated by dividing $s$ by half $s_{spat}$. The cutoff point for the DV ratio was at 1, since that value indicated equidistance of slope $s$ to the purely temporal and spatial slopes. The smaller a DV ratio when compared to 1, the more time-based responses were, since $s$ (the numerator) is closer to 0 than the denominator. And the larger the DV ratio, the more space-based responses were. Within each participant, we calculated the DV ratio for each motion time, in order to check whether different ratios were used simultaneously for the different motion times. When performing group analyses, we used the cutoff point of DV ratio = 1 to classify responses as mainly time or space-based. This participant grouping will be referred to as the temporal and spatial DV groups. Note the fact that each participant's responses were split by motion time. Thus, a participant's data could fall within both the temporal and spatial DV groups, depending on the main cue used to plan responses for a particular motion time. We followed this grouping procedure for every group analysis detailed in the results section.

## 3.4. Results

All values referring to temporal responses reported henceforth are scaled so that alignment occurs at $t = 0$. Negative values indicate responses before alignment, and positive values indicate responses after alignment.

Outliers accounted for 0.348% of trials. These were removed from subsequent analyses. We ran an initial series of Linear Mixed Models (LMM) to see how response times and their standard deviations evolved across blocks. The effect of target speed and motion time on reward and responses will be analysed thoroughly in the next sections. We chose LMMs because they allow for a better integration of continuous variables (e.g. block) and factors in the same design than ANOVA (Baayen, Davidson, & Bates, 2008). We used the lmer function implemented in the lme4 R package (v.1.0-6) (Bates, Mächler, Bolker, & Walker, 2015).

An LMM was conducted with response time as dependent variable and block and motion time as fixed effects. Participants were treated as random effects (random intercept and slope). The ANOVA on the LMM gave a significant main effect of motion time (*F*(2,22576) = 79.065, *p*<.001) and block (*F*(1,22576) = 8.572, *p*=.003), but no significant interaction. Response times were on average before alignment (initial block average per motion time, 0.8 s = -0.035 s; 1 s = -0.038 s; 1.2 s = -0.044 s). Across blocks, they became significantly later (average rate 0.0002 s/block, *t*= 2.906, *p*= 0.04, final block average per motion time, 0.8 s = -0.033 s; 1 s = -0.037 s; 1.2 s = -0.042 s). Concerning the variability, responses became more precise across blocks, with no difference among motion times. The SD decreased at a rate of 0.0004 s per block (t= -8.985, p<.001, initial block average = 0.031 s, final block average = 0.024).

*Reward and loss fraction*

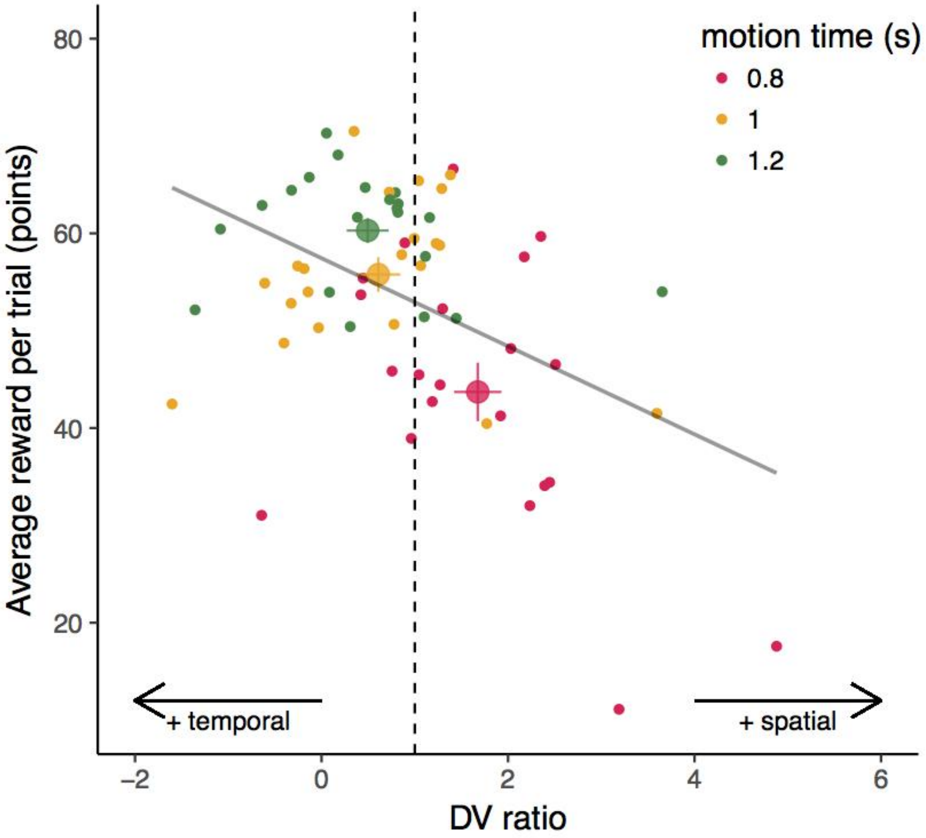


Figure 3.3. Average reward per trial (in points) as a function of the value of the DV ratio. The more towards the left a data point is located, the more time-based the response. Each participant is included in the plot with three data points, each corresponding to her responses within a certain motion time, as denoted by the color. Bigger points represent the centroid of each motion time's distribution, with associated error bars that show standard errors of the mean. The solid, grey line shows a linear fit of all the data. The dashed, vertical, black line shows the cutoff point of DV ratio = 1, used to binarily categorize a group of responses as time or space-based.

An LMM was run with average reward as dependent variable, DV ratio and motion time as fixed effects, and participant as random effects. Figure 3.3. shows the average reward as a function of the DV ratio (smaller being more time-based) for each motion time. As can be seen, the reward was clearly modulated by the type of information used to control the responses: smaller DV ratios resulted in larger reward (fixed effect of DV ratio: *F*(1,382.1)= 13.055, *p*<.001). In addition, average reward was also larger for longer motion times (fixed effect of motion times *F*(2,15927.1)=13.789, *p*<.001). Also the interaction between DV ratio and motion time reached significance (*F*(2,14913.8)= 3.092, *p*=.045). With the data collapsed across motion times, a Pearson's correlation between DV ratio and average reward per trial had an *r* = -0.453 (*p*<.001). Using the cutoff point of DV ratio = 1 to binarily divide participants into the DV groups, the % of participants classified into the temporal DV group was 28.571% when the motion time was 0.8 s, 61.905% for 1 s, and 76.190% for 1.2 s.

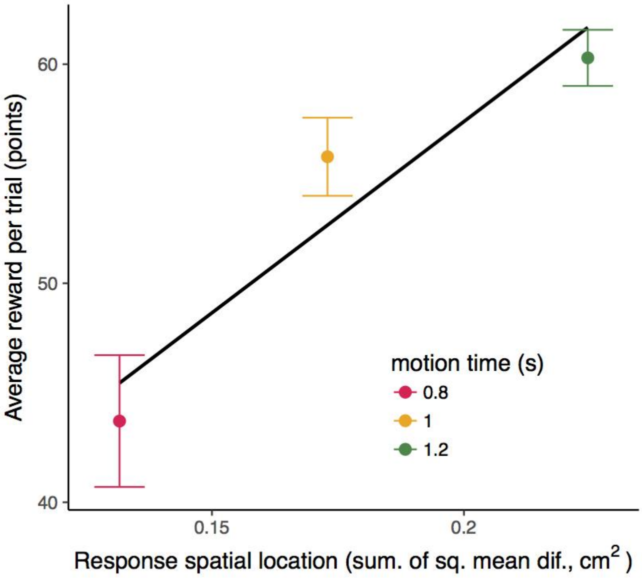


Figure 3.4. Average reward per trial (in points) as a function of the sum of the mean squared differences of response spatial locations (in $cm^2$) across speeds. Results are split by motion time, as denoted by the color, and error bars represent the standard error of the mean across participants. The solid, black line represents a linear fit of the data points.

One concern with the DV ratio analysis is that we implicitly assumed a relatively constant threshold (spatial or temporal) across the different blocks. However, as commented before, this is not a requirement for our predictions. Space-based overt responses will result in similar response spatial locations across speeds, irrespective of the remaining spatial distance for alignment used to trigger the response. Conversely, time-based responses will generate systematic differences of position across velocities. We conducted an additional and simple analysis, without the need to assume a unique threshold, in order to support our claim regarding an enhanced reward correlating with more time-based responses, and longer motion times facilitating them. We calculated the mean squared difference between the spatial location for each speed and the mean location pooled across speeds. We computed this quantity for each participant, block and motion time. The mean across participants and blocks is plotted against average reward in Figure 3.4. As can be seen, lower differences across speeds in response spatial locations were related to both lower average reward and to shorter motion times, this being consistent with all the previous analysis related to space-based responses. The average values were significantly different across motion times, as denoted by the non-overlapping error bars.

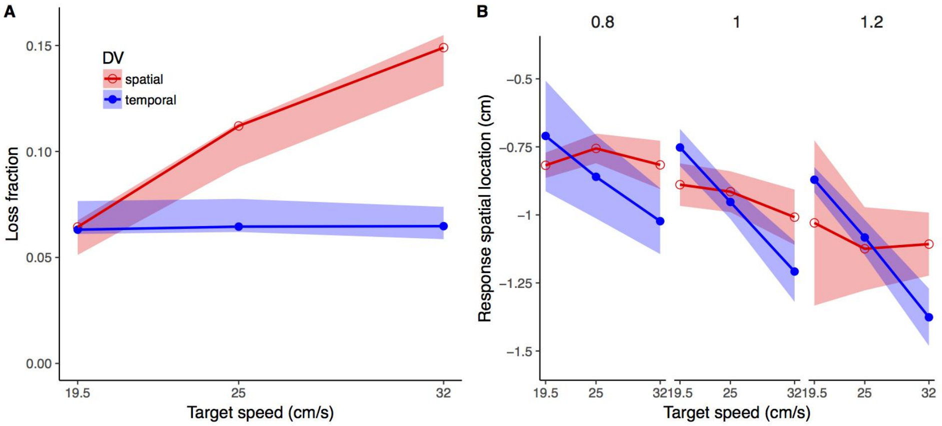


Figure 3.5. A. Average loss fractions between DV as a function of target speed. Shaded areas represent binomial 95% confidence intervals. B. Average response spatial locations between DV as a function of target speed. Shaded areas represent 95% confidence intervals. Each panel shows responses for a specific motion time.

To explain why more time-based responses enhanced performance, we explored a hypothesis already mentioned in the introduction: that participants exploiting more temporal information dealt better with faster target speeds. The precision by which we acknowledge the change of position of a moving object is limited by the temporal resolution of the visual system, with spatial uncertainty being larger for faster speeds. Thus, spatial variability increases for faster speeds, as depicted in Figure 3.1.A. In our task, for responses closer to the line that rely on a spatial position, the amount of penalized trials should thus be larger for faster speeds, whereas a temporal DV could allow for more stable responses across speeds. To test this, we grouped the whole dataset, split by each motion time within each participant, according to the type of information participants relied on more. As plotted in Figure 3.5.A, we calculated the loss fraction for the different target speeds. For each speed, the loss fraction was simply the proportion of trials for which responses were made after alignment, and thus were penalized. We conducted a general linear mixed model (binomial responses) with the loss fraction as dependent variable, target speed and DV as dependent variables (fixed effects) and participants as random effects (we allowed both intercept and slope to vary). Time-based responses did not lead to a significant increase of losses with increasing speed (slope= .006, *p*=.407), while space-based responses did result in a significant increase of losses (slope= .060, *p*<.001). The difference in percentage was not significant for the lowest speed ($\chi^2(1)= 0.042$, p=0.419), while it was for the two larger speeds (25: $\chi^2(1)= 26.781$, *p*<.001; 32: $\chi^2(1)= 80.634$, *p*<.001).

*Modulation of DV by motion time*

Figure 3.3. clearly shows how longer motion times (e.g. 1.2 s) were related to the use of more temporal information. This can be due to longer integration times leading to more reliable velocity estimates. One possibility, then, is that, in trials with longer motion times, velocity estimates can be used to update the target's position more rapidly. This, in its turn, would increase the presence of phenomena like motion extrapolation (De Valois & De Valois, 1991; Linares et al., 2007; Nijhawan, 1994; Whitney, 2002). Some evidence for this hypothesis can be seen in Figure 3.5.B, which plots the target's spatial position at response as a function of speed, split for the different motion times.

The trend differs is not equally clear in both DV groups. Those in the temporal group (in blue) generally respond earlier in space for faster speeds, but this seems to increase for longer motion times. For the spatial group, the most obvious change seems to be that longer motion times lead to overall earlier responses. And although the few number of spatial participants for the longest motion time may introduce a lot of variability, there seems to be a trend of responses getting earlier for faster speeds as motion time increases. In other words, spatial participants become more like temporal participants as motion time gets longer, and this would be congruent with the finding of DVs becoming more temporal. To explore these results, we ran an ANOVA on an LMM with position as a dependent variable, DV group (spatial or temporal), target speed and motion time as fixed effects, and participant as random effects. The analysis produced a significant interaction between DV group and target speed ($F(2,22563)=104.398$, $p<.001$), between DV group and motion time ($F(2,22573)= 3.546$, $p=.029$), and between target speed and motion time ($F(4,22563)= 7.110$, $p<.001$) consistent with the presence of motion extrapolation.

### *Response optimality*

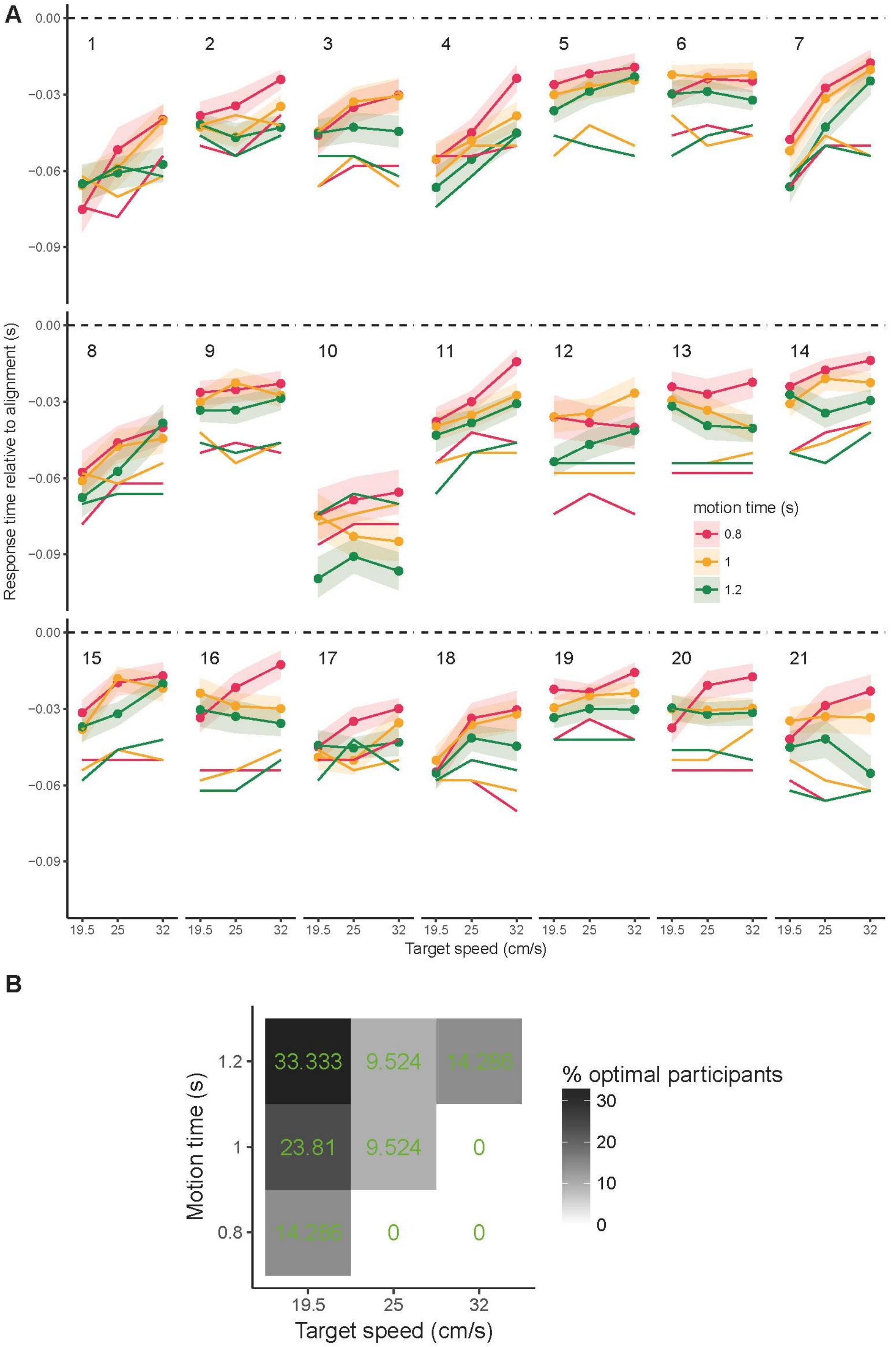

A
B
0.00
−0.03
−0.06
−0.09
Response time relative to alignment (s)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
motion time (s)
0.8
1
1.2
19.5
25
32
Target speed (cm/s)
Motion time (s)
1.2
1
0.8
33.333
9.524
14.286
23.81
9.524
0
14.286
0
0
19.5
25
32
Target speed (cm/s)
% optimal participants
30
20
10
0

Figure 3.6. A. Response times, split and colored by motion time, as a function of target speed. The dots connected by lines represent average observed data for the last 5 blocks, with the shaded area around them representing 95% confidence intervals calculated by non-parametric bootstrap. The solid lines with no points attached represent optimal responses. Each panel contains the data of one participant. The horizontal, dashed line marks alignment. B. Percentage of participants whose optimal response time fell within the confidence intervals of their observed response times, split by target speed and motion time. Percentages are both represented as a number and as the color of each tile within the gradient specified on the legend.

In order to see whether observed responses were optimal according to an expected reward maximization model, we calculated optimal responses by feeding the model each participant’s response standard deviation, split by target speed and motion time. As Figure 3.6.A depicts, average observed responses were almost always too close to alignment when compared to optimal ones, in a similar way as what has been reported in timing (Mamassian, 2008; Ota et al., 2015) and reaching (O’Brien & Ahmed, 2013). According to Mamassian (2008), this could be explained by participants underestimating their own motor variability, thus producing overconfidence. In some occasions, though, average response times were indistinguishable from optimal. Figure 3.6.B shows how the number of optimal participants changed across motion time and target speed. Even if optimality changed a little with motion time and target speed, in the best-case scenario only a third of the participants fell into the category of optimal responses. That was a product of participants typically responding earlier for longer motion times and slower speeds.

## 3.5. Discussion

Due to the spatiotemporal nature of interacting with moving objects, people can plan their actions by basing responses on temporal and/or spatial information. We investigated whether relying more on either type of information could improve performance, defined in our task as the average reward given after a response. We found that the more time-based responses were, the larger the average reward. We relate this result to a temporal DV being a better way to keep performance stable across speeds, while a spatial DV increases the number of penalized trials as target speed became faster. We suggest that this is a product of the mechanisms underlying the acquisition of these cues: space-based responses are made by continuously sampling the target's position in space until it reaches a spatial threshold, which then triggers the response. This planning is constrained by the temporal resolution of our visual system: we simply cannot increase the sampling rate when monitoring a position. Since, when compared to objects moving with a slower speed, those with a faster speed travel a larger distance in the same amount of time, acknowledging the position of an object that moves faster will be associated to a larger spatial error (Brenner et al., 2006). In our task, since participants responded close to the line, this meant that those trials with faster target speeds were more prone to being penalized. On the other hand, time-based responses are made by estimating the remaining time-to-contact (in this case, alignment) and establishing a threshold that is set somewhere along the estimated time (see López-Moliner & Bonnet, 2002). This would allow to circumvent the limitation imposed by the sampling rate of the visual system. Even if there is evidence for an integrated object-tracking system (Kwon et al., 2015), time-based responses would be related to prioritizing pure motion signals, in contrast with space-based responses mainly using spatial displacement, as has been shown with other type of visual stimuli (de la Malla & López-Moliner, 2010; Nakayama & Tyler, 1981). As stated in the Methods section, we were more focused on characterizing the nature of the response (from temporal to spatial) than revealing how subjects did estime a temporal threshold, which is a different research question in itself. Nonetheless, this issue deserves some discussion. One possibility is that exploiting temporal information is based on velocity estimates and participants combine this information with remaining distance as reflected in Equation (3.6). Alternatively, subjects could have used optical variables like the rate of contraction of the visual angle between the target and the line (Bootsma & Oudejans 1993). Normally, models based on optical variables do not make predictions on the uncertainty of respective estimates since neural processing, for example, is less emphasized. In any case, the present study did not aim at supporting one or the other claim by discussing the exact variables participants would use. However, the differences across different motion times cannot easily be accommodated by using optical variables alone and some reformulation of models only combining optical variables, probably by including uncertainty measures, would be necessary. This has already been proposed in optical models for catching within the context of the outfielder problem (Belousov, Neumann, Rothkopf, & Peters, 2016).

We found that the information used to plan responses could be modulated by motion time, understood as the total time the participant could see the target moving. Trials with a short motion time mainly originated space-based responses, while longer motion times elicited more time-based responses. Extracting motion information requires integrating evidence over time (Burr & Santoro, 2001; Neri et al., 1998). The shorter a motion time, then, the more difficult it is to obtain a reliable velocity estimate to establish the temporal response threshold. For short motion times responses are thus likely to be constrained to the use of spatial information.

More time-based responses – and larger average reward – may also be explained through the presence of position extrapolation. If longer motion times favour acquiring reliable velocity estimates, these estimates could be used to update position more rapidly. This motion information, though, would increase the presence of position extrapolation. This is consistent with findings showing that the amount of extrapolation depends on motion time (Linares et al., 2007) and aforementioned integrative accounts of position and motion (Kwon et al., 2015). Position extrapolation has been argued to have the adaptive advantage of compensating for the visual delays our visual system has when transmitting information along the visual pathways (Nijhawan, 1994). In our task, the advantage behind extrapolating the position (and more so for faster speeds) is the same as the advantage behind exploiting temporal information: reducing the risk of responding after the line and being penalized, thus increasing average reward. Nevertheless, this cannot be generalized to all situations involving interaction with moving objects, since it is not always the case that the outcome of responding a little later than the point of maximum reward (in our case, alignment) is much worse than responding a little earlier. Extrapolation could also be the reason why the DV ratio of some participants was considerably smaller than 0. A negative DV ratio can be interpreted as participants responding earlier (both in time and in space) for faster speeds. This is congruent with perceived position being more ahead of the true position for faster speeds. Plus, as Figure 3.3. shows, most negative DV ratios correspond to long motion times, and that supports the aforementioned findings of extrapolation increasing for longer motion times.

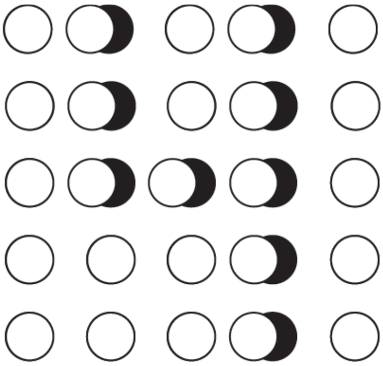
In our task, reward did not only depend on the participants' DV. When considering the correlation between DV ratio and average reward, the amount of variance that remains unexplained points to other contributing factors. An obvious one is where exactly the unified criterion for action is placed (see the optimality discussion below). Earlier thresholds would lead to less reward when the response was before the line, but also to reducing the chances of being penalized. An element directly related to this is how participants corrected their responses on a trial-to-trial basis. For instance, after a response after the line, some participants may decide to shift their criterion earlier so as to avoid future penalizations. It could also happen that, after a series of non-penalized responses, participants may be slowly responding later, in order to win more. These different dynamics are not analyzed here, but are important to characterize the nature of the responses and could represent an interesting future direction.

One important concern with our task could be that reward was given as a function of time. Thus, planning responses based on temporal information was more likely to improve reward. To address this issue, we conducted another experiment, with the only difference being that reward was given as a function of the remaining spatial distance to alignment (see supplementary experiment 1 in the appendix). Even if the manipulation slightly weakened the relationship, the average reward was still larger for more time-based responses, with the modulating effect of motion time also being replicated.

We defend that time-based responses improve performance. A direct test to our claim, then, would be to see whether, in situations where obtaining motion information is more difficult, performance is reduced. In another control experiment (see supplementary experiment 2 in the appendix), we hampered the motion system by removing those motion cues derived from a change in luminance in the retina (first-order motion). The available motion information, then, depended only on changes in contrast (second-order motion). As expected  (see Seiffert & Cavanagh, 1998), responses under this condition became more space-based, even for the longest motion times. This reduced performance, as measured by the average reward, which speaks to our main claim.

Even if the average reward was larger for more time-based responses and longer motion times, in terms of expected reward maximization (von Neumann & Morgenstern, 1944) responses were in general suboptimal. Even in those cases where observed behavior was indistinguishable from optimal estimates, though, it is unlikely that participants met the requirements often thought to be needed for optimal performance. In the sensorimotor instantiations of SDT we used (Körding & Wolpert, 2006; Trommershäuser et al., 2003b; Wolpert, 2007), optimal agents are expected to base decisions on knowledge of both their own variability and the task's reward function. Regarding the variability, for our data that would have meant that participants were aware of their variability only for the conditions of the stimuli where the model determined optimal performance (typically longer motion times and slower speeds), and not for the others. This unlikely scenario probably reflects the limitations of model when to explaining our data. Studies applying this model have normally featured pointing tasks where, if the stimulus (target) changed across trials, this did not imply a difference in the (always very low) sensory uncertainty. On the other hand, in our task both target speed and motion time changed. To better deal with our data, and to data involving moving objects in general, the current model should be adapted so that it accounts for the changing sensory uncertainty.

Something similar can be concluded if we take a look at the reward function. In our case, a temporal DV resulted in larger reward even when this reward was given as a function of space. Thus, exploiting temporal information was better than acting in accordance to the reward function’s domain (i.e. spatial DV for a spatially-defined reward). In line with recent evidence (Tsetsos et al., 2016), this shows how enhanced performance is sometimes achieved after ignoring some assumptions of economic decision theory. In our case, a physical limitation (the resolution of the visual system) affecting those displaying a spatial DV was at the root of the temporal DV leading to larger scores, and not differences in how participants used information about the reward function. Any normative model dealing with spatiotemporal situations susceptible to reproduce our findings should better decompose and specify the different visual information that can be used. It should also consider possible physical constraints that lead the use of some of this information to a better performance. In this sense, the present study also contributes to the body of research on sensorimotor decision-making within the framework of SDT by evidencing how such physical constraints should be taken into account before defining by which behavioral rules an optimal agent should abide. In sum, as others have claimed before (Jarvstad et al., 2014), the use of normative models that overlook essential aspects of the task will be of limited use to derive inferences of optimal performance.



# Study II - Can reward induce the use of temporal information when interacting with moving objects?

## 4.1. Summary

Actions relative to the course of a moving object can be planned by exploiting temporal or spatial information of the object. Our previous work claimed that the use of temporal cues leads to a better performance. Learning how to use these cues could allow for improvement in those tasks requiring interaction with moving objects. Studies of perceptual learning have commonly relied on reward to achieve their purposes. Here we designed a reward-based procedure in order to foster the use of temporal cues. We took our previous paradigm and manipulated reward after each trial, so that the more distant the last responses were in time, the more the reward was reduced. We also ran another condition where a similar procedure was implemented, but this time to foster the use of spatial cues. We found more participants exploiting temporal information in the condition where this was promoted. However, temporal information was exploited from the start, implying that learning did not happen. This made us rule out the effectivity of our reward manipulation. Future work should aim both at designing more successful methods and at investigating individual differences causing some participants to be more prone to exploit temporal cues.

## 4.2. Introduction

The experimental results described in the previous chapter showed how, when planning a movement in response to moving objects, exploiting temporal information can be advantageous. There is a certain flexibility in which types of proximal (retinal) information (e.g. pure motion versus changing of position (Smeets & Brenner, 1994)) could potentially be used to ascertain the real motion of an object in the scene. Thus, it may be that those who relied more on temporal information are simply more inclined to take advantage of pure motion cues. An interesting question would be whether the use of temporal information can also be learnt. This is relevant, since it could improve performance in tasks involving moving objects, such as ball sports.

The process of improvement in our sensory abilities after training is known as perceptual learning (Seitz & Watanabe, 2005; Watanabe & Sasaki, 2015). Changes in sensory abilities have been shown for many visual features, such as orientation (Schoups, Vogels, & Orban, 1995), motion direction (Ball & Sekuler, 1987), Vernier acuity (Fahle & Edelman, 1993), texture discrimination (Karni & Sagi, 1991) or speed discrimination (Huang, Lu, Zhou, & Liu, 2011; Saffell & Matthews, 2003). However, our case goes beyond fine-tuning the discrimination of a simple perceptual attribute. We are interested in knowing whether people can learn to use a specific visual information to plan an action. That would imply firstly sampling that information, and then establishing a decisional criterion for action that could be applied to all conditions of the moving object.

The aforementioned literature shows how, thanks to brain plasticity, our perceptual skills can get sharper after sustained exposure to a task. Sometimes, learners know their final goal in terms of behavior and, after each trial, they are shown the product of what they have done. Then, in the subsequent attempt, they try to correct for their previous error so as to be closer to the goal behavior. In some other cases, learners do not know what their best possible behavior should be, so they try to maximize some sort of reward function. Perceptual learning has been argued to be based on a process of reinforcement (Law & Gold, 2009;

Roelfsema, van Ooyen, & Watanabe, 2010), and reward can mediate this reinforcement (Seitz, Kim, & Watanabe, 2009). Even when sensory feedback is not available, explicit reward has proved to be useful in processes of sensorimotor learning and adaptation (Darshan, Leblois, & Hansel, 2014). In the present study, whose paradigm was very similar to Study I, we used explicit reward as reinforcement. Specifically, we manipulated reward to induce a change in the information used to plan responses. We did so by penalizing the use of a temporal DV when we wanted participants to use spatial information, and by penalizing the use of a spatial DV when we wanted them to use temporal information.

Our previous study found that the DV became more temporal when the trajectory of the target (motion time) was longer. In this study, we wanted the DV to change only in response to the penalizations applied to the reward function, and not as a consequence of a particular target's motion time. To ensure this, we only presented targets with one motion time. However, we did not know whether, in the previous experiments, the presence of multiple motion times could have had any influence in the different DVs: for instance, the DV becoming more temporal for shorter motion times when presented interleaved with longer motion times, as compared to a condition where only a short motion time was presented. We ran Experiment 1 to assess the DVs in a condition with just one motion time, so that later on we could see how these changed when introducing a penalization imposed on the reward function. We found that most participants exploited more spatial information. In Experiment 2, we introduced the penalization, so as to foster the use of one or another type of information. Although half of the sample had a temporal DV in the condition where this was promoted, the lack of a learning trend across blocks casts a doubt about the usefulness of this manipulation.

## 4.3. Experiment 1

Experiment 1 was planned to be compared with Experiment 2. This study aimed at introducing reward scaling to foster one or another DV variable, but first it was necessary to identify which changes in the DV arose from the reward manipulation, as opposed to those arising from the fact that, unlike in our previous study, only 1 target motion time was presented. To do that distinction, we ran Experiment 1, where targets had a single motion time, and reward penalization was not yet introduced. Still, it could be argued that the nature of the reward function could bias the adopted DV: if reward was defined as a function of temporal proximity to alignment, participants could have more time-based responses, and vice-versa (see Study I). In order to minimize this, when compared to the previous study, in this experiment we narrowed the differences among presented speeds. Instead of speeds of 19.5, 25 and 32 cm/s, targets could adopt speeds of 20, 23 and 26 cm/s. This made that, independently of the variable reward depended on, time-based and space-based responses would result into very similar rewards. Figure 4.1. shows how, when compared to the old speeds, the difference in reward among different DVs shrinks when using the new speeds. For instance, Figure 4.1.A depicts how, in a task where the reward function gives reward according to the time left for alignment, responding at the same point in space (spatial DV, dotted lines) translates into a different reward for the different speeds, but this is much less pronounced when the difference among speeds is smaller (Figure 4.1.C). Similarly, when giving reward as a function of the spatial distance to alignment, responding when the same time is left for alignment corresponds to a different reward for each speed (Figure

4.1.B), but less so when the difference among speeds is smaller (4.1.D). With this design, we could be more sure that results in this experiment would reflect the default DVs participants would use when dealing with moving objects with a single motion time, without the domain of the reward function influencing the DV. Results of Experiment 1 were meant to be used as a baseline for Experiment 2, so that any differences could be attributed to the reward manipulation introduced in the latter. We adopted one last measure to know whether, in spite of the minimal difference between conditions on how different decision variables would translate into rewards, the reward function still influenced the DV. We made half of the participants in Experiment 1 complete a condition where reward was given as a function of time, and the other half one where reward was given as a function of space. These will be referred to as the time condition and space condition, respectively.

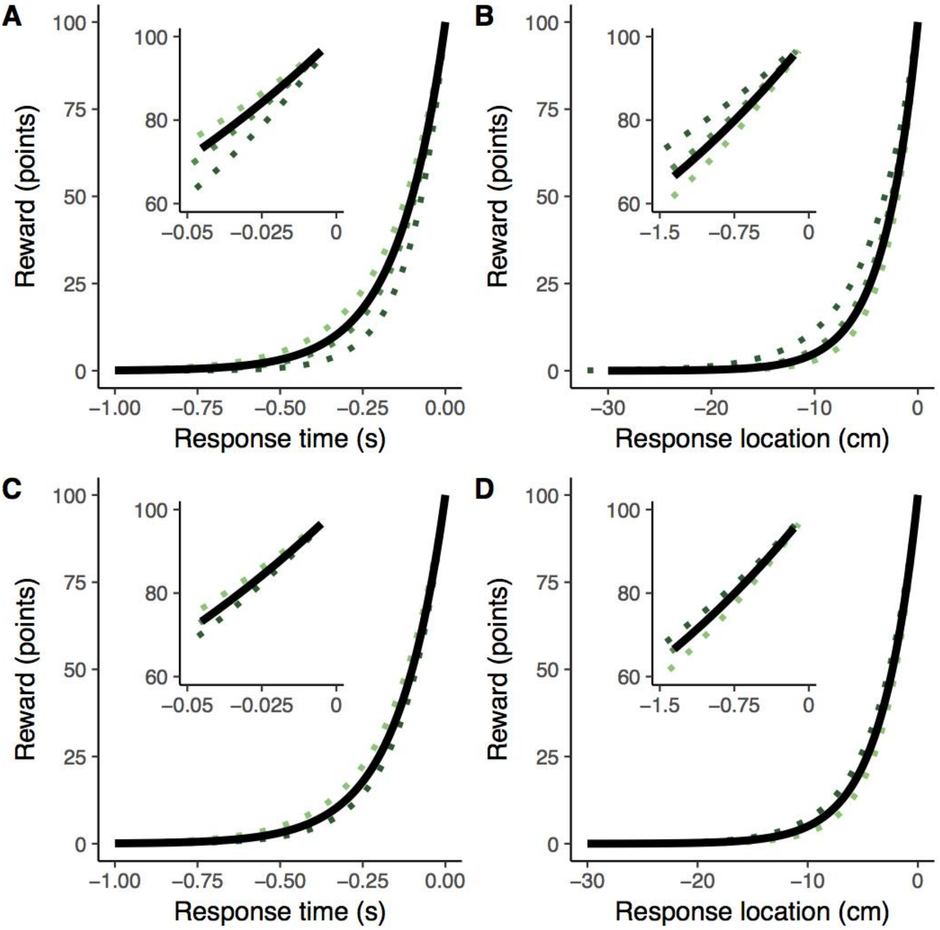


Figure 4.1. A. Non-penalized part of the reward function where reward is given as a function of time. 0 indicates alignment. The solid line represents a pure temporal DV: for all speeds, the response is executed when the time left for alignment is the same. In this case, equal response times translate to the same reward for all speeds. The dotted lines represent a pure spatial DV: for all speeds, the response is executed when the target is at the same spatial position. Since reward in this condition is given as a function of time, equal response locations would translate into a different reward for each target speed. The target speeds the dotted lines correspond to are those used in the experiments of Chapter 3: 19.5, 25 and 32 cm/s. The inset zooms in the part of the reward function closest to alignment. B. Non-penalized part of the reward function where reward is given as a function of space. 0 indicates alignment. The solid line represents a pure spatial DV. In this case, equal response locations translate into the same reward for all speeds. The dotted lines represent a pure temporal DV. Reward is given as a function of spatial

position, so equal response times correspond to a different reward for each target speed. As in A, the target speeds the dotted lines correspond to are also those used in Chapter 3. The inset zooms in the part of the reward function closest to alignment. C. Same as A, but now the dotted lines correspond to the target speeds used in Experiment 1 of the present chapter: 20, 23 and 26 cm/s. D. Same as B, but with the dotted lines corresponding to the target speeds used in Experiment 1 of the present chapter.

## 4.3.1. Methods

### 4.3.1.1. Participants

27 participants took part in the experiment after giving informed consent (17 women, all right-handed, age range 18-40). 14 completed the time condition and 13 the space condition. All had normal or corrected-to-normal vision and were naïve about the aim of the experiment. The study complied with the local ethics guidelines, in accordance with the declaration of Helsinki.

#### 4.3.1.2. Apparatus and stimuli

The room conditions and equipment were the same as those described in section 3.3.2.

A trial started with both a white vertical line and a white circular target appearing on a black background. The vertical line (10cm tall, 1px width) was positioned 15 cm right of the center of the screen and remained stable for the whole trial. The target had a radius of 0.3 cm and travelled from left to right. Three possible speeds (20 cm/s, 23 cm/s and 26 cm/s) matched a specific initial position (20 cm, 23 cm and 26 cm away from the vertical line, respectively) so that it always took 1 s until alignment. The target disappeared when the participant pressed the button or at a random point between 1.1 s and 1.2 sec after movement onset, always when the target had already completely crossed the line. Visual feedback was given when required (see below). As a clarifying point, the single motion time of 1 s was chosen because, out of the motion times used in our previous experiments, this was the value for which responses were more balanced between those binarily classified as time-based and space-based.

#### 4.3.1.3. Procedure

Participants completed two blocks of practice trials (90 trials per block including all speeds). In these practice trials, if their response time was within a temporal window of 200 ms centered on the line, visual feedback was provided. The aim of the baseline session was to familiarize participants with the experimental paradigm, especially with the timing until alignment.

After the practice trials, the main part of the experiment started. Everything remained the same except for the feedback after each trial. Reward was introduced, so that each trial was rewarded according to the response. For those participants in the time condition, this reward increased exponentially as a function of time, so that the closer the response time to alignment, the higher the reward was. For participants in the space condition, reward grew as a function of space and any response time beyond alignment was penalized. The reward function of the time condition was the same used in the main text of Study I, while the reward function of the space condition is that used in Supplementary Experiment 1 of the very same study (included in the Appendix). After a non-penalized trial, the reward won in that trial appeared on the screen, as well as the total reward won in that block until the moment. After a penalized trial, “-200” appeared next to the total reward. Blocks had 90 trials, composed by 30 randomly ordered trials per speed. Participants completed 12 blocks. They started each block with 5000 points, and they were instructed to finish each block with as many points as possible. When receiving the initial instructions, participants were shown a graphical depiction of the reward function. However, no reference was made to reward being related to time or space.

### 4.3.1.4. Data analysis

Disclosing the DV followed the exact same procedure as Study I. The only difference here was that the new target speeds were used. Thus, we obtained a DV ratio for each participant, and a binary categorization was made such that participants with a DV ratio below 1 were grouped into the temporal group, while those with a DV ratio above 1 were in the spatial group.

The method to identify and remove outliers was the same as in Study I. For every condition, participant, block and target speed, we eliminated those trials with response times and/or response locations were above or below 3 standard deviations.

## 4.3.2. Results

Subsequently reported response times and spatial locations are scaled, with alignment being a time value $t = 0$ for the time condition, and a space value $s = 0$ for the space condition. Negative values indicate responses before alignment, and positive values indicate responses after alignment.

Upon identifying as outliers 0.298% of the trials in the time condition and 0.342% of the trials in the space condition, these were removed before proceeding to the analyses described below.

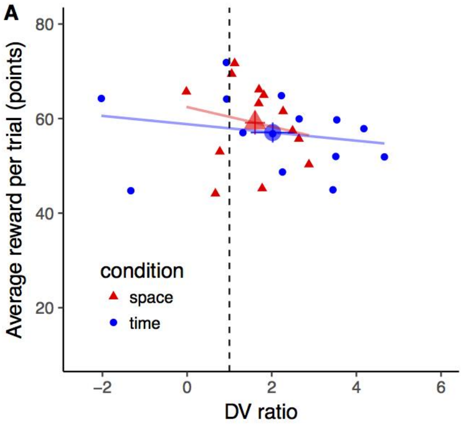


Figure 4.2. A. Average reward per trial (in points) as a function of the value of the DV ratio in Experiment 1. Results are split by condition, as denoted by color and shape. Bigger points represent the centroid of each condition's distribution, with associated error bars that show standard errors of the mean. Solid lines show linear fits of the data for each condition. The dashed, vertical, black line shows the cutoff point of DV ratio = 1, used to binarily categorize a group of responses as time or space-based. B. Average reward per trial (in points) as a function of trial block in Experiment 1. Results are split by condition. Data points are connected by lines and are surrounded by 95% confidence intervals calculated by bootstrapping.

Firstly, we focused on characterizing the DV of the participants, and whether this related to the average reward in the experiment. Using the cutoff point of DV ratio = 1 to categorize participants binarily as more time-based or more space-based, only 4 out of 14 participants were classified as time-based in the time condition, while the amount was 3 out of 13 in the space condition. A chi-squared test failed to reach significance for any differences between the two groups $\chi^2(1)= 0$, $p=1$). The correlation between DV ratio and reward was non-significantly different from 0 in any case (time condition: $r = -0.214$, $p = 0.462$; space condition: *r* = -0.197, *p* = 0.518). The lack of any clear trend can be seen in Figure 4.2.A.

We then aimed at exploring how different variables evolved across blocks. This was of special relevance here, since the present data were meant to be used for comparison with the next experiment, that would be testing effects of learning. Given the small number of temporal respondents, we did not split each condition according to its DV, comparing whole conditions instead. As in Study I, we fitted some LMM, performed an ANOVA on them, and explored with more detail any significant effects.

We started by focusing on the evolution of average reward across blocks. A LMM was fitted to our data, with average reward as the dependent variable. Block and condition were set as fixed effects, and participant as random effects. As seen in Figure 4.2.B, although participants in the time condition won less on the first block, the ANOVA on the LMM did not show a main effect of condition if this first block was excluded. As for the other main effect, reward increased across blocks (*F*(1,26620.6)= 34.401, *p*<.001), but at the slow average rate of 0.776 points / block.

Our focus then shifted to other relevant variables associated to the responses. We ran other LMMs, with these variables of interest as the dependent variable. Block, target speed and condition were set as fixed effects, and participant as random effects. The first of these variables was response times. These were not different between conditions, as supported by a non-significant main effect (*p*=.110), although response times were different among speeds (*F*(2,29031.1)= 25.986, *p*<.001), with later response times for faster speeds. Furthermore, response times became earlier across blocks (*F*(1,29031.1)= 237.695, *p*<.001) (initial response times and decrease per block, 20 cm/s: initial -0.035 s, -0.0009 s per block; 23 cm/s: initial -0.031 s, -0.0006 s per block; 26 cm/s: initial -0.029 s, -0.0004 s per block). The same could be said when replacing response times with response spatial locations, with significant main effects of target speed (*F*(2,29030.0)= 9.601, *p*<.001) and block (*F*(1,29030.0)= 220.783, *p*<.001), but not of condition (*p*=.108). Response locations became earlier across blocks, but they were closer to alignment for slower speeds (initial response locations and decrease per block, 20 cm/s: initial -0.698 cm, -0.018 cm per block; 23 cm/s: initial -0.717 cm, -0.014 cm per block; 26 cm/s: initial -0.769 cm, -0.011 cm per block). Finally, the SD of both response times and response locations was not different between conditions (*p*=.590 and *p*=.588, respectively), but it decreased across blocks (response times SDs: *F*(1,944.42)= 12.558, *p*<.001, initial average 0.024 s, -0.0001 per block; response locations SDs: *F*(1,943.00)= 10.818, *p*=.001, initial average 0.553 cm, 0.003 cm per block).

### 4.3.3. Discussion

Most participants of Experiment 1 planned their responses by mainly exploiting spatial information. This contrasts with the data of Study I, where more than half of the participants were closer to exploiting temporal information for those targets with a motion time of 1 s (as can be seen by a simple comparison of the yellow dots left and right of the dashed line in Figure 3.3.). The fact that we did not find a relationship between more time-based responses and larger reward could be due to the few participants with a clear temporal DV making the DV distribution across participants too homogeneous for that trend to be visible.

Responses and reward were very similar between conditions, with similar averages and evolution across blocks. Thus, the domain on which the reward was given did not greatly influence responses.

## 4.4. Experiment 2

Results from Experiment 1 show how, in both conditions, most participants mainly exploited spatial information to plan their actions. Considering that we previously showed how exploiting temporal information can be more advantageous, it would be particularly interesting if participants could somehow learn to use temporal information. In Experiment 2 we featured two different conditions, one where reward was given as a function of time to alignment and the use of temporal information was promoted (time condition), but also one where reward was given as a function of spatial distance to alignment and the use of spatial information was promoted (space condition). The particular mechanism we chose to foster the use of one or another information was through a penalization, where reward received after a trial was reduced if the past responses were not planned by using the relevant information. Specifically, in the time condition, the reward function linking responses to reward was scaled down according to how different average response times across speeds were so far in the past responses. For very similar response times, reward after each trial was close to the one received in the previous experiments. For more distant responses in time, the reward function was scaled down so that reward was reduced. This way, if participants wanted to increase their reward, they should respond with a more similar response time across speeds. The same happened in the space condition, but with response spatial locations.

Considering our design, in a condition where reward is given as a function of time, using a spatial DV by always responding when the target is at the same spatial location will yield different rewards for each speed. The same is true when using a temporal DV in a condition rewarding according to space (see Figure 4.1.). The wider the range of target speeds used, though, the more pronounced these differences in reward will be. This is why, in order for participants to realize whether they were using the appropriate DV, in parallel to the penalization procedure we widened the range of target speeds when compared to Experiment 1. In particular, we went back to those used in Study I.

To avoid any learning effect from Experiment 1 affect performance in Experiment 2, the new experiment was carried out with different participants.

Based on our past results, we expected larger reward for more time-based responses. If our method allowed people to learn how to exploit the appropriate cues, that meant that participants in the time condition would achieve larger reward. If learning was gradual, in this condition we should also find average reward increasing across trial blocks, at least for a while until it stabilized. In addition, the learning process should be captured by the DV ratio becoming more temporal. As for the space condition, most participants displaying a spatial DV should not be surprising, given that it seemed to be the default DV in Experiment 1. However, as in the time condition, both reward and DV might evolve across time.

## 4.4.1. Methods

### 4.4.1.1. Participants

26 new participants took part in the experiment after giving informed consent (21 women, 21 right-handed, age range 18-32). 13 completed the time condition and 13 the space condition. All had normal or corrected-to-normal vision and were naïve about the aim of the experiment. The study complied with the local ethics guidelines, in accordance with the declaration of Helsinki.

### 4.4.1.2. Apparatus and stimuli

All apparati used in experiment 2 were the same as in Experiment 1. The main change in stimuli is the fact that the target speeds in Experiment 2 were, as in Study I, 19.5 cm/s, 25 cm/s and 32 cm/s. To keep time to alignment at 1 second, targets started moving 19.5 cm, 25 cm and 32 cm away from the line, respectively.

### 4.4.1.3. Procedure

As in experiment 1, participants completed 2 blocks in the baseline session, and 12 blocks in the experimental session. However, the main change was that the reward given after each trial was modified in each condition in order to promote a spatial strategy in the space condition and a temporal strategy in the time condition (see below).

Scaling the reward function to foster the use of temporal or spatial information

In the experimental session of Experiment 2, in the time condition each trial's reward function depended on how close the last response times were across speeds. The more distant they were, the more that reward was scaled down.

$$u(t) = \beta \cdot exp(\frac{\alpha \cdot t}{1+\gamma}) \quad \text{if} \quad t \leq 1 \tag{4.1}$$

$$u(t) = -200 \quad \text{if} \quad t > 1$$

Where $\gamma$ is the difference, expressed from 0 to 1, between $\bar{t}_{max}$, the average response time of the past 5 trials with the maximum target speed (32 cm/s), and $\bar{t}_{min}$, the average response time of the past 5 trials with the minimum target speed (19.5 cm/s), all divided by $\bar{t}_{max}$.

$$\gamma = \frac{\bar{t}_{max} - \bar{t}_{min}}{\bar{t}_{max}} \tag{4.2}$$

Thus, $\gamma$ was updated after every trial where either the largest or the smallest target speeds had appeared. Consequently, if participants used a temporal strategy and responded at similar times for all speeds, reward would be close to the undistorted reward function used in Study I. Alternatively, using a spatial strategy meant responding at similar locations in space but different in time for each speed, and that led to reward being significantly scaled down. Likewise, in the space condition reward was scaled as a function of the closeness of the last response locations in space across speeds.

$$u(s) = \beta \cdot exp(\frac{\alpha \cdot s}{1+\gamma}) \quad \text{if} \quad s \leq 15 \tag{4.3}$$

$$u(s) = -200 \quad \text{if} \quad s > 15$$

Where $\gamma$ is the difference, expressed from 0 to 1, between $\bar{s}_{max}$, the average response time of the past 5 trials with the maximum target speed (32 cm/s), and $\bar{s}_{min}$, the average response time of the past 5 trials with the minimum target speed (19.5 cm/s), all divided by $\bar{s}_{max}$.

$$\gamma = \frac{\bar{s}_{max} - \bar{s}_{min}}{\bar{s}_{max}} \tag{4.4}$$

As can be seen in Figure 4.3.A. and 4.3.B., the larger the value of $\gamma$, the more reward was scaled down. For both conditions, at the start of each block reward did not suffer any distortion. Scaling started when four trials of both the largest and smallest speed had been presented, so that there were already some past responses to average. With this manipulation, we expected participants to respond at a similar point in the relevant reward domain across speeds, so that their reward was not diminished.

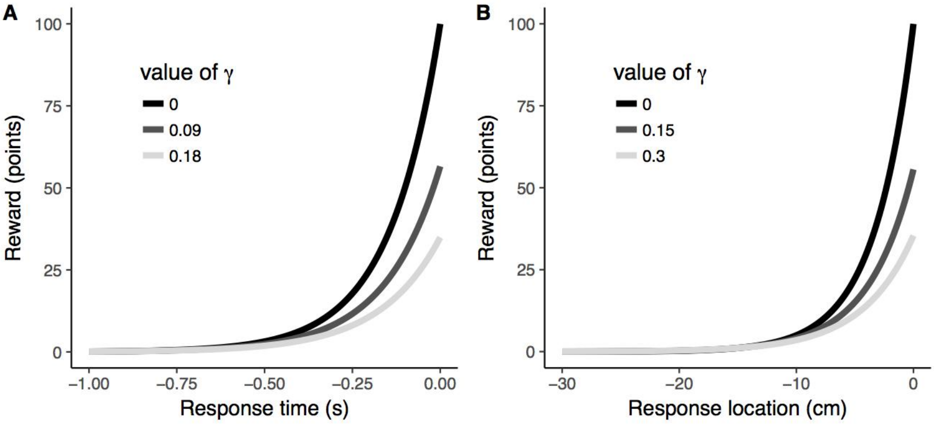


Figure 4.3. A. For the time condition of Experiment 2, reward as a function of the remaining time for alignment at response, depending on the value of $\gamma$. Larger $\gamma$ bring a larger penalization, achieved through downscaling the reward function. 3 example values

of $\gamma$ are shown. B. The same as in A, but for the space condition of Experiment 2, with 3 example values of $\gamma$.

### 4.4.1.4. Data analysis

The DV was obtained with the same procedure as in Study I. The criterion to identify and remove outliers was the same as the used in Study I and Experiment 1 of the present study.

## 4.4.2. Results

Upon identifying as outliers 0.298% of the trials in the time condition and 0.342% of the trials in the space condition, these were removed before proceeding to the analyses described below.

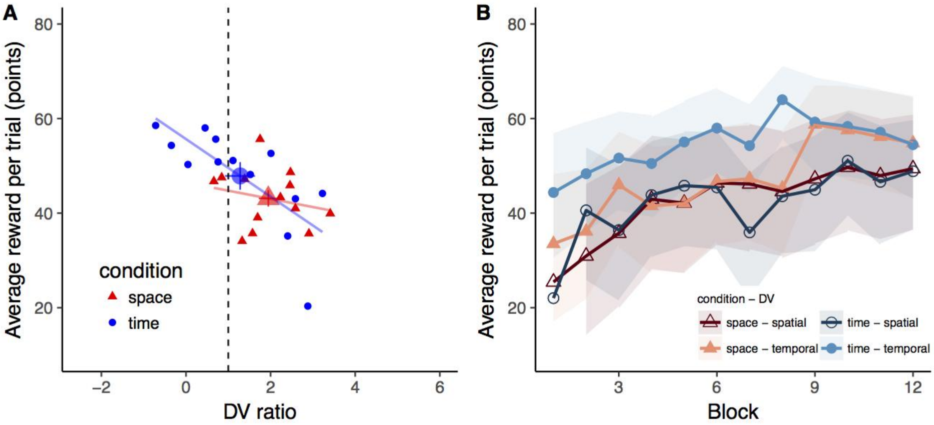


Figure 4.4. A. Average reward per trial (in points) as a function of the value of the DV ratio in Experiment 2. Results are split by condition and DV group, as denoted by color and shape. Bigger points / triangles represent the centroid of each distribution, with associated error bars that show standard errors of the mean. Solid lines show linear fits of the data for each condition. The dashed, vertical, black line shows the cutoff point of DV ratio = 1, used to binarily categorize a group of responses as time or space-based. B. Average reward per trial (in points) as a function of trial block in Experiment 2. Results are split by condition and DV group. Data points are connected by lines and are surrounded by 95% confidence intervals calculated by bootstrapping.

Binary categorization resulted in the time condition having 6 time-based and 7 space-based participants, while the space condition had 2 time-based and 11 space-based participants. Nevertheless, a chi-squared test determined that the number of participants in each DV group was non-significantly different between both conditions $\chi^2(1)= 1.625$, $p=.202$). In any case, the correlation between DV ratio and reward was significantly different from 0 in the time condition ($r$ = -0.737, $p$ = 0.004), but not in the space condition ($r$ = -0.225, $p$ = 0.459). As shown in Figure 4.4.A., in the time condition more time-based responses led to larger reward.

In this experiment there was a bigger presence of time-based responses. Consequently, we decided to ran group analyses subcategorizing each condition into each DV binary group. However, given the few number of temporal respondents in the space condition, we excluded that group from comparisons, although it will be displayed in the different figures. Thus, we compared those participants in the time condition with a mostly temporal DV with those in the time condition with a mostly spatial DV, and those in the space condition with a mostly spatial DV. We will refer to these groups as condition-DV groups.

### *Reward evolution*

In each condition, more use of the relevant information to plan responses translated into less downscaling of the given reward. As a result, we could have an initial insight on whether participants adapted to the use of the relevant information by taking a look at how reward evolved across blocks. We thus ran an LMM where average reward was the dependent variable, block and condition-DV group were the fixed effects, and participant the random effects. Figure 4.4.B shows how the average reward evolved for each condition-DV group, even for the one not included in the analyses. The ANOVA on the LMM showed significant main effects of block (*F*(1,25746.6)= 112.316, *p*<.001) and condition-DV group (*F*(2,36.6)= 8.278, *p*=.001), and both main effects produced a significant interaction (*F*(2,25746.6)= 3.007, *p*<.049). Thus, reward was different among condition-DV groups and, although it generally increased across blocks, the rate of reward increase per block was different (time condition, temporal DV: 1.065 points / block; time condition, spatial DV: 1.543 points / block; space condition, spatial DV: 1.898 points / block). Post-hoc contrasts were run to explore differences in final scores among these groups (all subsequent reported *p* are corrected for multiple comparisons with Hochberg's method). Temporal respondents in the time condition won more than both spatial respondents in the time condition (*p*=.003), and spatial respondents in the space condition (*p*<.001). Those with a spatial DV did not have any significant difference in reward between both conditions (*p*=.668). Temporal participants in the time condition, then, seemed to win more than the other groups. As the figure shows and the similar rates of reward increase seem to imply, this difference in reward were already present in the early part of the experiment. This was also supported by the fact that an ANOVA on the previous LMM but only considering the first 2 blocks of the experiment already had a significant main effect of condition-DV group (*F*(2,307.9)= 4.052, *p*=.018).

### *DV evolution*

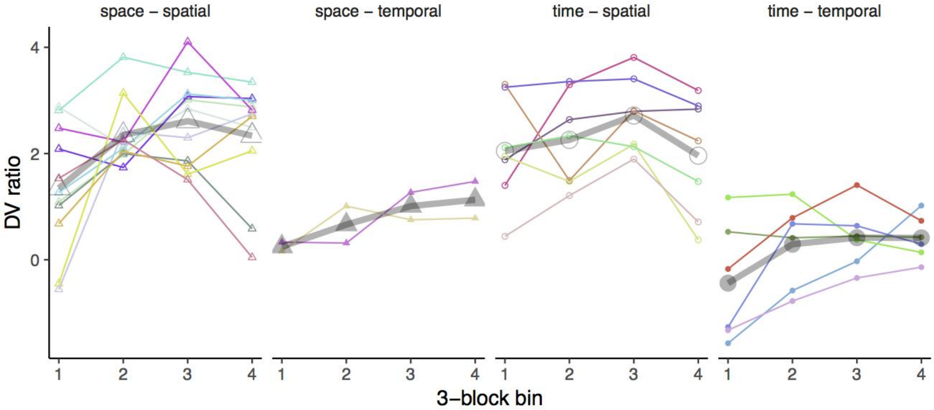


Figure 4.5. DV ratio across 3-block bins, for each participant of Experiment 2. Each panel groups participants with both the same condition and DV group, where the DV group is decided by classifying according to the DV of the whole experiment. Grey bigger points and thicker lines represent the average across participants.

Temporal participants in the time condition had a larger average reward already from the start. This somehow undermined the hypothesis that these participants owed their enhanced reward to the reward downscaling manipulation making them adapt to a more temporal DV. However, the most direct test of whether participants experienced a change in the information used to plan responses was taking a look at the evolution of their DV. We initially focused on the whole experiment by grouping blocks into 4 bins of 3 blocks each and calculating the DV ratio for every participant. Then, we saw how they evolved across bins in every condition-DV group. To do this, a LMM was fitted to our data with DV ratio as the dependent variable, bin and condition-DV group as fixed effects, and participant as random effects. The ANOVA on the LMM gave a significant main effect of bin (*F*(1,69.000)= 8.110, *p*=.006) and condition-DV group (*F*(2,79.683)= 10.480, *p*<.001), but no interaction (*p*=.174). However, a more detailed exploration of the evolution of the DV ratio in each condition-DV group showed that only the spatial participants in the space condition showed a significant change in their DV across blocks (DV ratio increase of 0.322 per bin, *F*(69.0)= 3.160, *p*=.002), while this evolution was marginally significant for those temporal participants in the time condition (DV ratio increase of 0.268 per bin, *F*(69.0)= 1.939, *p*=.057) and non-significant for spatial participants in the time condition (*p*=.879). As Figure 4.5. helps seeing, though, variability was very present and no condition-DV group presented a unified and clear trend across bins.

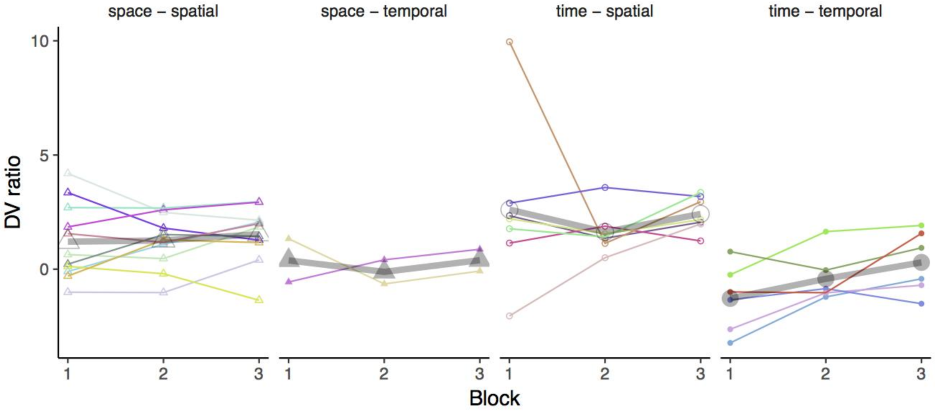


Figure 4.6. DV ratio across the first 3 blocks, for each participant of Experiment 2. Each panel groups participants with both the same condition and DV group, where the DV group is decided by classifying according to the DV of the whole experiment. Grey bigger points and thicker lines represent the average across participants.

As said before, the larger reward showed by participants with a temporal DV in the time condition was already present early on in the experiment. Thus, studying the DV across the whole experiment may not capture possible fast initial adaptations to the relevant DV. Considering this, we narrowed down the analysed data to the first three blocks of the experiment. A LMM, similar to the one performed before but replacing the block bins by individual blocks, gave a significant main effect of condition-DV group regarding differences in the DV ratio (*F*(2,65.999)= 6.754, *p*=.002), but neither the main effect of block (*p*=.146) neither the interaction between condition-DV group and block (*p*=.219) reached significance. Thus, as Figure 4.6. Helps appreciating, no early changes in the DV could be identified by this analysis.

*Response evolution*

The enhanced reward shown by those temporal participants in the time condition could be produced by many factors. Given that temporal respondents had displayed larger reward throughout the whole experiment, this could be influenced by a tendency to respond at a different moment than those with a spatial DV. To investigate this, we compared the responses of both condition-DV groups of the time condition: those with a predominantly temporal DV, against those with a predominantly spatial DV. In the LMM we ran, response times acted as the dependent variable, while block and DV group were set as fixed effects and participant as random effects. No significant main effect of DV group was found ($p$=.116): thus, there were no differences in the average response times for the whole experiment. However, the main effect of block ($F$(1,13955.4)= 7.629, $p$=.006) and the interaction between block and DV group ($F$(1,13955.4)= 52.286, $p$<.001) reached significance. Further analysis can link these results as temporal participants responding earlier in time in the first blocks (initial average: -0.042 s but then shifting their responses later in time across blocks (average 0.0007 s per block), while spatial participants started later (initial average: -0.035 s) and responded earlier in time as blocks went by (average -0.0003 s per block). Response locations followed the same trend, with the main effect of DV group being only marginally significant ($F$(1,11.6)= 4.242, $p$=.063) but block ($F$(1,13955.0)= 13.221, $p$<.001) and the interaction between the two ($F$(1,13955.0)= 60.599, $p$<.001) being significant. Temporal participants also started earlier (initial average: -1.145 cm) than spatial participants (initial average: -0.859 cm), but their responses evolved to be closer to alignment (average 0.020 cm per block), while spatial participants responded earlier in space as blocks went by (average -0.007 cm per block). Thus, although participants with a temporal DV and participants with a spatial DV followed different response trends, their average responses collapsed across blocks were very similar, and could not account for the difference in reward found throughout the whole experiment.

### *Loss fraction evolution*

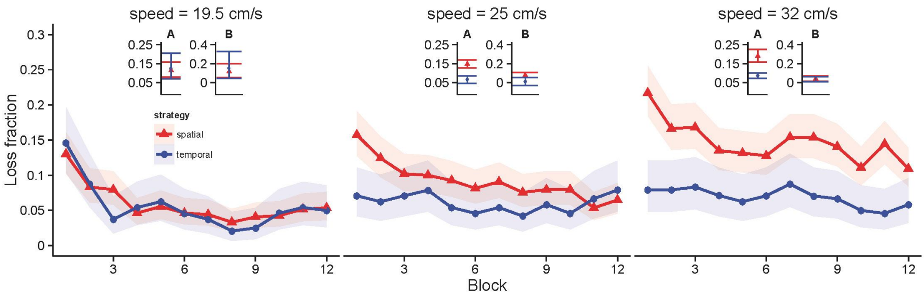


Figure 4.7. In Experiment 2, average loss fraction between strategies as a function of block. Shaded areas represent binomial 95% confidence intervals. Each panel corresponds to one of the three target speeds. Insets on top of each panel show the value and 95% confidence interval of A and B for both strategies.

We hypothesized that, regardless of how some participants displayed more time-based responses, the mechanisms mediating its enhanced reward would be the same as those unraveled in Study I. That is, a more stable performance across speeds. Nevertheless, we did not only aim at replicating our former results, but also at characterizing them across blocks. If those participants exploiting more time-based responses did so from the start of the experiment, their rate of penalized trials should be lower from the start. We thus compared the loss fraction (proportion of penalized trials) of the groups of temporal and spatial participants by making use of the cutoff point ot DV ratio = 1, collapsing data across conditions. To give more power to our analyses, this time we used the whole sample, so the temporal group included also the condition-DV group of temporal respondents in the space condition. We fitted an exponential decay model of the probability of losses with block (p(losses) = A*exp(-B*block)) for the different speeds. The model was fitted by minimizing mean least squares, and consisted of two parameters. The scaling parameter A informed about the initial value of the loss fraction, and the parameter B corresponded to the decay rate of the loss fraction across blocks. For each DV group and speed we obtained both parameter values using nonlinear regression (through the base R function *nls*), as well as their 95% confidence intervals. Figure 4.7. shows the average loss rate across blocks. While parameter A was similar between strategies when the speed was 19.5 cm/s, the confidence intervals did not overlap, but they did when the speeds were 25 cm/s (spatial strategy: A value: 0.065, confidence intervals: 0.047 - 0.086; spatial strategy: A value: 0.148, A confidence intervals: 0.128 - 0.170) and 32 cm/s (temporal strategy: A value: 0.087, confidence intervals: 0.073 - 0.102; spatial strategy: A value: 0.190, confidence intervals: 0.158 - 0.225). This shows that, for faster speeds, participants with a spatial strategy were penalized in more trials, what replicates our previous results. Parameter B was only different when the speed was 25 cm/s, with those with a temporal strategy (B value: 0.011, confidence intervals: -0.032 - 0.055) having a less steeper decay rate than those with a spatial strategy (B value: 0.080, confidence intervals: 0.056 - 0.105). The figure helps visualizing how the average loss rate was considerably flat across trials for the temporal group. And although the loss rate experiences an initial drop for the slowest speed, in the faster speeds the loss rate is flat from the very beginning. This strengthens the idea that those with an overall temporal DV displayed it from an early stage.

### 4.4.3. Discussion

In this experiment we introduced a method that introduced a penalization on reward. The more the participant deviated from the use of the cues we wanted to promote, the more reward was reduced. Had participants been able to identify and adapt to this penalization mechanism, that should be reflected in their DV. However, we found either a non-significant evolution of their DV or a small trend to become more spatial. Further analyses could not even find an early adaptation over the initial blocks of the experiment. Thus, there was no evidence for our manipulation making participants learn how to exploit the promoted cues.

Any trend concerning the evolution of the DV is put into doubt when taking a look at individual participants. As Figure 4.5. shows, variability is very present, so averages may not reflect a common DV trend. This lets us conclude something, though. Even in the unlikely case that the penalization system in the time condition had produced an initial adaptation that could not be captured by our analyses, participants could not keep their DV stable across blocks.

As it had happened in the experimental work presented in Study I, participants with more time-based responses had a larger average reward. Interestingly, most of these participants were part of the time condition, although a significantly difference of temporal participants between conditions was not statistically backed. The fact participants with a temporal DV displayed larger reward from the start of the experiment supports the claim that they did not learn how to exploit temporal cues after a learning process. As for the enhanced reward related to temporal DVs, this could not be explained by a difference in the times or spatial locations at which responses were made. Rather, as in the previous study, it was due to participants with more temporal DVs being more stable across speeds. This was reflected by the loss fraction (rate of penalized trials) not increasing for faster target speeds. Further exploration on how the loss fraction evolved across blocks revealed that, in general terms, it decayed at the same rate for both DV groups. However, temporal participants had a lower loss fraction already from the initial blocks, what further supports our previous claim.

## 4.5. General discussion

Study I showed that, when making decisions with moving objects, exploiting temporal information could be more advantageous. Spurred by these past results, we formulated the question of whether people could learn to use temporal cues, so that they could enhance their performance in this kind of tasks. Particularly, we wanted to explore whether reward could be used as the information to learn the appropriate DV. Our approach was to penalize reward after each trial relative to how much the story of past trials deviated from the use of the relevant cues. When compared to a baseline experiment, more participants adopted a temporal DV, but the lack of evidence for learning points at the reward manipulation not having a clear effect.

Two closely related questions remain regarding how the behavior of participants compares in the time condition of the two experiments. The first is why, in Experiment 2, more participants had a temporal DV than in Experiment 1. The second is why the relationship between more time-based responses and larger reward appeared in Experiment 2, but not in Experiment 1. Trying to speculate about the first question, one possibility is that the wider target speed range helped some participants identify the domain according to which the reward was given. Another is that, given a considerably faster maximum speed in Experiment 2 (32 cm/s) than in Experiment 1 (26 cm/s), participants could have been aware that responding at the same spatial location resulted in too many responses after the line for the fastest speed. In order to avoid this, people could have extracted velocity information and respond earlier in space for faster speeds. This would also relate to the second question formulated above: for an experiment with faster speeds, using a temporal DV would relate to a larger reward.

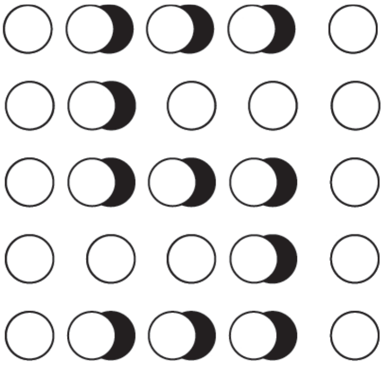
To act as a comparison for the time condition of Experiment 2, we ran the space condition, where we used the same mechanism to promote spatial cues. Results for this condition are at odds with the hypotheses introduced in the previous paragraph: that a wider range of target speeds accounted for more temporal participants. Although the space condition of Experiment 2 featured the same speeds as the time condition, very few participants had a mainly temporal DV. Despite a higher chance of responding after the line for faster speeds, the penalization system in the space condition could have made responding at the same spatial location to bring more reward in the long run. If this was so, the penalization structure may have had an effect after all. The contribution of the speeds could be clarified by an experiment where the target speeds of Experiment 2 were used, but no penalization was applied. The experiments in Study I match these criteria, although the multiple motion times interleaved may limit comparison between studies. In Study I, for motion times of 1 s, more than half of the participants were categorized as having a temporal DV even when reward was given as a function of space (see Figure S1 in the appendix). Thus, given the scarce number of participants with a temporal DV in Experiment 2 of the present study, we could think that, in the space condition, the penalization system could have kept participants spatial.

Going back to the ultimate goal of this study, the fact is that our reward manipulation could not make participants use temporal cues. It may simply have been too difficult for participants to grasp what reward depended on. In past experiments, people ignoring whether reward was given as a function of time or space was not translated into a severe penalization. Here, though, it was much more important to be aware of that, but the way of discovering it was keeping track of how close the past responses were in the relevant dimension, and realize reward was related to that. In the presence of such complexity, participants could have found reward not to be very informative to plan their responses. In consequence, they may have relied more on the place where they saw the target disappear. This connects our work with the literature that compares two learning systems: one based on reward prediction error, the other on sensory prediction error (i.e. Herzfeld & Shadmehr, 2014).

It may also be possible that, with much more training, participants could have finally extracted these rules and acquired a temporal DV. Perceptual learning experiments typically consist of several hours of training in order to improve just a basic perceptual attribute. And even within those attributes, speed discrimination, which may subserve a temporal DV, is learnt at a lower rate than other attributes like motion discrimination (Saffell & Matthews, 2003)

Since temporal participants were temporal from the start, individual differences may be at the root of some participants displaying a temporal DV. Identifying them would then be a good approach to design a method for learning how to exploit temporal cues.

Some other reward manipulation could make people learn how to use temporal cues. Nevertheless, there is the possibility that and easier way to induce this learning is by providing more direct feedback on the DV. This would imply giving explicit guiding on when to respond for each target speed, instead of manipulating the reward function. For instance, in a condition promoting the use of temporal information, feedback could consist in displaying the time left to alignment at the moment of response, so that participants could make this time be the same for all speeds. Future work should explore different ways of providing feedback.



# Study III - Which variability do people consider when making sensorimotor decisions with moving objects?

## 5.1. Summary

In many sensorimotor situations, when planning our actions we must take into account our variability so as to obtain the best possible outcome. Different work reported people being unable to plan their responses so that they maximize the expected reward of the task, as defined by normative models derived from statistical decision theory (SDT). The origin of this suboptimality, though, remained unclear. In our previous studies, participants used reward information to exert sudden corrections in their aimpoint so as to minimize penalizations. We hypothesized that the variability introduced by these corrections would not be considered when planning responses, and this could be the case of many other studies reporting suboptimality. To find support for our hypothesis, we took previous data and fitted a Kalman filter to it, which allowed estimating each participant's variability discounting that added by the corrections. Observed data was well explained by participants responding at the point that, according to SDT models, maximized expected reward for that reduced variability. There were also hints that pointed to participants' responses being most stable at reward regions that were optimal for that variability.

## 5.2. Introduction

Models derived from SDT have been extensively used to provide a normative account on how perceptual and sensorimotor decisions should be made. Different work has been devoted to explore the conditions under which optimality is achieved, and also those where it is not. Some endeavour has targeted highlighting deficiencies in the classic normative framework, as well as its derived models. Some claim that existent models may classify as optimal behavior what may not be so. For instance, Jarvstad et al. (2014) point out how commonly used optimal movement planning models (Trommershäuser et al., 2003b) may not offer a correct measure of optimality if people fail to maximize precision. Conversely, other work criticizes normative theories because they do not consider certain elements that, once taken into account, would turn non-optimal behavior into optimal. As an example, Tsetsos et al. (2016) defend that violating transitivity (one of the axioms of classical decision theory) can be optimal when the decision process is characterized by high neural noise.

In Study I we interpreted our data from the point of view of SDT-derived optimal movement planning models. There we already argued how the assumptions of these models limited their usefulness to assess optimality in our task. However, we also mentioned how our data seemed to reproduce a pattern of suboptimality previously identified by other authors. In some tasks with asymmetric reward functions, the action that maximizes expected reward requires some sort of response shift with respect to what, if the task had a symmetric reward function, would be the optimal response. For instance, in our task the maximum reward was given when responding at alignment, but since responding at any later point was penalized, it was better to respond on average before alignment, so as to avoid penalizations. According to the aforementioned models, the optimal response shift would depend on the person’s motor variability. The more variable someone is, the bigger the optimal shift has to be to maximize expected reward. In other words, the optimal response would be that which maximizes reward in the long term by finding the best trade-off between scoring as much as possible and minimizing penalizations. In some experiments, participants have been found to shift their responses, but still fall short of the optimal shift. Thus, their responses would be suboptimal.

This phenomenon has been interpreted in different ways. According to Mamassian (2008) (see 1.1.4.2. For more details), it would be revealing overconfidence. Participants in his timing task would underestimate their motor variability, and that would make them respond with an insufficient shift. Other authors who found similar results, both in timing (Ota et al., 2015) and reaching (O'Brien & Ahmed, 2013) tasks, labeled this behavior as risk-seeking. However, not much more was said to justify this interpretation. Interestingly, one finding seems to be recurrent. As in other decision-making experiments (i.e. Neyedli & Welsh, 2013a), one could expect that, with practice, a participant would eventually discover the response that maximizes her expected reward, and stick to it thereafter. However, in the past works that does not happen, even after days of practice (Ota et al., 2016). The present study tries to shed some light on the origin of this persistent suboptimal response shift. Our hypothesis points to people not planning their responses by considering their overall observed variability. They would rather act based on their measurement error, without taking into account the added motor variability derived from planned trial-to-trial corrections. Our basic rationale can be better understood explaining first the relevant concepts related to the tools we will use in this study, and then illustrate these concepts with an example involving our previous data.

In the kind of tasks we have been using, the overall observed response variability can be decomposed into several sources, following models of motor learning that explain how people make corrections (van Beers, 2009; van Beers et al., 2004). One type of variability would originate from planned trial-to-trial corrections. At any one trial, we would have a planned or aimed response time, and the observed response time would correspond to this planned time had the motor command not been corrupted by noise in the planning. However, different sources of noise further add to the observed variability. One clear source of variability is the sensory noise in our task. Unlike reaching static targets, moving objects generate more sensory uncertainty on the position, and this uncertainty scales with the speed of the object (see Study I). Suppose, for example, that one participant bases her response on a position threshold. Sensory noise will then affect the response variability. Finally, another source of variability is caused by what is usually called execution noise (Diedrichsen, Hashambhoy, Rane, & Shadmehr, 2005; van Beers, 2009). This noise refers to the pure motor noise added in motor-neurons.

To disentangle all the previous sources within the observed variability is very difficult (van Beers, 2009). Most models that cope with this problem are based on state space models or Kalman filters (i.e. Burge, Ernst, & Banks, 2008; Wei & Körding, 2010), which can estimate process (internal) noise and measurement noise. We will follow this approach, including the terminology, and decompose the variability in two, but first we will describe what each of the two variability categories most likely includes. In our scenario we can clearly assimilate the variability due to planned corrections to the process noise, since it is part of the aimpoint setting process. However, the process noise might well include other sources of variability like sensory noise. This is because sensory processes and decision processes about when to start a response are intimately related when dealing with moving objects (Eckhoff, Holmes, Law, Connolly, & Gold, 2008). The success of a state-space model to capture all the process noise accurately and separate it from the measurement noise depends on how well the process noise is specified.

If there is an incomplete specification, then the unexplained variability will add to the measurement noise. As for the measurement noise, in our case it clearly includes the pure motor noise, but we cannot discard that some sensory noise is somehow contributing to it. However, we take a simple approach and aim at discounting at least the variability due to corrections from the observed variability and assign it to the process variability. This way, we make sure corrections are not affecting the measurement variability. In conclusion, we believe our decomposition of the variability is good enough for our purposes. We can clearly distinguish each of these two variabilities in the next example, that uses real data coming from Experiment 1 in Study II.

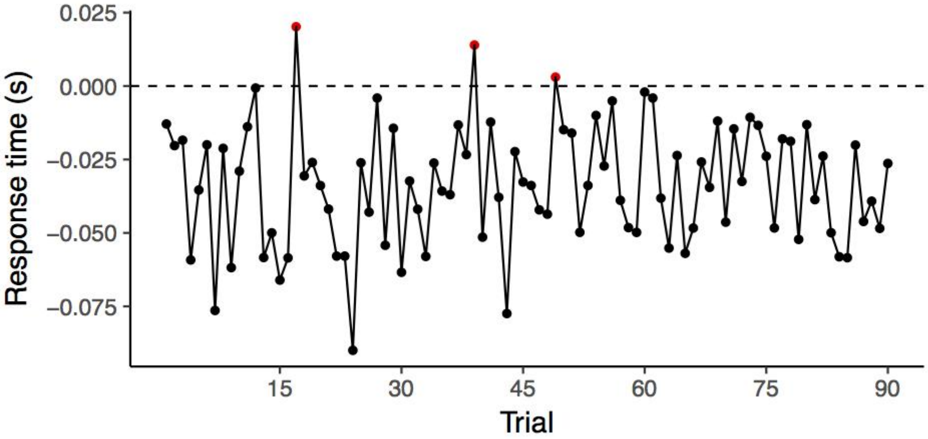


Figure 5.1. Response times relative to alignment as a function of trial number. The data reproduces block 12 from participant 12 in Experiment 1 from Study II. Points in red correspond to responses after alignment, which were penalized. Alignment is shown by a horizontal, dashed line.

Figure 5.1. displays the response times of a typical participant for a whole block. As can be seen, each trial's response is different from the previous one. Imagine a participant aim to respond at the same moment across different trials. Given all the sources of variability, consecutive

responses will inevitably vary to some extent. This is exemplified by the small jitters in trial-to-trial responses observed in the figure. We assume this captures what we call measurement variability. However, much larger differences can also be appreciated between some pairs of trials. Most of these differences in responses are product of planned corrections, where the participant intentionally aims at changing the moment of the response. A quick look at the figure shows that this typically happens after penalized trials (colored in red): that is, when the response has been made after alignment. In these cases, participants are most likely correcting their aimpoint in order not to incur into further penalizations. However, planned corrections are also happening after non-penalized trials: for instance, after responding just before the line, despite being highly rewarded, participants often made corrections. This could be probably linked to the near-losses phenomenon (Y. Wu, van Dijk, & Clark, 2015) in which the reward is attributed to high perceived luck. Importantly, the block reproduced here corresponds to the late part of the experiment, showing that, although responses can get less variable over time, corrections keep happening throughout the whole experiment.

Those studies reporting the suboptimal response shift (the most clear being Mamassian, 2008; O'Brien & Ahmed, 2013; Ota et al., 2015, 2016) share some common elements with our previous work. As a result, these similar aspects in the experimental design may have produced the existence of corrections. Firstly, the aimpoint where optimal responses should be placed is not explicit, so it is reasonable to assume that participants will explore, switching their responses back and forth as a consequence of the reward they receive. In addition, given the asymmetric reward function, small response differences can change the outcome from receiving the maximum reward to being penalized. Could, in all these experiments, suboptimality be related to participants not considering the variability generated by these corrections?

SDT models assume that ideal agents have a true representation of their own variability, and they use this knowledge to plan the optimal response, given the reward function of the task. Mamassian (2008) suggested that, from an ideal agent's point of view, participants were responding as if their variability was smaller than the one they displayed. Thus, suboptimality would be originated from a biased representation of their uncertainty. We go one step further and try to identify the variability participants have access to when planning their responses. What we suggest here is that, in our scenario, suboptimal responses could be partly attributed to participants using their measurement variability, without incorporating the variability arising from their planned corrections. Direct support for this hypothesis would come from observed behavior being close to optimal predictions if the latter were computed by using the participant's measurement variability, instead of the standard deviation of the observed responses, which is the way they are normally obtained.

In order to investigate the previous hypothesis, the different types of variability should first be disentangled. The approach we followed was to process our previous data by fitting them to a modified version of the Kalman filter (Kalman, 1960). This algorithm is commonly used in many fields to estimate the unknown true state of a system. To do that, the Kalman filter combines two elements in a Bayesian way. On one side, the last of a series of noisy measurements. On the other, a prediction coming from the previous measurements and prior knowledge about how the system behaves. Since the true state of the system is uncertain, both the measurement (or observation) and the prediction of the system (or process) are given as probability density functions, with a mean and a variance. In our case, the variance of the process would be equivalent to that coming from the response planning, including the corrections. The variance of the measurement would at least include the execution noise. Our use of the Kalman filter was directed at separating the process and the measurement variabilities, not at describing how participants estimated any uncertain state of the stimulus. Once the measurement variability was obtained, it was fed into one of the expected reward maximization models, to see whether the optimal responses they generated were similar to the observed responses.

## 5.3. Methods

### 5.3.1. Participants, apparatus and stimuli

This study consisted in a reanalysis from data previously presented in this thesis. Specifically, we recovered the whole dataset from Experiment 1 in Study II. This consists of 27 participants (14 for the time condition, 13 for the space condition), who ran 12 blocks of 90 trials each. Targets took 1 s from motion onset to reach alignment (single motion time), and they could adopt speeds of 20, 23 and 26 cm/s. To see the full information about the participants, apparatus, stimuli and the conditions of data collection, see section 4.3.1. We chose to analyze this particular dataset for one reason: of all the data we had collected, this was the most simple one. Firstly, because targets had a single motion time, so we reduced the effect that different integration times could produce on trials. Secondly, because the small range of speeds made the time and space conditions very similar in terms of reward, independently of the DV the participants were using. Thus, we could minimize these differences across participants and focus on trial-to-trial responses.

## 5.3.2. Procedure

We started by fitting our data to the Kalman filter. However, while the original Kalman filter assumes that the variance of the process is stable (stationarity assumption), in our task the existence of corrections leads us to think that the true state (aimpoint) is not stable but changing and therefore introducing changes in the process variance. This is why we used a version of a Kalman filter (Narain, van Beers, Smeets, & Brenner, 2013) that accounts for nonstationary processes. The distinguishing feature of this model is that it estimates process variance within a window of trials.

The Kalman filter consists in two basic steps, which are performed in a recursive fashion. The first of these steps is the prediction of the state of the world,

$$\hat{x}_{\bar{k}} = \hat{x}_{k-1} \tag{5.1}$$

At every time step, a prediction ($\hat{x}_{\bar{k}}$) is made based on the state estimate of the previous time step ($\hat{x}_{k-1}$).

$$\sigma_{\bar{k}}{}^{2} = \sigma_{k-1}^{2} + \sigma_{p_k}^{2} \tag{5.2}$$

The variance of the prediction ($\sigma_{\bar{k}}{}^{2}$) equals to the variance of the previous state ($\sigma_{k-1}^{2}$) plus the variance of the process ($\sigma_{p_k}^{2}$).

The second recursive step is the update of the state estimate. The estimate will be generated by combining the variance of the prediction with the variance of the measurement ($\sigma_m^2$). The proportion each variance contributes to the estimate is formalized by the Kalman gain ($K$).

$$K_k = \frac{{\sigma_{\bar{k}}}^2}{{\sigma_{\bar{k}}}^2 + \sigma_m^2} \tag{5.3}$$

Such that, the higher the Kalman gain, the more the estimate relies on the measurement. When the Kalman filter is used in standard estimation contexts, the Kalman gain indicates how much the model relies on previous information and how much on the current observation. Also, the Kalman gain can be regarded as the amount of correction in the planning process (van Beers, 2012), which is very adequate in our context. After certain trials, such as those where the response is penalized, we can expect the gain to increase, which would reveal giving less weight to the previous aimpoint and select a different one. This would predict a subsequent correction. On the other hand, a decreasing gain would reveal that responses are stabilizing around a common aimpoint, although very small corrections could also be present.

The variance of the measurement is assumed to be stable and it is the only free parameter of the model we use, and the value we want to estimate. Two more equations involved with the update step will be introduced. The first describes how the state estimate ($\hat{x}_k$) is updated before the next observation ($z_k$) is made.

$$\hat{x}_k = (1 - K_k)\hat{x}_{\bar{k}} + K_k z_k \tag{5.4}$$

The second describes the variance at a given state ($\sigma_k^2$).

$$\sigma_k^2 = (1 - K_k)\sigma_{\bar{k}}{}^2 \tag{5.5}$$

Finally, in this nonstationary model, the process variance is estimated from a fixed window of past observations. The predicted mean ($\bar{\mu}_k$) of the true process and its variance are updated in the following way:

$$\bar{\mu}_k = \frac{1}{T}\sum_{i=k-T}^{k-1} z_i \tag{5.6}$$

$$\sigma_{p_k}^2 = \frac{1}{T-1}\sum_{i=k-T}^{k-1} (z_i \tag{5.7}$$

Where *T* is the window size. We will fix *T* to 4, following the optimal value provided by Narain et al. (2013). Thus, process variance is estimated by taking into account the variance of the past 4 observations (in our case, trials). This dynamic variance is originally devised to capture rapid adaptation to changes in the environment's statistics. In our case, considering only a small amount of past trials allows us to capture the change in variance introduced by corrections. A consequence of this dynamic process variance is that the Kalman gain will also be dynamic. Focusing on our scenario, the Kalman gain will increase after a correction, until responses start to stabilize and the gain decreases again. Of note is that this model does not capture why the correction is made. In an experiment like ours, most corrections are probably due to the experienced reward after a trial. However, none of the equations incorporate a reward element. As previously said, we simply use the Kalman filter as a tool that helps us extract the measurement variability.

The method for fitting the model to the data is based on parameter optimization with respect to the participant's responses. We minimized the negative log likelihood of the Kalman filter to obtain the measurement variability that provided the best account of the data according to maximum likelihood estimation (Caines, 1988).

For each participant, we fitted the nonstationary Kalman filter to every block, and obtained a measure of the measurement variability in terms of variance, henceforth measurement variance. However, we also performed a fit with a stationary Kalman filter (Kalman, 1960) in which the process variance was constant and therefore not accounting for changes in the planned time. The likelihood of both models was compared so as to justify that the nonstationary model provided a better fit.

Once the measurement variance was estimated, we took its square root to obtain the standard deviation. Then, we proceeded to feed it into the expected reward maximization model previously used in Study I, to calculate the optimal response. In parallel, we did the same procedure but using the standard deviation from the observed responses. Finally, we determined whether any of the two optimal responses fell within the confidence intervals of the observed responses.

## 5.4. Results

In previous studies we filtered the data by eliminating those trials considered as outliers. However, since the Kalman filter needs the whole history of events to operate, we could not follow the same procedure here. This is why we chose to replace those outlier responses with the median of the responses for that participant. Outliers were identified by applying interquartile range differences (McGill, Tukey, & Larsen, 1978). Replaced outliers accounted for 1.680% of the trials. In all figures of this section, participants of the time and space conditions are easy to tell apart. In those figures where a number identifying each participant is shown, participants 1 to 14 are those of the time condition, while participants 15 to 17 are those of the space condition. However, due to non-existent significant differences between conditions (see figures in this section, but also section 4.3.2. in Study II), we did not engage into condition comparisons.

We started by statistically backing the nature of the corrections we suggested in the introduction. The most obvious corrections would happen after penalized trials, where participants would switch their aimpoints earlier. We ran a LMM with the response time of trial n+1 as the dependent variable, and the response time of trial n as fixed effects. Individual participants acted as random effects varying both slope and intercept. The data used for this LMM included only those cases where trial n had a response time after the line. The ANOVA on the LMM gave a significant main effect of the response at trial n ($F(1,17.354)= 8.634$, $p=.009$), with average responses at trial n+1 significantly earlier than alignment (average -0.036 s, $p<.001$). This supports the existence of corrections after penalized trials.

We hypothesized that corrections may have also happened after responses made shortly before the line. Despite the high reward given for those responses, participants could have thought their current aimpoint was too close to the line to avoid penalizations. This would reveal some sort of sensitivity to their variability. We ran a very similar LMM, but this time selecting the data where reward in trial n was of 85 points or more. The main effect of response at trial n+1 also reached significance (*F*(1,278.36)= 17.139, *p*<.001), with average responses at -0.036 s, significantly before alignment (*p*<.001). These results point to participants correcting even after non-penalized trials.

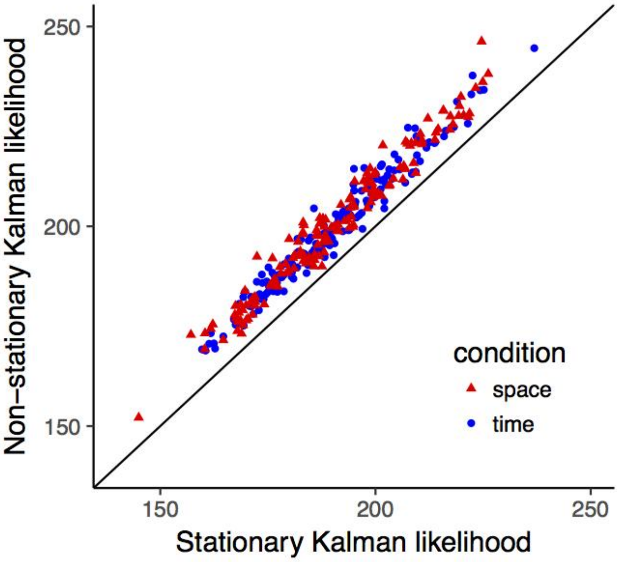


Figure 5.2. For every participant and block, non-stationary Kalman likelihood as a function of its respective stationary Kalman likelihood. Data points corresponding to participants from the time condition (reward given as a function of the time to alignment at response) are plotted as blue dots, while those corresponding to participants from the space condition (reward given as a function of spatial distance to alignment) are plotted as red triangles. The black, solid identity line represents the trend data should follow if both likelihoods were equal.

We proceeded to fit both the stationary and the nonstationary Kalman filter to our data, so that we could see whether the former offered a better fit. For every participant and block, the likelihood of one and the other models was obtained. Both likelihoods are plotted against one another in Figure 5.2. In all cases, the nonstationary model fitted the data better, as seen by the whole set of data points lying above the diagonal identity line.

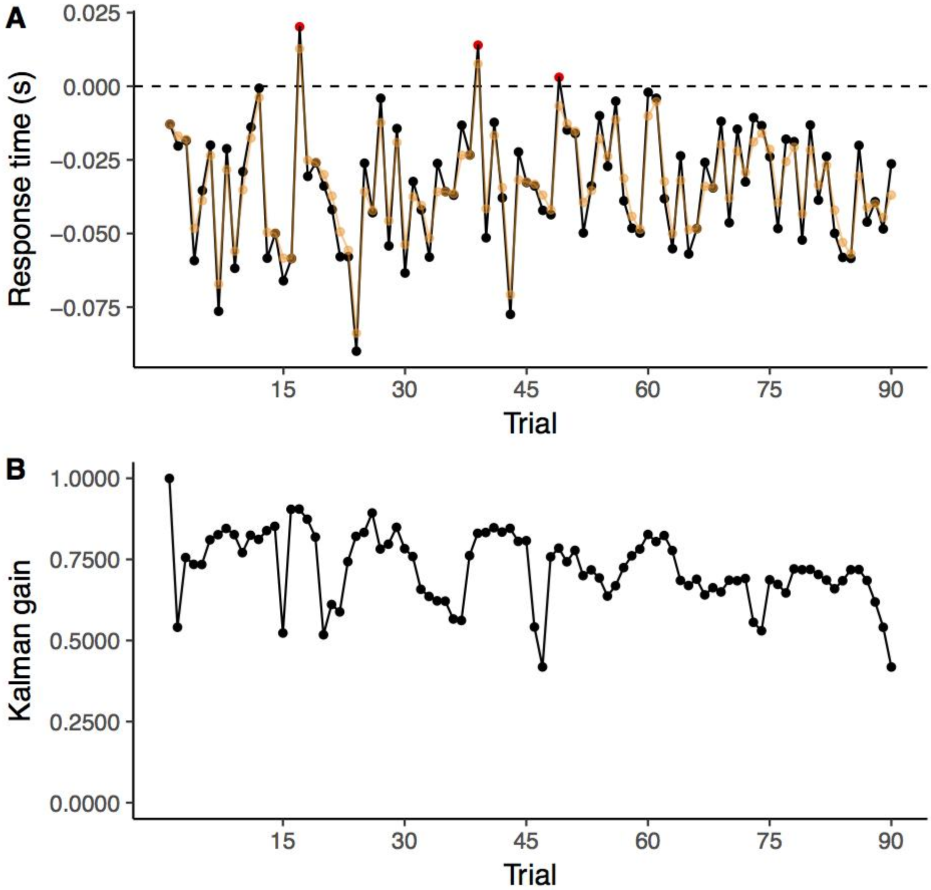


Figure 5.3. A. Response times relative to alignment as a function of trial number. The data reproduces block 10 from participant 12 in Experiment 1 from Study II. Points in red correspond to responses after alignment, which were penalized. Alignment is shown by a horizontal, dashed line. Superimposed orange points connected by orange lines

depicts posterior responses for each trial, as calculated by the nonstationary Kalman filter. Each posterior is used as the prior in the next trial. B. For the data shown in A, Kalman gain for every trial.

Figure 5.3. illustrates how the nonstationary Kalman fit adjusted to the data. In particular, we focus on the same participant and block depicted in Figure 5.1. In Figure 5.3.A, the posterior for each trial (orange points) is superimposed over the response times (black points). Each trial's posterior, which is obtained by combining that trial's observation with that trial's prior, is used as the prior for the next trial. It can be seen how this combination makes posteriors a smoothed version of the actual response time. Figure 5.3.B shows, for the data in A, the Kalman gain of the model. When comparing both panels of the figure, it can be appreciated how an increased gain predicts large corrections, while it goes down as responses stabilize.

The main reason for which we used the Kalman filter, though, was to obtain an estimate of the measurement variance. Before using it to calculate optimal responses, we analyzed possible changes throughout the experiment. We ran an LMM with the measurement variance estimated with the Kalman filter as dependent variable, block as fixed effect, and participant as random effects. The non-significant main effect of block (*p*=.109) produced by an ANOVA on the LMM determined that measurement variance did not change as the experiment went by.

We then calculated the optimal responses with the expected reward maximization model. However, we computed two versions of them. On one hand, the one derived from the measurement variance. On the other, the one where we used the standard deviation of the observed responses, without the data being processed by the Kalman filter.

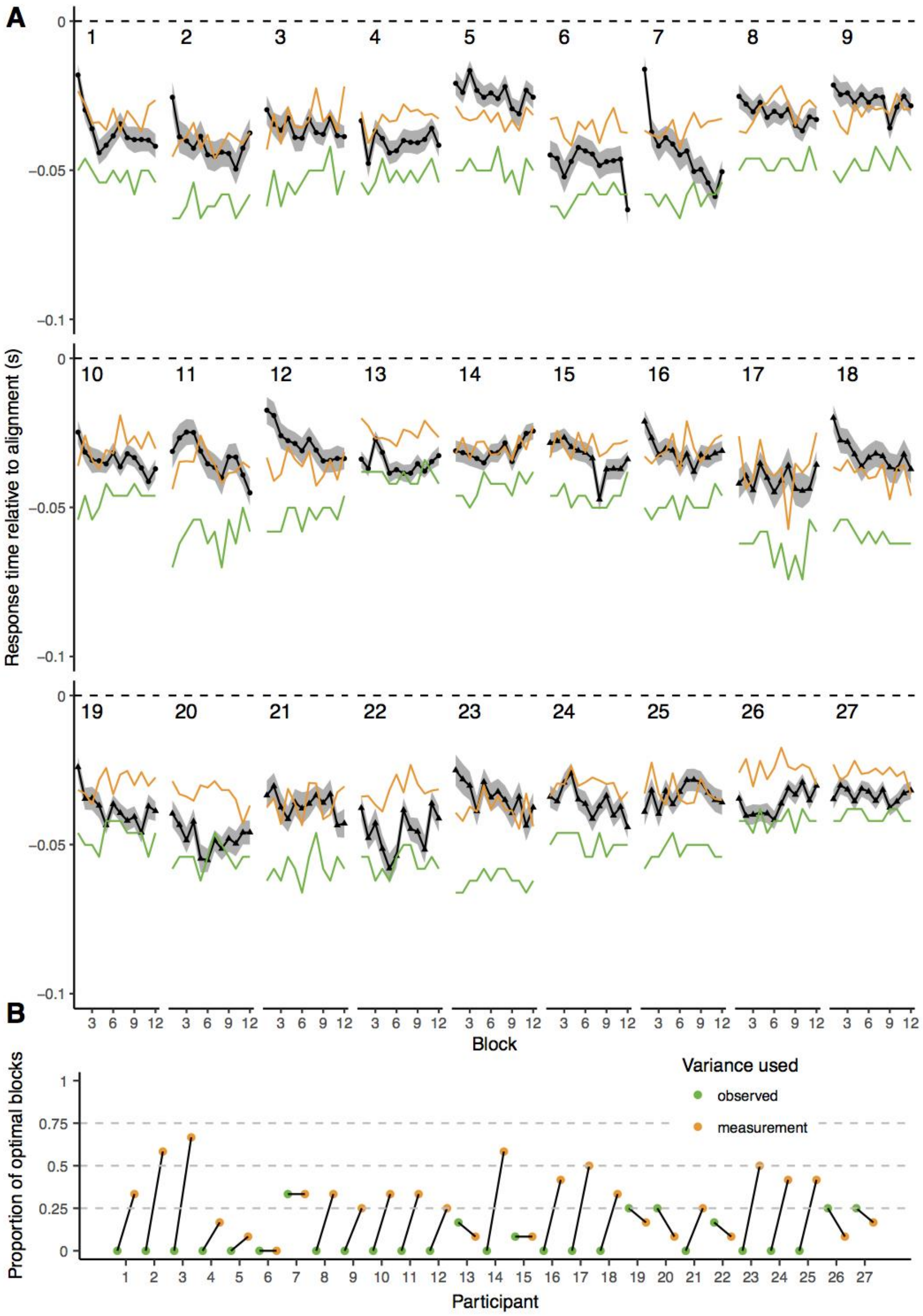


Figure 5.4. A. For each participant (denoted by a number), response time relative to alignment, as a function of trial block. Black dots connected by lines represent average

observed response times, surrounded by 95% confidence intervals calculated through bootstrap, in grey. Participants from the time condition are denoted by dots, while those from the space condition are denoted by triangles. Green and orange lines denote optimal responses as calculated by an expected reward model, but each use a different standard deviation. Green lines denote using the observed standard deviation, while orange lines denote using the standard deviation coming from the measurement variance, as obtained with the nonstationary Kalman filter. Black, dotted horizontal lines indicate alignment. B. For each participant, proportion of optimal blocks, where optimal blocks are defined as those where the optimal response lays within the confidence intervals of the observed average response. For each participant, the green point indicates the proportion when using observed SDs to calculate optimal responses, while the orange point reflects using SDs from measurement variance.

Figure 5.4.A depicts, for every participant, the average observed response times per block, surrounded by confidence intervals. The optimal responses made from the standard deviation of observed responses (in green) and those made from the measurement variance (in orange) are also shown. The orange lines are closer to alignment, given that the measurement variance is necessarily smaller than the overall observed variance. Nevertheless, the interesting question is whether they are within the confidence intervals of observed responses. A quick look allows to conclude that, although this is not overwhelmingly so, in most participants the orange lines are within the confidence intervals in more blocks than the green lines are. This can be easily seen in Figure 5.4.B, showing, for every participant, the proportion of blocks where each optimal response lays within the confidence intervals of the observed responses. The proportion of optimal blocks is clearly higher when using the measurement variance. This was backed by a Chi-square test ($\chi^2(1) = 54.804$, $p < .001$). We then checked whether, when obtaining optimal responses from measurement variance, there was any improvement as a consequence of practice. We compared the number of optimal blocks in the first and second half of the experiment

(6 blocks each), but a Chi-square test gave no significant differences ($\chi^2(1)= 1.214$, $p=.271$).

We can take a look at the general ways responses evolved across blocks (see Figure 5.4.A.). Many participants behaved as if taking into account only their measurement variability, even if most of them required some blocks to converge with that optimal, and others finished on the right trend but without having converged yet. Only a few seemed to be able to adapt their responses to the optimal from their overall variability. Interestingly, they got there after starting the experiment by responding at the optimal from the measurement variability.

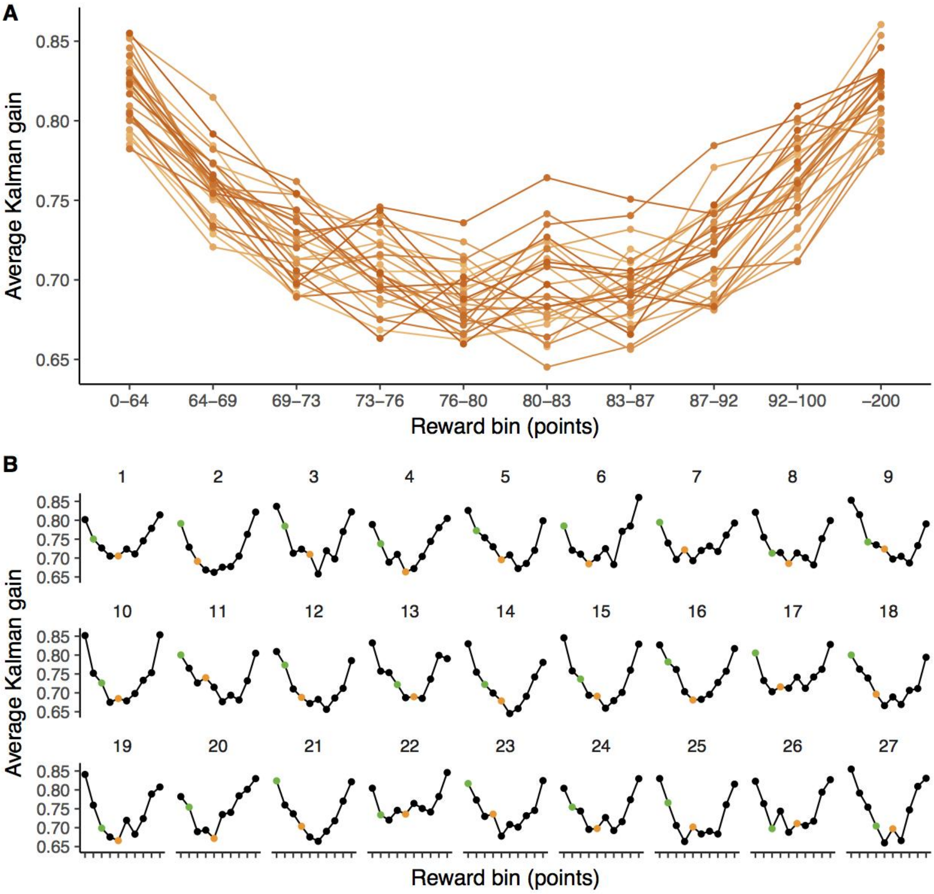


Figure 5.5. A. Average Kalman gain as a function of the reward given in that trial. Reward is grouped in bins that are chosen so that all contain the most similar number of trials. Each participant is plotted in a different shade of orange. B. Same as in A, but every participant is displayed on a different panel. Colored dots correspond to the bins containing the optimal reward, which is the reward that would be given after responding at the optimal response time. The color of the dot reflects how the optimal reward was calculated: the orange dot identifies the bin where the optimal reward falls, when this is calculated by using the standard deviation from each participant's measurement variance. The green dot denotes using observed standard deviations.

We have previously introduced the Kalman gain, explained how it should be interpreted in our context, and shown an example for a participant. We would expect an increased gain for those responses that would trigger a correction on the next trial. We have already showed that responses immediately before a correction are not only those that have been penalized, but also some that were very close to the line. We can use the Kalman gain to better map how relying on the current aimpoint changes as a function of the reward that has just been given. To do that, we took the whole dataset and split it in bins, according to the reward that had been given in that trial. Penalized trials formed a single bin, and the other bins were made so that they had a similar number of trials as the bin of penalized trials. For each participant and bin, we then calculated the average Kalman gain. As Figure 5.5.A shows, all participants followed the same trend. The Kalman gain was lower after a trial where reward was within a certain range of points, and as reward became smaller or bigger than that range, gain went up. This could be interpreted as responses rewarded with a value more distant from that range making a correction more likely. Bins with a low Kalman gain would point to more stable responses around that range of rewards, which could be due to participants aiming at keeping their responses within that area.

We next tried to see whether the previous analysis helped us find more support for our previous hypothesis: namely, that participants used their measurement error to plan responses. For each participant, we computed the optimal response time for the whole experiment. We then obtained the reward that would correspond to that optimal response time, given the task's reward function. Finally, we saw within which reward bin the optimal reward would fall. We did this calculating the optimal reward from each participant's measurement variance, but also from her observed variance, and we compared the average Kalman gain of both bins. Figure 5.5.B depicts, for each participant, the Kalman gain for every reward bin. The orange dot signals the bin containing the optimal reward calculated with the measurement variance. Its average Kalman gain tends to be lower than that of the bin where the observed variance optimal reward belongs (green dots). This was supported by a paired one-sided t-test (*t*(1)= -9.017, *p*<.001). The orange bins are, for all participants, among the bins with the lowest Kalman gain. Thus, participants kept their responses most stable at a point which brought a reward similar to the optimal reward, when this was obtained from the measurement variance. Although this descriptive analysis is no conclusive proof of participants planning their responses based on an estimate of their measurement variability, it is consistent with it. In any case, the Kalman gain for the orange bins was lower than that for the green bins, so, in general terms, the former hypothesis is more plausible than participants planning based on their overall variability.

In the introduction we made reference to most sensory noise probably contributing to the planning variability, at least in the particular version of the Kalman filter we used. We performed one analysis to know whether a significant part of it did also go to the estimated measurement variability. We know that sensory uncertainty scaled with target speed, with larger spatial variabilities for faster speeds. For each participant, block and speed, we calculated the standard deviation of the spatial location of the responses. We then fitted a linear model of the standard deviation as a function of each target speed, for each participant and block. The slopes of these linear models were a measure of how much sensory variability grew with speed. If we found that this slope was correlated with its respective measurement variance (per participant and block), we would have evidence for sensory variability to have gone both to planning and to measurement variability. Although significant (*p*=.048), the correlation had an *r* of 0.110. This weak relationship made us conclude that the measurement variability did not include a relevant portion of the sensory noise.

## 5.5. General discussion

In some timing and sensorimotor experiments (Mamassian, 2008; O'Brien & Ahmed, 2013; Ota et al., 2015, 2016), the literature classifies participants' behavior as suboptimal. According to a line of normative accounts of movement planning (Körding & Wolpert, 2006; Trommershäuser et al., 2003b), optimal behavior would require participants to shift their responses away from the moment that would hold maximum reward at a single trial, and choose the moment that maximizes expected reward. At its turn, this would require participants have a good representation of their own motor variability, since the optimal shift would need to be bigger for more variable participants. However, participants' observed responses normally fall short of this optimal shift.

Despite the different modalities of the tasks reporting this suboptimal shift, we defend this apparent suboptimality to be subserved by a common cause. We tried to find support for participants planning their responses by using the knowledge of their measurement variability, without taking into account the variability added by trial-to-trial response corrections. We found that average observed responses were much closer to optimal responses when the latter were obtained from the participants' measurement variability, estimated by the Kalman filter, than when their whole observed variability was used. And although for some participants responses were optimal from the start, for many more repeated practice was necessary so that observed responses reached the optimality from measurement variability, and for some others some more practice would have been needed. These results provide evidence for an important amount of our participants slowly forming a representation of their measurement variability, which would be used to plan responses. Eventually, they would shift their responses enough so that they were made on average at the point that, if observed variability was reduced to that coming from measurement variability, would maximize expected reward.

Participants could have scored more if they had integrated the variability coming from trial-to-trial corrections. However, a few seem to have been able to successfully do so. They behaved as if they started responding by only considering their measurement variability, but then started taking the rest of their variability into account. Consequently, they ended up responding at a point that matched optimal responses calculated from their whole observed variability. In order to shed some light into why some subjects did account for their whole variability we can resort to standard models of motor learning (Pearce & Hall, 1980; Thoroughman & Shadmehr, 2000). These models postulate corrections based on a learning rate or fraction of the committed error, but do not keep track of previously made errors and, therefore, the corresponding corrections. Interestingly, a more recent model (Herzfeld, Vaswani, Marko, & Shadmehr, 2014) does indeed postulate a memory of past errors depending on the task environmental conditions. Although speculative, some participants could have been able to keep a memory of errors and consider the previous corrections. More dedicated experimental designs in future works should address this important question.

From the point of view of reward processing, the present work offers an insight on how people use reward as an information that, when combined with the representation of our sensorimotor variability, can signal risk. We have shown how not only penalized trials triggered changes in the planned aimpoint. Corrections were also made after trials with a high reward, simply because the response was a few milliseconds away from being penalized. Participants, then, could have used the received reward to assess the risk that, given their variability (whichever they were using), aiming there would suppose. Of course, it could be argued that participants did not use reward, but the point at which the target disappeared. However, this last strategy is less useful. Izawa & Shadmehr (2011) studied to what extent people base their motor adaptation on sensory feedback and to what extent on reward prediction errors. They found that, as the quality of the sensory feedback decreases, they would make more use of reward information. In our task, which involved moving objects, there was a gap between the time the participant committed to pressing the button and the time she pressed it. For faster speeds, this gap was translated into the target travelling a longer distance (Brenner et al., 2006). If we consider that the target did not appear again to unequivocally show the place where it had been stopped, it is reasonable to think that reward was a more accurate information.

Related to the previous point, here we show how the Kalman filter can be a useful tool to analyze corrections from reward. In particular, the Kalman gain revealed how likely was a correction after being given reward within a particular range. Interestingly, if we calculated the optimal response time by using only the measurement variability, and obtained its corresponding reward, this tended to be in a range after which corrections were least likely. Although we recognize that our analyses in this respect are not enough to make a strong claim, it could be that participants stabilized their aimpoint to obtain on average this optimal reward.

Despite many participants showing a progressive evolution towards some sort of optimality, and even a mild support for participants preferentially keeping their responses around a certain reward range, corrections never ceased to happen. However, if a participant believed she had identified the point that maximized expected reward, she should have kept her aimpoint stable, even after an occasional penalization. Corrections after penalized or nearly penalized trials could be interpreted as sudden aimpoint changes made in order to avoid the risk of penalizations, even if it was at the cost of winning very little for the next trials. Since the aim of the task was to maximize reward, though, responses would soon have been brought closer to the line. This behavior could be described in terms of risk aversion or seeking, as others do in sensorimotor tasks (Grau-Moya, Ortega, & Braun, 2012; Nagengast, Braun, & Wolpert, 2011; S.-W. Wu et al., 2009). Our study did not aim at fully characterizing the nature of these corrections, but future work could investigate whether risk aversion is linked to corrections. This could be done by altering the number of points subtracted when applying a penalization and see how that affected the magnitude of the corrections. In any case, it must be pointed out that labeling behavior as risk-seeking or risk-averse is just descriptive, and it does not imply any mechanism. The discovery of the computational processes underlying corrections would be a greater contribution to our understanding of this phenomenon.

# General discussion

An essential part of every scientific work is to relate its results to its original objectives. That is reserved to the next chapter. Before doing that, here we will discuss our findings in different ways. First we will make clear what our main contributions are to the fields we tried to bridge: sensorimotor decision-making and interaction with moving objects. We next present some limitations of our work. Future directions are then introduced. We finish with possible applications of our results.

## 6.1. Main contributions

Every scientific work intends to bring an advancement to its field of research. When describing the aims of this thesis we stated that we mainly planned on contributing to two fields: sensorimotor decision-making, and the study of interactions with moving objects. Some of our contributions mainly apply to one of those fields. The particularity of our work is that, in some cases, contributions to one field are made by borrowing elements from the other. And in others, since both fields have a common perception-action denominator, results can help advance the two. The following lines try to identify the most important contributions of this doctoral thesis.

We will first discuss a contribution that concerns the field of sensorimotor decision-making, especifically with regard to how optimality has been investigated. Decades of literature have repeatedly described the many ways in which human choices deviate from the axioms that, according to normative theories, should guide an optimal decision-maker. That work, often compiled into the so-called descriptive theories, often missed why people fail to reach optimality. When sensorimotor tasks were reframed within the principles of SDT, the emphasis was put in challenging previous work and presenting humans as optimal decision-makers. Yet the same lack of explanation was carried over. At the most, the limits of optimality were traced, leaving those situations outside it with little research. Luckily, recent accounts have started to go beyond the simple classification between optimal and suboptimal. For instance, the efficient coding hypothesis (Summerfield & Tsetsos, 2015) has been proposed as a way to give a normative explanation to why people fail to adhere to classic optimal behavior. This view defends that, given its limited capacity, our brain may have evolved to avoid redundancy and represent information with the minimum number of resources. A way it could do this is by coding information about value in a relative rather than in an absolute way, and this would explain why the firing rate of neurons is dynamically adjusted to the most informative stimuli at each moment. However, in rapidly changing environments, our system would take time to adjust its dynamic range, and this could create behavior that looks suboptimal from the normative point of view. But actually, violating normative axioms would sometimes be the only way our computationally-limited systems could enhance performance (Tsetsos et al., 2016). Another line of research is that investigating the role of suboptimal inference (Beck, Ma, Pitkow, Latham, & Pouget, 2012; Drugowitsch, Wyart, Devauchelle, & Koechlin, 2016). This account points to internal noise (single-cell variability, noise in motoneurons, ...) and external noise (variability in the stimuli) not being enough to explain the behavorial variability often shown in perception. An important part of the variability, even the most important one in complex problems, would mainly being caused by the low precision of the computations carried out by the nervous system. Following the analogy of the optics, an image recognition system may perform badly not because of the bad quality of the optics of the lens, but to the algorithm that processes the picture taken. Similarly, in perceptual decision-making tasks we may have really fast integrators, but we may perform some computations in a flawed way: for instance, by giving the incorrect weight to different pieces of information. Without aiming at being so influential, we believe that part of our work shares the motivation of both the efficient coding and the suboptimal inference hypotheses. In this sense, our work contributes to the field of sensorimotor decision-making by digging into the causes of suboptimality. We do so in Study III, by showing evidence of an apparent suboptimality being caused by the variability participants use to plan their responses. Similarly, in Study I, instead of simply classifying behavior as optimal or suboptimal, we challenge classic normative assumptions and investigate what causes an improvement in performance. As a criticism to our work, it is also true that we left some phenomena at the descriptive level. For instance, a satisfactory conclusion for Study II would have involved unraveling why some people naturally exploit temporal information.

Studies I and III can also make another contribution on the theoretical side within perceptual and sensorimotor decision-making, but also formulate a caveat for anybody within the field of moving objects who wants to study it from the decision-making point of view. Namely, these studies teach us a similar lesson on how to implement normative models: obtaining optimal estimates should be preceded by identifying the relevant sources of variability for the task. In Study I we evidence the need for tasks with moving objects to consider two sensory elements: the sources of visual information that can be used to plan the response, and the different states of the stimulus. On the other hand, Study III focuses on the motor variability: the use of the models is only useful once whe have discounted the planning from the measurement variability. And while Study III does not decompose the sensory variability, we chose the data for which this was more reduced. An integrated normative account of movement planning with moving objects should, though, explicitly consider both the sensory and the motor uncertainty, as other frameworks already do. For instance, Hoppe & Rothkopf (2016) present a computational model that tries to account for how humans adjust the timing ratio between saccades and fixations so that their eye movements are optimal in order to detect important events in the environment according to the statistical structure of these events. This model is decomposed into the ideal observer model, that considers the relevant sources of sensory variability and the ideal actor model, that does the same with motor variability. By learning the temporal regularities of the stimuli, the agent is able to decide how to trade-off the sensory and the motor uncertainties: for instance, a by establishing a relationship between the probability of detecting a change if eye movements are switched at a faster rate, and the motor cost of this increase in speed.

A way to expand the knowledge of how we interact with moving objects is by placing the weight on the decisional component, and this is how we tried to contribute to the field. An obvious advantage of this approach the possibility to borrow all the models and experimental paradigms that perceptual, sensorimotor and general decision-making has long used. To study the decisional aspect, some situations may be better than others. The title of our thesis implies no particular focus within sensorimotor decisions with moving objects. However, in practice all our work focuses on one particular set of situations. Let us think about sensorimotor decisions in the broader sense. When throwing a dart at a target composed of concentric circles, the best choice is to aim at the center. However, a golfer may have a bigger trouble selecting his aimpoint when the hole is partially surrounded by a sand bunker. As in the last example, in many situations the optimal policy is not always explicitly clear, and the best outcome can be very close to the worst one. The literature on sensorimotor decision-making has long dealt with such situations (for a review of some of it, see Trommershäuser et al., 2008). However, the study of sensorimotor interaction with moving objects has normally had the focus on other aspects: mainly, the mechanisms subserving interception. These tasks involve some uncertainty on how to trade-off sensory and motor phases (Battaglia & Schrater, 2007): for example, when to start the movement, and deciding where along the object's trajectory should the catch be in order to move by minimizing jerk and effort (Todorov, 2004). Nevertheless, the task's objective is always very clear, and although some intrinsic reward may be present (for instance, the amount of effort minimized), it does not place the focus on making participants try to look for the action that maximizes expected reward. We believe that, by creating the paradigm and adapting the normative framework, our work sets the ground to study interaction with moving objects from that perspective.

Finally, the field of sensorimotor decision-making can benefit from exploring situations with moving objects. As we illustrate, this new terrain can contribute to the notion of performance in sensorimotor decision-making tasks. Within this field, different approaches have different meanings for the term performance. For instance, for signal detection theory (Green & Swets, 1988) performance is a synonymous of accuracy. In our experiments, since participants were instructed to win as much as possible, performance was assessed as the accumulated reward. However, performance was affected by many variables, and these variables interacted with each other. To give a full perspective, we next identify how the different factors affecting performance depended on each other. The first one is variability. Even if practice could make participants less variable, some were still more variable than others. More variable participants deviated more from their aimpoint, and were more likely to be penalized. Thus, we see how variability was linked to the selected aimpoint. However, Study III showed how aimpoint changes also affect the observed variability, so the relationship goes both ways. The link between variability and endpoint is obvious, to the extent that normative accounts expect the aimpoint to be selected according to how variable the decision-maker is. Given the unavoidable variability and the structure of the task, setting the aimpoint closer to the line increased the average reward of non-penalized trials, but also increased the probability of being penalized. The opposite happened when setting the aimpoint earlier. The third factor, and our main contribution here, is how participants use different visual cues to plan their responses: that is, what we call DV. Somebody with a temporal DV will look very variable if we analyze the spatial dimension of her responses, and it could look as if a different aimpoint was being used for each speed. However, if we we analyze the temporal dimension, the precision will increase, and the unique aimpoint will be revealed. The opposite will happen for somebody with a spatial DV. Of course, this was a product of our design, and these differences increased for a broader range of speeds. Nevertheless, it illustrates how important is, in situations where responses can be planned by using more than one variable, to identify which one (or to what degree each of them) participants are using. This insight is a valuable contribution that the case of interaction with moving objects can make within the field of sensorimotor decision-making.

## 6.2. Limitations

Our experiments consisted in laboratory situations that reduced much of the complexity that involves daily interactions with moving objects. Although this may have made our tasks look less like real world problems, our way of proceeding was motivated by the need to control some undesired effects. Although our reasons were justified, any attempt to extrapolate our results to real life must be cautious. Particular situations should be studied in order to see to what extent the complexity we reduced could affect our claims holding. Some of the elements our experimental paradigm disregarded are commented next.

It has already been said that the motor action our tasks required was a simple button press. This was so because previous studies (Hagura et al., 2017; Marcos et al., 2015) have shown that the motor cost of an action may bias choices. However, interception nearly always requires to move our arm, and often locomotion is required to catch the target. The presence of a sophisticated motor process can complicate the continuous collection of sensory information.

The fact that our experimental paradigm required the line made us work with fronto-parallel motion. This made us overlook motion in depth, which is the type of motion we experience in catching situations when an object approaches us. Temporal and spatial cues are also present in this kind of motion (e.g. López-Moliner & Keil, 2012). Nevertheless, they are different from those present in lateral displacements.

Our tasks constrained target motion to one dimension. This made the reward function easier, as it depended only on the distance from alignment along the x axis. Furthermore, for the sake of simplicity we worked only with constant speeds. But if we imagine a baseball player trying to catch the ball hit by a batter, we can see how objects usually move in three dimensions, and accelerate or decelerate, in many cases due to gravity.

Perception and action sometimes have to be traded off. In the example of the last paragraph, when the baseball player starts to run, she will probably lose track of the moving object at some point. Our stimuli did not disappear until the participant responded with a simple keypress, thus reducing the need for trading off sensory and motor phases.

Finally, instead of reducing complexity, it could seem that one of our manipulations increased it. Indeed, our experimental design interleaved different target speeds and motion times. It could be reasoned that suboptimality in our tasks could partly be due the stimuli varying too much across trials, and that more stable situations could lead to clear optimality. Our manipulation had some practical uses, such as revealing the DV used, or preventing strategies based on interval estimation from motion onset. However, we also want to highlight the reason of ecological validity. As much as it could be interesting to see whether we are undubiously optimal in very restricted situations, it is also important to think about the minimum level of complexity that happens in plausible scenarios. In real life moving objects are rarely presented always with the same speed, and their trajectory seldom lasts the same amount of time.

## 6.3. Future directions

Our studies have many possible continuations. Some of them could simply introduce elements present in real-life situations that were left out by our experimental design, as noted in the previous section. The following lines aim at introducing more types of possible future work. The first is a follow-up that could clarify some of our results. The others are modifications of our tasks that could tap into other relevant topics in perception, action and decision-making.

This thesis leaves unexplained why some people tend to exploit temporal information. While we mention possible individual differences, this does not provide any answer. Individual differences can be of many origins. For instance, let us consider the possibility that motion perception may be different in those participants with temporal DVs. They could possibly range from different neurotransmitter concentrations in the brain (Takeuchi, Yoshimoto, Shimada, Kochiyama, & Kondo, 2017) to specific sports training (Nakamoto, Mori, Ikudome, Unenaka, & Imanaka, 2014). If we were to stick to the behavioral level of study that has characterized this thesis, some psychophysical experiments could be conducted. For instance, there may be differences in the velocity discrimination threshold. Alternatively, the asymptote for integrating motion information may be lower. As a last example, motion extrapolation could be assessed in tasks where a moving object disappears and the participant has to report the position where that happened.

In Study III we showed that many participants ended up converging with those optimal responses calculated by only considering their measurement variability. However, most participants needed extensive practice with the task, and some others would have needed more of it to finally converge. Consequently, even if participants were only considering part of their observed variability, representing it accurately took time. This finding is not new, since previous movement planning tasks have already required experience before optimality was reached (Neyedli & Welsh, 2013a). Yet this opens the door to asking whether the time spent acquiring these representations could be reduced in future occasions. This relates to savings, a common topic in the sensorimotor literature (Ebbinghaus, Ruger, & Bussenius, 1913; Krakauer, 2005; Zarahn, Weston, Liang, Mazzoni, & Krakauer, 2008), understood as a faster rate of readaptation compared to the initial rate of adaptation. However, in real life no two situations are identical. Thus, an interesting focus would be to what extent this already formed representation could transfer to different tasks. After getting used to one experimental setting, participants in pointing tasks quickly adapted to other configurations (Neyedli & Welsh, 2013b; Trommershäuser et al., 2003a). Interestingly, also in pointing, when variability was artificially added, participants adapted in less than 120 trials (Trommershäuser et al., 2005), but this could not be replicated in a coincidence timing task (Mamassian, 2008). In our case, where participants could be representing only part of their variability, it would be worth studying whether this same representation rapidly transfers to other tasks, and if it remains unchanged.

Another future study could come from manipulating the statistics of the presented stimuli. In Experiment 1 of Study II we found how most participants adopted a spatial DV. A reason we proposed for that was the limited range of target speeds. The extra computational cost of using temporal cues could not be worth it if using a spatial DV barely supposed any difference in terms of penalized trials for the fastest speeds. On the other hand, adopting a temporal DV could be promoted by including speeds for which using a spatial DV supposed a serious risk of incurring into penalizations. Now let us suppose we ran a condition where a wide range of speeds was featured, but very fast speeds were seldom presented. In the case of an unequal presentation probability, the brain could build internal models representing the environmental statistics, and adapt behavior so as to expect more some speeds than others (Kwon & Knill, 2013). If that was so, we could expect most participants to still use an overall spatial DV, due to the low probability of fast speeds making them unimportant for the internal model, in a similar fashion of what is known as robust averaging (de Gardelle & Summerfield, 2011). In any case, predictions could be tested with Bayesian Decision Theory (Berger, 1985), giving the opportunity to study how likely would be for participants to acquire different DVs.

We will describe a final possible next study. TTC estimation has been proved to benefit from integrating speed with visual timing cues (Chang & Jazayeri, 2017). Precision can also be improved from multisensory cue combination, with the visual cue being integrated with an auditory cue (Mendonça, Santos, & López-Moliner, 2011; Wuerger, Meyer, Hofbauer, Zetzsche, & Schill, 2010). This research opens the way to test whether, when interacting with moving objects, planning responses by using temporal visual information could be enhanced by other cues.

## 6.4. Practical applications

We believe our findings can have helpful practical applications, particularly how exploiting temporal information could help improve performance when interacting with moving objects. However, in order for this to be applied, a previous step would be needed. Namely, achieving what we could not in Study II: a method to successfully learn how to base sensorimotor responses on the use of specific visual cues. Once designed, the method could be easily implemented. Virtual reality or augmented reality are idoneous platforms, with some other sort of videogame also being a possibility. And once with the general implementation, some adaptations could be made to induce learning in different scenarios.

The perfect candidate is ball sports: for instance, training a goalkeeper or a baseball batter to time her actions. But also other sports could benefit from exploiting temporal cues. We have previously mentioned clay pigeon shooting. This situation features lateral motion, so it is very similar to our experiments. Finally, navigation offers an interesting context. Whether it is training drivers or jet pilots, the agent can be considered the moving object. This would probably require to explore how responses are planned by exploiting temporal information extracted from optic flow, and how our movement can be discounted from that of the other moving objects (i.e. other cars).

# Conclusions

Individual studies of this thesis contain a discussion section where we expose how well our experiments or modelling can answer the specific questions we formulated in that particular study. In this section we take the results from all our studies in order to conclude to what extent we can fulfill the thesis' main objectives.

*Exploiting temporal information vs exploiting spatial information*

Our first objective was to study how, when interacting with moving objects, our performance depends on the visual information we exploit to plan our responses.

We found evidence for an advantage of using temporal cues over spatial cues. This advantage manifested in our results as a more stable performance across the different speeds the moving object could adopt. We claim that this was subserved by using temporal estimates circumventing the constraint imposed by the limited resolution of the visual system. When exploiting spatial information, based on tracking the position of the moving object, this constraint would be seen as increased spatial variability when dealing with faster speeds.

We also showed that the extent to which each type of information was used depended on the time that the object could be seen moving: longer motion times were related to a greater use of temporal cues. We suggest that this is produced by an increased integration time allowing better velocity estimates.

Finally, our results point to some participants being more prone to exploiting temporal information, as revealed by the fact that they did not require practice to adopt this decision variable. As for those who exploit spatial information by default, we could not design a method that made them rely more on time. The fact that in some situations most of our participants exploited spatial information can be related to this approach being a more straightforward method of planning responses, whereas using temporal cues requires the extra step of making reliable velocity estimates to compute the time that the object will take to reach a certain point.

### *Statistical decision theory as a normative reference for interaction with moving objects*

Our second objective was assessing to what extent existing SDT-derived optimal movement planning models could be applied to situations of interaction with moving objects. The importance behind this endeavor lays in the usefulness of being able to define an optimal behavior: by having an upper bound on performance, we could know if participants were doing as good as they possibly could.

The main limitation of the current models was that, in our situations, they did not decompose the relevant sources of uncertainty. These models assume that optimal decision-makers integrate an accurate representation of their variability with knowledge about the reward function of the task. However, they do not specify how different variabilities affect performance. This made that, when trying to interpret our results, these were at odds with some of the assumptions of the models.

We found that some participants were classified as optimal for some conditions of the stimuli, while for others they were not. Recovering the assumptions of these models, for our data this would have meant that participants could successfully represent and integrate their variability with the reward function for only for some conditions of the stimuli. This unlikely scenario has a cause. Unlike pointing tasks with static targets, interaction with moving objects includes a relevant component of sensory variability. This variability increases for shorter motion times and faster speeds. Without the model capturing this uncertainty, it will fail to provide a valid measure of optimality.

Our participants did better when exploiting temporal information, even if the reward function was linked to space. Thus, we see how improved performance was found even if violating the domain of the reward function. This evidences how, when different sources of information can be used to plan the responses, they should be taken into account by the model. This is the case for interaction with moving objects.

Once we separated the variabilities, SDT models proved useful for our purposes. This happened when investigating which variability participants were using when planning their responses. Specifically, we computed the optimal responses in two different ways: on one hand, with the whole observed variability of each participant; on the other, discounting that variability originated from large corrections and leaving only measurement variability, which was roughly equivalent to execution variability. We concluded that our data was more congruent with participants using a representation of their measurement variability to plan their responses.

### *Using reward to plan future actions*

The last main objective of this thesis consisted in investigating the way reward information was used in order to plan subsequent responses.

In one study, we deliberately tried to make participants use reward to exploit a specific type of visual cue (temporal or spatial). While we do not have evidence for our penalization system to work for our purposes, that does not mean that other methods cannot successfully make reward shape the decision variable in tasks with moving objects.

Our data leads us to conclude that participants use reward information to control the aimpoint of their responses. This was evidenced by the finding that penalized responses triggered corrections in order to prevent from further penalizations. However, and more interestingly, these corrections were also present after nearly penalized trials. This points towards participants using their reward not only as a reinforcer, but also to assess the risk of being penalized if they maintained their current aimpoint. Further analyses allow us to hypothesize that participants may be stabilizing their aimpoint around reward areas that are optimal when only considering their measurement variability.

# References

Addams, R. (1834). LI. An account of a peculiar optical phænomenon seen after having looked at a moving body. *Philosophical Magazine Series 3*, *5*(29), 373–374.

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, *2*(2), 284.

Adelson, E. H., & Bergen, J. R. (1986). The extraction of spatiotemporal energy in human and machine vision. In *Proceedings of the Workshop on Motion: Representation and Analysis* (pp. 151–155).

Anstis, S., Verstraten, F. A., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences*, *2*(3), 111–117.

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition*, *48*(1), 66–79.

Aschersleben, G., & Prinz, W. (1998). Timing in perception and action: Introductory remarks. *Psychological Research Psychologische Forschung*, *61*(1), 1–3.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Badcock, D. R., & Derrington, A. M. (1985). Detecting the displacement of periodic patterns. *Vision Research*, *25*(9), 1253–1258.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*(6), 953–965.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Battaglia, P. W., & Schrater, P. R. (2007). Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(26), 6984–6994.

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron*, *74*(1), 30–39.

Belousov, B., Neumann, G., Rothkopf, C. A., & Peters, J. R. (2016). Catching heuristics are optimal control policies. *Advances in Neural Information Processing Systems*, 1426–1434.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*.

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica: Journal of the Econometric Society*, *22*(1), 23.

Berry, M. J., 2nd, Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*(6725), 334–338.

Blackwell, D., & Girshick, M. A. (1954). *Theory of games and statistical decisions*. Wiley.

Block, R. A., & Grondin, S. (2014). Timing and time perception: A selective review and commentary on recent reviews. *Frontiers in Psychology*, *5*, 648.

Bootsma, R. J., & Oudejans, R. R. (1993). Visual information about time-to-collision between two objects. *Journal of Experimental Psychology. Human Perception and Performance*, *19*(5), 1041–1052.

Brenner, E., & Smeets, J. B. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Research*, *40*(13), 1645–1648.

Brenner, E., van Beers, R. J., Rotman, G., & Smeets, J. B. J. (2006). The role of uncertainty in the systematic spatial mislocalization of moving objects. *Journal of Experimental Psychology. Human Perception and Performance*, *32*(4), 811–825.

Burge, J., Ernst, M. O., & Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *Journal of Vision*, *8*(4), 20.1–19.

Burr, D. C. (1981). Temporal summation of moving images by the human visual system. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society* , *211*(1184), 321–339.

Burr, D. C., & Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Research*, *41*(15), 1891–1899.

Burr, D. C., & Thompson, P. (2011). Motion psychophysics: 1985-2010. *Vision Research*, *51*(13), 1431–1456.

Caines, P. E. (1988). *Linear stochastic systems*. Wiley.

Chang, C.-J., & Jazayeri, M. (2017). An integral role for timing in interception. https://doi.org/10.1101/155531

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America. A, Optics and Image Science*, *5*(11), 1986–2007.

Darshan, R., Leblois, A., & Hansel, D. (2014). Interference and shaping in sensorimotor adaptations with rewards. *PLoS Computational Biology*, *10*(1), e1003377.

Dean, M., Wu, S.-W., & Maloney, L. T. (2007). Trading off speed and accuracy in rapid, goal-directed movements. *Journal of Vision*, *7*(5), 10.1–12.

de Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences*, *108*(32), 13341–13346.

de la Malla, C., & López-Moliner, J. (2010). Detection of radial motion depends on spatial displacement. *Vision Research*, *50*(11), 1035–1040.

de la Malla, C., & López-Moliner, J. (2015). Predictive plus online visual information optimizes temporal precision in interception. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(5), 1271–1280.

Derrington, A. M., & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research*, *25*(12), 1869–1878.

Derrington, A. M., & Henning, G. B. (1987). Errors in direction-of-motion discrimination with complex stimuli. *Vision Research*, *27*(1), 61–75.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research*, *31*(9), 1619–1626.

Diedrichsen, J., Hashambhoy, Y., Rane, T., & Shadmehr, R. (2005). Neural correlates of reach errors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(43), 9919–9931.

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., & Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of

Human Choice Suboptimality. *Neuron*, *92*(6), 1398–1411.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science*, *287*(5460), 2036–2038.

Ebbinghaus, H., Ruger, H. A., & Bussenius, C. E. (1913). *Memory: A contribution to experimental psychology*.

Eckhoff, P., Holmes, P., Law, C., Connolly, P. M., & Gold, J. I. (2008). On diffusion processes with variable drift rates as models for decision making during learning. *New Journal of Physics*, *10*, nihpa49499.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Archiv Für Die Gesammte Physiologie Des Menschen Und Der Thiere*, *11*(1), 581–602.

Fahle, M., & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, *33*(3), 397–412.

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews. Neuroscience*, *9*(4), 292–303.

Ferguson, T. S. (1967). *Mathematical statistics: a decision theoretic approach*. Academic Pr.

Gepshtein, S., Seydell, A., & Trommershäuser, J. (2007). Optimality of human movement under natural variations of visual-motor uncertainty. *Journal of Vision*, *7*(5), 13.1–18.

Grau-Moya, J., Ortega, P. A., & Braun, D. A. (2012). Risk-sensitivity in Bayesian sensorimotor integration. *PLoS Computational Biology*, *8*(9), e1002698.

Gray, R., & Regan, D. (1998). Accuracy of estimating time to collision using binocular and monocular information. *Vision Research*, *38*(4), 499–512.

Green, D. M., & Swets, J. A. (1988). *Signal Detection Theory and Psychophysics*. Peninsula Pub.

Grondin, S. (2010). Unequal Weber fractions for the categorization of brief temporal intervals. *Attention, Perception & Psychophysics*, *72*(5), 1422–1430.

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(4), 880–890.

Hagura, N., Haggard, P., & Diedrichsen, J. (2017). Perceptual decisions are biased by the cost to act. *eLife*, *6*. https://doi.org/10.7554/eLife.18422

Hatsopoulos, N., Gabbiani, F., & Laurent, G. (1995). Elementary

Computation of Object Approach by a Wide-Field Visual Neuron. *Science*, *270*(5238), 1000–1003.

Herzfeld, D. J., & Shadmehr, R. (2014). Motor variability is not noise, but grist for the learning mill. *Nature Neuroscience*, *17*(2), 149–150.

Herzfeld, D. J., Vaswani, P. A., Marko, M. K., & Shadmehr, R. (2014). A memory of errors in sensorimotor learning. *Science*, *345*(6202), 1349–1353.

Hoppe, D., & Rothkopf, C. A. (2016). Learning rational temporal eye movement strategies. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 8332–8337.

Huang, X., Lu, H., Zhou, Y., & Liu, Z. (2011). General and specific perceptual learning in radial speed discrimination. *Journal of Vision*, *11*(4), 7.

Hudson, T. E., Maloney, L. T., & Landy, M. S. (2008). Optimal Compensation for Temporal Uncertainty in Movement Planning. *PLoS Computational Biology*, *4*(7), e1000130.

Hudson, T. E., Wolfe, U., & Maloney, L. T. (2012). Speeded reaching movements around invisible obstacles. *PLoS Computational Biology*, *8*(9), e1002676.

Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Computational Biology*, *7*(3), e1002012.

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex. *The Journal of Physiology*, *556*(Pt 3), 971–982.

Jarvstad, A., Hahn, U., Warren, P. A., & Rushton, S. K. (2013). Perceptuo-motor, cognitive, and description-based decision-making seem equally good. *Proceedings of the National Academy of Sciences*, *110*(40), 16271–16276.

Jarvstad, A., Hahn, U., Warren, P. A., & Rushton, S. K. (2014). Are perceptuo-motor decisions really more optimal than cognitive decisions? *Cognition*, *130*(3), 397–416.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *International Journal of Engineering, Transactions A: Basics*, *82*(1), 35.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(11), 4966–4970.

Keil, M. S., & López-Moliner, J. (2012). Unifying time to contact estimation and collision avoidance across species. *PLoS Computational Biology*, *8*(8), e1002625.

Kelly, D. H. (1972). Flicker. In *Handbook of Sensory Physiology* (pp. 273–302).

Kerzel, D., & Gegenfurtner, K. R. (2003). Neuronal Processing Delays Are Compensated in the Sensorimotor Branch of the Visual System. *Current Biology: CB*, *13*(22), 1975–1978.

Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2013). Motion-based prediction explains the role of tracking in motion extrapolation. *Journal of Physiology, Paris*, *107*(5), 409–420.

Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The Flash-Lag Effect as a Motion-Based Predictive Shift. *PLoS Computational Biology*, *13*(1), e1005068.

Körding, K. P., & Wolpert, D. M. (2006). Bayesian Statistics and Utility Functions in Sensorimotor Control. In *Bayesian Brain* (pp. 299–319).

Krakauer, J. W. (2005). Adaptation to Visuomotor Transformations: Consolidation, Interference, and Forgetting. *Journal of Neuroscience*, *25*(2), 473–478.

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research*, *39*(16), 2669–2679.

Kwon, O.-S., & Knill, D. C. (2013). The brain uses adaptive internal models of scene statistics for sensorimotor estimation and planning. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(11), E1064–73.

Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(26), 8142–8147.

Lappe, M., & Krekelberg, B. (1998). The position of moving objects. *Perception*, *27*(12), 1437–1449.

Law, C.-T., & Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, *12*(5), 655–663.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*(4), 437–459.

Linares, D., Holcombe, A. O., & White, A. L. (2009). Where is the moving object now? Judgments of instantaneous position show poor temporal precision (SD = 70 ms). *Journal of Vision*, *9*(13), 9.1–14.

Linares, D., López-Moliner, J., & Johnston, A. (2007). Motion signal and the perceived positions of moving objects. *Journal of Vision*, *7*(7), 1.1–7.

López-Moliner, J., & Bonnet, C. (2002). Speed of response initiation in a time-to-contact discrimination task reflects the use of eta. *Vision Research*, *42*(21), 2419–2430.

López-Moliner, J., Field, D. T., & Wann, J. P. (2007). Interceptive timing: prior knowledge matters. *Journal of Vision*, *7*(13), 11.1–8.

López-Moliner, J., & Keil, M. S. (2012). People favour imperfect catching by assuming a stable world. *PloS One*, *7*(4), e35705.

López-Moliner, J., Supèr, H., & Keil, M. S. (2013). The time course of estimating time-to-contact: switching between sources of information. *Vision Research*, *92*, 53–58.

Maloney, L. T., & Zhang, H. (2010). Decision-theoretic models of visual perception and action. *Vision Research*, *50*(23), 2362–2374.

Mamassian, P. (2008). Overconfidence in an objective anticipatory motor task. *Psychological Science*, *19*(6), 601–606.

Marcos, E., Cos, I., Girard, B., & Verschure, P. F. M. J. (2015). Motor Cost Influences Perceptual Decisions. *PloS One*, *10*(12), e0144841.

Maus, G. W., Fischer, J., & Whitney, D. (2013). Motion-dependent representation of space in area MT+. *Neuron*, *78*(3), 554–562.

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of Box Plots. *The American Statistician*, *32*(1), 12.

McKee, S. P., & Welch, L. (1985). Sequential recruitment in the discrimination of velocity. *Journal of the Optical Society of America.*

*A, Optics and Image Science*, *2*(2), 243–251.

Melcher, D., Crespi, S., Bruno, A., & Morrone, M. C. (2004). The role of attention in central and peripheral motion integration. *Vision Research*, *44*(12), 1367–1374.

Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience*, *6*(8), 877–881.

Mendonça, C., Santos, J. A., & López-Moliner, J. (2011). The benefit of multisensory integration with biological motion signals. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *213*(2-3), 185–192.

Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature*, *355*(6358), 344–346.

Nagengast, A. J., Braun, D. A., & Wolpert, D. M. (2011). Risk-sensitivity and the mean-variance trade-off: decision making in sensorimotor control. *Proceedings. Biological Sciences / The Royal Society*, *278*(1716), 2325–2332.

Nakamoto, H., Mori, S., Ikudome, S., Unenaka, S., & Imanaka, K. (2014). Effects of sport expertise on representational momentum during timing control. *Attention, Perception & Psychophysics*, *77*(3), 961–971.

Nakayama, K. (1985). Biological image motion processing: a review.

*Vision Research*, *25*(5), 625–660.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research*, *21*(4), 427–433.

Narain, D., van Beers, R. J., Smeets, J. B. J., & Brenner, E. (2013). Sensorimotor priors in nonstationary environments. *Journal of Neurophysiology*, *109*(5), 1259–1267.

Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience*, *9*(9), 1186–1192.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, *395*(6705), 894–896.

Neyedli, H. F., & Welsh, T. N. (2013a). Optimal weighting of costs and probabilities in a risky motor decision-making task requires experience. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(3), 638–645.

Neyedli, H. F., & Welsh, T. N. (2013b). People are better at maximizing expected gain in a manual aiming task with rapidly changing probabilities than with rapidly changing payoffs. *Journal of Neurophysiology*, *111*(5), 1016–1026.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*(6487), 256–257.

Nishida, S. ’ya. (2011). Advancement of motion psychophysics: review 2001-2010. *Journal of Vision*, *11*(5), 11.

O’Brien, M. K., & Ahmed, A. A. (2013). Does risk-sensitivity transfer across movements? *Journal of Neurophysiology*, *109*(7), 1866–1875.

Olsson, H. (2014). Measuring overconfidence: Methodological problems and statistical artifacts. *Journal of Business Research*, *67*(8), 1766–1770.

Ota, K., Shinya, M., & Kudo, K. (2015). Motor planning under temporal uncertainty is suboptimal when the gain function is asymmetric. *Frontiers in Computational Neuroscience*, *9*, 88.

Ota, K., Shinya, M., & Kudo, K. (2016). Sub-optimality in motor planning is retained throughout 9 days practice of 2250 trials. *Scientific Reports*, *6*, 37181.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552.

Regan, D. (2002). Binocular information about time to collision and time to passage. *Vision Research*, *42*(22), 2479–2484.

Regan, D., & Hamstra, S. J. (1992). Dissociation of orientation discrimination from form detection for motion-defined bars and luminance-defined bars: effects of dot lifetime and presentation duration. *Vision Research*, *32*(9), 1655–1666.

Regan, D., & Vincent, A. (1995). Visual processing of looming and time to contact throughout the visual field. *Vision Research*, *35*(13), 1845–1857.

Reichardt, W. (1957). Autokorrelations-Auswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschrift Für Naturforschung B*, *12*(7). https://doi.org/10.1515/znb-1957-0707

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences*, *14*(2), 64–71.

Saffell, T., & Matthews, N. (2003). Task-specific perceptual learning on speed and direction discrimination. *Vision Research*, *43*(12), 1365–1374.

Savelsbergh, G. J., Whiting, H. T., & Bootsma, R. J. (1991). Grasping tau. *Journal of Experimental Psychology. Human Perception and Performance*, *17*(2), 315–322.

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *The Journal of Physiology*, *483 ( Pt 3)*, 797–810.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli.

*Vision Research*, *38*(22), 3569–3582.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, *61*(5), 700–707.

Seitz, A. R., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, *9*(7), 329–334.

Seydell, A., McCann, B. C., Trommershäuser, J., & Knill, D. C. (2008). Learning stochastic reward distributions in a speeded pointing task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(17), 4356–4367.

Shapiro, A., Lu, Z.-L., Huang, C.-B., Knight, E., & Ennis, R. (2010). Transitions between central and peripheral vision create spatial/temporal distortions: a hypothesis concerning the perceived break of the curveball. *PloS One*, *5*(10), e13296.

Smeets, J. B., & Brenner, E. (1994). The difference between the perception of absolute and relative motion: a reaction time study. *Vision Research*, *34*(2), 191–195.

Smeets, J. B. J., & Brenner, E. (1995). Perception and action are based on the same visual information: distinction between position and velocity. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(1), 19–31.

Smeets, J. B. J., Brenner, E., Trébuchet, S., & Mestre, D. R. (1996). Is

Judging Time-to-Contact Based on “Tau”? *Perception*, *25*(5), 583–590.

Smith, A. T., Snowden, R. J., & Milne, A. B. (1994). Is global motion really based on spatial integration of local motion signals? *Vision Research*, *34*(18), 2425–2430.

Smith, M. R., Flach, J. M., Dittman, S. M., & Stanard, T. (2001). Monocular optical constraints on collision control. *Journal of Experimental Psychology. Human Perception and Performance*, *27*(2), 395–410.

Snowden, R. J., & Braddick, O. J. (1989). The combination of motion signals over time. *Vision Research*, *29*(11), 1621–1630.

Snowden, R. J., & Braddick, O. J. (1991). The temporal integration and resolution of velocity signals. *Vision Research*, *31*(5), 907–914.

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends in Cognitive Sciences*, *19*(1), 27–34.

Sundberg, K. A., Fallah, M., & Reynolds, J. H. (2006). A motion-dependent distortion of retinotopy in area V4. *Neuron*, *49*(3), 447–457.

Sun, H., & Frost, B. J. (1998). Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons. *Nature Neuroscience*, *1*(4), 296–303.

Sweet, A. L. (1953). Temporal discrimination by the human eye. *The*

*American Journal of Psychology*, *66*(2), 185–198.

Takeuchi, T., Yoshimoto, S., Shimada, Y., Kochiyama, T., & Kondo, H. M. (2017). Individual differences in visual motion perception and neurotransmitter concentrations in the human brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1714). https://doi.org/10.1098/rstb.2016.0111

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, *407*(6805), 742–747.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, *7*(9), 907–915.

Todorov, E. (2006). Optimal Control Theory. In *Bayesian Brain* (pp. 268–298).

Tresilian, J. R. (1990). Perceptual information for the timing of interceptive action. *Perception*, *19*(2), 223–239.

Tresilian, J. R. (1994). Approximate information sources and perceptual variables in interceptive timing. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(1), 154–173.

Tresilian, J. R. (1995). Perceptual and cognitive processes in time-to-contact estimation: analysis of prediction-motion and relative judgment tasks. *Perception & Psychophysics*, *57*(2), 231–245.

Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., &

Banks, M. S. (2005). Optimal compensation for changes in task-relevant movement variability. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(31), 7169–7178.

Trommershäuser, J., Landy, M. S., & Maloney, L. T. (2006). Humans rapidly estimate expected gain in movement planning. *Psychological Science*, *17*(11), 981–988.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003a). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *20*(7), 1419–1433.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003b). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision*, *16*(3-4), 255–275.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in Cognitive Sciences*, *12*(8), 291–297.

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(11), 3102–3107.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and*

*Uncertainty*, *5*(4), 297–323.

Tyldesley, D. A., & Whiting, H. (1975). Operational timing. *Journal of Human Movement Studies*.

van Beers, R. J. (2009). Motor learning is optimally tuned to the properties of motor noise. *Neuron*, *63*(3), 406–417.

van Beers, R. J. (2012). How does our motor system determine its learning rate? *PloS One*, *7*(11), e49373.

van Beers, R. J., Haggard, P., & Wolpert, D. M. (2004). The role of execution noise in movement variability. *Journal of Neurophysiology*, *91*(2), 1050–1063.

van Beers, R. J., Wolpert, D. M., & Haggard, P. (2001). Sensorimotor integration compensates for visual localization errors during smooth pursuit eye movements. *Journal of Neurophysiology*, *85*(5), 1914–1922.

Van Essen, D. C. (2004). Organization of visual areas in macaque and human cerebral cortex. In L. M. Chalupa & J. S. Werner (Eds.) (pp. 507–521). MIT Press.

Van Essen, D. C., & Drury, H. A. (1997). Structural and functional analyses of human cerebral cortex using a surface-based atlas. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *17*(18), 7079–7102.

van Santen, J. P., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America. A, Optics and Image Science*, *2*(2), 300–321.

von Neumann, J., & Morgenstern, O. (1944). Theory of Games and Economic Behaviour. *Journal of the Royal Statistical Society*, *107*(3/4), 293.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, *32*(5), 931–941.

Watanabe, T., & Sasaki, Y. (2015). Perceptual Learning: Toward a Comprehensive Theory. *Annual Review of Psychology*, *66*(1), 197–221.

Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America. A, Optics and Image Science*, *2*(2), 322–341.

Wei, K., & Körding, K. (2010). Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Frontiers in Computational Neuroscience*, *4*, 11.

Westheimer, G., & McKee, S. P. (1977). Perception of temporal order in adjacent visual stimuli. *Vision Research*, *17*(8), 887–892.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends in Cognitive Sciences*, *6*(5), 211–216.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience*, *1*(8), 656–657.

Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, *26*(4), 511–524.

Wuerger, S., Meyer, G., Hofbauer, M., Zetzsche, C., & Schill, K. (2010). Motion extrapolation of auditory–visual targets. *An International Journal on Information Fusion*, *11*(1), 45–50.

Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences*, *106*(15), 6088–6093.

Wu, S.-W., Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2006). Limits to human movement planning in tasks with asymmetric gain landscapes. *Journal of Vision*, *6*(1), 53–63.

Wu, Y., van Dijk, E., & Clark, L. (2015). Near-wins and near-losses in gambling: a behavioral and facial EMG study. *Psychophysiology*, *52*(3), 359–366.

Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., & Krakauer, J. W. (2008). Explaining savings for visuomotor adaptation: linear time-invariant state-space models are not sufficient. *Journal of Neurophysiology*, *100*(5), 2537–2548.

Zhang, H., Daw, N. D., & Maloney, L. T. (2013). Testing whether humans have an accurate model of their own motor uncertainty in a speeded reaching task. *PLoS Computational Biology*, *9*(5), e1003080.

# Resum en català

## Introducció i objectius

El moviment és una part essencial de la nostra existència. És a la base de les conductes més elementals de qualsevol ésser viu, com ara l'alimentació, l'autoconservació o la reproducció. En moltes ocasions, aquest moviment es produeix com a resposta a certa informació recollida pels nostres sentits. Entre aquests sentits, la visió té un paper preponderant, fins al punt que com a mínim un 30% del nostre cervell està destinat a processar informació visual. Ara bé, la manera com representem el món que ens envolta no és perfecta: la nostra capacitat per representar informació visual és limitada, i a això se li ha de sumar el soroll que es produeix als diversos nivells del processament visual. Els nostres moviments també inclouen variabilitat, tant en la planificació com en l'execució, i en conseqüència hi haurà una diferència entre el moviment que inicialment volem fer i el que finalment acabem fent. Així doncs, sembla que tant percepció com moviment tenen un grau d'incertesa inherent. Davant d'aquest escenari convé preguntar-nos de quina manera gestionem aquesta incertesa quan planegem un moviment a partir d'informació visual.

Quina és la millor manera de prendre les decisions? Les anomenades teories normatives s'han centrat a respondre aquesta pregunta. Una de les més famoses és la statistical decision theory (Teoria de Decisió Estadística, SDT, Blackwell & Girshick, 1954), que tracta d'explicar com hauríem de decidir en presència d'informació incompleta (incertesa).

Aquesta teoria resumeix els elements que intervenen en les decisions de la següent manera: hi ha una sèrie d'estats del món i, donat cert estat, la la nostra percepció el correspondrà amb un estat sensorial. En base a cert estat sensorial, nosaltres podem dur a terme una acció, i aquesta acció donarà lloc a una conseqüència, que es pot definir com un altre estat del món. L'element més rellevant és el fet que nosaltres duem a terme accions pensant en les conseqüències. Assumint que volem una conseqüència el més favorable possible, la SDT s'aboca a estudiar com podem prendre la decisió òptima. Aquesta sovint es defineix com aquella que ens porta a maximitzar el guany o recompensa esperats, que no és més que la recompensa promig que obtindríem a llarg termini si la situació es presentés repetidament i sempre prenguéssim la mateixa decisió. Per exemple, si la situació requerís decidir entre guanyar 5€ amb una probabilitat de 0.5, o 3€ amb una probabilitat de 0.8, tot el que caldria fer per obtenir la recompensa esperada d'una d'aquestes "loteries" és multiplicar la recompensa per la seva probabilitat, i després senzillament caldria escollir la loteria amb el valor esperat més alt (en el nostre cas, la primera).

En clar contrast amb les teories normatives, l'objectiu de les quals és establir com hauria de ser el comportament òptim, les teories descriptives s'han dedicat a observar com els humans prenem decisions. La conclusió general a la que s'ha arribat és que el comportament observat difereix substancialment de l'òptim normatiu. Ara bé, les teories descriptives, la més famosa de les quals és la Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), es basen en situacions decisionals similars a l'exemple de les loteries de l'anterior paràgraf. Diverses investigacions en tasques de naturalesa sensoriomotora semblen arribar a conclusions totalment diferents.

L’any 2003, Julia Trommershäuser i col·legues (Trommershäuser et al., 2003b) reformularen la planificació de moviment en termes de SDT. D’acord als seus resultats, els humans som capaços de seleccionar el moviment que maximitza la recompensa esperada de la situació. Aquest plantejament teòric inaugurà una influent línia de recerca amb resultats diversos. Malgrat molts estudis trobessin evidència d’optimalitat, d’altres conclogueren que el comportament era subòptim. En les situacions sensoriomotors estudiades, múltiples causes s’han donat com a causa de la suboptimalitat. Per exemple, falta d’experiència amb la tasca. També una incorrecta representació de la variabilitat motora (que equivaldria a la incertesa), de la funció de recompensa (que descriu la relació entre cada possible acció amb la seva conseqüència en termes de recompensa), o de les dues. Part dels esforços en aquest camp també han estat dedicats a adreçar crítiques als models, ja que en ocasions aquests classifiquen com a òptim comportament que és realment subòptim.

La línia de recerca acabada d’esmentar s’ha limitat sobretot a situacions consistents en fer moviments simples envers objectius estàtics. Hi ha un camp interessant dins del terreny sensoriomotor que aquest punt de vista ha ignorat. Aquest camp és la interacció amb objectes en moviment. Els estudis en percepció visual i en tasques sensoriomotores (sobretot d’intercepció) ens han proveït de valuós coneixement respecte a com percebem els objectes en moviment, com realitzem accions per afectar la seva trajectòria, quina és la relació entre el moviment i la posició percebuda, … Malgrat tot, el focus principal d’aquestes investigacions no ha sigut el component decisional en termes de les conseqüències de les nostres accions. Si es fa un esforç, però, resulta senzill esboçar una situació amb objectes en moviment en termes de teories decisionals com la SDT. Imaginem-nos un pilot de curses intentant decidir quan girar el volant per estalviar temps en un revolt, un competidor de tir al plat aguaitant per prémer el gallet, o un porter de futbol intentant capturar una pilota que s'aproxima. En totes aquestes situacions hi ha certs elements en comú. Primerament, hi ha una acció que cal fer relativa a l'objecte en moviment. A més, actuar en un moment no portarà a les mateixes conseqüències que actuar en un altre moment. Finalment, l'acció requerirà una decisió prèvia sobre quan actuar. Aquesta decisió estarà basada en diferent informació: l'estimulació visual que ens arriba, el coneixement de com cada acció està lligada a una conseqüència, i el coneixement de la nostra incertesa quan duem a terme l'acció planificada. Mitjançant un plantejament similar al que s'acaba de fer, aquesta tesi es proposa unir els camps de la presa de decisions sensoriomotora i el de l'estudi de la interacció amb objectes en moviment. Els objectius generals de la mateixa es presenten a continuació.

- Quan planegem una acció amb un objecte en moviment, podem fer-ho basant-nos en diverses fonts d'informació visual que l'objecte ens proporciona. Per una banda, informació espacial. Per altra, informació temporal. El primer objectiu de la tesi consisteix en saber si fer servir més un d'aquests tipus d'informació pot fer augmentar el nostre rendiment en aquest tipus de situacions.
- Els models de planificació de moviment òptima que s'han generat a partir de la SDT han sigut d'utilitat, ja que han permès establir un comportament òptim amb el que comparar el comportament observat en diverses tasques. El segon objectiu de la tesi és estendre aquests models a les situacions d'interacció amb objectes en moviment. A la vegada, s'intenten

identificar limitacions a l'hora d'aplicar els models en aquest tipus de situacions.

- El tercer objectiu de la tesi se centra en estudiar com les persones fan servir informació sobre la recompensa després de cada acció per a guiar accions futures. Això s'investiga tant a curt termini, per veure com una recompensa influeix la següent acció, com a llarg termini, influint la informació que s'explota en general per prendre decisions.

**Estudis experimentals**

Els objectius de la tesi es van abordar a través de tres estudis diferents. El que segueix és un resum dels mateixos.

Estudi 1 - Explotar informació temporal millora el rendiment en la presa de decisions sensoriomotora amb objectes en moviment

Quan interactuem amb objectes en moviment podem basar les nostres accions en diversa informació visual. Per exemple, en informació espacial: podem seguir la posició de l'estímul fins que arriba a cert punt en l'espai, i llavors iniciar l'acció. Per altra banda, es pot emprar informació sobre moviment i velocitat per realitzar una estimació del temps que l'objecte trigarà a arribar a cert punt, i basar la decisió en aquesta informació temporal. També hi ha la possibilitat de combinar ambdós tipus d'informació. L'objectiu que es va fixar per aquest estudi va ser conèixer, en aquestes situacions, fins a quin punt el rendiment depèn del tipus d'informació emprada per planificar la resposta. També es pretenia saber si el temps de visualització de l'objecte en moviment pot afectar l'ús del tipus d'informació. Finalment, ens proposàrem implementar models provinents d'SDT per esbrinar fins a quin punt són una bona eina per interpretar el comportament en aquest tipus de tasques.

El paradigma experimental que es va dissenyar consistí en un target, que podia diferir en velocitat a cada assaig, desplaçant-se lateralment cap a una línia vertical. Els participants premien un botó per aturar l'objecte, i eren recompensats en funció de la proximitat de l'objecte a la línia en el moment de resposta. La recompensa augmentava exponencialment a mesura que el temps de contacte entre target i línia disminuïa, amb un màxim en el moment d'alineament entre ambdós. Qualsevol resposta després de l'alineament era fortament penalizada. El rendiment de la tasca es va quantificar a partir de la recompensa promig aconseguida en els múltiples assaigs de l'experiment. A partir d'analitzar les respostes a través de les velocitats, vam identificar la variable decisional (DV) dels participants: és a dir, si la resposta estava basada més en claus temporals o espacials.

Vam descobrir que fer més ús de claus temporals portà a un major rendiment, i que temps de visualització més grans afavoriren fer més ús d'informació temporal, augmentant el guany. Podem explicar aquests resultats amb dos mecanismes compatibles. En primer lloc, la resolució del sistema visual és limitada, de tal manera que basar la nostra decisió en un seguiment de la posició portarà a major variabilitat espacial com més gran sigui la velocitat de l'objecte. En canvi, fer servir informació temporal elimina aquest problema, pel que el rendiment és més estable en totes les velocitats, la qual cosa en la nostra tasca significà no incrementar les penalitzacions per les velocitats altes. En segon lloc, un major temps de visualització del target es tradueix en més integració d'informació de la velocitat, la qual cosa pot afavorir l'ús d'informació temporal. Cal dir que això es va poder veure afectat per una possible presència d'extrapolació de moviment, consistent en percebre la posició d'un objecte per davant de la seva posició real. Aquesta extrapolació creix per velocitats més altes i més temps d'integració, pel que podria haver influït en els nostres resultats. Un parell d'experiments reforçaren les nostres conclusions: en primer lloc, els resultats de l'experiment principal foren replicats en una tasca on la recompensa es donà en funció de la distància espacial. Així, no es pot dir que el major rendiment fos causat perquè el nostre experiment principal donés la recompensa en funció del temps. En segon lloc, realitzàrem un experiment on obtenir informació temporal era més difícil, doncs l'única informació de moviment de l'estímul era de segon ordre (a partir de contrast, i no de luminància). En coherència amb les nostres hipòtesis, aquest escenari va portar a DV menys temporals, i això va portar a un descens del rendiment.

Pel que fa a l'aplicació dels models d'SDT, segons aquests les respostes de la mostra foren clarament subòptimes. Malgrat això, es van posar de relleu clares limitacions d'aquests models a l'hora d'explicar les nostres dades. Aquests models assumeixen que les decisions òptimes es prenen integrant una representació acurada de la variabilitat motora amb el coneixement de la funció de recompensa de la tasca. Tanmateix, els models actuals no especifiquen les diverses fonts de variabilitat. En tasques amb objectes en moviment, l'incertesa sensorial és rellevant, pel que aquesta cal que sigui tinguda en compte. També cal considerar com varia aquesta incertesa, per exemple quan l'objecte pot adoptar diferents velocitats o temps de visualització. Finalment, els models han d'incorporar el fet que, en tasques com la nostra, les respostes es poden planificar fent servir diverses fonts d'informació.

### Estudi 2 - Pot la recompensa induir l'ús d'informació temporal en la interacció amb objectes en moviment?

L'estudi anterior demostra que, quan s'nteractua amb objectes en moviment, fer servir informació temporal per planificar les accions pot ser avantatjós. Així doncs, aprendre a fer servir aquestes claus temporals podria portar a una millora en aquest tipus de situacions.

Els estudis que han intentat aconseguir aprenentatge perceptual sovint han implementat sistemes de recompensa per a que, a més de ser un reforç, aquesta es fes servir com a informació. En aquest estudi es va agafar aquesta idea per crear un procediment basat en la recompensa, amb l'objectiu que els participants poguessin aprendre a fer servir claus temporals. En concret, es va modificar el paradigma de l'estudi anterior manipulant la recompensa després de cada assaig. La recompensa sofrí un escalament de tal manera que aquesta era més reduïda com més distància temporal hi havia entre les respostes anteriors. Per tal d'exercir de comparació, es va crear una altra condició amb un procediment similar, però aquest cop per promoure l'ús de claus espacials. Els resultats obtinguts foren una major presència de participants que van explotar informació temporal en la condició on aquesta es va promoure. Tanmateix, aquests participants feren servir les claus temporals des d'un principi. Això implica absència d'aprenentatge, pel que possiblement les diferències individuals van determinar qui va explotar informació temporal. Això fa concloure la poca eficància de la nostra manipulació de la recompensa. Malgrat tot, no es pot descartar que altres maneres de manipular la recompensa puguin portar a un aprenentatge de com fer servir claus temporals.

### Estudi 3 - Quina variabilitat considerem quan fem decisions amb objectes en moviment?

En moltes situacions sensoriomotores, si volem que les nostres accions es tradueixin en el millor resultat possible, aquestes han de ser planificades tenint en compte la nostra variabilitat. Com ja s'ha apuntat abans, l'SDT és una teoria normativa que estableix que la decisió òptima és aquella que produeix l'acció que maximitza la recompensa esperada de la situació. En moltes situacions s'ha conclòs que les persones són subòptimes. En multitud de casos, però l'origen d'aquesta suboptimalitat roman sense ser esclarida. La hipòtesi de partida d'aquest estudi és que, en algunes d'aquestes situacions, els participants podrien haver planificat les respostes considerant només part de la seva variabilitat observada. En els nostres estudis previs (i probablement també en diversos experiments fets per altres investigadors) els participants feren servir informació sobre la recompensa per realitzar correccions sobtades en la seva resposta planificada, per tal de minimitzar les penalitzacions. Nosaltres apuntàrem a que la variabilitat introduïda per aquestes correccions és la que no hauria sigut tinguda en compte.

Per tal de trobar suport per la nostra hipòtesi, vam utilitzar dades recollides anteriorment i vam ajustar-hi un Kalman filter. Això ens permeté estimar, per a cada participant, la variabilitat de mesura, que correspondria a la variabilitat resultant un cop descomptada aquella afegida per les correccions. Aquesta variabilitat de mesura equival aproximadament a la variabilitat d'execució motora. A partir dels models de SDT vam agafar aquesta variabilitat i vam generar les respostes òptimes per cada participant. Això vol dir que aquestes respostes correspondrien al punt òptim que maximitzaria la recompensa esperada si només s'hagués tingut en compte la variabilitat de mesura. Tal i com havíem hipotetitzat, aquestes respostes òptimes coincidiren bastant bé amb les respostes observades. Malgrat tot, la majoria de participants necessità temps per convergir amb aquestes respostes òptimes. Això reflecteix que es necessità experiència perquè els participants poguessin adquirir una representació de la seva variabilitat, encara que aquesta representació no correspongués a tota la variabilitat observada.

Per altra banda, es va obtenir informació rellevant sobre l'ús que els participants van fer de la informació donada per la recompensa. Com ja s'havia apuntat anteriorment, els participants feren servir la recompensa després d'una resposta per tal d'estimar el risc de penalització derivat de mantenir la planificació en aquell punt. A més, també es trobà evidència preliminar que apunta a que els participants podrien haver estabilitzat les seves respostes en punts que, tenint en compte únicament la variabilitat de mesura, portarien a la recompensa òptima.

## Conclusions

Els resultats dels estudis individuals ens porten a poder dir fins a quin punt s'ha acomplert cadascun dels objectius de la tesi.

En primer lloc, les nostres dades demostren que, en situacions d'interacció amb objectes en moviment, planificar les accions en base a informació temporal pot portar a una millora del rendiment. Això es deuria a que aquest tipus de planificació pot eludir una limitació que ve donada per la resolució del sistema visual. També es conclou com aquest ús d'informació temporal és facilitat quan l'objecte en moviment es pot veure durant més temps, segurament gràcies a una major integració de la informació de velocitat. Tanmateix, la informació usada també depèn de diferències individuals.

Els models actuals de planificació de moviment òptima derivats de SDT han demostrat una utilitat limitada per ser aplicades a tasques d'interacció amb objectes en moviment. Les limitacions provenen de no considerar les diferents variabilitats rellevants per la tasca, així com de no tenir en compte el fet que més d’una variable pot ser emprada per planificar les accions. Un cop separades les variabilitats, aquests models van resultar útils. En concret, per tal d’investigar la variabilitat que els participants consideren per respondre. La comparació de les estimacions òptimes amb les respostes observades donaren més suport a una representació de la variabilitat de mesura, que difereix de la variabilitat total observada en no incorporar aquella variabilitat afegida per les correccions en la planificació.

Finalment, tot i que el mètode dissenyat perquè els participants aprenguessin a explotar claus temporals no va complir el seu propòsit, resultats d’altres experiments mostren l’ús que la informació proporcionada per la recompensa va tenir per tal de planificar futures accions. Aquesta informació permeté programar canvis en la planificació, tant en forma de correccions sobtades per evitar penalitzacions, com per delimitar regions on estabilitzar les respostes.

# Appendix

## SUPPLEMENTARY EXPERIMENT 1: SPACE-BASED REWARD

The present study mainly claims that, in sensorimotor decision-making with moving objects, responses based on temporal information enhance performance. In the experiment in the main text, we assessed performance by looking at the reward won after each trial. And this reward was given as a function of the time remaining to alignment when the response was made. Finding a better performance for using temporal information in a task that defined reward in a temporal fashion, then, could be thought to arise entirely from our specific task design. Since this could undermine our claim, we decided to conduct a control experiment. The task remained the same except for the way reward was given. Instead of depending on the temporal proximity to alignment, reward depended on the spatial proximity to alignment. If we found that, in a task rewarding as a function of the spatial location of the responses, we still replicated enhanced performance for time-based responses. These results support our previous findings.

### Participants

21 participants of a completely new sample took part in the experiment after giving informed consent (9 women, all right-handed, age range 18-40). All had normal or corrected-to-normal vision and were naïve about the aim of the experiment. The study complied with the local ethics

guidelines, in accordance with the declaration of Helsinki, and was approved by the University of Barcelona's Bioethics Commission.

### Apparatus and stimuli

All apparati used here were the same used in the main experiment. However, the reward given after each trial was based on the remaining spatial distance to alignment. The reward function was:

$$u(s) = \beta \cdot exp(\alpha \cdot s) \quad \text{if} \quad s \leq 15 \qquad \text{(S1)}$$
$$u(s) = -200 \quad \text{if} \quad s > 15$$

And the corresponding parameters were $\beta = 1.1109$ and $\alpha = 0.3$, so that the reward was also 100 points at alignment that was centered at position *s* = 15 (for the line was positioned 15 cm right of the center of the screen). Later responses were penalized with -200 points. The value of these parameters was selected so that, for the medium speed, both the time and space condition had the same mapping between time / space and reward.

### Procedure

As in the main experiment, participants completed two blocks of practice trials and twelve blocks of the experimental task. However, in the practice trials, feedback was provided if their response time was within a spatial window of 0.25 cm centered on the line.

## RESULTS

0.401% of trials were identified as outliers. Since reward was given as a function of space, we analyzed response spatial locations. An ANOVA on the LMM gave a significant main effect of motion time (*F*(2,22563)= 71.926, *p*<.001) and block (*F*(1,22563)= 85.578, *p*<.001), but no interaction. For all motion times, response locations started before alignment (initial average per motion time, 0.8 s= -0.900 cm; 1 s= -1.059 cm; 1.2 s= -1.197 cm) but got closer to it across blocks (average rate .013 cm/block, *t*= 9.154, *p*<.001) . As for variability, SDs evolved across blocks (*F*(1,730)= 110.366, *p*<.001) and differed among motion times (initial average per motion time, 0.8 s= .734 s; 1 s= .775 s; 1.2 s= .798 s, *F*(2,730)= 3.769, *p*=.024), but no interaction was produced. On average, SDs decreased at a rate of -.014 cm / block (*t*= -10.45, *p*<.001).

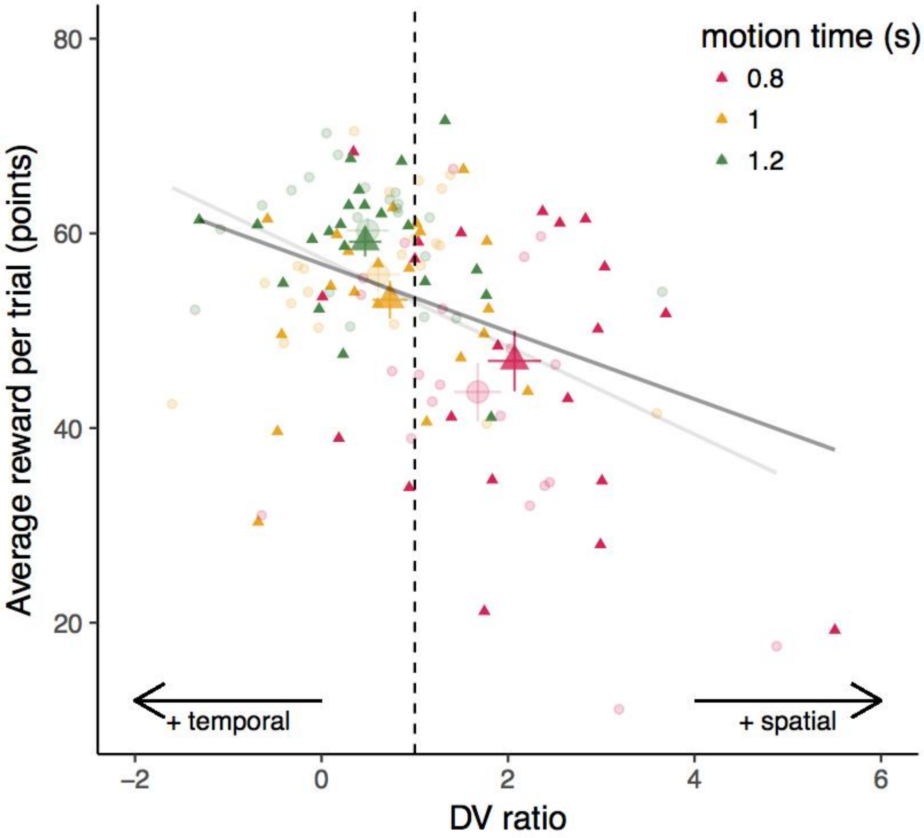


Figure S1. Average reward per trial (in points) as a function of the value of the DV ratio. Each participant is included on the plot with three data points, each corresponding to her responses within a certain motion time, as denoted by the color. Bigger points represent the centroid of each motion time's distribution, with associated error bars that show standard errors of the mean. Strongly colored triangles correspond to the control experiment, pale points correspond to the main experiment. The grey lines show linear fits of the data, with the dark line for the control experiment and the light line for the main experiment . The dashed, vertical, black line shows the cutoff point of DV ratio = 1, used to binarily categorize a group of responses as time or space-based.

As in the main experiment, we ran LMM with the average reward won per trial as a dependent variable, DV ratio and motion time as main effects, and participant as a random effect. Although the ANOVA on the LMM did not show a significant main effect of DV ratio, motion time (*F*(2,19544.7)= 22.947, *p*<.001), and the interaction between DV ratio and motion time (*F*(2,10581.0)= 5.882, *p*=.003) reached significance. This means that, as seen in Figure S1., in this condition a larger average reward was still produced for more time-based responses and longer motion times. Taking all the data together, the correlation between DV ratio and average reward per trial gave an *r*= -0.369 (*p*=.003). When classifying participants into the DV groups, the % in the temporal DV group was 19.048% for a motion time of 0.8 s, 57.143% for 1 s, and 76.190% for 1.2 s.

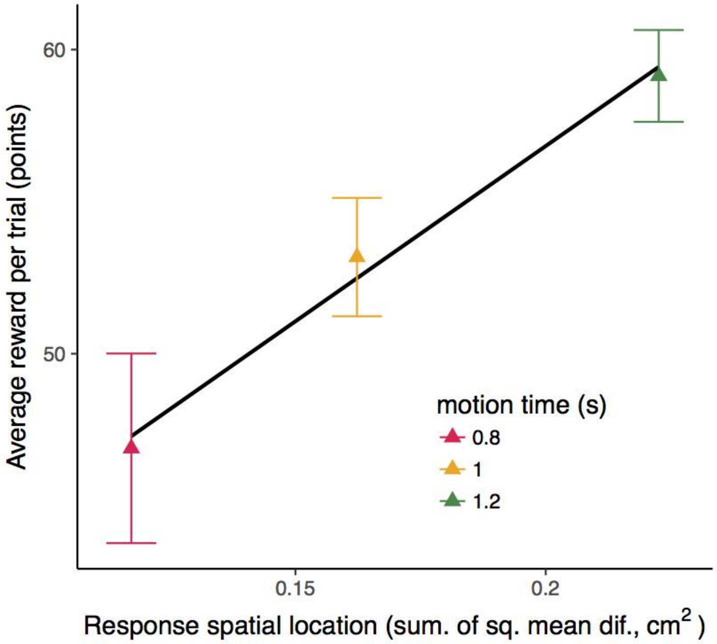


Figure S2. For the control experiment where reward was given as a function of space, average reward per trial (in points) as a function of the sum of the mean squared differences of response spatial locations (in cm$^2$) across speeds. Results are split by motion time, as denoted by the color, and error bars represent the standard error of the mean across participants. The solid, black line represents a linear fit of the data points.

As an additional analysis to the previous one, and following the same procedure from the main experiment, we calculated the mean squared difference between the spatial location for each speed and the mean location pooled across speeds. Figure S2. shows this against average reward. Again, the lower the difference across speeds in response spatial locations, the lower average reward. And lower values were associated to shorter motion times. Non-overlapping error bars show significantly different average values across motion times.

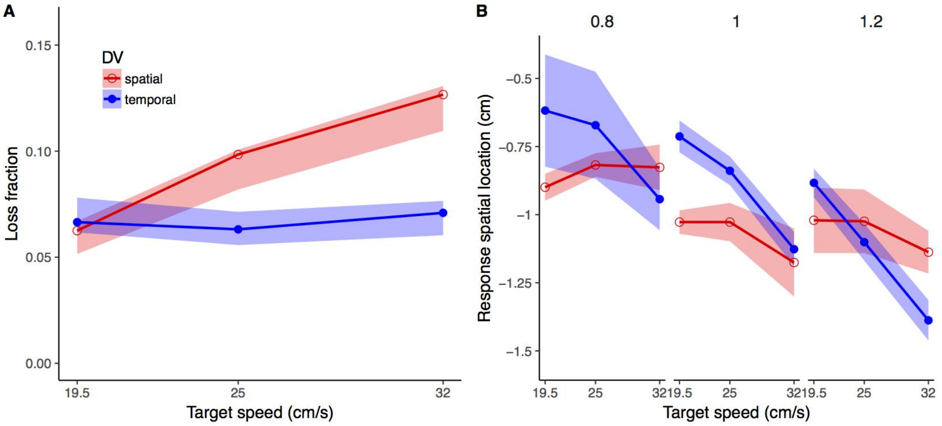


Figure S3. A. For the control experiment where reward was given as a function of space, average loss fractions between DV as a function of target speed. Shaded areas represent binomial 95% confidence intervals. B. For the control experiment where reward was given as a function of space, average response spatial locations between DV as a function of target speed. Shaded areas represent 95% confidence intervals. Each panel shows responses for a specific motion time.

To check how the loss fraction evolved across target speeds, we ran a GLM with loss fraction as the dependent variable, DV and target speed as fixed effects, and participant as random effect. As a quick glance at Figure S3.A can help to see, results from the main experiment were replicated. The increase in the loss fraction across speeds was significant for space-based responses (slope= .056, *p*<.001) but not for time-based responses (slope= .005, *p*=.514). Plus, the difference in percentage was not significant for the lowest speed ($\chi^2(1)$= 1.759, p=.908), but was significant for the two larger speeds (25: $\chi^2(1)$= 16.803, *p*<.001; 32: $\chi^2(1)$= 45.847, *p*<.001).

Finally, we compared spatial locations between DVs. We explored the ANOVA on the output of an LMM with the spatial location of the target at response as the dependent variable, DV group, motion time and target speed as fixed effects, and participant as random effect. As Figure S3.B depicts, space-based responses were still closer in space across speeds than time-based responses, that had earlier responses in space for faster speeds. Moreover, the difference in responses across speeds became more evident for longer motion times, thus replicating the results of the main experiment (interaction between DV group and target speed, *F*(2,22551)= 88.236, *p*<.001)). However, in contrast with the main experiment, space-based responses were on average earlier in space than time-based responses, with the exception of the longest motion time, where this trend reversed (interaction between DV group and motion time not significant; triple interaction among DV group, target speed and motion time marginally significant (*F*(4,22551)= 2.016, *p*=.089)).

## SUPPLEMENTARY EXPERIMENT 2: SECOND-ORDER MOTION

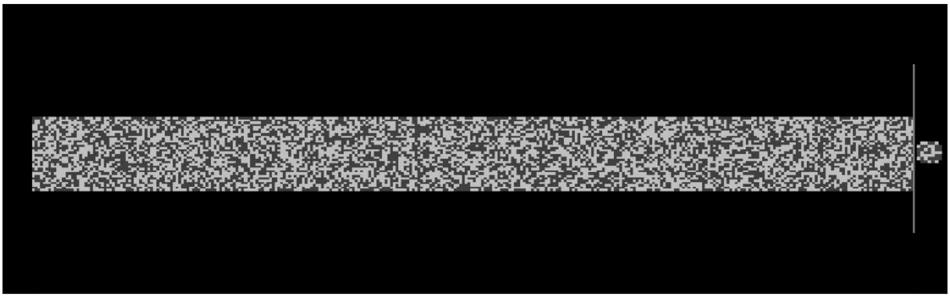


Figure S4. Depiction of the stimuli shown in the second-order motion control experiment. The target moved rightwards towards a line, with the displacement being produced over a horizontal stripe made of constantly changing random dots. The target had the same texture as the stripe, and is shown here beyond the line.

In this control experiment, 9 completely new participants went through the same procedure as in the main experiment. The only difference concerned the visual stimuli presented to the participant, shown in Figure S4. The target's lateral displacement was produced over a horizontal stripe (30 cm width and 3 cm height) consisting of a contrast-defined random dot texture (black and white dots). The size of the texture elements was 0.12x0.12 cm2 and the dots were updated every 80 msec. The moving target (circle of 1 cm diameter) was also composed of the same texture elements and updated at the same rate. With such stimuli, participants had to obtain motion information from a change in contrast, not in luminance. Figure S5. shows the average reward per trial as a function of the DV ratio. Colored dots, corresponding to the second-order motion control, are shifted towards the right with respect to the black dots, which reproduce the data from the main experiment. This shift means that responses were more space-based, which is congruent with the enhanced difficulty to obtain motion cues from second-order motion stimuli. Interestingly, more space-based responses were also translated into a lower average reward.

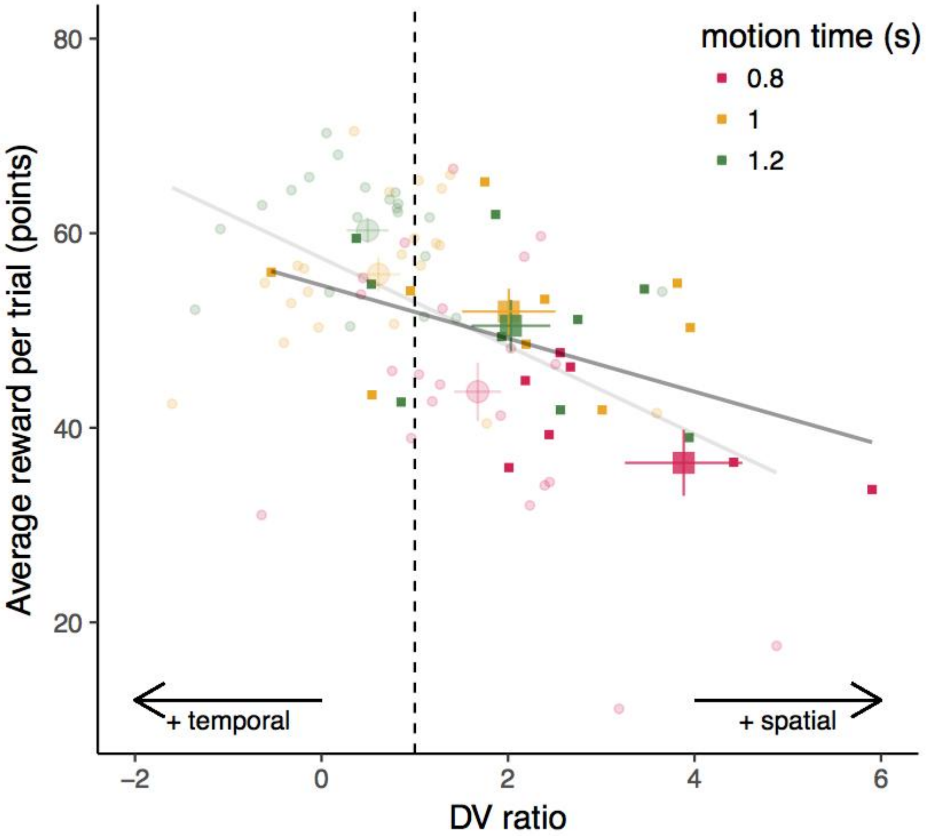


Figure S5. Average reward per trial (in points) as a function of the value of the DV ratio. Each participant is included on the plot with three data points, each corresponding to her responses within a certain motion time, as denoted by the color. Bigger points represent the centroid of each motion time's distribution, with associated error bars that show standard errors of the mean. Strongly colored squares correspond to the SO control experiment, pale points correspond to the main experiment. The grey lines show linear fits of the data, with the dark line for the control experiment and the light line for the main experiment . The dashed, vertical, black line shows the cutoff point of DV ratio = 1, used to binarily categorize a group of responses as time or space-based.