Embodied decision making and its neural substrate

Encarni Marcos Sanmartín

TESI DOCTORAL UPF / 2014

Director de la tesi

Prof. Paul F. M. J. Verschure

Departament de Tecnologies de la Informació i les Comunicacions

UNIVERSITAT POMPEU FABRA

Legal deposit: September 22nd, 2014

To my grandparents
Mariano and Encarna

Nothing is more difficult,
and therefore more precious,
than to be able to decide

Napoleon Bonaparte

Acknowledgements

This thesis has been possible thanks to the encouragement, expertise and support of many people.

My deepest gratitude to my supervisor Paul Verschure, who gave me the opportunity to start all this. Thanks for teaching me to be critical with myself and for all the guidance and inspiration that made this thesis possible.

I would like to thank all the external collaborators that let me learned from them and that contributed to part of this work. Especially to Stefano Ferraina, Emiliano Brunamonti, Gustavo Deco, Paul Cisek and Ignasi Cos.

I would also like to express my gratitude to the Department of Information and Communication Technology of the University of Pompeu Fabra (UPF) for giving me the opportunity to be part of such an excellent institution. This work has been funded by UPF and by European Community FP7/2007-2013 Grant 217148 - SF and Grant 270108 - Goal-Leaders. This financial support is gratefully acknowledged.

I am also very grateful to all my colleagues in SPECS that helped me with their collaborations, advice, and supportive attitude. Special thanks to Mireia Mora, Carme Buisan and Christian Moscardi for making our lives much easier. I would also like to thank Lydia García for her help with the PhD paper work.

A very special thanks to all the people that made my time in Barcelona easy and happy. To Martin Inderbitzin for his constantly positive thinking. To César Rennó-Costa for our trips in USA and around Europe. To Nina Valkanova for the first year in the lab and for all the others that came later. To Sytse Wierenga for being there from the beginning till the end of this thesis. To Inma Silla, for saying "it's done, just finish it" at the right time giving me that little push that I needed, one year ago! To Anna Carreras for her encouragement during these years. To María Piles, for being like my family in Valencia, Lund and then in Barcelona. To María Espinosa for her emotional support. To Felipe and Manolito for their company in the long working summer days.

I would also like to thank my friends in Elche. Specially to Alicia Brotons, Nuria Galvañ and Patricia Campello for always being there and for having shared with me part of my life in Barcelona.

Finally, a million of thanks to my parents, Juan and Enriqueta, for their unconditional trust and support. To my brother and sister, Juan and Josefina, for being, together with them, the most important pillar of my life.

Abstract

Decisions are the result of a deliberative process that evaluates the suitability of specific options. In most cases, decisions are reported through actions that allow for an interaction with the physical world. Studies about decision making have been mainly conducted by using restricted tasks in which humans or animals are requested to discriminate between two options, such discrimination being based either on a perceptual property of a stimulus or on the voluntary control of motor responses. However, the influence that factors related to embodiment, such as experience during a task or motor cost, might have on this process has frequently been ignored. In this thesis, we adopt a combined experimental and theoretical approach to examine the effect that such factors have on decision making, even when optimal decisions do not depend on them. Our results confirm an important bias of behavior and neural activity resulting from factors related to embodiment that are external to the goal of the task itself. In our studies, we enhance existing computational models of binary decision making to account for this bias that, in turn, shed some light on the neural mechanisms producing it. The thesis concludes with the presentation of a single model that integrates all the findings presented and that could be used as a new framework for the conducting of future research. Altogether, the results included here translate into significant progress in the understanding of embodied decision making, providing new insights into neural mechanisms and theoretical models.

Resumen

Las decisiones son el resultado de un proceso de deliberación que evalúa la idoneidad de opciones específicas. En la mayoría de los casos, las decisiones son comunicadas a través de acciones que permiten una interacción con el mundo fsico. Los estudios acerca de la toma de decisiones han estado principalmente dirigidos usando tareas restringidas en las que, a humanos o animales, se les pide escoger entre dos opciones. La elección se basa en una propiedad perceptual de un estímulo o en el control voluntario de respuestas motoras. Sin embargo, la influencia que factores relacionados con la corporificación de la toma de decisiones podrían tener en este proceso se ha ignorado frecuentemente. En esta tesis, adoptamos un enfoque experimental y teórico combinado para examinar la influencia que estos factores tienen en la toma de decisiones, incluso cuando las decisiones óptimas no dependen de ellos. Nuestros resultados confirman un importante sesgado del comportamiento y de la actividad neuronal causados por factores que son externos al objetivo de la tarea en sí. En nuestros estudios, realzamos modelos computacionales existentes de tomas de decisiones binarias para interpretar este sesgado que, a su vez, nos da una intuición del mecanismo neuronal que los está produciendo. La tesis concluye con la presentación de un único modelo que integra todos los hallazgos presentados y que podría utilizarse como nuevo marco teórico para investigaciones futuras. En general, los resultados incluidos aquí se traducen en un significante progreso en la comprensión de la toma de decisiones corporificada, aportando nuevos conocimientos sobre los mecanismos neuronales y modelos teóricos.

Resum

Les decisions són el resultat d'un procés de deliberació que avalua la idoneïtat d'opcions específiques. A la majoria dels casos, les decisions són comunicades a través d'accions que permeten una interacció amb el món físic. Els estudis sobre la presa de decisions han estat principalment dirigits fent servir tasques restringides a les quals, a humans o animals, se'ls demana escollir entre dues opcions. L'elecció es basa en una propietat perceptual d'un estímul o al control voluntari de respostes motores. No obstant això, la influència que factors relacionats amb la corporificació de la presa de decisions podrien tenir en aquest procés s'ha ignorat freqüentment. En aquesta tesi, adoptem un enfocament experimental i teòric combinat per tal d'examinar la influència que aquests factors tenen en la presa de decisions, fins i tot quan les decisions òptimes no depenen d'ells. Els nostres resultats confirmen un important esbiaixat del comportament i de l'activitat neuronal degut a factors externs a l'objectiu de la tasca en sí. Als nostres estudis, realcem models computacionals existents de preses de decisions binàries per tal d'interpretar aquest esbiaixat que, a la vegada, ens dóna una intuïció del mecanisme que l'està produint. La tesi conclou amb la presentació d'un únic model que integra tots els descobriments presentats i que podria utilitzar-se com a nou marc teòric per a recerques futures. En general, els resultats inclosos aquí es tradueixen en un significant progrés en la comprensió de la presa de decisions corporificada, aportant nous coneixements sobre els mecanismes neuronals i models teòrics.

Publications

Included in the thesis

Peer-reviewed

Marcos, E., Pani, P., Brunamonti, E., Deco, G., Ferraina, S., and Verschure, P. Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. *Neuron*, 78(2):249–255, 2013b

Marcos, E., Duff, A., Sánchez-Fibla, M., and Verschure, P. F. M. J. Generalization of integrator models to foraging: a robot study using the DAC9 model. In Prescott, T. J., Lepora, N. F., Mura, A., and Verschure, P. F. M. J., editors, *Biomimetic and Biohybrid Systems*, volume 7375 of *Lecture Notes in Computer Science*, pages 156–167. Springer Berlin Heidelberg, 2012b. ISBN 978-3-642-31524-4

Marcos, E., Sánchez-Fibla, M., and Verschure, P. F. M. J. The complementary roles of allostatic and contextual control systems in foraging tasks. In Doncieux, S., Girard, B., Guillot, A., Hallam, J., Meyer, J-A., and Mouret, J-B., editors, *From Animals to Animats 11*, volume 6226 of *Lecture Notes in Computer Science*, pages 370–379. Springer Berlin Heidelberg, 2010b. ISBN

978-3-642-15192-7

Marcos, E., Ringwald, M., Duff, A., Sánchez-Fibla, M., and Verschure, P. F. M. J. The hierarchical accumulation of knowledge in the distributed adaptive control architecture. In Baldassarre, Gianluca and Mirolli, Marco, editors, *Computational and Robotic Models of the Hierarchical Organization of Behavior*, pages 213–234. Springer Berlin Heidelberg, 2013c. ISBN 978-3-642-39874-2

In preparation

Marcos, E., Carland, M., Thura, D., Cisek, P., and Verschure, P. F. M. J. Decision-making depends on a task-dependet urgency. *In prep.*, 2014a

Marcos, E., Cos, I., Cisek, P., Benot, G., and Verschure, P. F. M. J. The cost of embodiment in decision-making. *In prep.*, 2014b

Carland, M., Marcos, E., Thura, D., and Cisek, P. Perceptual decisions are better explained by urgency-gating than by sensory accumulation. *In prep.*, 2014

Other publications and abstracts

Marcos, E., Cos, I., Cisek, P., Girard, B., and Verschure, P. F. M. J. Biomechanical costs of reaching movements bias perceptual decisions. *BMC Neuroscience*, 14(1):P408, 2013a

Marcos, E., Carland, M., Thura, D., Cisek, P., and Verschure, P. F. M. J. Decision-making depends on an urgency signal modulated by context. In *2012 Society for Neuroscience meeting (SfN 2012)*, New Orleans, US, 2012a

Marcos, E., Pani, P., Brunamonti, E., Ferraina, S., and Verschure, P. F. M. J. Behavioral and neural modulation during motor decision making in a

countermanding task. In *2011 Society for Neuroscience meeting (SfN 2011)*, Washington, US, 2011

López, L., Vouloutsi, V., Escuredo Chimeno, A., Marcos, E., Bermúdez, S., Ziyatdinov, A., Mathews, Z., Perera Lluna, Alexandre, and Verschure, P. Moth-like chemo-source localization and classification on an indoor autonomous robot. In Pramatarova, Assoc. Lilyana D., editor, *On Biomimetics*, pages 453–466. InTech, 2011. ISBN SBN 978-953-307-271-5

Rennó-Costa, C., Luvizotto, A., Marcos, E., Duff, A., Saáchez-Fibla, M., and Verschure, P. F. M. J. Integrating neuroscience-based models towards an autonomous biomimetic Synthetic Forager. In *2011 IEEE International Conference on Robotics and Biomimetics*, pages 210–215, Phuket, Thailand, December 2011. IEEE. ISBN 978-1-4577-2138-0

Duff, A., Rennó-Costa, C., Marcos, E., Luvizotto, A.L., Giovannucci, A., Sánchez Fibla, M., Bernardet, U., and Verschure, P.F.M.J. Distributed adaptive control: A proposal on the neuronal organization of adaptive goal oriented behavior. In *From Motor Learning to Interaction Learning in Robots*, pages 15–41. In: Sigaud, O., Peters, J. (eds.) From Motor Learning to Interaction Learning in Robots, 2010. ISBN 978-3-642-05180-7

Marcos, E. and Verschure, P. F. M. J. Optimal memory acquisition and recall in real-world foraging. In *FENS Forum 2010 (FENS 2010)*, Amsterdam, Netherlands, 2010

Marcos, E., Sánchez-Fibla, M., and Verschure, P. F. M. J. From continuous analog to discrete symbolic representations of the world in optimal foraging: a robot study. *4th International Conference on Cognitive Systems (CogSys)*, 2010c

Contents

List of Figures

Chapter 1

Introduction

A problem can't be solved with the same level of thinking that created it

Albert Einstein

Even in an apparently simple case of behavior, such as that of a rat exploring and exploiting a new environment, the brain must properly perceive, learn, act and remember. But how are body, brain and mind connected to achieve simple and complex patterns of behavior? Finding an answer to this question has been the central goal of psychology, physiology and neuroscience over the past centuries. However, it seems that interest has normally been focused on problems related to the mind and brain and their unilateral communication with the body, while the potential influence of the body on the mind-brain processes has been usually ignored. Over the course of the last centuries, the brain and mind problem has been addressed from many perspectives, whereas the body has mainly been seen as the end point of their interactions. It is commonly believed by modern scientists that the mind is a process emerging from the brain and, in turn, providing an answer to part of the ancient question. But what is the role of the body in behavior? Is it a pure receptor of motor commands or does it also bias the selection of actions, thus determining the way we act and the flow of information we perceive, learn and remember? Recent neuroscience research

has invested great efforts into investigating the influence of embodiment in decision making. However, there is still much to be discerned. In this thesis, we will focus on the implications that some factors related to embodiment have in neural activity and behavior during decision making.

In neuroscience, decision making has been widely studied using specific tasks, usually in a controlled environment such as a lab. These tasks generally consist in problems which are less complex than those confronted by humans or animals every day. However, there is a common belief that these initial results will provide the basis for future research about more complex decision-making situations. To cite an example, today we get a good understanding of how the brain makes decisions under well-constrained cases of constant sensory information or motor costs presented in isolation (see Section 1.3 and 1.4). Activity of neurons recorded in certain brain regions has shown to be correlated with decisions, and some models have been proposed to account for this correlation and behavior (see Section 1.2). This way, experimental and theoretical neuroscience creates a closed loop in which experimental data is explained by models, and models predict neural data and behavior to motivate future experimental research. In this thesis, we contribute to the investigation of decision making; more concretely, we address the question of how factors related to the embodiment of an agent, including experience within a task or motor cost, could influence decision making. First, we use a specific artificial embodied system to investigate how acquisition of information in memory could be influenced by decision making, and how it could change the way in which it is later retrieved from it (Chapter 2). Next, we describe the neural correlates of the behavioral bias caused by memory, context and motor cost in previous published tasks (Chapter 3 and Chapter 4) as well as in our own experimental paradigm (Chapter 5).

Using robot experiments, we equip a mobile agent with a cognitive architecture that enables it to explore, learn and exploit novel environments to survive. The cognitive system is provided with one of the best-known and most broadly accepted decision-making models, an integrator model

(see Section 1.5). In this study, our aim was to investigate the generalization of such a model in a real-world situation and to assess the possible implications that it might have, in terms of behavioral performance and acquisition of knowledge. In Chapter 2, we start by comparing an integrator model with a model that does not accumulate information to make decisions. This comparison leads to the detection of a fundamental difference in the way that information is gathered from the environment and might be further used. As opposed to the non-accumulator model, that stores and retrieves discrete actions from memory, the integrator model gathers information from goals and the actions to reach them are self-generated. Next, we extend the cognitive architecture with a reactive control system that is able to regulate the internal drives of the artificial agent. We use this new extended system to investigate the integrator model with an even more realistic system. Several questions arise from this theoretical study; among them are the following: what is the neural mechanism that allows the retrieval of information from memory in its specific temporal order? How do the physical constraints of an agent influence action selection? How are actions represented in the brain? Does context itself influence the way we act? To what extent are internal drives attached to the physical body? We approach some of these questions in subsequent chapters.

In order to succeed foraging, the artificial agent needed to solve one fundamental problem (among others): to perform the appropriate behavior based on the sequentiality of the external events. In the cognitive architecture (DAC, see Section 1.7), this is achieved by a proper chaining of the memory space based on perceptual information. Hence, information stemming from both, memory and perception, is integrated in the service of the goal oriented action of the agent, causing a bias in its behavior. The specific neural mechanism underlying this biasing, however, is not well understood. This is the question that we intend to address in Chapter 3. To this purpose, we begin by examining the behavioral bias provoked by trial history that has been previously observed in monkeys performing a motor decision-making task (see Section 1.4) in which two kinds of trials requiring opposite behavior (cancel a movement or move) can occur. Then, we

investigate the neural substrate (related to memory) causing that bias in behavior during decision formation. Interestingly, we find that the firing rate of neurons in the dorsal premotor cortex (PMd) involved in decision making do not exhibit any modulation, whereas the across-trial variability of their response does. To further understand the mechanism causing such modulation of behavior and neural activity we use a mean-field approximation of a cortical decision-making model. The computational study predicts the existence of a system that monitors trial history and biases competition between populations of neurons. This bias in the competition generates the aforementioned modulation in the behavior and in the across-trial response variability, while the mean-firing rate does not vary.

Once we have provided a neural substrate of the bias that experience within a task has in decision formation, we assess the way in which the context (task) itself can also affect experience within a task and, subsequently, decisions. In Chapter 4, we comprehensively expound two decision-making experiments that have shown to provide apparently contradictory results (see also Section 1.3) when subjects or monkeys are presented with stimuli that change over time. Using the same decision-making model as in Chapter 3 we provide a plausible mechanism by which the behavioral policy adopted in each context might operate. Both experimental datasets can be explained through the same neural mechanism: the apparently opposed results can be captured with a context-dependent signal that evolves differently within each task, possibly due to an optimization of the speed-accuracy trade-off that requires a different policy in each context. This study is pioneering in the sense that, for the first time, it interprets and unifies the results of two separate decision tasks that seemed to be paradoxical until now. We also add experimental data recorded at the University of Montreal that verifies one of the predictions of our model, thus proving the validity of our approach.

The final contribution of this thesis consists of both experimental and modeling work. First, we use an experimental paradigm to assess the influence that motor cost of actions might have on perceptual decision making.

To this end, we manipulate the cost of actions required to report perceptual decisions. Our results show that human subjects exhibit a significant bias towards the actions requiring less biomechanical cost, even when it reduces the overall performance accuracy. Next, we use a spiking neural model of binary decision making, sharing the same principles of our previous models, to reproduce the obtained experimental data. We predict that the motor cost of each action is represented in the weight of the lateral connectivity of the population of neurons involved in decision formation, and that this weight is learned over life experience. To account for trials in which subjects simply give up reporting that they are unable to make a decision, we use the response variance that emerges from the neural populations. In our model, the giving up occurs when the variance is above a predefined threshold. In this way, we also elaborate on a plausible mechanism to interpret the role of the neural response variability in decision making.

In summary, this thesis has meant significant progress in the research of decision making due to the description of neural substrates correlated with behavior and the resulting advance in the development of biologically constrained decision-making models that might lead to the design of future research. In what follows, we will provide a brief overview of the different ways in which the mind, brain and body issue has been addressed in the past. Next, we provide the necessary background for the reader to fully understand the basis of our studies and their particular contributions. We review, although non-exhaustively, the state of the art in decision making by delineating some of the experimental paradigms and models used, as well as some of the neural correlates that have already been shown. This introduction finishes with the definition of embodied decision making and how we approach it in this thesis.

1.1 Historical perspectives

During the Middle Ages and the early Renaissance, the work of the Roman physician Galen of Pergamon (AD 129 - AD 210/216) was the most influential in the fields of anatomical and medical belief and practice (Ergil,

1997; Said, 1975). Galen proposed that mind and body could not be separated, and that the complex actions performed by the body were caused by material interactions between its parts and by forces produced by the mind. From Galen's point of view, sensations were collected by the body and sent to the mind for further analysis. After this process, the mind sent the appropriate instructions back to the body in order to activate the corresponding nerves and muscles that would trigger complex behaviors. He stated that the link between body and mind is the brain, as opposed to what Aristotle had suggested centuries before: that the heart was responsible for this connection. Galen also provided comprehensive and accurate (in comparison with the ones preceding it) descriptions of the anatomy of the brain, as well as of the structure and function of the ventricular system (Rocca, 1997).

Prior to Galen's work, the ancient Greek philosophers Epicurus and Democritus had argued that the world is composed of small tiny particles, which Democritus called *atoms* (Kirk, 1983), and that the behaviors that we observe are caused by the physical interactions of those particles (like the duck that Vaucanson built in 1739, see Fig. 1.1). Therefore, they postulated that human or animal behavior is deterministic, -just as any other physical phenomena- and that free will does not exist. It was Plato who, by presenting his *Theory of forms*, declared that the world that we see is not the real one but a "copy" of it, and that the true world only exists in a metaphysical level that we cannot penetrate (Cooper and Hutchinson, 1997). It is in that world, Plato proposed, that mind and behavior are connected. A contemporary Greek physician, Hippocrates, also suggested a compromise between the material and nonmaterial views, and so he inferred that humans are composed of material and nonmaterial processes. He suggested that the physical behaviors produced by the human body are governed by a nonphysical process, namely the mind (Grammaticos and Diamantis, 2008). He was also the first known person to believe that psychopathologies were caused naturally, due to factors such as living habits, and not by external entities such as gods. Galen interpreted and extended Hippocrates' theories, hence becoming the most important reference during

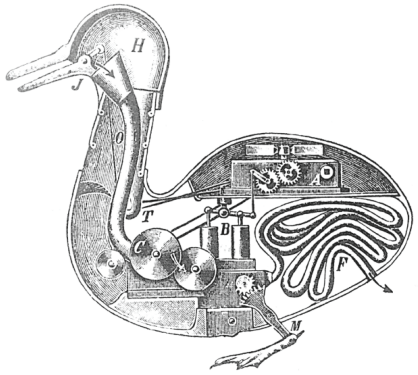

Figure 1.1: Vaucanson's digestive duck. The duck is an automaton created by Jacques Vaucanson in 1738. Once activated, it was difficult to tell the difference between it and a real duck. It could flap its wings, move and digest grain. The duck was able to produce a complex pattern of behavior just by means of physical interactions of its parts. Sketch of how the digestive apparatus of the automaton duck might look like, done from an American observer (From Chapuis and Gélis (1928))

the rise of medieval universities in Europe. To Galen, the body was a complex physical machine and the mind the nonmaterial unit sending causal forces to it. For several centuries, the work of Galen was largely accepted without question. It was in the sixteenth century, though, when Andrea Vesalius revised and corrected some of the anatomical work done by Galen, thus marking the beginning of modern anatomical science (O'Malley, 1964).

In the seventeenth century, René Descartes played a critical role in defining the way in which brain and behavior should be studied, and provided neural science with an important impetus and a shift of direction. He was a firm believer that our body actions as well as those of animals could be understood in terms of physical cause and effect (Aristotle theory of efficient cause). To account for involuntary, reflexive behavior of humans and animals, Descartes proposed that a physical chain of cause and effect takes place unintentionally: a sensory stimulus activates specific organs that connect with the brain which, in turn, prompts muscle response (Clarke, 982). Later, Pavlov described this finding as the starting point of his experiments about classical conditioning (Pavlov, 1927). Descartes believed that

animals were purely driven by reflexive behavior and that, therefore, they are merely biological automata. From Descartes point of view, voluntary actions can only be performed by humans because they possess a mind that is independent from the body (mind-body or Cartesian dualism) and that connects with it at a specific location in the brain, known as the pineal gland (De Rosa, 2009).

By the end of the eighteenth century, the functioning of the nervous system had been thoroughly detailed, and the speculation that the brain might be functionally divided appeared for the first time. Moreover, neurophysiological studies had revealed that nerve cells produced electricity, and that this electricity was used by cells to communicate between themselves. Advances in the development of technology for microscopy along the early 1800s provided the necessary tools for the histology of the nervous system that culminated with the investigations of Camillo Golgi and Santiago Ramón y Cajal (Dröscher, 1998). Golgi developed an impregnation method that makes all the parts of a neuron -the body, the dendrites and the axon- visible, thus allowing neuroanatomists to track connections between neurons (Golgi, 1873). This is the technique that Ramón y Cajal used to gain new insights into the organization of the nervous system, therefore postulating that it is made up of many individual signaling elements: the neurons (Fig. 1.2), as opposed to being a mass of fused cells as it was believed until then (Ramon y Cajal, 1977). Ramón y Cajal is considered the father of the modern neuroscience by many.

Although the study of brain, mind and behavior has been of great interest for a long time, most of the progress towards their understanding has been achieved during the last two centuries. One of the reasons for this was the lack of technology that limited the study of the brain to behavioral observations and speculations until that time. Another reason was the absence of a proper scientific methodology, which made theories weak in some cases and difficult to integrate. What neuroscience aims to do is precisely to provide a common scientific methodology, so that the nervous system can be investigated and related to cognition and behavior. To that end, it splits

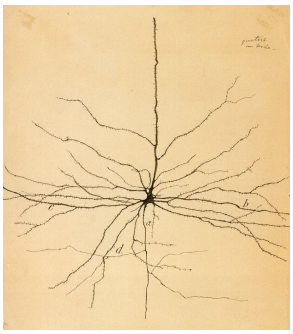

Figure 1.2: Drawing of a retinal neuron by Ramón y Cajal (1899)

the problem into subproblems or levels of analyses that range from molecular to cognitive types. In the last century, neuroscience has become one of the leading scientific fields thanks to the advances attained in molecular biology and electrophysiology. The new technology has provided a wide range of new possibilities that allows, for instance, the detailed study of complex processes occurring within a single neuron. In this way, the interest has focused on the activity of neurons forming the brain and their correlation with behavior. Out of this emerging field, decision making has been of particular interest and a great effort has been made towards the description of its basic underlying neural mechanisms. Decisions are seen as the ability to select a proper action in order to reach specific goals which, in most cases, culminate in a physical movement. Accordingly, human movements are not always just a mere result of deterministic physical interactions between tiny particles without free will, but the consequence of intended behavior.

1.2 Decision making

A decision can be defined as a choice between alternatives resulting in an action or proposition. A deliberative process assessing the suitability of each alternative option is known as the decision-making process. To study the neural mechanisms underlying this process and the resulting behavioral outcome, research has followed two approaches: theoretical and/or experimental research. Experiments have mainly focused on perceptual and motor

decision making in which humans or animals have to make decisions based on the discrimination of a perceptual feature of a stimulus or on the selection between reaching movements, respectively. Neurophysiological data have shown that the activity of neurons in areas such as the superior colliculus(Munoz et al., 2000; Ratcliff et al., 2007; Shen and Paré, 2007), the lateral intraparietal area (Roitman and Shadlen, 2002; Leon and Shadlen, 2003), the frontal eye fields (Gold and Shadlen, 2003), the prefrontal cortex (Kim and Shadlen, 1999), the supplementary motor area (Scangos and Stuphorn, 2010) and the dorsal premotor cortex (Mirabella et al., 2011) is modulated during the decision-making process. Theoretical research focused on computational models has helped to explain the observed behavioral and neural data, and has been very useful to inspire new experiments, therefore producing a closed loop between experimental and theoretical studies.

Both experimental and theoretical approaches have been evaluated by using controlled situations in which the decision usually depends solely on one variable. However, in real-world situations, this is unlikely to be the case. First, to make a perceptual decision reported with an action, one cannot easily neglect the potential influence that the physical constraints of our body might have in the assessment of each alternative, probably biasing the value of each option (Sabes et al., 1998; Cos et al., 2011, 2012). Second, the decision-making process might also be influenced by past experience with an adaptation of preferences that depend on previous outcomes (Emeric et al., 2007). And last, the context or situation in which decisions are made might also have an impact in the process. That is, even if the decisions should solely be based on one variable (for instance, one of the visual properties of a stimulus), other factors such as memory or motor cost should not be ignored.

In order to explain the experimental data, many models have been proposed. A branch of well-accepted models are, for example, the "sequential sampling" or "integrator" models (Stone, 1960; Laming, 1968; Ratcliff, 1978; Usher and McClelland, 2001; Mazurek et al., 2003; Bogacz and Gurney, 2007; Link and Heath, 1975). These models propose that sensory in-

formation is sampled and accumulated in favor or against each alternative option. When a certain level of activity is reached, the decision towards the option is made. The dynamics of these models follow the trend observed in the activity of neurons recorded in brain areas involved in the decision process. Alternatively, more biologically plausible models known as attractor models have been widely used to understand the connections, dynamics and interactions between neural populations (Wang, 2002; Albantakis and Deco, 2009). Although all these models have been of great help in the development of an accurate explanation for behavioral and neural data in constrained tasks, their generalization to real-world situation has not been proven yet.

1.3 Perceptual decision making

Deciding on the basis of noisy perceptual information is a common situation with which the brain is confronted every day. A proper classification of the noisy sensory input is fundamental to guide behavior in a daily basis. This is the reason why tasks in which humans or animals have to discriminate a perceptual property of a stimulus and to report it with an action (sensory-motor tasks) have become increasingly popular in decision-making research (Mountcastle et al., 1990; Britten et al., 1992; Maunsell and Van Essen, 1983; Uka and DeAngelis, 2006; Allred et al., 2005; Baylis et al., 2003; Dolan et al., 1997; Freedman et al., 2002, 2003; Grill-Spector et al., 2000; Op de Beeck et al., 2001; Rainer et al., 2004; Uchida and Mainen, 2003). One example of these tasks is the vibrotactile frequency discrimination task, developed by Mountcastle et al. in the 60s. Subjects are required to compare the frequencies of two tactile stimuli separated by a time gap and decide which of the two is the greatest. Neurons in the primary somatosensory areas have shown to increment their firing rate with increasing tactile frequency. Another example is the face/object discrimination tasks in which subjects are asked to classify whether a noisy image contains a face or not. Responses of neurons in the inferior temporal cortex are highly correlated with complex visual stimuli such as faces (Afraz et al., 2006). In

the following sections, we review a well-established task for visual discrimination, the so-called "Random-dot Motion (RDM) Discrimination", which is considered to be a standard in the study of decision making and is the basis for our studies in Chapter 4 and Chapter 5.

RDM task and variations

The RDM task has been widely used to study binary decision making (Britten et al., 1992, 1993; Roitman and Shadlen, 2002; Shadlen et al., 1996). In this task, subjects are required to detect the net direction of motion (they commonly have to discriminate between two possible motion directions) of some dots displayed in a screen (Fig. 1.3). Only a restricted amount of dots move coherently in the same direction, whereas the remaining dots change direction randomly. The difficulty of the trial depends on the percentage of dots that move coherently towards the same direction. Generally, the direction identified has to be reported either with a saccade or a reaching movement towards one of the two targets that are placed on opposite sides of the screen (commonly right/left or bottom/up).

Experiments have been typically conducted using a "reaction time" (RT) version of the task (Fig. 1.3A) (Roitman and Shadlen, 2002), or a "fixed duration" (FD) version of it instead (Fig. 1.3B) (Shadlen et al., 1996; Shadlen and Newsome, 2001). In the first case, subjects or monkeys have to report the detected motion as fast as possible. On the contrary, in the FD version, subjects or monkeys report their choice when they are instructed to do so (for instance, by means of a visual cue). The FD task allows researchers to study the accuracy of choices and its neural correlates whereas the RT task also provides an additional behavioral measurement: the speed in the selection of choices. It has been shown that RT and accuracy of choices have an inverse relationship, i. e. RT decreases as the difficulty of the trial increases, while the opposite occurs with accuracy (Fig 1.4A). An additional advantage that the RT task has over the FD one is that the decision formation process can be more easily separated from any motor planning activity.

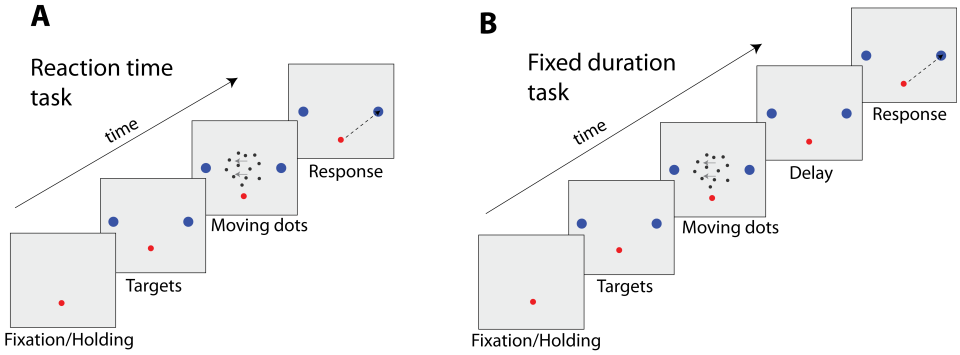

Figure 1.3: Illustration of a RDM task. The trial starts with the appearance of a central visual stimulus. Human subjects or monkeys are requested to hold the stimulus with their arm or fixate their gaze (reaching/saccade task). After some random delay, two targets indicating the possible directions of motion appear on the screen, followed by the moving dots. **(A)** Reaction time task. Humans or monkeys have to report their choice by selecting one of the two targets as fast as they can. **(B)** Fixed duration task. After a fixed duration, the moving dots disappear from the screen informing humans or monkeys that they must report their choice.

The activity of neurons in the middle temporal area (MT/V5) has shown to be tuned to the direction of visual motion and causally linked with task performance (Ditterich et al., 2003; Newsome and Paré, 1988; Salzman et al., 1990, 1992). MT and medial superior temporal (MST) areas project to the lateral intraparietal cortex (LIP) that is connected to the frontal eye fied (FEF) and the superior colliculus (SC) (Andersen et al., 1990, 1992; Asanuma et al., 1985; Blatt et al., 1990; Fries, 1984; Lewis and Van Essen, 2000). Because of its anatomical placement between the sensory and the motor areas, LIP has been the focus of most neurophysiological research in decision making. LIP is involved in high order processes such as selection of saccade targets, working memory or representation of elapsed time (Chafee and Goldman-Rakic, 2000; Friedman and Goldman-Rakic, 1994; Janssen and Shadlen, 2005; Leon and Shadlen, 2003). Using both the FD and RT versions of the RDM task with monkeys, Shadlen et al. (Shadlen et al., 1996; Shadlen and Newsome, 2001; Roitman and Shadlen, 2002) provided evidence of the implication of LIP in decision formation. Single neurons and target

locations were selected based on the response fields (RF) of the neurons; i. e. targets were arranged in such a way that the activation of a neuron under study indicated the monkeys' choice. Neurons with their RF in the selected target increased their activity during the decision formation, reached a peak of maximum activation some milliseconds before movement execution, then declining again to baseline levels. Same neurons showed a decrease in their activity when the opposite target was selected. Moreover, the slope of the activity build-up depended on the amount of motion coherence (motion strength) present on the stimulus (Fig. 1.4B). Difficult trials implied longer RTs and a smoother slope in the build-up activity as opposed to easy trials. The time course of the activation of LIP neurons suggests that sensory evidence is integrated before a decision is made. When the accumulation of evidence reaches a decision bound, the selection is made.

In the recent years, several variations of the RDM task have been proposed. On the one hand, Churchland et al. (2008) and Niwa and Ditterich (2008) augmented the difficulty of the task by increasing the number of possible motion directions. In particular, Churchland et al. (2008) used both a two-choice RDM task and a four-choice RDM task with the aim to compare behavioral and neural results between them and to generalize the observations of binary decision making to decision making with multiple alternatives that are more similar to what a living organism generally faces. Similar to what was observed in the binary RDM task, the behavior of monkeys directly depended on the motion strength in terms of speed-accuracy, and neurons in LIP exhibited a similar modulation of their activity.

Huk and Shadlen (2005) studied the temporal integration of visual evidence by perturbing the strength of motion during short periods of time (100ms) within a trial (motion pulses). Behavioral results showed that the motion pulses exerted an effect over the RTs and the accuracy of monkeys performing the task. Positive pulses (pulses that favored the net base direction of motion) increased the probability of the monkeys of selecting the correct target, whereas negative pulses (pulses that favored a direction opposite to the base direction of motion) had the opposite effect. Moreover,

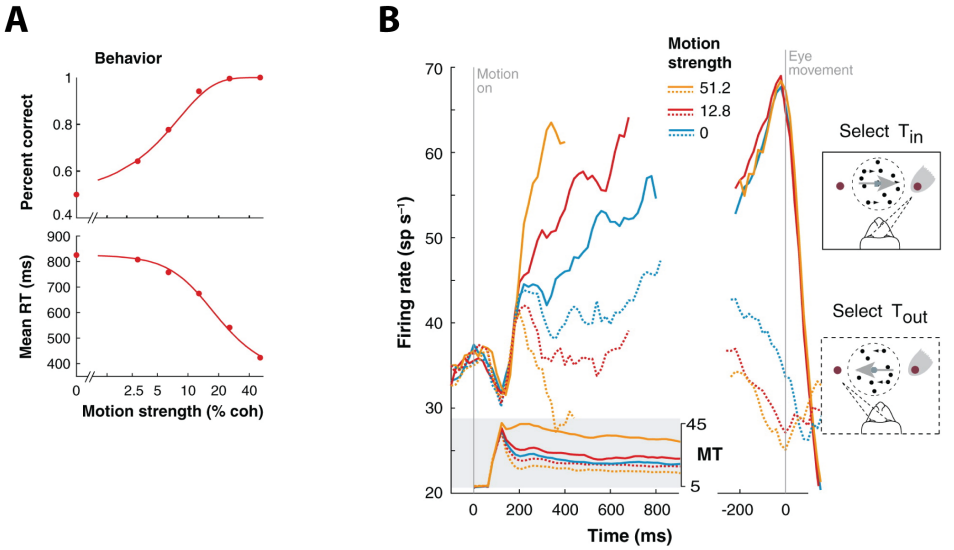

Figure 1.4: Behavioral and neural modulation in the RDM task. **(A)** Percentage of correct responses and mean RTs for one monkey along different values of motion coherence (strength). **(B)** Activity of neurons in LIP during decision formation. The plot shows the average firing rate of 54 neurons for three different values of motion coherence. *Left panel*, responses are aligned to the onset of the moving dots. Shaded box shows activity of neuron in MT for preferred (solid lines) and anti-preferred (dashed lines) directions. *Right panel*, responses are aligned to movement onset for same values of motion strength as in Left panel. Activity of LIP neurons resembles an integration of the difference between MT neurons response for preferred and anti-preferred directions. Adapted from Gold and Shadlen (2007)

monkeys responded faster in trials with positive pulses in comparison with trials with negative pulses. Neural activity in LIP showed a sustained biased due to the presence of the pulse, thus supporting the idea that neurons in LIP integrate sensory evidence. Interestingly, early pulses had a greater effect in performance and neural activity than late pulses.

A recent study by Thura et al. (2012) questioned the observed effect that early and late information seemed to exert on the decision process. The authors conducted an experiment that they called the variable coherence motion discrimination task (VMD). As it is the case with the RDM task, subjects are required to either detect or predict the current or final

direction of the majority of moving dots presented on a screen. The main difference with respect to the RDM task is that the strength of motion increases (positive pulse) or decreases (negative pulse) many times during the course of a trial. The results suggest that, in this case, subjects are influenced to a greater extent by late information than by early information, contrary to what was observed by Huk and Shadlen (2005). In a subsequent study where a task similar to the VMD was used (Thura and Cisek, 2014), neurons of monkeys recorded from the dorsal premotor cortex (PMd) and the primary motor cortex (M1) reflected the time course of sensory evidence. In Chapter 4, we propose a mechanism by which the two paradoxical results with early/late information can be explained.

1.4 Motor decision making

The implication of the motor cortex in decision making is of particular relevance. One of the reasons is that cognitive abilities were necessarily preceded by motor skills, and, consequently, they contributed to the development of the specific brain structure that we possess nowadays (Butler and Hodos, 1996; Redgrave et al., 1999). Motor decision making has been mainly studied by looking at two main issues: the ability to suppress a movement already planned and the influence that the intrinsic properties of our body have on the selection of movements.

One of the experimental paradigms most commonly used to study the ability to cancel a planned movement is the countermanding task (Logan and Cowan, 1984; Verbruggen and Logan, 2008). It investigates the behavior and neural correlates of movement suppression instructed by an infrequent Stop signal. The task consists of two kinds of trials: Go trials (Fig. 1.5A) and Stop trials (Fig. 1.5B). Both kinds of trials start with a cue signal (Go signal) that instructs subjects or monkeys to execute a movement. In some trials (generally a third of the total number of trials), after the Go signal, a Stop signal appears indicating that, in order to correctly perform the trial, the movement has to be cancelled. The performance during this task depends on the delay between the Go and Stop signals (Stop Signal

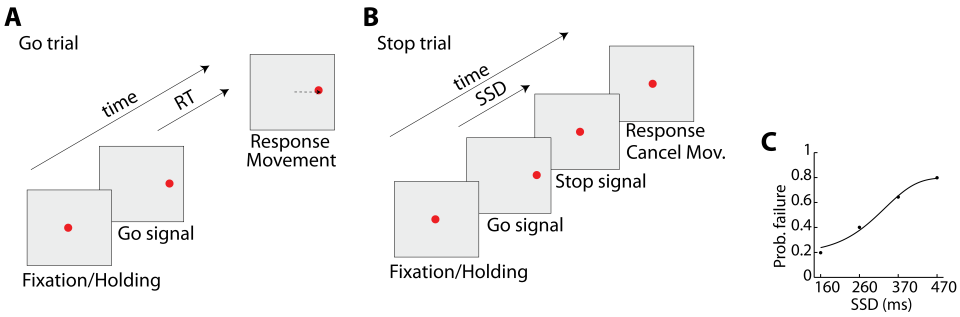

Figure 1.5: Illustration of a countermanding task. **(A)** and **(B)** Temporal sequence of Go (A) and Stop (B) trials. A trial begins with the presentation of a central stimulus. After a variable fixating/holding time (500-800 ms.) it disappears and, simultaneously, a target, i.e. Go signal, appears on one side of the screen. In Go trials, the human subject/monkey has to accelerate the saccade/finger movement in order to reach the target. In a small fraction of trials (commonly 33%), the central visual stimulus reappears after a Stop Signal Delay (SSD). In these trials, the monkey has to stop the planned movement and continue fixating/holding the central stimulus. **(C)** Illustrative example of the probability of failure for different values of SSD.

Delay, SSD). Fig. 1.5C shows an illustration of the probability of failure generally observed during the performance of the task for different values of SSDs. As perceived, the probability of failing to cancel the movement is higher for longer values of SSDs than for shorter values. Neural data recorded in a countermanding task requiring saccade movements revealed the involvement of the frontal eye field (FEF) (Hanes et al., 1998) and the superior colliculus (SC) (Paré and Hanes, 2003) in the control of saccade cancellation. In the arm reaching version of the task, neural recordings have been made from the supplementary motor area (SMA) and pre-SMA of monkeys (Scangos and Stuphorn, 2010; Chen et al., 2010). Activity of neurons in these regions did not seem to control movement initiation but could contribute to movement cancellation. An investigation of neural recordings from the dorsal premotor cortex (PMd) by Mirabella et al. (2011) showed a modulation in the activity of reaching related neurons that might be related to the suppression of planned movements. Neurons exhibiting this modulation were classified into two types (Fig. 1.6). The most common

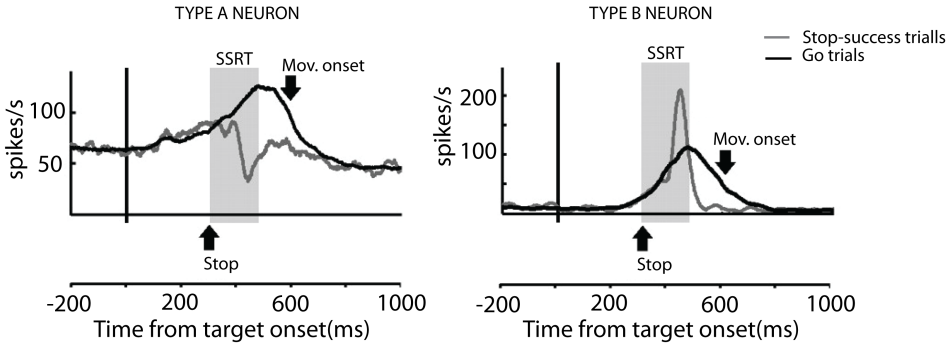

Figure 1.6: Example of activity of two neurons (Type A and Type B neurons) recorded from PMd during a countermanding task. The activity is aligned to the target onset (Go-signal). Shaded areas indicate the estimated SSRT period. Both neurons increase their firing rate after the presentation of the Go-signal and behave similarly during Go trials. Neuron Type A decreases its firing rate after the appearance of the Stop signal, whereas neuron Type B increases its activity. Adapted from Mirabella et al. (2011)

neurons were called type A neurons, and they showed a decrease in their activity in correct Stop trials before the end of the estimate of movement cancellation (stop signal reaction time, SSRT) with respect to the activity recorded during Go trials. Neurons classified as type B neurons exhibited a temporary increase of their activity when compared to their response during Go trials. Type A neurons followed a similar pattern of activation to those observed in FEF and SC (Hanes et al., 1998; Paré and Hanes, 2003).

Recent studies have revealed that behavioral performance during the countermanding task does not only depend on the SSD in a current trial, but also on the previous sequence of trials that has been experienced (Rieger and Gauggel, 1999; Mirabella et al., 2006; Emeric et al., 2007; Verbruggen and Logan, 2008; Nelson et al., 2010). In general, reaction times tend to be longer and the probability of properly suppressing the planned movement higher when many Stop trials have recently been experienced than when many Go trials have been encountered. In Chapter 3, we use the neural data recorded by Mirabella et al. (2011) to investigate the response modulation

of neurons in PMd caused by trial history and its correlation with the behavioral outcome.

Other studies have looked at the influence of factors such as arm biomechanics in decisions involving a motor response. Sabes and Jordan (1997) and Sabes et al. (1998) provided evidence that arm morphology and its impedance parameters are taken into account to plan movements around obstacles. However, the question of whether these factors are considered before or during movement execution remained unclear. Recently, Cos et al. (2011) and Cos et al. (2012) extended this study to investigate the issue. The investigation reveals that human subjects are able to predict the cost of potential movements and use this information to bias their movement choices (Cos et al., 2011, 2012). Specifically, when subjects are asked to freely choose between two potential actions, they are more likely to select the one that requires less biomechanical cost. Fig. 1.7 shows an example that illustrates two kinds of trials: Transverse, T1-Major and T1-Minor trials. The biomechanical cost of each action required to reach T1 and T2 targets mainly depends on the alignment of the trajectory with the axes of the ellipse of movement (minor or major), being less costly a movement with its trajectory aligned with the major axis of the ellipse than with the minor one. Therefore, in this case, subjects select significantly more often T1 when the targets are aligned following the T1-Major configuration than when they are arranged in the T1-Minor configuration. These two trials are the basis for our experimental paradigm, which is described and used in Chapter 5, and where we study the influence of motor cost in perceptual decision making.

1.5 Decision-making models

Sensory-motor decision making is thought to be a form of statistical inference (Rao, 1999; Tenenbaum and Griffiths, 2001). Signal detection theory (SDT) provides a well-established formalism to study perception. It classifies a single observation of sensory noisy data into a categorical choice (Gold and Shadlen, 2007). Decisions are the result of a competition be-

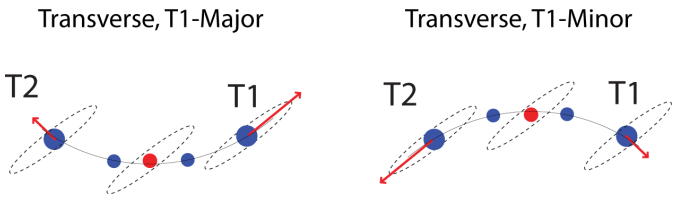

Figure 1.7: Free motor choice task (Cos et al., 2011). Example of target arrangement in two different trials. The trajectory to reach a target requires less biomechanical cost when its end is aligned with the major axis of the ellipse of movement (T1 in Transverse T1-Major and T2 in transverse T1-Minor).

tween hypotheses that need to be inferred from sensory noisy data, e. g. h_1, h_2,... h_n, with n=2 for binary decision making. Information favoring each hypothesis is called "evidence" (e). Within the context of conditional probability ("likelihood"), $P(e|h_i)$ is the probability of observing e given that h_i is true. Because the sensory information is noisy, the value of e is obtained from a distribution with mean h_i and variance σ_i. According to SDT, in order to make a decision, a decision variable (DV) has to be constructed from e. For binary decisions, DV is defined as the ratio of the likelihood (LR_{12}) for each option as:

$$DV(e) = LR_{12} = \frac{P(e|h_1)}{P(e|h_2)} \tag{1.1}$$

Following Bayes' theorem, the conditional probability $P(e|h_i)$ can be described as:

$$P(e|h_i) = \frac{P(h_i|e)P(e)}{P(h_i)} \tag{1.2}$$

where $P(h_i)$ is the prior probability of h_i, $p(e)$ is the total probability of e and $P(h_i|e)$ is the probability that h_i is true given e. SDT states that a decision is made when a criterion is satisfied, for instance, h_1 is selected when $DV(e) \geq 1$ when the prior probabilities of h_1 and h_2 are the same. On the contrary, if the prior probabilities of the alternatives are not the same, but the reward is the same, then the criterion to select h_1 to maximize

accuracy is $DV(e) \geq \frac{P(h_2)}{P(h_1)}$. Further mathematical analyses of this criterion are reported in Green and Swets (1966).

The sequential analysis (SA) extends the SDT to the case of multiple samples of evidence that are available over time (like, for instance, in the RDM task). By assuming that the samples e_1, e_2,... e_m are independent, the "likelihood" ratio can be calculated as:

$$DV(e) = logLR_{12} = log\frac{P(e_e, e_2, ...e_m|h_1)}{P(e_1, e_2, ...e_m|h_2)} = \sum_{i=1}^{m} log\frac{P(e_i|h_1)}{P(e_i|h_2)} \tag{1.3}$$

DV(e) is updated with each new sample. The sequential probability ratio test (SPRT) proposes that the decision process ends when DV(e) is greater than a specific positive or negative decision bound (Laming, 1968; Stone, 1960; Wald, 1947). The SPRT solves the problem of optimality in some cases, such as ensures the shortest time for a given success rate or the best success rate for a given time (Wald, 1947).

Following the SPRT procedure, models called "sequential sampling" or "integrator" models have been widely accepted and employed to analyze the behavioral and neural data observed in previous research (Stone, 1960; Laming, 1968; Ratcliff, 1978; Usher and McClelland, 2001; Mazurek et al., 2003; Bogacz and Gurney, 2007; Link and Heath, 1975). Although many different integrator models have been proposed, all of them generally follow the same dynamics:

$$x_i(t) = g\int_{t_0}^{t_f} e_i(t)dt \tag{1.4}$$

where x describes the state of the process, g is the gain of the integration process and e is the sensory evidence provided to make a decision. The difference between variations of decision-making models lies on the way they define these three variables. For instance, the "drift-diffusion" model (Stone, 1960; Laming, 1968; Ratcliff, 1978; Ratcliff et al., 2003; Smith and

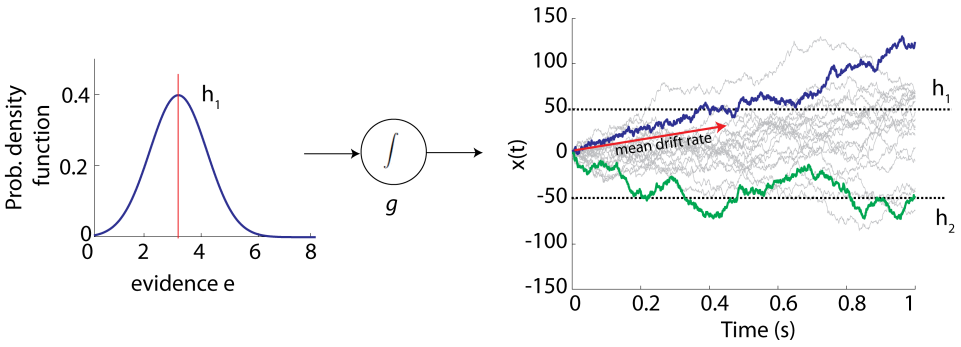

Figure 1.8: Simulation of the drift-diffusion model, using 20 trials. For this example, we used the parameters: $g = 0.01$, a decision bound of 50 (for option h_1) and -50 (for option h_2), and a Gaussian noise with mean 0 and standard deviation of 1 added to the sensory evidence. *Left panel*, probability density function of the distribution from which the sensory evidence (e) is sampled. *Right panel*, accumulated evidence over time. When infinitesimal temporal steps are used to sample the evidence and accumulate it, the process is called drift diffusion. One correct trial is marked in blue and one incorrect trial is marked in green.

Ratcliff, 2004) suggests that there is only one variable $x(t)$ that accumulates sensory evidence e in favor or against each alternative. An example of this model is illustrated in Fig. 1.8. The decision-making process is considered to be terminated when the accumulated process reaches a predefined decision bound. Another variation of this model considers that two independent variables $x_1(t)$ and $x_2(t)$ accumulate sensory evidence separately and that the decision is made when the difference between the two variables reaches a decision bound. Yet another variation of an integrator model is called the "leaky competing accumulator" model (Usher and McClelland, 2001), which proposes that evidence is also independently accumulated by using two variables that mutually inhibit each other and that contain a leaky term so that information is "forgotten" over time.

The "urgency" models (Ditterich, 2006) can be considered as alternatives to the integrator models. These models propose that there is also a build-up activity that governs the decision formation, though the increase of activity is actually due to a multiplication between sensory evidence and an urgency signal that grows over time as follows:

$$x_i(t) = ge_i(t)u(t) \tag{1.5}$$

where $x_i(t)$, g and $e_i(t)$ represent the same as in the integrator models and $u(t)$ is a function of time that is not related to the sensory evidence but to some internal urgency (timer) to make a decision (Ditterich, 2006; Churchland et al., 2008). Although integrator and urgency models propose different mechanisms by which decisions are made both theories lead to equivalent predictions when the sensory evidence is constant over time (Cisek et al., 2009). However, recent studies with variable sensory evidence (Huk and Shadlen, 2005; Thura et al., 2012) have shown that the two branch of models make clearly different predictions when this is not the case. For instance, in the VMD task introduced in Section 1.3, the drift-diffusion model predicts an influence of early information in reaction time, as opposed to a prediction of no influence by urgency models. Therefore, urgency models could properly explain the experimental data in this case, but conversely, they failed to do so with the RDM task with pulses (Huk and Shadlen, 2005) (see also Section 1.3). We will further elaborate on this issue in chapter 4, where we propose an alternative model to explain both: tasks with constant and with variable sensory evidence.

The "independent race" model (Logan and Cowan, 1984) is also an integrator model. It has mainly been used to explain the countermanding task stating that, to decide to either move or cancel a movement, there is a race between two accumulator variables ($x_{go}(t)$ and $x_{stop}(t)$) that integrate evidence in favor of each alternative in an independent way. When one of the two process, which are called *Go* and *Stop*, reaches a threshold, the decision is made (Fig. 1.9). Boucher et al. (2007) extended this model to account for neurophysiological studies (Munoz and Schall, 2003) asserting that the neural systems that control movements are formed by layers of inhibitory interactions between neurons that, in turn, are responsible for movement initiation and movement inhibition. In this version of the race model, which is called the "interactive race" model, there is a competition between the *Go*

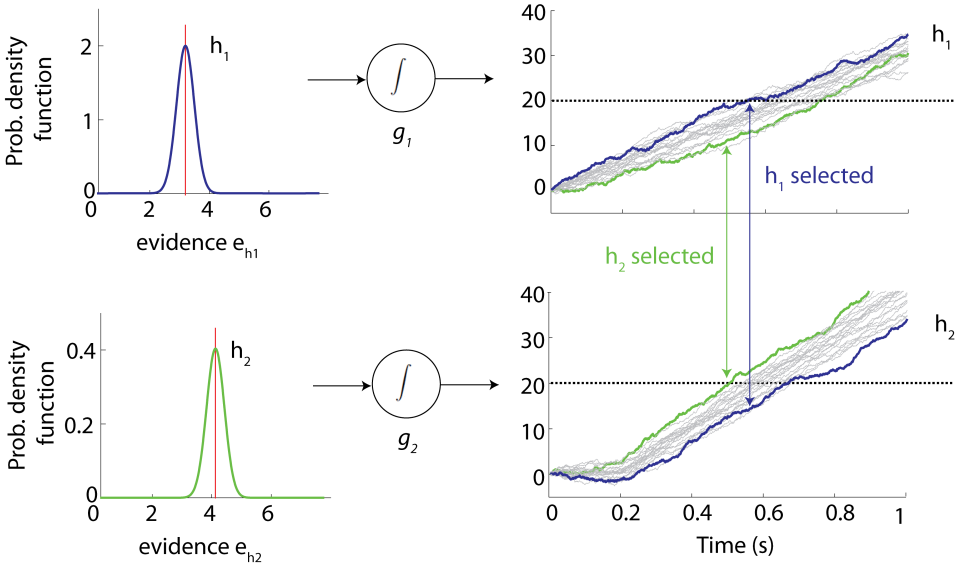

Figure 1.9: Simulation of the race model, using 20 trials. For this example, we used the parameters $g_1 = 0.01$, $g_2 = 0.015$ and Gaussian noise added to the sensory evidence with mean 0 and standard deviation of 0.2. *Left panel*, probability density function of the distribution from which the sensory evidences (e_{h_1} and e_{h_2}) are sampled. *Right panel*, two independent processes accumulate evidence in favor of h_1 or h_2 over time. Whenever one of the two processes reaches a threshold, the decision is made towards the hypothesis associated with that process. One trial in which h_1 was selected is marked in blue and one in which h_2 was selected is marked in green.

and *Stop* processes and a leaky term within each that prevents the activity to grow boundless.

A biophysically constrained spiking neural model that successfully explains behavioral and neural data in the RDM task was proposed by Wang (2002). It is based on attractor dynamics and a competition between two neural populations representing each alternative option. The advantage of this model over the integrator and urgency models lies in its constraints of parameter values and dynamics, which provide a very specific network to make accurate neural predictions and to explore the cellular and circuit mechanisms that result in long integration times along the decision process. This model is the basis for our theoretical work in Chapter 5.

However, the large number of neurons, interaction and nonlinear dynamics make the spiking neural model a difficult network to analyze and understand. Consequently, mean-field models have been proposed as an alternative to the high dimensional spiking neural model when the interest of the study is mainly focused on the mean response of a population of neurons and on the behavioral outcome of the network. Wong and Wang (2006) created a reduced version of the spiking neural network of Wang (2002) with only two dynamical variables able to reproduce most of the behaviors of the original spiking neural network. In Chapter 3, we used a well established and reduced mean-field rate model based on the Wilson and Cowan equations (Wilson and Cowan, 1972). This version cuts down the complexity of a spiking neural model to two differential equations. It differs from the mean-field approach proposed by Wong and Wang (2006) mainly in that it does not account for synaptic dynamics, which in our case could be simplified.

1.6 Role of variance in decision making

A method frequently used in neuroscience to characterize the neural mechanisms that cause specific activations in the brain consists in repeatedly activating it with same stimuli, so that many observations are obtained out of the same process. Since behavior is stochastic, even if the repetition is exactly the same, varying responses are obtained. To lessen the problem, the approach most commonly employed is to calculate the average response from the recorded neural activity. With this approach, the neural response variability is considered to be just noise and non informative about the underlying neural mechanisms. Although this might be true in some occasions, it can also be informative in others. To give an example, the neural response might become more or less variable before or after the presentation of a stimulus, and this might suggest a signature of additional processes not explained by the mean neural response.

In recent research, it has become increasingly more common to expand the study of neural correlates to across-trial variability, in addition to the

mean firing rate. In particular, it has been of great interest to investigate whether the measured mean response of neurons comes from a narrow (low variance) or wide (high variance) distribution of firing rates. Two measures of across-trial response variability as a function of time have been commonly applied to calculate the spike count variance of the neural data: the Fano Factor (FF) and the Variance of the Conditional Expectation (VarCE). FF and VarCE seem to be qualitatively consistent and to give, in most cases, equivalent results (Churchland et al., 2011). The fundamental advantage of VarCE over FF is that it is principled and, therefore, more robust than FF.

Both measurements of variability intend to remove any variance caused by the neuron's intrinsic variability in spike generation and, in that way, completely isolate the variance attributed to the underlying firing rate on each trial. The FF achieves this by dividing the total calculated variance of each neuron by its mean firing rate. The scientific concept behind this approach is that the generation of spikes by cortical neurons follows a Poisson like process and thus, its variance scales linearly with the mean. If the variance is totally attributable to the variability of the spike generation, the FF would be 1; in accordance, any value of FF greater than 1 means that there is some variability in the underlying mean firing rate that is not caused by the Poisson process. An analysis of many recorded datasets by Churchland et al. (2010) revealed a stimulus driven decrease in FF in the across-trial response of cortical neurons, despite the definite stimulus or behavioral state. The authors refer to this observation as a general property of the cortex that is non-specific to the task.

In the case of VarCE, the spike variance caused by the Poisson process is estimated and subtracted from the total calculated variance as follows:

$$VarCE = Var[N_i] - \phi\overline{N_i} \tag{1.6}$$

where a scaled value (ϕ) of the mean spike count represents the approximated Poisson variance. ϕ represents the minimum FF of a neuron to ensure that VarCE never gets negative values (see Chapter 3 for details). Churchland et al. (2011) applied this measurement to neurons recorded in

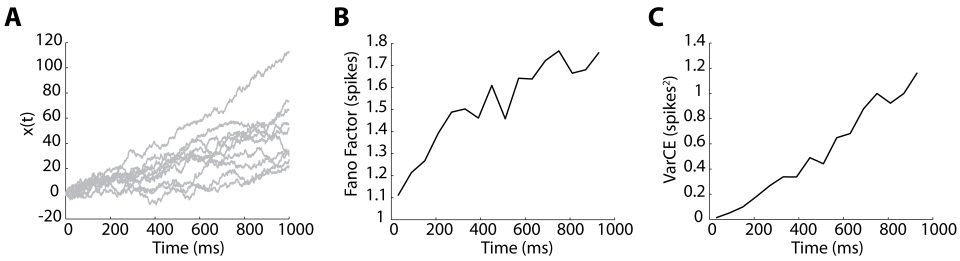

Figure 1.10: Fano Factor and VarcCE in the drift-diffusion model. **(A)** Few trials simulating a Drift-diffusion model with $g = 0.08$, mean $e_1 - e_2 = 3.2$ and an added Gaussian noise with a mean of 0 and a standard deviation of 8. **(B)** Fano Factor and **(C)** VarCE aligned to stimulus onset and calculated over 1,000 trials using a temporal window of 60ms. Spikes are generated following a Poisson process from the mean rate obtained from the model.

LIP while monkeys performed a two and four-choice RDM task with saccade responses. VarCE was shown to increase during decision formation and to decline at the end of the decision process. Therefore, VarCE provides a signature of decision formation. In general, decision-making models exhibited the same pattern of across-trial response variability (Churchland et al., 2011), thus proving their validity to describe neural responses.

An example of FF and VarCE is illustrated in Fig. 1.10. Both FF (1.10B) and VarCE (1.10C) show and increase of their value after the stimulus onset showing a signature of decision formation, as it was observed in monkeys data (Churchland et al., 2011). Qualitatively, both measurements show the same effect, they only differ in their specific values. This is due to the method employed to remove the variability generated by the Poisson process, i. e. a neuron presenting variability solely because of its Poisson spike generation would have a FF of 1, but a VarCE of 0. Values of FF greater than 1 mean that there is an across-trial variability resulting from a process which is different from that generating spikes. The same is true for values of VarCE greater than 0. In our Fig. 1.10 example, FF and VarCE show an across-trial variability that is independent from the Poisson process.

It is commonly believed that across-trial variability originates from

two sources of noise: internal and external noise. The internal noise relates, among others, to the noise associated with neurons or sensory areas (Faisal et al., 2008). The external noise refers to the variability of the external world. Besides, a third cause of across-trial variability has been lately attributed to the deterministic approximations in the complex computations performed by the nervous system (Beck et al., 2012). Anyhow, even if the reasons for neural response variability are not well settled, it seems that its correlation with behavioral variability has been highly proven (Osborne et al., 2005), hence becoming a great focus of attention.

As we have seen, across-trial variability can be useful to describe neural processes more thoroughly. However, since it is a measure obtained from a large number of trials, the question of how this information might be used by the brain in a single trial -if this were the case- remains unclear. One of the possible solutions to overcome this limitation would be to analyze the neural data from a population of neurons recorded simultaneously. Nevertheless, this is normally not the case since most of the recordings are done over single units. In this thesis, we contribute to a better understanding of the role of variance in decision making in two ways. First, in Chapter 3, we investigate the possible signature that trial history could have in the across-trial variability of single unit recordings of neurons from dorsal premotor cortex during a decision-making task. Later on, in Chapter 5, we use a spiking neural model to suggest a mechanism by which response variance of a population of neurons might be used by other areas of the brain to form decisions. With this second study, we overcome the problem of multiple unit recordings through a realist decision-making model.

1.7 Embodied decision making

The central idea of embodied decision making is that actions -or, in general terms, the motor apparatus- influence decision making even when the decision itself does not depend on them. Most studies on decision making ignore this issue and focus exclusively on specific features of a task, such as the perceptual discrimination tasks introduced above (Section 1.3), treat-

ing them in isolation and completely neglecting the possible effect that the action to report the decision might have over the process itself. Most of the models introduced above (see Section 1.5) consider decision making a serial process: firstly, the perceptual discrimination is made, and then the action is selected. Classical views on cognitive functions (Fodor, 1983; Pylyshyn, 1984) support the idea that cognitive processes and sensorimotor control are functionally segregated. However, recent neurophysiological and behavioral studies have shown growing evidence that areas involved in decision (at least, when the decision is reported by an action) are also part of the planning and execution of the subsequent action (Cisek and Kalaska, 2005; Gold and Shadlen, 2007; Pesaran et al., 2008; Hernandez et al., 2002; Romo et al., 2002, 2004). For instance, Wallis and Miller (2003) demonstrated that, when monkeys were asked to release or hold a lever based on the match or non-match of a sequence of visual stimuli, neurons in the premotor regions showed stronger and earlier correlation with the behavioral rule than those in the prefrontal cortex. Similarly, neurons in the LIP (Dorris and Glimcher, 2004; Gold and Shadlen, 2007; Yang and Shadlen, 2007), the FEF (Coe et al., 2002) and the SC (Carello and Krauzlis, 2004; Horwitz et al., 2004; Thevarajah et al., 2009) seemed to be correlated with both the formation of the decision and the execution of eye movements. Analogous functional organization has also been found in the motor system (Hernandez et al., 2002; Romo et al., 2002, 2004). In short, these results suggest that both processes, perceptual decision making and action selection, do not necessarily occur in a serial manner but in an integrative loop (Pezzulo et al., 2011; Barca and Pezzulo, 2012).

Contrary to what happens in most laboratory experiments, it is highly unlikely to find a situation in which perception and action are detached in the real world. The general practice of simplifying the problem by using experimental paradigms (as introduced in Section 1.3 and 1.4) that are highly controlled in the lab has the advantage of providing an isolation of one specific problem in such a manner that it can be deeply analyzed and understood in detail. Nonetheless, even if these studies are the basis for further general cases, we cannot ignore that, in order to solve more com-

plex problems, the different isolated elements need to be integrated. Even so, the experimental cases that would allow for the study of this integration are, in most of the cases, difficult to control and therefore experimentally intractable. Embodied systems have been used to bypass this limitation and have been very useful to provide a theoretical framework to make predictions, and to quantify behavior and neural activity in real-world situations. To cite an example, Verschure et al. (2003) used a mobile robot to examine the interaction between behavior and perception, and showed that perceptual learning progressively structures behavior. This shaping of behavior causes a bias in the sampling of sensory inputs that eventually results in a macroscopic feedback loop. The cognitive architecture that enabled researchers to investigate these dynamics is the Distributed Adaptive Control (DAC) architecture (Verschure et al., 2003; Duff et al., 2010; Marcos et al., 2013c). In this thesis, we use this architecture to study the generalization of an integrator decision-making model when an embodied system is used in foraging tasks.

DAC is a multi-layered architecture that has been examined by means of formal approaches (Verschure and Coolen, 1991) and robots (Verschure et al., 1993, 2003), and has proven to be a suitable framework to study biological systems (Verschure and Althaus, 2003; Verschure et al., 2014). To properly solve a task, it needs to acquire information from the environment and select the proper actions based on perceptual evidence. Both perceptual decision making and action selection processes interact and influence each other, following the principle of embodied decision making (Verschure et al., 2003).

DAC is a robot based neuronal model of classical and operant conditioning (Verschure and Coolen, 1991; Verschure and Althaus, 2003). Classical conditioning is a form of associative learning (Pavlov, 1927) where the presentation of a neutral stimulus (conditioned stimulus CS) together with a significant stimulus (unconditioned stimulus, US) leads to an association of the initially neutral stimulus to a, so called, conditioned response (CR). In one of the interpretations, the CS substitutes the US because the behav-

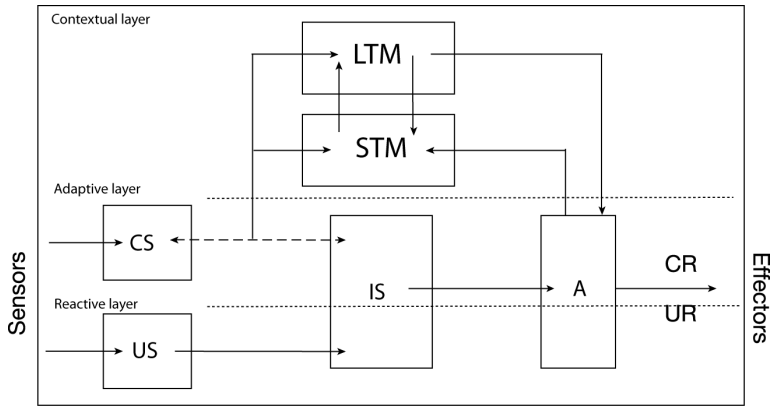

Figure 1.11: Schematic representation of the DAC architecture. It is based on the assumption that behavior results from three tightly coupled layers of control: reactive, adaptive and contextual. Squared boxes stand for neuronal groups; arrows stand for static (solid) and adaptive (dashed) synaptic connections between cell groups. Abbreviations mean: *US*, unconditioned stimulus; *CS*, conditioned stimulus; *IS* , internal states; *A*, action group; *UR*, unconditioned response; *CR*, conditioned response; STM, long-term memory; LTM, long-term memory

iorally significant stimulus (*US*) triggers an unconditioned response as an innate automatic response that forms a template for the CR. If *CS* and *US* are repeatedly paired, the two stimuli become associated and the organism begins to behaviorally respond to the presence of *CS* alone. One typical paradigm is eyeblink conditioning, where an air-puff (*US*)-after which inevitably the animal reacts with an eyeblink response- is paired with the presentation of a tone (*CS*). After a number of trials, the animal begins to react to the *CS* with a *CR* similar to the *UR*, even if the air-puff (*US*) is not present anymore (Mackintosh, 1990). Operant, or instrumental, conditioning is also a form of associative learning. However, the association is not always as direct as it is in classical conditioning, since series of actions are needed to reach a reward or punishment (Thorndike, 1911). These actions are weighted with different values depending on the *US* resulted from an action -i.e. with an appetitive or aversive US-, so the ones that leaded to a reward will occur much more frequently that the ones that were paired with punishment. DAC proposes that Classical and Operant Conditioning reveal

a fundamental scaffolding of learning that advances through three stages. First, sensor statistics based on *p*erceptual learning provide a 'neutral' representation of the state space. Second, mechanisms underlying classical conditioning provide for a biasing of this state space representation with respect to its immediate survival value (construction of *CS* representations), plus the shaping of discrete actions (tuning of the amplitude time course of the *UR* to define the *CR*). Subsequently, operant conditioning builds on the representational building blocks provided by the preceding two stages to construct plans for actions, apart from settling a foundation for cognition and problem solving.

Chapter 2

Generalization of integrator models to foraging

A robot is made not born
Tony Stark

We begin by studying the generalization of a well-established accumulator model of decision making, the race model (see Section 1.5), to real-world foraging. To this purpose, we use an embodied artificial agent, equipped with a cognitive architecture (DAC), that is required to explore and exploit novel environments to reach specific goal positions on them. In Section 2.1, we investigate the implications of the race model in these tasks by comparing it with a non-accumulator model of decision making. Our results show a fundamental difference in the way that information is represented in memory and might be further retrieved from it. Specifically, the race model suggests that goals are stored in memory and the actions are self-generated by the agent when this information is retrieved from memory. In Section 2.2, we investigate the implications of this self-generation of actions by studying how reactive and contextual control systems complement each other. Actions triggered by the reactive control systems are egocentric whereas the ones triggered by the contextual control system are generated during recall from memory. We demonstrate that reactive behav-

ior is enough when a task is not very demanding, but contextual control is fundamental when task complexity increases.

2.1 The influence of the decision-making process in the representation of information in memory

This sections reproduces the paper "Generalization of integrator models to foraging: a robot study using the DAC9 model" published in the proceedings of the Living Machine conference (Marcos et al., 2012b). The abstract reads:

Experimental research on decision making has been mainly focused on binary perceptual tasks. The generally accepted models describing the decision process in these tasks are the integrator models. These models suggest that perceptual evidence is accumulated over time until a decision is made. Therefore, the final decision is based solely on recent perceptual information. In behaviorally more relevant tasks such as foraging, it is however probable, that the current choice also depends on previous experience. To understand the implications of considering previous experience in an integrator model we investigate it using a cognitive architecture (DAC9) with a robot performing foraging tasks. Compared to an instantaneous decision-making model we show that an integrator model improves performance and robustness to task complexity. Further we show that it compresses the information stored in memory. This result suggests a change in the way actions are retrieved from memory leading to self-generated actions.

Introduction

Binary perceptual tasks have been widely used to study the neural mechanisms underlying decision making (Gold and Shadlen, 2007; Smith and

Ratcliff, 2004). This kind of task involve a simple decision about a feature of a stimulus that is expressed as a choice between two alternative options. Many models have been proposed to explain this decision-making process predicting the relationship between reaction time and accuracy (Logan and Cowan, 1984; Ratcliff and Rouder, 1998). Most of them explain decision making as an accumulation process that takes place over time until a decision bound is reached. These models are known as integrator models and have been generally accepted as an explanation for decision making in perceptual tasks where learning is not required to successfully perform the task. Here, we investigated the interaction between an integrator model and memory in foraging tasks using a well establish cognitive architecture as a framework (Verschure et al., 2003).

One largely used perceptual experimental paradigm is defined by a random-dot motion (RDM) task where humans or monkeys have to select between two possible stimulus categories, such as leftward or rightward motion (Shadlen and Newsome, 2001). Integrator models, such as race models (Logan and Cowan, 1984) and drift-diffusion models (Ratcliff and Rouder, 1998), provide a straightforward account of the speed-accuracy trade-off. These models suggest that evidence is accumulated over time until this accumulation reaches a bound, i. e. criterion level, and a decision is made. As the RDM task, many of the experimental paradigms used to study the decision-making process are simple perceptual tasks where the correct performance of a trial depends exclusively on the current perceptual information, e. g. color. The proposed integrator models assume that the alternative options are known a priori and therefore learning during the task is not required. However, this would not be the case in more realistic foraging tasks where the information about different targets have to be acquired from the environment and many alternative choices might be available at each decision point. Therefore, a two-fold problem has to be solved during foraging: the appropriate learning of the environment and actions and the appropriate retrieval of information to achieve goal states (targets), i. e. sequences of perception and action need to be learned and retrieved to reach goal positions in an environment.

To study the interaction between decision making and memory, we worked in the framework of the Distributed Adaptive Control (DAC) architecture (Verschure and Althaus, 2003; Verschure et al., 2003). The decision making in DAC follows the Bayesian principle (Bayes, 1763; Verschure, 2012). We extended the architecture with an integrator decision-making model (DAC9; see (Verschure and Althaus, 2003; Duff and Verschure, 2010; Mathews et al., 2012) for details about previous versions of DAC), based on the race model, to investigate its implications during foraging tasks and we further compared it with the decision making in DAC (instantaneous model). We show that the integrator model resulted in a new mechanism of storing and recalling information from memory suggesting that the actions are not stored in memory but self-generated during retrieval of information. In a previous study (Marcos et al., 2010a), we assessed the impact of these two decision-making models in the learning of event order and interval in a sequence in two foraging tasks. In the current study, we go one step further quantifying (1) the scalability of the two models with task complexity in five different foraging tasks and (2) the implications on the information stored in memory and proposing (3) a new working memory mechanism that accounts for a continuous action space.

Materials and Methods

Cognitive Architecture

The DAC architecture has already proven its suitability to study the problems encountered in biology helping to investigate perception, cognition and behavior in foraging situations in which the access to real neuronal and behavioural data is difficult (Verschure et al., 2003). DAC is based on the assumption that learning consists of the interaction of three layers of control: reactive, adaptive and contextual, as illustrated in Fig. 1.11. The reactive layer provides pre-wired responses that allows for a simple interaction with the environment and accomplish simple automatic behaviours. The adaptive layer provides mechanism for the classification of the sensory events (internal representations) and the shaping of responses in simple tasks as in

classical conditioning (Pavlov, 1927). The internal representations (prototypes; see (Duff and Verschure, 2010) for details) generated by the adaptive layer are stored in the contextual layer as couplets of sensory-motor states and used to plan future behaviour, as in operant conditioning (Thorndike, 1911).

In this study, we mainly focused on the contextual layer of DAC that provides mechanisms for memorizing and recalling information. It consists of two memory structures: the short-term memory (STM) and the long-term memory (LTM), for permanent storage of information. During learning, pairs of prototype-action are stored in the STM as the robot interacts with the environment. When a goal state is reached, i. e. reward or punishment, the content of the STM is copied into the LTM and the STM is reset. The LTM has sequences of pairs of prototype-action that lead the robot to goal states. The prototype-action pairs that form a sequence are called segments. During the recall process, the prototypes stored in LTM are matched against the generated prototypes from ongoing sensory events. The degree of matching of segment l in sequence q determines the input to its, so called, *collector* unit, c_{lq}:

$$c_{lq} = (1 - d(e, e_{lq}))t_{lq} \tag{2.1}$$

where $d(e, e_{lq})$ is calculated as the Euclidean distance between stored *prototype* e_{lq} and current *prototype* e and t_{lq} is called *trigger*. The trigger value biases the sensory matching process of the segments and allows chaining through a sequence, i. e. its default value is 1 and it is set to a higher value if the previous segment $l - 1$ is activated.

The activity of the collectors contribute to the action proposed by the contextual layer. We only consider the collectors' activity that satisfy both conditions: (1) its activity is above a certain threshold (θ^C), (2) its activity is inside a predefined percentage range from the maximum collector's activity, i.e. the collectors compete in an E%-Max Winner Take All (WTA)

mechanism (de Almeida et al., 2009). The actual action proposed by the contextual layer (a_c) is calculated as:

$$a_c = \sum_{l,q \in LTM} \pm \frac{c_{lq} H(c_{lq} - \theta^C)}{\delta_{lq}} a_{lq} \tag{2.2}$$

where $H(.)$ is a step function that is 0 for values lower than θ^C and is 1 for values higher than θ^C, δ_{lq} is the distance measured in segments between the selected segment l and the last segment in the sequence, i.e. the distance to the goal state and a_{lq} is the action stored in segment l of sequence q. By doing this division the segments closer to the goal state have more impact on the contextual action. The sign is positive if the segment belongs to an appetitive sequence and negative if it belongs to an aversive sequence.

The actions triggered by each of the three different layers are filtered by priority, giving more priority to reactive actions (a_r), then to contextual actions (a_c) and finally to adaptive actions (a_a). The one that takes the control of the motor is stored in STM and afterwards in LTM (see Supplementary Material and Methods for further details).

Integrator models

Many integrator models have been proposed, but mainly, in all of them, the change in the accumulation of evidence in favor of one alternative ($x_i(t)$) can be described as:

$$\frac{dx_i}{dt} = \mu E_i(t) + \xi \tag{2.3}$$

where μ is the growth rate of the accumulation, $E_i(t)$ is the internal estimate of evidence at time t and ξ is a Gaussian noise with mean of zero and variance of σ^2. The proposed models consider the variables $x_i(t)$, $E_i(t)$, μ and ξ in a different manner. We implement a rise-to-threshold model based on the race model. The race model (Logan and Cowan, 1984) suggests that there are separate variables $x_i(t)$ for each option that accumulate evidence

independently until one of them reaches a decision bound and a decision is made.

Our implementation of the race model consisted of a number of independent variables that compete to take the control of the robot. Each variable accumulated evidence in favor of one action, such as right or left. When the value of a variable grew above a criterion level, i. e. decision bound, the action associated with it was performed by the robot. The change in the activity of the variables within a time step dt was defined as:

$$da_i(t) = \begin{cases} dt(\mu_r a_{r_i} + \mu_a a_{a_i} + \mu_c a_{c_i} + \xi) & \text{, if } t - t_{la} > T_{ref} \\ 0 & \text{, if } t - t_{la} \leq T_{ref} \end{cases} \tag{2.4}$$

where a_r, a_a, a_c are the actions triggered by the reactive, adaptive and contextual layer respectively, $i \, \epsilon \, \mathbb{N}^N$ and it is the subindex of the N different possible actions, μ_r, μ_a, μ_c are the mean growth rates of the variables units, ξ is a Gaussian noise term with a mean of zero and a variance of σ^2, t_{la} is the time at which the last action was executed and T_{ref} is the refractory period. In our experiments $dt = 1ms$ and $\xi = 0$. When the value of a_i reaches a predefined threshold the associated action is executed. In biology, the refractory period is the amount of time a excitable membrane needs to be ready for a second stimulus once it returns to the resting state. Consistent with this, the T_{ref} term referred to the amount of time necessary to start again the competition between actions after one of them was executed.

An illustration of the dynamics of the two deision-making models is shown in Fig. 2.1. As observed, in the case of the instantaneous model many actions might be executed for same visual input whereas in the integrator model only one is executed.

Foraging tasks

The mobile agent was simulated in C++ and wSim (Wyss, 2003) using the 3D Open Graphics Library approximating a Kephera robot [1]. Different

[1] K-Team, Lausanne, Switzerland

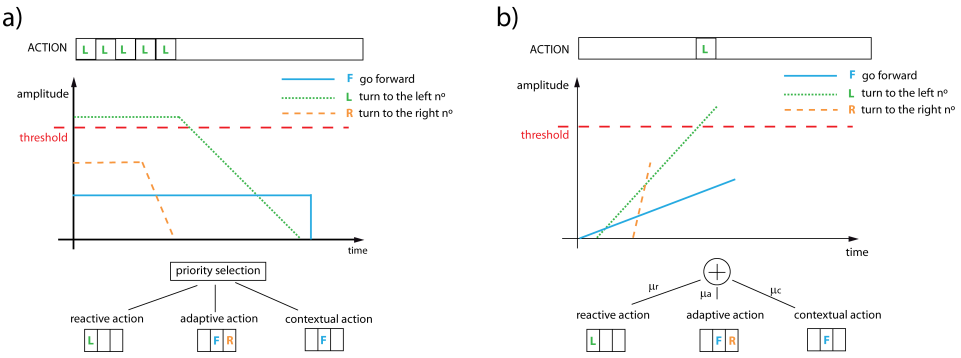

Figure 2.1: Decision-making dynamics. (a) Decision making in the instantaneous model. The different actions generated by the reactive, adaptive and contextual layers are filtered by priority as in a Winner Take All competition. Only one of them defines the final action. The robot performs an action during the interval that the selected action is above a predefined threshold. (b) Decision making in the integrator model. The different actions generated by the reactive, adaptive and contextual layer are continuous in time. When the sum of the actions over time reaches a predefined threshold an action is performed. The three layers contribute to the integration with different growth rates (μ_r, μ_a, μ_c). Time integration leads to only one action per sensory state triggered by a colored patch in the environment.

previous studies have proven the validity of this simulated robot with respect to a real one (Wyss, 2003; Wyss et al., 2006). The robot has a radio of 5.5 cm and 8 proximity sensors and 8 light sensors. The values captured by both light and proximity sensors decay exponentially. The proximity sensors measure the distance to obstacles while the light sensors measure the intensity of light sources. The robot is equipped with a color camera with a visual angle of 45 deg. of amplitude. The image from the camera is color separated such that there are three channels: red, green and blue, each of them with a resolution of 36x36 pixels. Except otherwise specified the camera is always pointing to the floor with a tilt angle of -60 deg. with respect to the horizontal axis. The robot translates with a speed of 0,1 $\times$ robot radius and it rotates with a speed of 10 deg.

To study the interaction between an integrator model and memory we defined a number of foraging tasks where not only perceptual but also memory information was essential to achieve a performance about chance.

The tasks had different rated complexity to assess how the decision-making models scaled to it (Fig. 2.2). In all the environments the goal of the task was to go to the light source, i. e. reward. Every trial started from one of the positions shown in Fig. 2.2, randomly selected. The trial ended when the robot hit the light or collided with the wall. A successful trial ended when the light was hit. The environments contained colored patches that served as cues. The light was detected by the light sensors of the robot. However, the light was not strong enough to trigger a reactive action from the side patches. The adaptive layer used reactive layer sub-threshold activity to generate the prototypes and to learn the associations between prototype-action. Once the prototypes were stable the contextual layer started storing sequences of prototype-action that leaded to a goal state. The goal state occurs when the robot reaches the light or collides with the wall. When a collision occurred it was stored as an undesirable state in memory and had a negative influence on the action proposed by the contextual layer.

The complexity of the tasks was rated taking into account the number of patches and how ambiguous they were as follows:

$$TC = \frac{n_p}{n_c} n_a \tag{2.5}$$

where n_p is the number of patches, n_c is the number of different colors and n_a is the number of different turning angle amplitudes needed to be learned. This measure was then useful to compare the robot performance in each of the tasks for the two proposed models. The complexity of the task 1 is 3, the complexity of the task 2 is 4.5 and 5, 7.5 and 11.7 for the tasks 3, 4 and 5, respectively.

The first task was an unambiguous restricted open arena foraging task, i.e. no context information was needed because the location of the target was uniquely predicted by the color patches (Fig. 2.2a). Therefore, this task could be correctly solved by the adaptive layer, but still we tested the performance at the contextual layer level. The rest of the four tasks consisted of ambiguous restricted open arenas, because, in all cases, con-

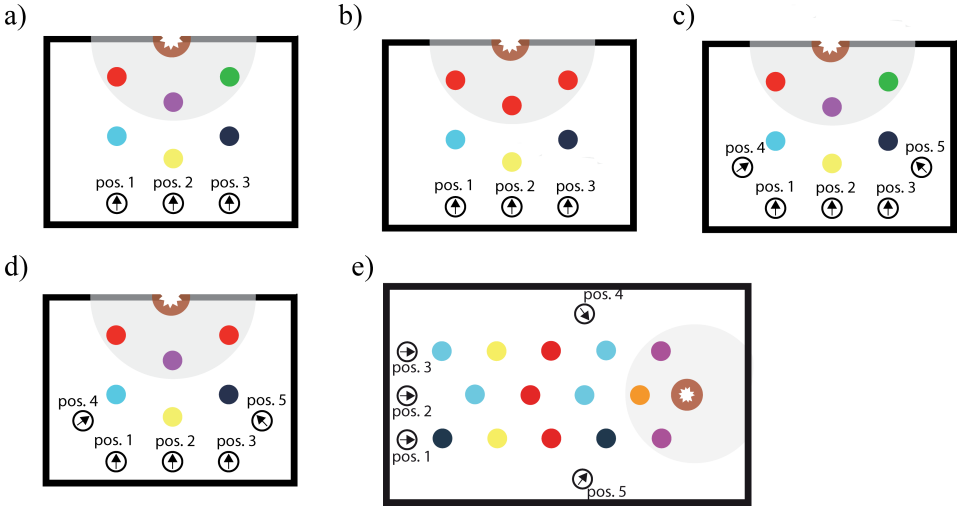

Figure 2.2: Foraging tasks ordered by task complexity. (a) Task 1. Unambiguous restricted open arena, $TC = 3$. (b) Task 2. Ambiguous restricted open arena, $TC = 4.5$. (c) Task 3. Ambiguous restricted open arena, $TC = 5$. (d) Task 4. Ambiguous restricted open arena, $TC = 7.5$. (e) Task 5. Ambiguous restricted open arena, $TC = 11.7$.

text information was needed to reach a performance above chance. The actions associated with the patches closest to the light were not unique but depended on the previous context (Fig. 2.2b, 2.2c, 2.2d and 2.2e) and therefore the problem could only be solved at the contextual layer level. These experiments allowed us to study the detailed performance of each model and its dynamics as well as to evaluate the results in tasks where different kind of actions were required. In all the foraging tasks, when the contextual layer was enabled, the actions from the reactive and the adaptive layers were deactivated to avoid any influence they could have on the results.

To test the system for robustness we added 5% of noise to the motors, following a Gaussian random distribution and we varied the initial position of each trial according to a two dimensional normal distribution with mean 0 and variance $0,1\times$ robot radius. Moreover, to assess the impact of the camera noise in the information stored in memory, in Task 2 (Fig. 2.2b and

2.2e), we added noise to the hue sensed by the camera from 0% to 10% in steps of 1%, following a Gaussian random distribution. For every condition, we ran 10 experiments with 1000 trials each.

To investigate what was the impact on memory of the interaction between the decision-making models and memory itself we calculated the degree of compression of information in memory through the entropy of the stored information as follows:

$$E_M = -\sum_{s \epsilon S} p(s) log_2(p(s)) \tag{2.6}$$

where s is one segment in memory and $p(s)$ is the probability that the segment is selected in a current experiment. This measurement allows us to assess the amount of information needed to encode a visual stimulus in memory.

Results

In this study, we investigated the generalization of an integrator model in foraging tasks. We designed a number of foraging tasks with increased complexity to assess the generalization of the integrator model in more realistic tasks. The results in these tasks suggested a new mechanism to store and recall information from memory. We further tested the implications of this new mechanism in a foraging task and we show that it resulted in a more optimal way of learning and exploiting the environment

Foraging tasks

In all the tasks, we recorded the performance of the robot after LTM acquisition. As shown in Fig. 2.3, as task complexity increased the performance of the robot decreased dramatically in the case of the instantaneous model where it dropped to a mean value of 0.55 for the most complex task. It kept stable in the integrator model, maintaining a mean value of performance above 0.9.

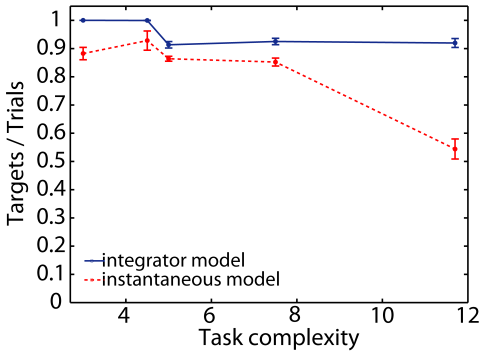

Figure 2.3: Robot performance along different task complexity. Ratio targets/trials with instantaneous and integrator models as task complexity increases. Bars represent means ± sem

In order to evaluate the impact of the camera noise in the information stored in memory due to the influence of each decision-making models, we used Task 2 ($TC = 4.5$) because it was the simplest one that requires the use of the contextual layer. For clarity we also report here the performance of the robot with varying camera noise (Marcos et al., 2010a). As previously reported in (Marcos et al., 2010a), the performance of the robot decreased as the camera noise increased in both models (Fig. 2.4a). The difference between the performance of the two models was significantly different along the different values of camera noise (KolmogorovSmirnov test, $p < 0.01$). From 0% of camera noise to 6% the instantaneous model was incrementally more affected by the noise than the integrator model. However, from 6% to 10% the noise had an important impact on the integrator model, resulting in a smaller difference in performance with respect to the instantaneous model. Once the noise of the camera started to critically affect the sequentiality of the actions the performance decayed in both models with a similar slope (Fig. 2.4a).

To assess the impact of both models at the memory level we calculated the entropy of the stored information, E_M (see Eq. 2.6). As shown in Fig. 2.4b, E_M with the integrator model was higher along the different camera noise compared to the instantaneous model (K-S test, $p < s0.001$). More-

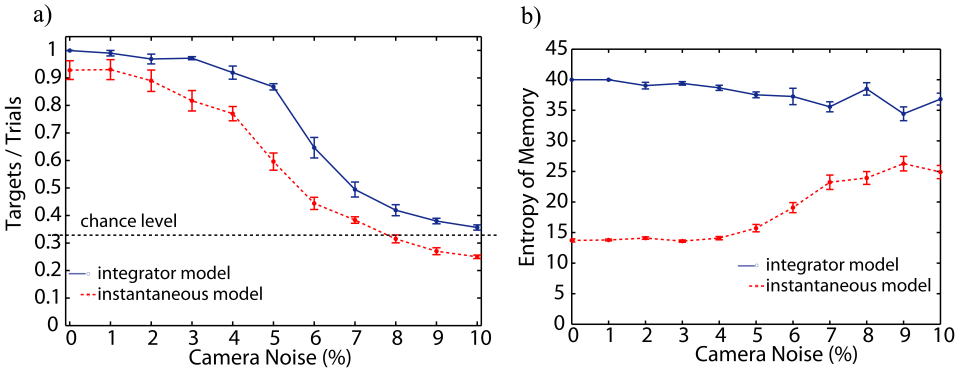

Figure 2.4: Performance and entropy along different camera noise. (a) Ratio Targets / Trials distribution for instantaneous and integrator models. (b) Entropy of the memory along different camera noise. In both figures bars represent means ± sem.

over, the dynamics in both cases were opposite: E_M decreased as camera noise increased in the integrator model whereas it increased as camera noise increased in the case of the instantaneous model. Low values of E_M means that segments of memory respond to a small fraction of the stimuli resulting in a higher number of segments in memory. The opposite occurs for high values of E_M. Therefore, the integrator model compressed the memory and less number of segments were necessary to encode same stimulus. As a drawback, explicit representation of time in memory, i. e. the number of steps needed to cross a patch, is lost.

Self-generated actions

The compression of information in memory due to the use of the race model changes the way information is stored in memory suggesting a new mechanism to recall it. Instead, of a recall of actions from memory it suggests the recall of goals. Consequently, we hypothesize that the actions are self-generated rather than stored in memory. During the recall period, visual information is retrieved from memory and actions are performed depending on the position of perceptual target with respect to the robot. When the information is selected from memory, we distinguish between two different

recall methods: (1) the next prototype in the sequence is retrieved, i.e. sub-goals are progressively achieved; (2) the prototype associated with the final goal is retrieved. In both cases, the retrieved information is stored in working memory and the robot searches for it. To do so, it moves the tilt angle of the camera from -60 deg. to -20 deg. and rotates over its own axis. In this way, the robot can see the visual cues that are far away from its current position. Once the robot sees the sub-goal or final goal it moves again the tilt angle from -20 deg. to -60 deg. and goes straight to the goal, i. e. self-generating the actions. We test this new way of retrieving information from memory in the Task 1 and compare the results to the non-self-generated actions investigated in the section. As shown in Fig. 2.5a and 2.5b, the robot follows a different path depending on the mechanism it uses to recall information from memory. We observed that the ratio targets divided into travelled distance was significantly higher in the case of the final goal search mechanism of self-generated actions (Fig. 2.5c; K-S test, $p < 0.001$), i. e. the robot follows a shorter path to hit the target. On the contrary, there is no significant difference between the sub-goal search and the non-self-generated actions (K-S test, $p > 0.05$). This result shows an optimal way of using the compressed information of the memory when the actions are not stored but self-generated during the retrieval of information from memory. It results in more flexibility in the actions to be taken and allows to account for a continuous action space.

Conclusions

We tested the implications of an integrator decision-making model in sequence learning tasks with multiple alternatives using a cognitive architecture that we called DAC9, evolving from previous implementations (Verschure and Althaus, 2003). We compared the results with a Bayesian decision-making model which is thought to be optimal for action selection. As a framework we used a robot based architecture which allowed us to understand the behavioural and architectural implications of these alternative models during foraging tasks. We showed that the race model has a more robust task-related performance when perceptual noise is added (Marcos

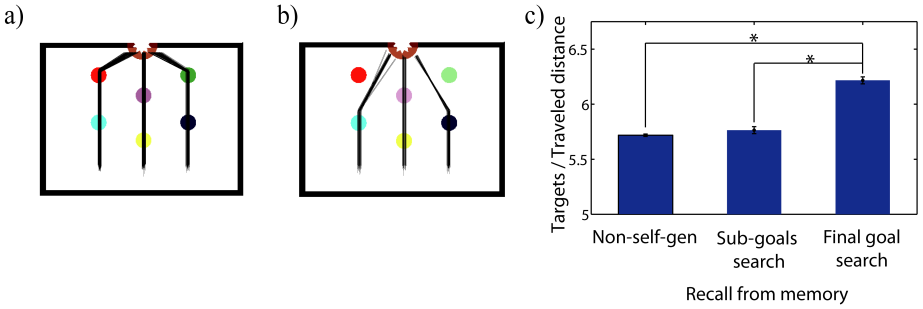

Figure 2.5: Robot trajectories and performance. (a) Trajectories generated during sub-goals search. (b) Trajectories generated during final goal search. (c) Ratio Targets / Travelled distance for the integrator model with non-self-generated actions recall, sub-goals search recall and final goal search recall. Bars represent means ± sem.

et al., 2010a) and when task complexity increases compared to the Bayesian model. Moreover, the race model also implied a compression of information in memory suggesting an alternative way of storing information, i. e. only perceptual information is acquired and the actions are self-generated during recall. The self-generation of actions during the retrieval from memory shows a mechanism able to account for a continuous action space.

In a previous study (Marcos et al., 2010a), we reported the differences in the storage of information in memory due to both models and the impact they have in performance. Here, we quantified the difference in the information stored in memory by calculating the memory entropy. We showed that the entropy is higher in the integrator model than in the instantaneous model. In the instantaneous model the actions are continuously recalled and performed. Therefore, in this case, the robot executes a number of actions, generally greater than one, each time it crosses a visual cue. In the case of the integrator model, we proposed a new mechanism to optimally use the information from memory. We implemented a goal oriented mechanism that retrieves visual cues from memory instead of actions. Once the visual cue, i. e. goal, is selected from memory the robot searches for it in the environment. Whenever the robot sees the goal it goes towards it. This new mechanism can be seen as the storage of an abstract object in memory,

i. e. a door. If a person wants to leave a room she/he has to first localize the door and then go towards it.

So far the integrator models have been used to explain simple decision-making tasks, such as RDM task (Shadlen and Newsome, 2001) or the countermanding task (Hanes and Schall, 1996a). The implementation of the integrator model was based on the race model. Generally, the race model has been mainly used to explain behavior in a countermanding task (Hanes and Schall, 1996a), predicting probability of failure and reaction time. Here, we showed the implication of this decision-making model in a more general framework. We observed that it has an important impact on how the memory is constructed and therefore on how the information is used later on.

The main assumptions we made in our proposal of self-generated actions during the recall from memory is that visual cues can always be seen from the current position of the robot. However, in wide open field environments, when this is not the case, our assumption would fail. In those situations, we would rely on head direction accumulator (Mathews et al., 2009) cells. The heading direction information would be stored in memory together with the visual information. During the recall from memory the actions would be also self-generated. Similar to the search of visual prototypes tested in this study, the robot would rotate around its own axis until its current head direction is equal or close enough to the retrieved head direction.

Physiological studies have shown that granular and pyramidal cells in the hippocampus encode information with high sparsity (low entropy), i. e. neurons respond to a small fraction of stimuli (Jung and McNaughton, 1993). In contrast, cells in the PFC have shown to be selective to particular cues with less sparsity (higher entropy) than the hippocampus and also with distinct temporal profile (Asaad et al., 1998). We observed that these two mechanisms of encoding memory have some similarity with the implications shown in this study due to the two decision-making models, i. e. higher entropy in the integrator model compared to the instantaneous

model. One could speculate that there is a distributed control system for sequence learning involving the hippocampus and the PFC connected to an external area which accumulates evidence, as found in the superior colliculus (Ratcliff et al., 2007), the lateral intraparietal area (Roitman and Shadlen, 2002), the frontal eye fields (Gold and Shadlen, 2003) and the PFC itself (Kim and Shadlen, 1999).

2.2 The complementary roles of allostatic and contextual control systems in foraging tasks

This sections reproduces a paper entitled "The complementary roles of allostatic and contextual control systems in foraging tasks" published in the proceedings of From Animals to Animats, Simulation of Adaptive Behavior (SAB) conference (Marcos et al., 2010b). The abstract reads:

To survive in an unknown environment an animal has to learn how to reach specific goal states. The animal is firstly guided by its reactive behavior motivated by its internal needs. After exploring the environment, contextual information can be used to optimally fulfill these internal needs. However, how a reactive and a contextual control system complement each other is still a fundamental question. Here, we address this problem from the perspective of the Distributed Adaptive Control architecture (DAC). We extend DAC's reactive layer with an allostatic control system and integrate it with its contextual control layer. Through robot foraging tasks we test the properties of the allostatic and contextual control systems and their interaction. We assess how they scale with task complexity. In particular, we show that the behavior generated by the contextual control layer is of particular importance when the system is facing conflict situations.

Introduction

One of the main challenges an animal faces when exploring a novel environment is how to learn about it and exploit it. Firstly, reactive behaviors drive animal exploration motivated by the animal's internal needs. Reactive behaviors also allows the exploration and acquisition of the state space. Once the states of the environment are learned and appropriate behaviors shaped, the animal is able to use this information to reach goal states, e.g. food. However, how these low and high level systems complement each other when the task difficulty increases is not clear yet. Here, we exploit this question in the basis of a combined robotics and computational neuroscience approach.

We investigate a robot model of self-regulatory processes based on the behavior of rodents. Rodents are optimal real-world foragers that can smoothly regulate complex sets of behaviors (Drai et al., 2000) based on their internal motivation, maintaining a dynamic stability with the environment while learning about it. Our model tackles exactly these two issues: self-regulation and learning about the environment.

Self-regulation is provided by a reactive layer that is based on the concept of *allostasis* (McEwena and Wingfield, 2003). This reactive layer drives the robot behavior while information about the environment is acquired and retained in a long-term memory. This memory is part of the contextual control (CC) system which will be capable of driving the robot's behavior based on the robot previous experience.

The allostatic control system (AC) of our model allows the robot to not only explore the environment but also to acquire its salient states. The robot locally senses different reward gradients present in the environment and can reach its desired values in the gradient by performing instantaneous reactive motor actions. The CC system picks up information about the environment. Both systems need each other and are fundamental for solving navigation tasks. The CC system will be able to solve tasks when the cues, e.g. gradients, that guide the behavior of the reactive layer is incomplete or contains conflicts. In addition, it can optimize the content of the long-term

memory system thanks to the states that are classified and the behaviors triggered by the AC system.

Many models deal with the problem of realizing an artificial rodent (Meyer et al., 2005; Sheynikhovich et al., 2009). Generally, it is usually tackled in a bottom up approach solving the navigation problem using a cognitive map. Our approach differs from these models in that our first building block integrates many regulatory subsystems and on top of it we add a cognitive system able to learn about the environment.

The AC and CC systems integration will be made in the context of the biomimetic Distributed Adaptive Control architecture (DAC) (Verschure and Althaus, 2003; Verschure et al., 2003; Duff et al., 2010). In this paper we make two new contributions. Firstly, we will augment the DAC reactive layer in order to support self-regulation on the basis of the physiological principle of *allostasis*. The AC orchestrates different homeostatic subsystems achieving stability at a meta-level (see Section 2.2). However, during this study, we will use only one homeostatic subsystem at a time thus bypassing the question of how multiple homeostatic subsystems affect optimal performance. The second contribution of this study is that we extend the contextual layer to be able to exploit the different internal states of the agent in its memory structures. The long-term memory stores sequences that belong to different goal states, i. e. desired values in the gradient. In the recall phase, information is retrieved based on the internal motivation of the robot, e.g. hunger (see Section 2.2). We integrate these two control systems and investigate the main implications of their integration (see Section 2.2). Our results show a successful integration which gives rise to realistic foraging in a variety of benchmark tasks using a simulated robot (see Section 2.2). In addition, it also indicates how low level predefined behavior control systems of the brain can be integrated with more advanced neuronal systems.

Methods

To understand how AC and CC systems complement each other we work with the DAC architecture. DAC distinguishes three coupled layers that interact between each other: reactive, adaptive and contextual. The reactive layer contains a pre-wired repertoire of reflexes, which creates a behavior that allows an interaction with the environment. Originally, this reactive layer implements collision avoidance and light appetitive mechanisms. We will extend this layer to provide it with an AC system. The adaptive layer processes and classifies the sensory input. This classification together with the actions executed by the robot are sequentially stored in the contextual layer which is equipped with a short and a long-term memory. These representations are used to plan ongoing behavior, and have been shown to be compatible with formal Bayesian models of decision making (Verschure and Althaus, 2003). In the original DAC, the contextual layer stores positive or negative sequences that lead to goal states defined by reward or punishment respectively. We extend the contextual layer to equip it with labeled information where the content of the memory can lead to different kinds of rewards or punishments.

We will test our model in different foraging tasks where gradients are projected into the environment and visual cues are placed on the floor. The gradients are gaussian functions sensed by the robot. The AC system will steer the motors of the robot depending on the desired and actual value of the gradient. Sequences of sensory-motor contingencies are learned by the CC system from the robot's interaction with the environment. We will show that AC system alone is not always sufficient to reach goal states when the gradient has a conflict information.

The allostatic control system

In our self-regulation model approach, different simpler homeostatic subsystems coexist (Sanchez-Fibla et al., 2010). Each homeostatic subsystem is associated to one reward gradient and has access to an actual (V_a) and desired (V_d) value in that gradient. The actual value is determined by the

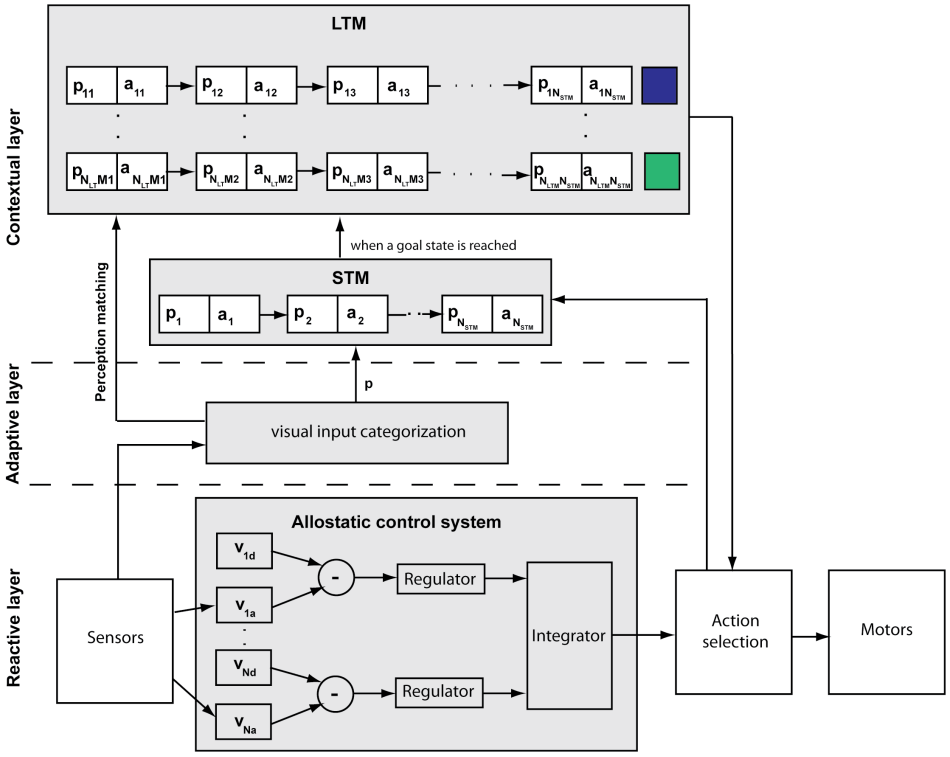

Figure 2.6: Contextual and allostatic integration in the framework of DAC. Abbreviations mean: p, perception; a, action; v, value. Blue and green colors at the end of the sequences mean different reward associated with them. Action selection is done by priority: contextual layer actions have more priority than reactive layer actions.

actual position of the robot in the gradient and the desired value represents the goal state in the gradient. The homeostatic subsystem acts in a closed loop trying to bring close the actual to the desired value and thus achieving stability (see Fig. 2.6). An integrator mechanism orchestrates the different homeostatic subsystems. All the subsystems compete to control the robot. In this study, only one homeostatic subsystem is activated at a time.

Imagine, as an example, an animal placing itself in an optimal distance to a heating source in order to achieve a desired temperature (Iizuka and Di Paolo, 2008). The gradient would correspond to the temperature map

which would have its highest peak around the source. The agent is able to sample the temperature gradient locally. The homeostatic subsystem could bring the agent closer to or further away from the heating source depending on the difference between the desired and actual value.

Short and long term memory

The contextual layer of DAC contains a short-term memory (STM) and a long-term memory (LTM). Sequences of sensory-motor contingencies are learned from the robot's interaction with the environment. Perception-action associations are stored in STM. When a goal state is reached, the sequences of associations are copied into the LTM labeled with the reached goal state (see Fig. 2.6). When the robot is exposed to a perception, it is classified by the adaptive layer and compared with the perceptions stored in LTM. The stored perceptions that match this comparison, and belong to a sequence with a goal state coincident with its current internal motivation, are selected and an action is executed. For further explanation about the memory structure in DAC see (Verschure et al., 2003).

Allostatic and contextual integration

We integrate AC and CC systems using the framework of DAC (see Fig. 2.6). The reactive layer of DAC is provided with an AC system. This AC system steers the motor of the robot driven by the gradient in the environment and the internal motivation of the robot. This results in egocentric actions executed. These egocentric actions are converted into allocentric ones by the CC system. To do the conversion, we added a path integration computation that calculates the vector between visual perceptions using the head orientation of the robot. In this way, the information in memory contains visual cues and the vector connecting two visual cues or visual cue to a goal state.

Foraging tasks

We test the integrated model in foraging tasks using 3D environment with a mobile agent. The simulated agent is implemented in C++ and wSim (Wyss et al., 2006) using the Open Graphics Library approximating a Khepera robot [2] widely used for behavioral modeling. The validity of the simulated robot with respect to a real one has been demonstrated in several studies (Wyss et al., 2006). The robot has a radius of 5.5 cm and it is equipped with three blocks of eight light sensors and eight proximity sensors. The sensors integrate an exponential decay function with respect to the distance to the light sources or to the obstacles respectively. The robot is also equipped with a color camera pointing to the floor (with an angle of 45). Therefore, the 3rd dimension of the environment is limited to the walls. The action group from the architecture is connected to the motor group of the robot. Each cell of the motor group maps a direction of movement. A winner-take-all (WTA) takes place at the motor map level and selects the neuron with highest activity. The default movement of the robot is to go forward.

Our aim is to test the model in tasks with increasing difficulty to understand how AC and CC systems scale with task complexity, i. e. tasks where gradients and visual cues have coherent information and tasks where this information is contradictory. This will allow a better understanding on when AC system would be enough to fulfill the robot's internal motivation and when CC system would be necessary to optimally fulfill this internal motivation.

We run experiments in three different environments. Every environment contains visual cues and rewards. The visual cues are patches on the floor whereas the rewards are gradients. The internal motivation of the robot is set to the highest value in a way that the desired value is reached at the center of the gradient. The first foraging task is an open field environment with one kind of gradient, a light. We vary the weight of the gradient from 0 to 1 in steps of 0.1 (see Fig. 2.7A). When the robot reaches the reward, i. e. its desired value in the gradient, the gradient is turned

[2] K-Team, Lausanne, Switzerland

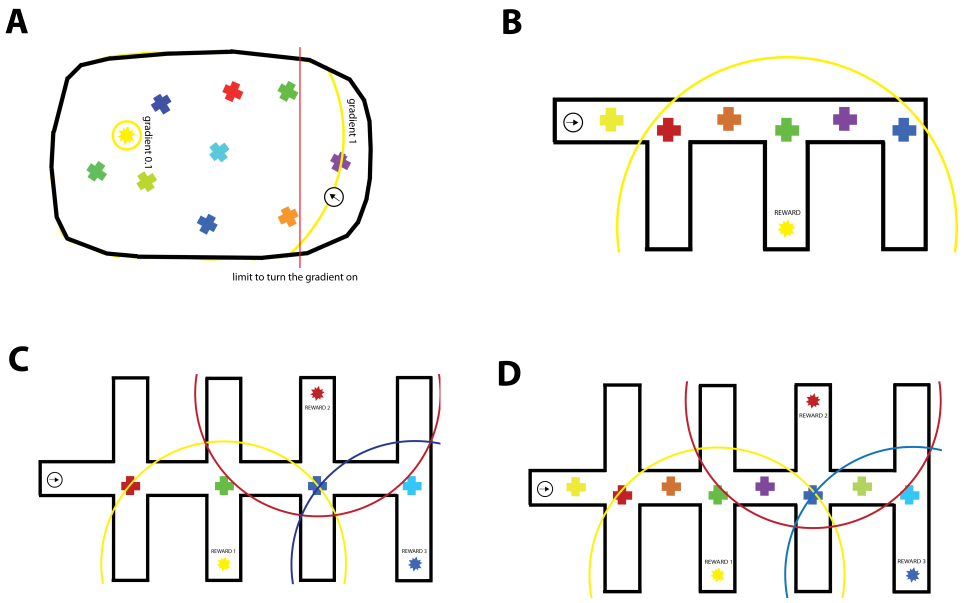

Figure 2.7: Foraging tasks environments. **(A)** Open field task. Yellow lines indicate the lowest and the highest value of the gradient. Red line indicates the limit the robot has to cross so the gradient is activated again after reaching the reward. **(B)** Maze with one reward. Yellow line indicates the area covered by the gradient of the reward. **(C)** and **(D)** Maze with three different rewards. Yellow, red and blue lines indicate the area covered by each of the reward gradients.

off to avoid that the robot stays next to the reward during the whole experiment. It is activated again when the robot is far away from it. For high weights of gradient it covers almost the whole foraging space, therefore we expect a similar performance in both AC and CC systems, since gradient and patches have coherent information. Secondly, we set an environment also with one kind of gradient, a light, with obstacles that do not allow a direct path between the initial position of the robot and the reward (see Fig. 2.7B). This environment is remotely based on the Tolman maze (Tolman and Honzik, 1930) as in (Hartland et al., 2009). We expect to have a significantly better performance with the CC system, due to its capability of learning sequences of perception-action to reach goal states. In this case, every trial finishes when the robot reaches the reward and starts again from same spatial position and a random orientation selected from

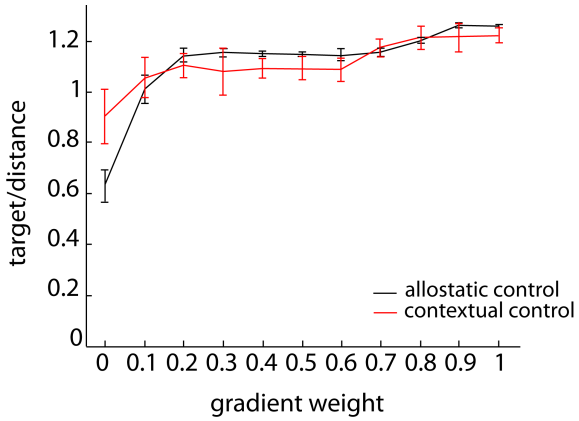

Figure 2.8: Allostatic and contextual performance. Ratio targets/distance for the allostatic and contextual control systems for different gradient values.

a two dimensional normal distribution $N(0, 1)$. As a final step, we set a third environment with three different gradients (see Fig. 2.7C and Fig. 2.7D). The robot searches for one reward or other depending on its internal motivation which is randomly selected at the beginning of every trial. We expect a significantly better performance in the case of the CC system due to the presence of obstacles as in the previous task. To see the influence in performance in the CC system due to the number of visual cues we run experiments with four and eight patches in the environment. In the three tasks, we keep constant the size of the memory, with 40 sequences of a maximum of 120 perception-action associations.

To simulate real conditions, we added 5% of noise to the motors of the robot. For every condition, we run 10 experiments with 20000 cycles each of them.

Results

Open field task

To compare performance between AC and CC systems we record data only when the AC system is activated. Later on, we activate the CC system and when the memory is full we record data again. We look at the ratio between

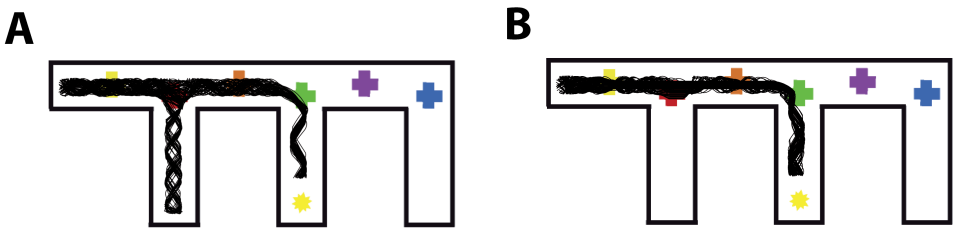

Figure 2.9: Trajectory plots. **(A)** Trajectories of the robot with allostatic behavior. **(B)** Trajectories of the robot for contextual behavior.

the number of targets reached and the distance explored by the robot (see Fig. 2.8). We observe that when the gradient is not present, i. e. the weight of the gradient is 0, the CC system performance is significantly higher than the AC system performance (Wilcoxon rank sum test, $p < 0.001$). However, as the weight of the gradient increases the AC system performs better than the CC system. This is expectable since the gradient gradually occupies the whole environment and the AC system can optimally calculate next action to perform. Thus, the CC system improves performance for a low value of the gradient but for values higher than 0.3 the AC system performance is significantly better than the CC system performance (Wilcoxon rank sum test, $p < 0.01$). Therefore, when there is not conflict in the gradient information the CC system is not fundamental and the AC system can properly perform the task.

Maze one gradient task

Firstly, ee test the AC system performance. We observe that the gradient of the reward drives the robot to a wrong path in some occasions (see Fig. 2.9A). When the CC system is also activated, we observe that the robot is not driven to the wrong path (see Fig. 2.9B). This difference in the paths selected in both cases is translated to a significant increase in the number of cycles needed by the robot to reach the reward with the AC system in comparison to the CC system (mean in the AC system is 222 cycles Vs. 141 in the CC system).

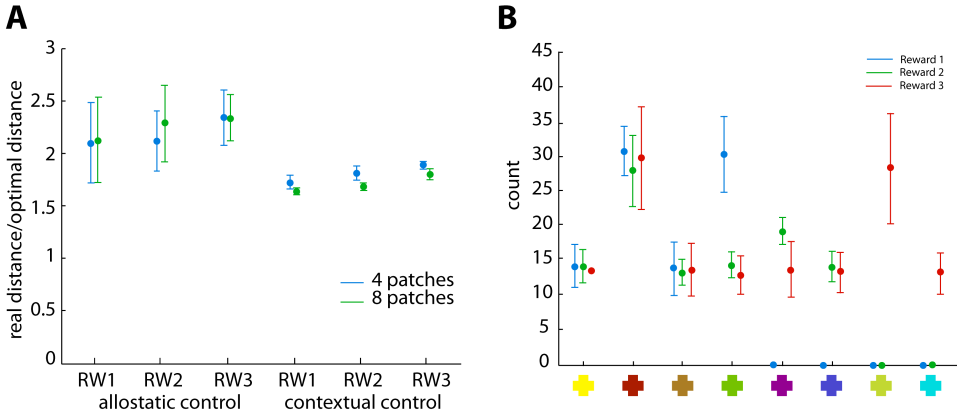

Figure 2.10: Allostatic and contextual performance II. **(A)** Ratio real distance/optimal distance for allostatic and contextual control systems for each of the three rewards present in the environment: RW 1, reward 1; RW 2, reward 2, RW 3, reward 3. **(B)** Number of times that each patch is stored in memory for each of the three rewards.

Maze three gradients task

We start testing the integration model with 4 patches in the environment. As in previous tasks, we first record data with the AC system alone and then we activate the CC system. We calculated the optimal distance the robot should cover to reach each of the rewards (see Fig. 2.7C and 2.7D). In Fig. 2.10A, we plot the ratio between the real distance covered by the robot and the optimal one for each of the rewards with the AC and CC systems. We observe that the CC system performance is significantly better for the three rewards (Wilcoxon rank sum test, $p < 0.001$). The results also show that the performance decreases as the distance to the reward increases. This is to be expected since the probability to leave the optimal path increases. However, we observe that the slope of the performance for the three rewards is lower in the CC system than in the AC system suggesting that the performance will decrease faster in the AC system as the distance to rewards increases.

To test the influence of the number of patches in the CC system performance we increase the number of the patches in the environment, from 4 patches to 8. We observe in Fig. 2.10A that the performance at the AC

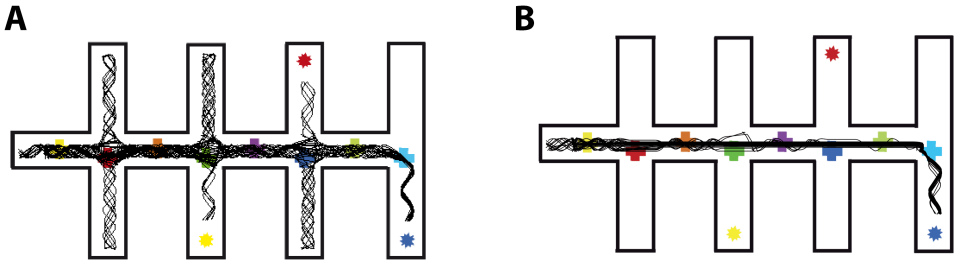

Figure 2.11: Trajectory plots II. **(A)** Trajectories of the robot with allostatic behavior. **(B)** Trajectories of the robot for contextual behavior.

system level keeps constant since the patches are not used by it. However, we observe an increase in performance in the CC system. Due to the increase in the number of patches the CC system has a more accurate knowledge of the environment, i. e. greater number of internal representations leading to a greater number of actions that the memory might trigger.

To better understand how the CC system picks up information about visual cues, we look at the content of the memory. In Fig. 2.10B we plot the number of times each visual cue appears in memory for the different rewards. We observe that the last two patches are used only for reaching the third reward whereas the rest of patches are part of the sequences corresponding to the three different rewards. As an example, in Fig. 2.11 we see the trajectories followed by the robot with the AC and CC systems when the internal motivation of the robot is to reach the reward 3. The AC behavior guides the robot to a wrong path in some occasions. The CC system trajectories are more accurate than the AC system ones.

Conclusions

We investigated how allostatic and contextual control systems complement each other. In order to do that, we extended the DAC architecture. On the one hand, we equipped its reactive layer with an allostatic control capability. On the other hand, we extended the memory content allowing the storage and retrieval of information related to different goal states. Furthermore, we integrated both systems converting the egocentric actions from the allo-

static control system in allocentric ones for the contextual control system. We tested the model in a variety of foraging tasks. We show that allostatic control system is able to succeed fulfilling the robot's needs when the information from the reward does not contain conflicts. When obstacles in the environment do not allow a direct path to the reward a contextual control system is needed. Our results show that in these cases the contextual control system is able to pick up salient information from the environment and use it to fulfill the robot's needs.

The model integration proposed here can be reminiscent of taxon Vs. route navigation strategies (Redish, 1999). On the one hand, the allostatic control system is capable of solving taxon task, i. e. tasks where the goal is visible. On the other hand, as the complexity of the task increases, i.e. the goal is not visible, the contextual control system is capable of solving it through route navigation, i. e. chaining taxon strategies. A number of robotics and artificial intelligence algorithms have been proposed to solve the taxon chaining problem (Kuipers, 2000; Mallot and Basten, 2009). Our approach differs from them in that our first building block is self-regulation and on top of it we add a contextual control system able to learn about the environment. Different internal states leading to different goal states can be handled achieving self-regulation.

Further experiments would need to be done in order to better understand how allostatic and contextual systems interact. For instance, the possibility of multiple homeostatic subsystems activated at the same time might influence the information learned by the contextual control system. The implications of this influence in the memory content should be tested.

To integrate allostatic and contextual systems in a model allow us to better understand how they complement each other and how they scale with task complexity. Moreover, we propose that these implications might be extended to the biological brain and its multi-level architecture.

Supplemental Experimental Procedures

DAC architecture

The DAC architecture distinguishes between three different layers of control: reactive, adaptive and contextual layers. The most basic behavior is generated at the reactive layer and allows interaction with the environment while acquiring information that is furthered used and processed by the higher adaptive and contextual layers.

Reactive and adaptive layers The reactive layer provides predefined automatic responses (UR) triggered when an unconditioned stimulus (US) is encountered. Static and prewired connections from sensory stimuli to actions implement the reactive behavior, i. e. aversive or appetitive. This reactive behavior allows the sampling of data from the environment which are learned and classified in terms of their valence and associated actions by the adaptive layer. The adaptive layer creates internal representations of the data, i. e. transforms data into information, and learns to associate these internal representations with actions, i. e. learns to associate between conditioned stimuli (CS) and unconditioned responses (UR) triggered by the unconditioned stimulus (US). Later on the anticipation of the adaptive response to a conditioned stimulus in the absence of US will be called conditioned response (CR). Therefore, the interaction between reactive and adaptive layers implements classical conditioning, where there is a transformation from information into knowledge. The knowledge created at the adaptive layer is accumulated in the contextual layer.

Following the abbreviations from Fig. 2.12, the actions generated by the reactive (a_r) and adaptive (a_a) layers are defined as:

$$a_r = U^T r H(r - \theta^A) \tag{2.7}$$

$$a_a = U^T y H(y - \theta^A) \tag{2.8}$$

where r stands for the contribution of the US to $IS \in \mathbb{R}^K$, y stands for the contribution of the CS to $IS \in \mathbb{R}^K$, U is the weight matrix from the

IS to the *A* cell group and *H(.)* is the Heaviside or step function [3]. Their mathematical expression is: $r = V^T s$ and $y = W^T x$, where V is the weight matrix from *US* to *IS* cell group $\in \mathbb{R}^{NxK}$ and W is the weight matrix from *CS* to *IS* cell group $\in \mathbb{R}^{NxK}$. The weight matrix W changes following a, so called, predictive Hebbian learning rule (Verschure and Pfeifer, 1992). The changes in the associations between *CS* dependent states and internal states depend on the difference between actual *CS* x and predicted *CS* e. e is the backwards projection of y to the *CS* cell group. It is described as $e = Wy$ and and it is the sensory information stored in the internal state population of the adaptive layer. With these definitions W changes as (Duff and Verschure, 2010):

$$\Delta W = \eta(x - \gamma e)((1 - \zeta)y + (1 + \zeta)r)^T \tag{2.9}$$

where the parameter η is the learning rate and ζ, $\zeta \in [-1, 1]$, balances the influence of the xy^T and xr^T, i. e. ζ balances the influence of behavioral and perceptual learning. The negative normalization term $-e$ depresses the weights and assures convergence, while the γ controls the influence of this normalization term. This learning rule directly captures the Rescorla and Wagner laws of associative competition that essentially state that animals only learn when events violate their expectations (Rescorla and Wagner, 1972).

Contextual layer The contextual layer provides mechanism for memorizing and recalling behavioral sequences. It consists of two structures: short-term memory (STM) and long-term memory (LTM).

The contextual control is based on the following assumptions:

- Memorize:
  - STM stores sensory-motor events generated by the adaptive layer.

[3] $H(x)$ is 1 if $x \geq 1$ and 0 if $x < 0$

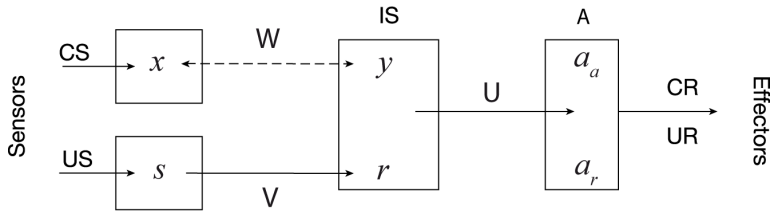

Figure 2.12: Reactive and adaptive layer. Squared boxes stand for neuronal groups. Arrows stand for static (solid) and adaptive (dashed) synaptic connections between cell groups. Abbreviations mean: *US*, unconditioned stimulus; *CS*, conditioned stimulus; *IS*, internal states; *A*, action; *UR*, unconditioned response; *CR*, unconditioned response; *W*, *V* and *U* are connection matrix.

  - When a goal state is reached the content of the STM is stored in the LTM and the STM is initialized.

- Recall
  - The content of LTM is compared with ongoing sensory events.
  - Matching elements contribute to action selection.
  - Chaining through sequences is achieved by biasing LTM matching.

The STM is a ring buffer with fixed size N_S. Every element of this memory is called a *segment*. A series of consecutive segments is called *sequence*. The STM is formed by one sequence of N_S segments. At each moment the generated CS prototype e and the action a executed by the robot are stored in the STM. When a goal state is reached, the sequence stored in the STM is copied into the LTM and the STM is reset. The LTM has a N_L number of sequences. The size of the LTM is, therefore, N_LxN_S. The sequences in the LTM are defined by their different goal states (e.g. -1 for an aversive events such as a collision, +1 for an appetitive event such as a reward).

The contextual layer integrates the representations of sensor states and actions formed at the level of the adaptive layer and the activation of its segments depends on the matching between the predicted *CS* prototype e

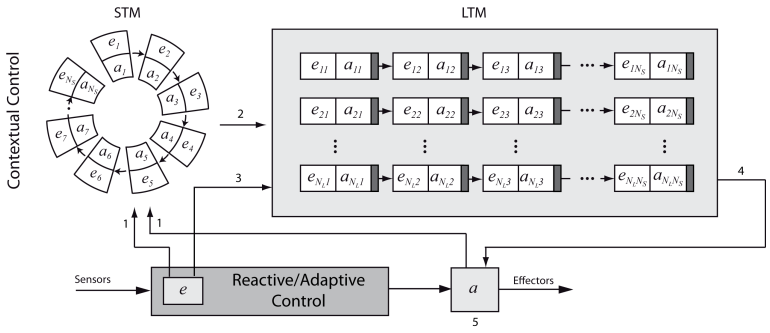

Figure 2.13: Contextual layer. (1) The *CS* prototype e and the executed action a are stored in the STM as a segment. (2) When a goal state is reached the content of the STM is copied in the LTM as a sequence and the STM is reset. (3) The values of the generated *CS* prototype e are matched against those stored in the LTM (4) The A population receives as an input the action response calculated as a weighted sum over the memory segments. (5) The actions proposed by reactive, adaptive and contextual layers compete in a priority selection mechanism to control the robot.

generated by the adaptive layer and the actual *CS* prototype x. This quality of matching is defined by an internally generated discrepancy measure (D) that is running an average distance between the prototypes *CS* x and *CS* e.

$$D(t+1) = \alpha_D D(t) + (1 - \alpha_D) d(x, e) \tag{2.10}$$

where α_D defines the integration time constant and the distance *d(x,e)* between actual *CS* x and estimated *CS* e prototypes is calculated as:

$$d(x, e) = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{x_j}{max(x)} - \frac{e_j}{max(e)} \right| \tag{2.11}$$

Initially only the reactive and adaptive layers are active. The contextual layer is activated when D falls bellow a certain *confidence threshold*.

During the recall all the *CS* e prototypes stored in the LTM are matched against the generated *CS* e prototype. The degree of matching of segment l in sequence q determines the input to its, so called, *collector*:

$$c = (1 - d(e, e_{lq}))t_{lq} \tag{2.12}$$

The collector determines the contribution of the segment to the action selection. Its activity depends on the distance *d(.)* of the generated *CS e* prototype to the *CS e* prototype stored in the segment and on a, so called, *trigger value t*.

The trigger value biases the sensory matching process of the segments and allows chaining through a sequence. Its value depends on the previous activity of neighbouring segments inside the sequence. Its default value is 1 and does not bias the collector value. When a segment *l-1* in sequence q is activated the trigger of segment l is set to a value higher than 1. This means that a segment, following a previously effective one, will be given higher priority in future decision making. This trigger decreases its value to 1 asymptotically with a defined time constant according to:

$$t_{lq}(t + 1) = \alpha_t + (1 - \alpha_t)t_{lq}(t) \tag{2.13}$$

where $\alpha_t \in [0; 1]$. The trigger of a selected segment resets its value to 1.

The activity of the collectors will contribute to the action proposed by the contextual layer. We will only consider the collectors' activity that satisfy that (1) its activity is above a certain threshold (θ^C) and (2) its activity is inside a predefined percentage range from the maximum collector's activity, i.e. the collectors compete in an E%-Max Winner Take All (WTA) mechanism (de Almeida et al., 2009), where at each moment the maximum collector's activity and all collectors' activity distance to this maximum activity is calculated. The percentage range from the maximum collector's activity is a parameter that reflects the certainty or uncertainty of the robot, i. e. when the robot is not enough certain about which action to perform more suggestions from collectors are listened. The collectors selected from the WTA will contribute to the contextual action. All collector units are connected to the action group A. The actual action proposed from the contextual layer is calculated as:

$$a_c = \sum_{l,q \in LTM} \pm \frac{c_{lq} H(c_{lq} - \theta^C)}{\delta_{lq}} a_{lq} \tag{2.14}$$

where δ_{lq} is the distance measured in segments between the selected segment l and the last segment in the sequence, i.e. the distance to the goal state. By doing this division the segments closer to the goal have more impact on the contextual action. The sign is plus if the segment belongs to an appetitive sequence and minus if it belongs to an aversive sequence. After summing up all the selected actions and having the calculated action (a_c) if it results in a negative action it is not triggered and therefore only final positive actions are triggered. This avoids having backwards actions.

The action selection that defines the final action performed by the agent in order of priority executes one of the actions triggered by the reactive (a_r), adaptive (a_a) or contextual layer (a_c), giving more priority to reactive actions, then to contextual actions and finally to adaptive actions. The actions generated by the contextual layer are those that find the most support by the knowledge available in LTM and its matching to sensory events.

Chapter 3

Neural response variability in premotor cortex is modulated by trial history and predicts behavioral performance

The problem is, they're variable

Steve Rhode

In the previous chapter, an embodied artificial agent, provided with a cognitive architecture, was able to gather information from the environment and use it to reach specific goals. In order to do so, the behavior of the robot was biased by the temporal order in which external events were encountered. The DAC architecture achieves this by biasing the retrieval of information from memory based on previous perceptual inputs. However, the neural mechanisms underlying this bias are not well known yet. To elucidate this issue, we investigate the neural correlates causing a bias in behavior due to experience (trial history) during a motor decision-making task. Our main result shows that behavior in a given trial is influenced by the recent trial history and that the across-trial response variability of neurons recorded from the dorsal premotor cortex (PMd) closely correlates with it, whereas

the mean firing rate does not. This suggests that the mean firing rate reflects the strength of the perceptual input, which is the same for all trials, on the decision-making process while the across-trial variability reflects the variable influence of memory. To validate this, we perform an additional test with a theoretical model of binary decision-making and we observe that, if the perceptual input was fixed, the modification of a memory related signal directly caused changes in the variability of the across-trial response, while the mean firing rate did not exhibit significant variation. This result raises two questions: whether there really is a monitoring signal capturing previous history and whether the origin of it is internal or external to PMd.

In this chapter we reproduce a manuscript called "Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance" published in Neuron (Marcos et al., 2013b). The abstract reads:

> In the study of decision making, emphasis is placed on different forms of perceptual integration, while the influence of other factors, such as memory, is ignored. In addition, it is believed that the information underlying decision making is carried in the rate of the neuronal response, while its variability is considered unspecific. Here we studied the influence of recent experience on motor decision making by analyzing the activity of neurons in the dorsal premotor area of two monkeys performing a countermanding arm task. We observe that the across-trial variability of the neural response strongly correlates with trial history-dependent changes in reaction time. Using a theoretical model of decision making, we show that a trial history-monitoring signal can explain the observed behavioral and neural modulation. Our study reveals that, in the neural processes that culminate in motor plan maturation, the evidence provided by perception and memory is reflected in mean rate and variance respectively.

3.1 Introduction

Two-choice perceptual and motor tasks have been widely used to explore the neural mechanisms underlying decision-making processes (Logan and Cowan, 1984; Smith and Ratcliff, 2004; Gold and Shadlen, 2007; Verbruggen and Logan, 2008). Neural activity in parietal and frontal cortical areas has been shown to be correlated with behavioral performance of monkeys trained in specifically designed tasks (Platt and Glimcher, 1999; Gold and Shadlen, 2000, 2007; Cisek and Kalaska, 2005; Mirabella et al., 2011). In the last years, these binary simple tasks have been extended to account for multiple choices (Churchland et al., 2008; Albantakis and Deco, 2009). Although there has been progress in the understanding of the decision-making process in these tasks, little is known about how the recent history of the task influences the neural mechanisms underlying this process. In a previous theoretical investigation of the dynamics of working memory in optimal decision making, we have proposed that the integration of information from perception and memory requires temporal integration, supporting perception, and dynamic modulation of this temporal integration, serving memory (Verschure et al., 2003). A specific neural mechanism explaining such memory biasing, however, has not yet been described. One could argue that the across-trial variance of the neuronal response could reflect effects of task history. It has been proposed that variance of neuronal responses is correlated with the progress of motor preparation (Churchland et al., 2006) and that it is a general feature of cortical dynamics that is nonspecific with respect to the behavioral task at hand (Churchland et al., 2010). Here we investigate the possible signature of recent trial history in the variance of neuronal responses by analyzing the single-unit activity recorded in the dorsal premotor (PMd) area of two macaque monkeys performing a countermanding arm task (Mirabella et al., 2006).

The countermanding task has been extensively used to study motor decision mechanisms. It evaluates the ability to cancel a planned cued movement in response to the presentation of an infrequent Stop signal presented at variable delays (Stop signal delay, SSD; Fig. 3.1A) from the time of

presentation of the visual target Logan and Cowan (1984); Verbruggen and Logan (2008). The overall behavioral performance in this task has been explained with the so-called race model (Logan and Cowan, 1984). The race model proposes that the behavioral outcome of the countermanding task is the result of a competition between a Go and a Stop process that evolves, driven by the accumulated sensory evidence, toward a decision threshold. Neuronal correlates of the movement generation process, as predicted by the race model, have been found in the modulation of firing rate (FR) of single-unit activity in the frontal eye field (FEF) and the superior colliculus (SC) for countermanding saccade tasks (Hanes and Schall, 1996b; Paré and Hanes, 2003) and in the supplementary motor area (SMA) and PMd for countermanding arm tasks (Scangos and Stuphorn, 2010; Mirabella et al., 2011). However, all these results ignore the role of trial history in the task. After each trial in which a Stop cue is delivered (Stop trials), subjects increase their movement reaction time (RT), purportedly reflecting an increase in uncertainty about the current trial (Rieger and Gauggel, 1999; Mirabella et al., 2006; Emeric et al., 2007; Verbruggen and Logan, 2008; Nelson et al., 2010).

Here, using the data set reported in Mirabella et al. (2011), we investigate the behavioral adaptation and the modulation of the activity of reaching related neurons dependent on the temporal order of a trial in a sequence, i.e., the recent history of a trial. We observed that both behavior and variability of the neuronal responses were modulated by trial history. Using a computational model, we show that these effects can be explained in terms of a competitive process that is modulated by a monitoring signal.

3.2 Results

To quantify the biasing of the neuronal response due to the history of a trial, we calculated the mean FR and the across-trial spike variability during Go trials that were sorted by different history conditions. We observed a significant and systematic difference in RT and neural response variability that held over a wide range of trial history conditions. These results suggested

that, other than perceptual signals, neurons in PMd are also influenced by an additional input related to the history of the trial, i.e., memory. To validate this hypothesis, we studied the response of a mean-field approximation of a spiking neural model (Wilson and Cowan, 1972) in a simulated countermanding task. We observed that an additional monitoring-related signal can directly account for the observed changes in the neural response variability and the behavioral performance.

Behavioral Responses

We analyzed the behavioral responses of the monkeys looking at their RT in Go trials and probability of failure to cancel a planned movement in Stop trials. Consistent with previous work (Emeric et al., 2007; Pouget et al., 2011), we observed that the mean RT of the monkeys increases when the current Go trial was preceded by a Stop trial (Fig. 3.1B), in contrast to when it was preceded by a Go trial. This confirms that performance is modulated by trial history. In addition, the SD of the RT was higher when a Go trial was preceded by a Stop trial than when preceded by a Go trial (see Fig. 3.5). Moreover, a longer RT was associated with a lower probability of failure in the following trial (Fig. 3.1C), i.e., successful cancellation was more likely in a Stop (t) trial that followed a sequence of Go (t-1) and Stop (t-2) as opposed to a sequence comprising two Go trials.

Neural Correlate of the Decision Process

To assess the neural correlate of the decision process, we analyzed the modulation of the mean FR of the neurons and their across-trial spike variability, as measured by the variance of conditional expectation (VarCE) (Churchland et al., 2011) during motor preparation. For this analysis we used only Go trials from the time of the presentation of the Go signal until arm movement onset. We sorted the data with respect to the type of trial that was preceding the current Go trial: a Go or a Stop trial. We observed that after the presentation of the Go signal, both the FR and the VarCE increased until they reached a peak value at about 150 ms before movement onset

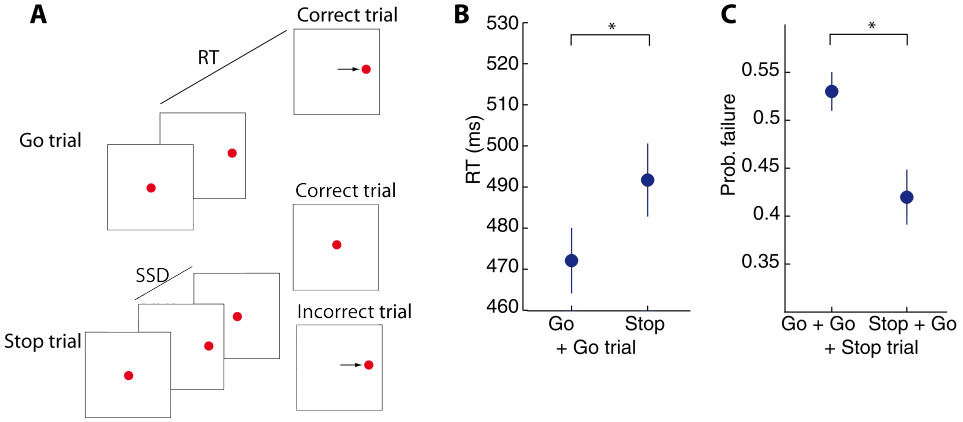

Figure 3.1: The countermanding task. **(A)** In a countermanding task, visual cues are presented either to induce movements (Go trial) or to prevent movements initiated by the Go signal (Stop trial). A central stimulus signals the start of the trial and the monkey is required to touch this cue with its hand. The start cue is followed (usually after 500800 ms) by a visual cue (Go signal) that indicates the location to which a movement must be made. In the Go trials, the monkey has to execute a speeded reaching movement toward this peripheral target. During the Stop trials (33% of the trials), the central stimulus reappears after a variable delay, or Stop signal delay (SSD), instructing the monkey to withhold the planned movement, keeping the hand on the central stimulus. (B and C) Behavioral performance in a countermanding task relative to task history (53 experimental sessions). Error bars indicate SEM. **(B)** RT in Go trials when preceded by a Go or a Stop trial (Go + Go trial, 16,060 trials; Stop + Go trial, 6,671 trials; Kolmogorov-Smirnov test, $* = p < 0.01$). **(C)** Probability of failure in a Stop trial when preceded by Go + Go trial or Stop + Go trial (Go + Go + Stop trial, 4,601 trials; Stop + Go + Stop trial, 2,116 trials; Kolmogorov-Smirnov test, $* = p < 0.01$). See also Fig. 3.5.

(Fig. 3.2A and Fig. 3.2B). After this peak, the mean FR and the VarCE gradually decreased to their baseline (right panels of Fig. 3.2A and Fig. 3.2B). The mean FR in the analysis epoch did not significantly differ between the two conditions, i.e., whether the current Go trial was preceded by a Stop or a Go trial. In contrast, VarCE displayed a strong modulation by the task history and was significantly higher in case the preceding trial was a Stop as opposed to a Go trial (Fig. 3.2B). Single-unit analyses showed a consistent effect across the whole population (Fig. 3.6A). We also tested the correlation of task history with VarCE during Stop trials in two different

contexts: when a Stop trial was preceded by Go (t - 1) and Stop (t - 2) trials or by two consecutive Go trials. We observed the same modulation in VarCE by task history (Fig. 3.6B).

Interestingly, the difference in VarCE between both conditions disappeared about 70 ms after the presentation of the Stop signal. This latency is consistent with the average processing delay of visual information in PMd (Cisek and Kalaska, 2005).

In a next analysis, we assessed the relationship between task history, VarCE, and performance (Fig. 3.2C, 3.2D, and 3.6C). This analysis revealed that mean and SD of RT closely mirrors the effect of task history on VarCE over a wide range of task history conditions. The three factors, mean RT, SD of RT, and VarCE, increased with an increase in the number of previous Stop trials, while they decreased with an increase in the number of preceding Go trials. Moreover, changes in mean RT over a range of trial history conditions are due to systematic shifts of the entire RT distributions (Fig. 3.6D). We observe that the mean RTs are very well correlated with VarCE (Fig. 3.3A) and that RT and VarCE distributions seem to have similar shape (Fig. 3.3B). The mean FR for the same conditions did not show any variation (Fig. 3.6E). Interestingly, the modulation of VarCE also depends on the difficulty of the previous trial (Fig. 3.6F), so that its value increased as the SSD in the Stop trial preceding the Go trial increased. Thus, these results suggest that the influence of task history is reflected in the variance of neuronal activity in PMd and that both variables, VarCE and trial history, are linearly correlated with performance.

Mean-Field Approximation

In order to understand the neural mechanisms causing the observed behavioral and across-trial neuronal response variability differences due to varying trial history conditions, we used a mean-field approximation (Wilson and Cowan, 1972) of a biophysically based binary decision-making model (Fig. 3.4A). The model receives two segregated inputs: perceptual evidence provided by the visual cues (Stop and Go signals) and a task history signal

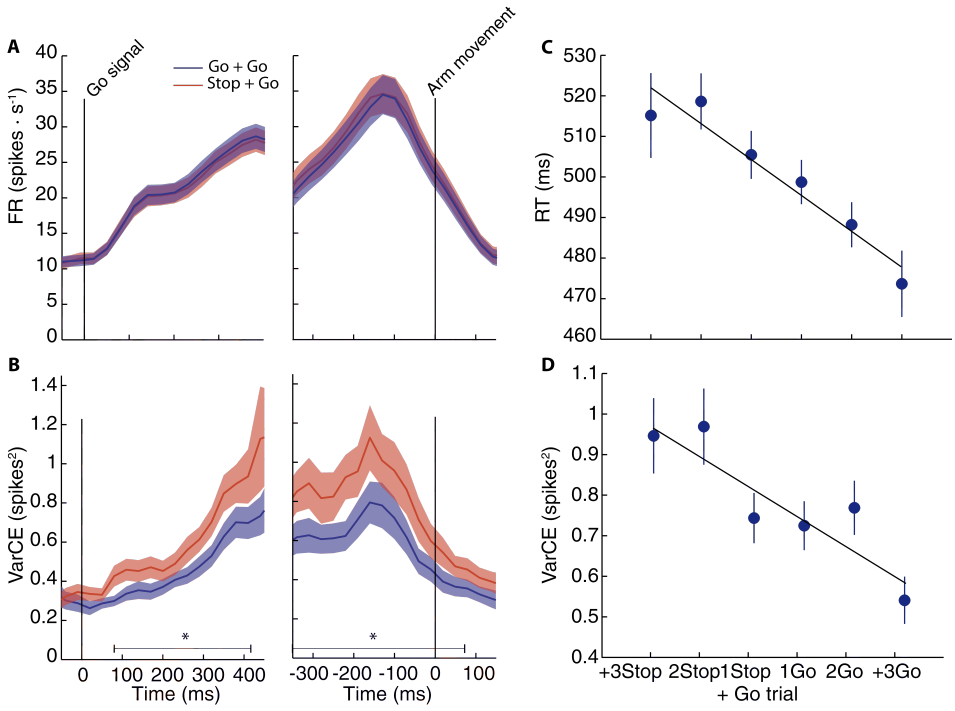

Figure 3.2: Neuronal dynamics during the countermanding task. **(A)** and **(B)** Neural activity is aligned to Go signal (left) and arm movement (right) onsets. Shaded areas indicate SEM. **A)** Mean FRs in Go trials when preceded by a Stop (red) or Go (blue) trial (Kolmogorov-Smirnov test, $p > 0.05$). Results are obtained from 142 neurons (Go + Go trial, 16,060 trials; Stop + Go trial, 6,671 trials). **(B)** VarCE for the same two conditions as in (A). VarCE is significantly different from 80 to 410 ms after the onset of the Go signal (Kolmogorov-Smirnov test, $* = p < 0.01$) and from -350 to 70 ms when aligned to movement onset (Kolmogorov-Smirnov test, $* = p < 0.01$). **(C)** and **(D)** RT **(C)** and VarCE **(D)** during a Go trial in six different trial history conditions when it was preceded by three or more (+3) Stop trials (850 trials), two Stop trials (1,353 trials), one Stop trial (4,468 trials), one Go trial (4,578 trials), two Go trials (3,257 trials), and three or more (+3) Go trials (8,225 trials). Error bars indicate SEM. Data are obtained from same 142 neurons. See also Fig. 3.6.

provided by a monitoring system. The model has two populations of excitatory neurons: one population is sensitive to the appearance of the Go signal (λ_{go}; Go pool), while the other population is sensitive to the appearance of the Stop signal (λ_{stop}; Stop pool). The two populations mutually

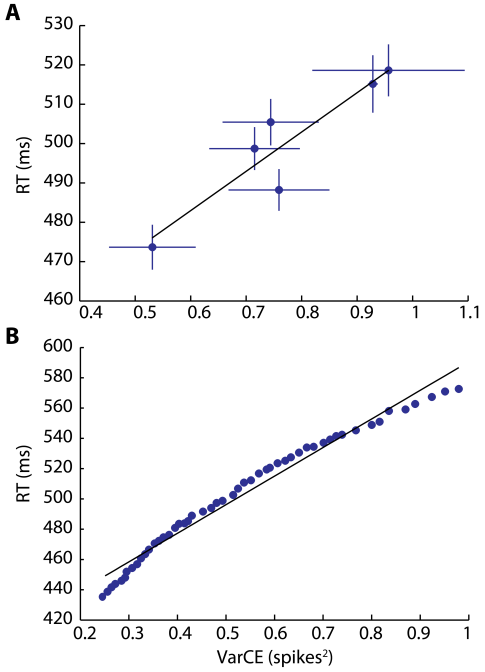

Figure 3.3: The relationship between performance and VarCE. **(A)** Mean RT is fitted by VarCE using a linear regression ($R = 0.93$ and $p < 0.01$). Data points are as in Fig. 3.2C and 3.2D. Error bars indicate SEM. **(B)** Quantile-Quantile plot of the interquartile range of RT and VarCE distributions formed by pooled data from all conditions. The data points are linearly correlated ($R = 0.99$ and $p < 0.01$), suggesting a high similarity in the shape of the distributions.

inhibit each other. In the absence of any of the two visual signals, both λ_{go} and λ_{stop} are equal to 0. A monitoring process modulates the strength of the input (λ) to each group of neurons simulating different trial history conditions: λ increases its value as the number of Stop trials preceding a Go trial increases and decreases its value as the number of Go trials preceding a Go trial increases (Fig. 3.4B). We observe that the model reproduces the same relationship between the probability of failure and SSDs as observed during the countermanding task, i.e., the probability of failing in the Stop trials increases as the SSD increases (Mirabella et al., 2006) (Fig. 3.7A).

To compute decision times in the simulations, we considered that the decision process was terminated when the difference in activity between Go

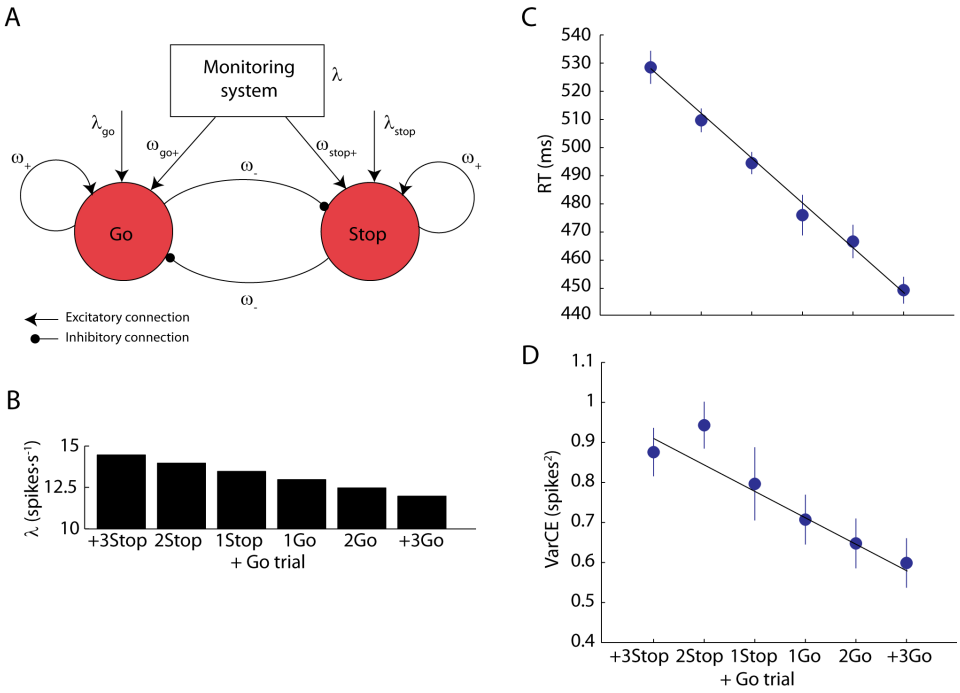

Figure 3.4: Mean-Field approach. **(A)** Network structure of the binary decision model. The 'Go' pool is selective for the Go signal (λ_{go}), while the 'Stop' pool is selective for the Stop signal (λ_{stop}). The two pools mutually inhibit each other (ω_-) via inhibitory pools (not represented) and have self-excitatory recurrent connections (ω_+). The 'Monitoring system' is connected with the two selective pools. The synapses that connect the 'Monitoring system' with the 'Go' and 'Stop' pools have different strength ω_{go+} and ω_{stop+}. **(B)** Firing rate value of the signal provided by the monitoring system. The value of this signal depends on the history of the current trial. The value of λ increases as the number of Stop trials preceding a Go trial increases and decreases as the number of Go trials preceding a Go trial increases. **(C)** and **(D)** RT **(C)** and VarCE **(D)** of the response of the 'Go' pool sorted by the recent history of a trial. Simulation results were obtained from 10,000 trials grouped in ten sessions of 1,000 trials each. Error bars indicate SEM. See also Fig. 3.7.

and Stop pools was above a fixed threshold (Roxin and Ledberg, 2008). The RT was calculated by adding 150 ms to the decision time, consistent with the peak in FR observed 150 ms before movement onset in the physiological data (Fig. 3.2A). The mean and SD of RT obtained from the simulations (Fig. 3.4C and Fig. 3.7B) exhibit the same trend as observed in the physiology

of PMd (Fig. 3.2C and Fig. 3.6C): the mean and SD of RT in a Go trial are longer/shorter as the number of preceding Stop/Go trials increase.

Consistent with our analysis of the physiological data, the different simulated trial history conditions have a similar impact on the variability of the Go pool response (Fig. 3.4D). This impact of the monitoring signal λ on RT and VarCE can be intuitively understood in terms of the competition between the two neuronal pools Stop and Go through mutual inhibition (Fig. 3.4A). The model is tuned such that the firing rate of the Go pool is not affected by this neuronal competition (Fig. 3.7C), as observed in the response of the neurons we have analyzed (Fig. 3.6E). We observe that, given these assumptions that reflect the physiological properties of PMd, the addition of the monitoring signal leads to the modulation of the effect that the Stop pool has on the dynamics of the overall network, leading to a change in the mean RT. In addition, when the influence of the Stop pool on the dynamics is increased, the intrinsic noise of the system starts to have a larger impact on the performance and dynamics of the network, resulting in an increase in VarCE and RT variability. Indeed, it has been demonstrated that the neural response variability changes with the strength of the input to this model, due to a shift in the distance from the working point of the system to the bifurcation point (Deco and Hugues, 2012; Roxin and Ledberg, 2008). Here we exploit this effect through the monitoring signal. Hence, perceptual input defines the mean rate, while the history dependent monitoring signal defines a modulation around this rate expressed in VarCE. These results confirm that the response variability we observed in PMd can be seen as a signature of trial history and predict the existence of a system that both monitors the recent history of a task and modulates competition between pools of neurons dedicated to Go and Stop.

3.3 Discussion

We have investigated the hypothesis that perceptual cues and memory of trial history are integrated in the decision-making process underlying the countermanding task. Our analyses of the responses of neurons in PMd of

monkeys performing a countermanding arm task show the influence of recent trial history on both the performance of monkeys and on the variability of neuronal responses in PMd. We show that the behavior of the monkeys becomes increasingly more conservative (longer RT) when a Go trial was recently preceded by one or more Stop trials and increasingly hastier (shorter RT) when it was recently preceded by one or more Go trials, as previously reported (Rieger and Gauggel, 1999; Emeric et al., 2007; Verbruggen and Logan, 2008; Nelson et al., 2010; Mirabella et al., 2006). We show that the behavioral performance is linearly correlated with changes in the variability of the neural response. To validate the possible signature of trial history in neural response variability, we performed an additional theoretical study using a mean-field approximation of a spiking neural model. We show that changes in the strength of a modulatory input that reflects trial history accounts for the observed changes in behavior and neural response variability, suggesting the existence of a trial history monitoring system in the brain. Our study provides a neural correlate for task history and its impact on the neuronal substrate of decision making and is a further example of how adaptive behavior is monitored and orchestrated in the brain (Walton et al., 2004; Ito et al., 2003).

One of the weaknesses of using VarCE as a measurement of the across-trial variability lies in the estimation of the scaling factor f. We computed it separately for each neuron (see Experimental Procedures), and the obtained distribution of the values of ϕ was consistent with the ones previously reported for the neocortex (Fig. 3.6G) (Shadlen and Newsome, 1998; Nawrot et al., 2008). To check the robustness of our results to variations in the value of f, we repeated our analyses (Fig. 3.2B) but setting the same value of f for each neuron. We observed that the difference in VarCE between history conditions is independent on the value of f used (Fig. 3.6H).

Similar to VarCE, the Fano Factor (spike count variance divided by spike count mean) has been used to calculate the across-trial variability of neural responses. Although in most cases both measurements are considered to be equivalent, for significant changes in mean FR, the VarCE has shown

to be more robust than the Fano Factor (Churchland et al., 2011). However, our conclusions hold for both the Fano Factor and the VarCE (see Fig. 3.6I and 3.6J) and are further supported by the equivalent histogram obtained from the interspike interval observed in a Go trial preceded by different sequences of trials, i.e., with different history (see Fig. 3.6K). Hence, our results are robust with respect to the specific method used to obtain a measure of variability.

Our results suggest that the observed change in strategy during the task might be due to an increase or decrease in the uncertainty about Stop cue appearance in the current trial, suggesting a relationship between trial history and uncertainty. Under this interpretation, one might speculate that the degree of the monkeys uncertainty is updated based on the trial history and increases as a function of the number of Stop trials. Subsequently, this relationship implies a direct link between uncertainty and variability: higher uncertainty is related to a higher variability in the neural response and a longer and more variable RT. Our simulation predicts the existence of a system that monitors either trial history itself or uncertainty based on trial history and updates its value according to new incoming information, i.e., actions and their outcome in a new trial. This definition of uncertainty is consistent with previous work in which uncertainty is defined in terms of the accuracy to predict the possible consequences of actions (Huettel et al., 2005; Yoshida and Ishii, 2006). For instance, in the countermanding task, after a Stop trial both humans and monkeys increase their expectation about the probability of a next trial including a Stop signal (Emeric et al., 2007).

The use of a mean-field approximation of a realistic network of integrate-and-fire neurons (see Experimental Procedures and Supplemental Experimental Procedures) allows us to study the dynamics of the decision-making process from the perspective of the neuronal substrate. We have shown that the biasing of the neural responses and the consequent changes in the behavioral strategy during different trial history conditions could be caused by a signal coming from a system that monitors the recent history of a

trial and that directly changes the strength of the competition between the neural populations involved in the decision making. This modulation in the competition influences the variability of the across-trial average activity, while the average response of the population correlated with the execution of movement (Go pool) is the same due to the balance in the excitatory and inhibitory connections of the network.

Changes in the behavioral strategy could be explained with the same mechanism, i.e., due to a modulation in the strength of the competition between neuronal populations, a suprathreshold difference in their activity will take varying amounts of time to be generated. Hence, according to our proposal, VarCE is a derived measure caused by a difference in the strength of the competitive process with different trial history conditions. Because our neural data are based on single-unit recordings, it is difficult to conceive how VarCE could be read out. However, areas like primary motor cortex, posterior parietal cortex, SMA, and cingulate cortex (Johnson and Ferraina, 1996; Johnson et al., 1996) that read information from PMd would have access to the population and, in this case, an instantaneous measure of variability could be possible by trading off temporal integration for spatial integration. This would raise the question of whether this redundant representation of trial history would be necessary. The answer to this question is, however, out of the scope of this study. Changes in the initiation of activity accumulation in FEF and SC have shown to be correlated with task history-dependent changes in performance (Pouget et al., 2011). We did not observe, at the population level, any modulation of firing rate in PMd after adaptive response time adjustment. A possible explanation is that the functional organization of the neural network controlling eye movements is very different of that controlling limb movements (see also Discussion in Mirabella et al. (2011). We exclude that the modulation of FEF could be a source of the neural response variability we observed. In fact, our recording region included the more rostral portion of PMd but not supplementary eye fields (Mirabella et al., 2011). Only this last portion receives input from FEF, while the rostral PMd is preferentially connected with dorsolateral prefrontal regions (Luppino et al., 2003). A monitoring

signal could be provided by the connection of PMd with cingulate cortex (Johnson and Ferraina, 1996; Luppino et al., 2003). The anterior portion of cingulate cortex has been shown, in humans, to display trial history modulation of baseline activity (Domenech and Dreher, 2010). Further studies are needed to clarify all these aspects in detail.

Our study shows a key role of the across-trial variability of the firing rates as a signature of trial history during decision making, confirming an earlier theoretical prediction (Verschure et al., 2003) and adding an extra variable to be considered in future experimental and theoretical studies. In the context of the countermanding arm task, the information provided by perception and memory to the decision-making process is reflected in different aspects of the neuronal activity: mean FR and acrosstrial variance respectively. We have shown that the latter is linearly related to the RT and the trial history experienced by the monkeys. Our results imply that there is a continuous monitoring of trial history that, combined with the current perceptual evidence, is used to make a decision. An important question is now whether the origin of this monitoring process is internal (Domenech and Dreher, 2010) or external (Zandbelt and Vink, 2010) to the PMd and its immediate cortical efferent and afferent areas.

3.4 Materials and Methods

Behavior and Physiology

Two adult male rhesus macaques (Macaca mulatta; monkey S and monkey L) weighing 78 kg were used. Details of the experimental procedures have been provided in Mirabella et al. (2011). Monkeys were trained to perform a countermanding reaching task. It consists of a random mix of 67% Go trials and 33% Stop trials. All trials began with the appearance of a stimulus at the center of a touch screen (Fig. 3.1A). Monkeys were required to touch the stimulus with their fingers, within 2 s, and hold it for a variable period of 500800 ms. Thereafter, in the Go trials, the central stimulus disappeared and, simultaneously, a target appeared (Go signal) randomly at one of two

possible opposite peripheral positions. To get a juice reward, monkeys had to reach the target within a maximum time, named upper reaction time (to discourage monkeys from adopting the strategy of excessively slowing down the RTs), and to maintain their fingers on it for 300 ms. Stop trials differed from the Go trials because at a variable delay (SSD) after the Go signal was presented, the central stimulus reappeared (Stop signal). In these instances, to earn the juice, the monkeys had to inhibit the pending movements, holding the central target for 300 ms. Monkeys were given an auditory feedback when their responses in either Go or Stop trials were correct. A countermanding session consisted of 480 trials. In the Stop trials, the successful inhibition of the planned movement critically depends on the duration of SSD. Cancelling the movements becomes increasingly more difficult as the SSD is larger. In the two monkeys, we used different values of SSDs (see (Mirabella et al., 2011) for details) with the goal to obtain a good performance, i.e., an average probability of successful suppression of the movement close to 0.5.

Data Analysis

Behavioral Performance

Probability of failure and RT distributions were calculated from the mean values obtained for each experimental session. The SD of RT distribution was obtained from the SD of RT for each experimental session.

Estimation of Mean Firing Rate

Starting from the original data set (Mirabella et al., 2011), we selected 142 neurons obtained from 53 experimental sessions in the two monkeys. Neurons selected are those with reaching-related modulation, i.e., their average FR in the RT was significantly higher (Tukey Kramer test, $p < 0.05$) than the activity measured 400 ms before target appearance. We computed mean FR responses (Fig. 3.2A) using windows of 60 ms over trials with same recent history. All references to time correspond to the midpoint of the window. Varying the size of the window did not result in significant

changes (data not shown). The significance test (Kolmogorov-Smirnov test) was computed using a 60 ms nonoverlapping window.

Estimation of Neural Variability

To calculate the across-trial variability of the neural response, we follow the method in (Churchland et al., 2011) in which the total calculated variance is approximated as the sum of the VarCE and the point process variance (PPV). VarCE is then estimated ($S^2_{\langle N_i \rangle}$) by subtracting an estimated value of PPV from the total calculated variance:

$$S^2_{\langle N_i \rangle} = Var[N_i] - \phi\overline{N_i} \tag{3.1}$$

where N_i and N_i are spike counts and the mean spike counts in epoch i for one neuron and trial history condition and f is a scaled factor that corresponds to the minimum value of the calculated Fano Factor for each neuron (see (Churchland et al., 2011) and Supplemental Experimental Procedures for details). To compute VarCE, we used the same time window as in the estimation of the mean.

To calculate VarCE in the six history conditions shown in Fig. 3.2D, we averaged the value of VarCE in the interval between 80 ms and 410 ms after the Go signal onset. We used this range because it is the time interval in which VarCE in a Go trial is significantly different when preceded by a Go trial than when preceded by a Stop trial (Fig. 3.2B). The significance test (Kolmogorov-Smirnov test) was computed using a 60 ms nonoverlapping window.

Model and Simulations

We used a standard neuronal model proposed by (Wilson and Cowan, 1972). It is a mean-field approximation of a realistic complex network of spiking integrate-and-fire neurons. The dynamics of the network can be described through two differential equations each of them referring to each population (pool) of neurons (in our case 'Go' and 'Stop' pools):

$$\tau \frac{dU_{go}(t)}{dt} = -U_{go} + f(\omega_{go+}\lambda + \lambda_{go} + \omega_{+}U_{go} - \omega_{-}U_{stop}) + \sigma\xi(t) \quad (3.2)$$

$$\tau \frac{dU_{stop}(t)}{dt} = -U_{stop} + f(\omega_{stop+}\lambda + \lambda_{stop} + \omega_{+}U_{stop} - \omega_{-}U_{go}) + \sigma\xi(t) \quad (3.3)$$

where U stands for the average firing rate of a pool, u stands for the different weight of the connections, λ defines external inputs to the network, and the function $f(.)$ is a sigmoidal function defined as:

$$f(x) = \frac{F_{max}}{1 + \exp\frac{-(x-\theta)}{k}} \quad (3.4)$$

where F_{max} denotes the firing rate value to which the population of neurons will saturate independently of the strength of the external input signal. In this study, we have used the values of: $\tau = 20$ ms, $\omega_{go+} = 0.70$, $\omega_{stop+} = 1$, ω_{+}=1, $\omega_{-} = 1.5$, $F_{max} = 40$ spikes/s, $k = 22$ spikes/s, $\theta = 15$ spikes/s, and $\lambda_{go} = 7.3$ spikes/s when the appearance of the Go signal is simulated, $\lambda_{stop} = 0$, and λ linearly varies its value from condition to condition following the trend in Fig. 3.4B. It can be described by the equation: $\lambda = -0.35x + 18$, where x goes from 1 to 6 to describe the trial history conditions: +3Stop, 2Stop, 1Stop, 1Go, 2Go, and +3Go. The decision was considered to end when the difference between Go and Stop pools response was above 15 spikes/s.

The fluctuations of the network are modeled by the term ξ, which adds an additive Gaussian noise (with mean 0 and variance 1) to the average firing rate. This noise represents the effects of a finite number of neurons in the network. The term $\sigma = 2$ spikes/s in our simulations.

VarCE of the simulated response of the network was calculated by estimating the spike counts from the mean firing rate of the Go pool. The spike counts were estimated by using a scale factor of 12, which depends on the population size, following a standard procedure (Albantakis and Deco, 2009; Wang, 2002). We did this scaling in order to fit quantitatively the experimental data.

Supplemental Experimental Procedures

Behavior and Physiology

Before starting the training, under aseptic surgical conditions, a head holding device, and a scleral eye coil (Robinson, 1963) were implanted. Antibiotics and analgesics were administered postoperatively. In each monkey, at the end of the training period, again under general anesthesia, a recording cylinder (18mm in diameter) was implanted in the frontal lobe in order to allow recordings over the arm representation of the dorsal premotor cortex (PMd; see (Mirabella et al., 2011)). The location of the neural recordings was confirmed by structural MRI on one monkey (monkey S) and visual inspections of the anatomical landmarks, such as the central (CS) and the arcuate sulcus (AS), on the other monkey (monkey L), after dura removal. In the present paper we will deal only with recordings obtained from PMd (see (Mirabella et al., 2011)).

During recording, animals were placed in a darkened, sound attenuated chamber and seated in a primate chair, with their head fixed in front of 21" PC monitor (CRT non interlaced, refresh rate 85 Hz, 800x600 resolution, 32 bit color depth; distance monitor-eye: 21 cm), equipped with a touch screen (MicroTouch, sampling rate 200 Hz) for touch positions monitoring. A non-commercial software package, CORTEX [1], was used to control stimuli presentation, behavioral responses and to collect neural (1000 Hz) and eye-movement (200 Hz) data. Eye movements were unconstrained but monitored by using a magnetic search coil technique (Robinson, 1963). Neural activity of single units was recorded extracellularly using a seven channel multielectrode system (Thomas Recording, Giessen, Germany). Electrodes were quartz insulated platinum-tungsten fibers (80m diameter, 0.8-2.5M impedance) and were inserted transdurally, one at a time, using microdrives (Thomas Recording, Giessen, Germany). Electrical signals were amplified, filtered (Thomas Recording, Giessen, Germany), and single unit were isolated on-line exploiting a dual time amplitude window discriminator (BAK electronics, Mount Airy, MD).

[1] www.cortex.salk.edu

Animal care, housing, and surgical procedures were in conformity with European (Directive 86/609/ECC) and Italian (D.L. 116/92) laws on the use of nonhuman primates in scientific research and were approved (no. 58/2005-B) by the Italian Ministry of Health.

Neural variability (VarCE)

To estimate the variability of the neural response we followed the method proposed by (Churchland et al., 2011). The total calculated variance from the neural response is decomposed into two components: the variance of the conditional expectation (VarCE) and the point process variance (PPV):

$$\sigma^2_{N_i} = \sigma^2_{\langle N_i \rangle} + \langle \sigma^2_{N_i|\lambda_i} \rangle \tag{3.5}$$

where N_i is the count of spikes in epoch i and λ_i is the mean firing rate in that epoch. In the right part of the equation the first term is VarCE and the second term is PPV. The term we are interested in is the VarCE ($spikes^2$) because it defines the across-trial variability of the count of spikes. To compute its value we first have to estimate the value of PPV, i.e. the within trial variability, that can be defined as the average over variances of each λ_i. To estimate its value from the observed neural response we based our calculations on the renewal theory and assume that the variance of λ_i is proportional (scaled by a factor of ϕ) to the mean spike count observed in that condition and epoch (Geisler and Albrecht, 1995; Nawrot et al., 2008). The value of VarCE was calculated by subtracting from the total measured variance the estimated value of PPV (Churchland et al., 2011):

$$s^2_{\langle N_i \rangle} = s^2_{N_i} - \phi \bar{N}_i \tag{3.6}$$

where s is used to refer to an estimation of the value of σ. The value of phi is specific to each neuron and corresponds to the highest value that ensures that VarCE is always positive. As in (Churchland et al., 2011), we calculate

the value of ϕ for each neuron and it corresponds to the minimum value of the calculated Fano Factor (Eq. 3.8) for that neuron.

The VarCE is the variance of the spike counts for each neuron and trial history minus the scaled average of the mean spike count for that neuron and trial history (PPV estimate):

$$s^2_{\langle N_i \rangle} = Var[N_i] - \phi \bar{N}_i \tag{3.7}$$

where N_i and $\bar{N}_i$ is the count and mean count for a given neuron, trial history and epoch i. Similar results (not reported) were obtained when we calculated VarCE using the union of normalized spike counts from all neurons sharing the same condition (Churchland et al., 2011).

Using the same notations as in Eq. 3.7, the Fano Factor was computed as:

$$FF_{\langle N_i \rangle} = \frac{Var[N_i]}{\bar{N}_i} \tag{3.8}$$

Mean-field approximation

Simulation of a network of integrate-and-fire neurons allows the study of the dynamics of neuronal activity, but are computationally expensive and their results probabilistic, which makes them rather unsuitable for systematic parameter explorations. Thus, in this study we adopt a well established reduced mean field rate model. The essence of the mean-field approximation is to simplify the integrate-and-fire equations by replacing after the diffusion approximation, the sums of the synaptic components by the average DC component and a fluctuation term. (Tuckwell, 1988) The dynamics of each population can be described by the population transfer function, which provides the average population rate as a function of the average input current (Brunel and Wang, 2001; Amit and Brunel, 1997; Del Giudice et al., 2003; Renart et al., 2003). Therefore we used a mean-field approximation

to explore how the different operational regimes of the network depend on the values of certain parameters.

Rate based models are used to approximate the mean response of a population of spiking neurons with similar properties. This approximation is supported by the fact that in different areas of the brain neurons are organized in populations comprising neurons with similar properties. Generally, the neurons from the same population receive similar external inputs and are coupled by synapses with a similar weight (Wilson and Cowan, 1972; Abbott, 1991). Therefore, the rate model approach must exclusively be used when describing the response of neurons of the same type and co-localized in a small volume. This small volume is called a computational unit or neuronal pool.

To obtain the equations describing the average response of a neuronal pool one has to start from the individual spiking neurons. Therefore, for simulating large-scale processes the population code is the relevant computational unit. From the spikes of a single neuron we derive a dynamical model that describes the mean activity of each pool of neurons. In a pool of neurons the mean activity $U(t)$ is defined by the number of spikes (n_{spikes}) that occur in a small time interval (Δt) divided by the number of neurons M contained in that pool and by that time interval Δt:

$$U(t) = lim_{\Delta t \to 0} \frac{n_{spikes}(t, t + \Delta t)}{M \Delta t} \tag{3.9}$$

Note that the concept of mean activity of a neuronal population is very different from the mean firing rate of a single neuron. The computation of mean firing rate of a single neuron requires that the input changes its value in a way comparable to the time window used to calculate the firing rate of the neuron. To use the mean activity of the pool avoids this problem and allows a rapid adaptation to rapid changes in the input. The meanfield approximation adds to this equation the average DC component of the synapses and a fluctuation term, simplifying the integrate-and-fire equations. The activity of the pool can then be described with a transfer

function that provides the average rate of the group of neurons as a function of their average input current. The changes in the activity of the population without external input can be described as:

$$\tau \frac{dU(t)}{dt} = -U(t) + f(U(t)) + \sigma\xi(t) \tag{3.10}$$

where $U(t)$ is a decay term, $f(U(t))$ is the activity of the neuronal pool due to the recurrent connections of the network and $\sigma\xi(t)$ is a fluctuation term that simulates the noise of the population affected by its finite size.

3.5 Supplemental figures

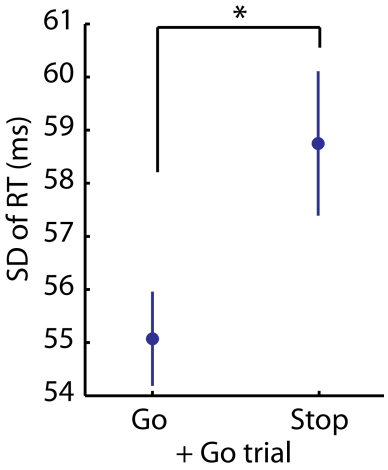

Figure 3.5: Behavioral adaptation. Standard deviation (SD) of RT in the two conditions shown in Fig. 3.1B: Go trial when preceded by a Go trial (left) or by a Stop trial (right). SD of RT is significantly different (Go + Go trial, 16,060 trials; Stop + Go trial, 6,671 trials; 53 experimental sessions for each condition; Kolmogorov-Smirnov test, $* = p < 0.01$). Error bars are SEM.

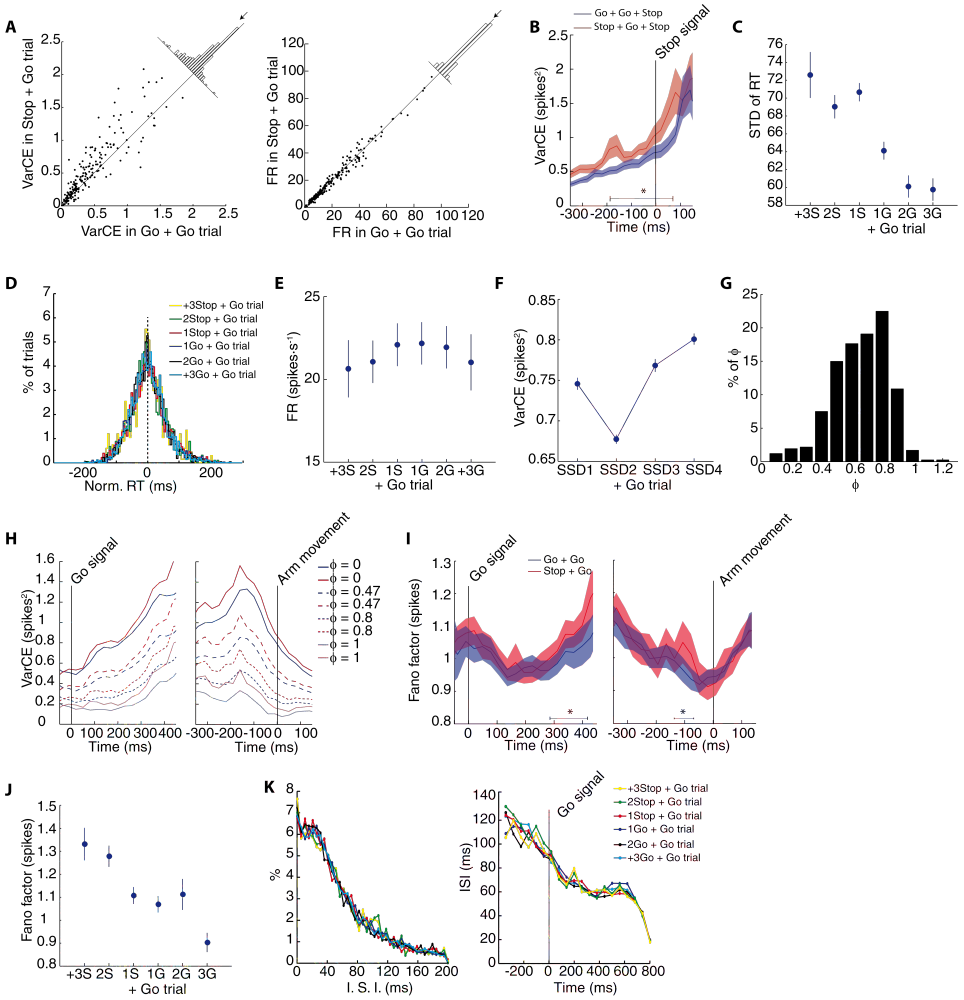

Figure 3.6: Neural response dynamics and properties. **(A)** Comparison of single unit VarCE (left) and FR (right) for 142 neurons during a Go trial when it was preceded by either a Go (x-axis) or a Stop trial (y-axis). Data points are based on the interval between 80ms and 410ms after presentation of the Go signal. The arrows indicate the mean difference in VarCE and FR between the two trial history conditions. *Left panel*, VarCE was higher when a Stop trial was preceding a Go trial, most of the data points are above the diagonal. *Right panel*, mean FR was not different in both conditions, most of the data points are on the diagonal. This figure confirms that our main results are not due to an artefact of outliers or to changes in the mean FR for each condition, but instead it shows a clear effect for the whole population in VarCE and no effect in FR.

Figure 3.6: (B) VarCE in the two conditions displayed in Fig. 3.1C of the main text: VarCE in a Stop trial when it was preceded by Go + Go trials or by Stop + Go trials. Neural activity is aligned to the Stop-signal onset. VarCE was calculated counting the spikes in a sliding window of 60 ms. Same 142 neurons, number of trials and experimental sessions as in Fig. 3.1C of the main text were considered. Shaded areas are SEM. VarCE is significantly different in the range from 190 ms before and 70 ms after the stop-signal appearance (Kolmogorov-Smirnov test, $* = p < 0.05$). Since Go trials and Stop trials are undistinguishable until stop signal presentation, the result shows that VarCE, influenced by trial history in the Stop trials, is also informative about the recent presence of a stop signal. In fact, after a stop signal is presented, the VarCE difference is not evident anymore. **(C)** Standard deviation (SD) of RT in the same six trial history conditions as in Fig. 3.2C and 3.2D. The SD of RT follows the same trend along history conditions as VarCE: the RT in a Go trial is more variable as the number of Stop trials preceding that Go trial increases as opposed to an increase in the number of Go trials preceding that Go trial. Error bars are SEM. **(D)** Distribution of normalized RT for the same trial history conditions as shown in Fig. 3.2C and 3.2D of the main text. The distributions consist of the RT minus the mean RT of the history condition and session they belong to. A Kruskal-Wallis test across conditions showed no-significant difference between the distributions ($p = 0.98$). This result shows that the changes in mean RT reported in Fig. 3.2C are caused by a shift of the whole distribution and not by a change in the shape of the distribution, i. e. the skewness, or outlier values. **(E)** Mean FR for six different kinds of Go trial sequences (as shown in the main text for VarCE, Fig. 3.2D). Error bars are SEM. The FR was averaged in the period between 80ms and 410ms after the presentation of the Go-signal. Contrary to the modulation found in VarCE, the mean FR is not modulated by the task history. This result validates our main finding: RT changes due to different task history conditions are not reflected in FR but rather in VarCE. **(F)** VarCE during a Go trial when it was preceded by a Stop trial with different SSD. The SSDs were classified in four groups based on its relative duration inside an experimental session. The averaged values across sessions are: SSD1: 234ms; SSD2: 275ms; SSD3: 307ms; SSD4: 347ms. This figure shows that VarCE in a Go trial depends on the difficulty of the previous stop trial with larger VarCE for more difficult preceding trials (longer SSD in the Stop trial preceding the Go trial). VarCE for the shortest SSD (SSD1) might be seen as a special case since it has been shown that the race model fails to accurately predict RT in this kind of trials when they are incorrect (Logan and Cowan, 1984; Mirabella et al., 2011). Error bars are SEM.

Figure 3.6: (G) Histogram of the values of ϕ calculated as the minimum value of the Fano Factor that satisfies that VarCE is positive for a neuron over all trial history cases. This is the value used to estimate PPV for each neuron and, therefore, to scale the value of the total measured variance. **(H)** Robustness of the difference between VarCE for two different conditions: Go + Go and Stop + Go (Fig. 3.2B in the main text) for different values of ϕ. When $\phi = 0$, VarCE is equal to the total measured variance of the normalized spike count (PPV is zero). Except for the values of ϕ the calculations are identical to those reported in Fig. 3.2B in the main text. *Left panel*, data is aligned to the Go signal. *Right panel*, data is aligned to arm movement. **(I)** Fano Factor (variance of the spike count divided by its mean) for Go + Go and Stop + Go conditions (as in Fig. 3.2A and 3.2B in the main text) measured in sliding windows of 60ms. The difference between both conditions was not affected by varying the size of the sliding window. Kolmogorov-Smirnov test, $* = p < 0.05$ from 290 ms to 410 ms when the data are aligned to the go-signal onset and $* = p < 0.05$ from 130 ms to 70 ms before the movement onset when the data are aligned by the arm movement. Consistent with the VarCE measure the value of the Fano Factor is higher in condition Stop + Go as compared to Go + Go over a wide time interval both when the neuronal responses are aligned to the Go signal (*Left panel*) or to the movement onset (*Right panel*). VarCE estimates the specific component of the across-trial variable responses removing its non-specific contributions due to point process variance (PPV). In contrast, the Fano Factor also considers the PPV when estimating the across-trial variability. Therefore, although both results are consistent in our case, the VarCE is a more principled measurement of the specific across-trial neural variability (Churchland et al., 2011). *Left panel*, when aligned to the Go-signal we observe a decrease in the value of the Fano Factor after Go-signal onset in both conditions. *Right panel*, when aligned to the movement onset the Fano Factor decreases before the movement onset. Similar patterns have been observed in the parietal area LIP during a decision task (Churchland et al., 2011). Shaded areas indicate SEM. **(J)** Fano factor calculated in a Go trial for same history conditions as in Fig. 3.2C and 3.2D of the main text. The Fano factor was calculated as the average Fano Factor in the time interval from 290ms to 410ms, that corresponds to the period in which the Fano factor is significantly different for Go + Go trial and Stop and Go trial conditions. Error bars are SEM. **(K)** Inter-spikeinterval (ISI) for six different kinds of task sequences (as shown in the main text for VarCE, Fig. 3.2C and 3.2D). *Left panel*, histogram of quantities of ISI found. *Right panel*, temporal dynamics of the ISI. Both panels are obtained with the data aligned to Go signal. These results show that there is no difference in the ISI distribution between the six task conditions supporting the main finding that the difference in neural variability expressed in the VarCE cannot be due to changes in the within-trial variability.

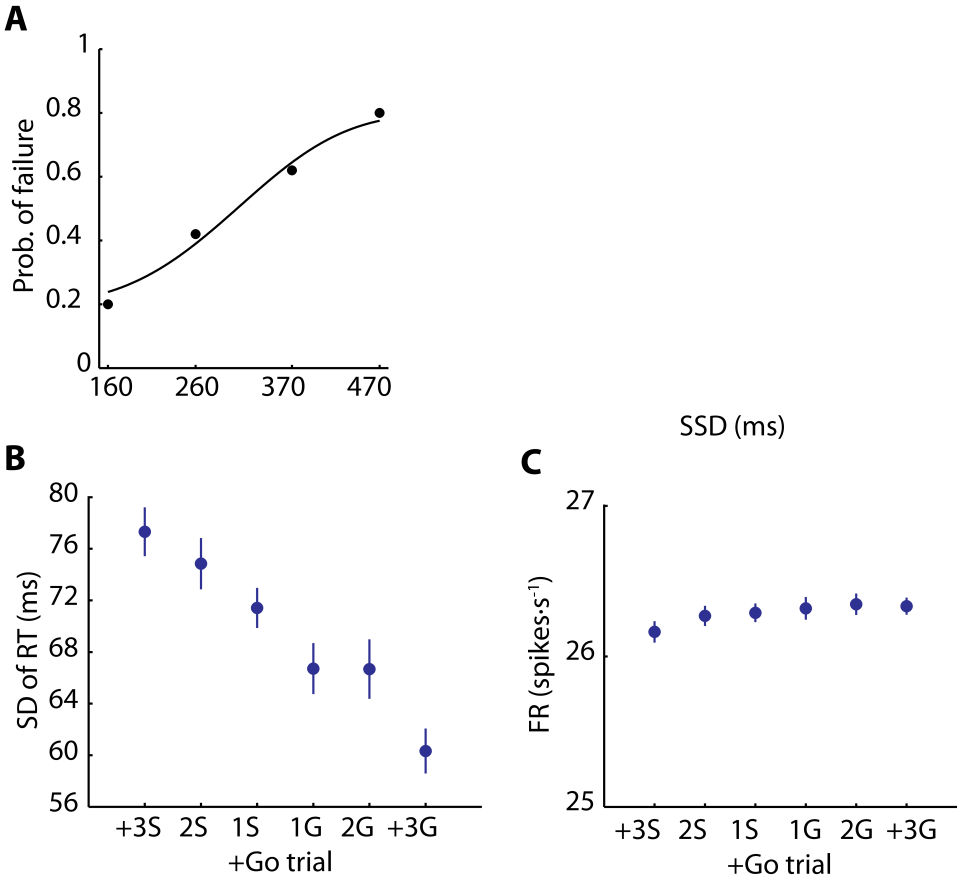

Figure 3.7: Simulated behavioral results. **(A)** Inhibition function, i.e. the probability of failure to cancel a movement for different stop signal delays (SSD). The data points are from the simulations and the curve is the best fit using a Weibull function. The results are the average over 1,000 Stop trials preceded by: 3 or more Stop trials (+3S), 2 Stop trials (2S), 1 Stop trial (1S), 1 Go trial (1G), 2 Go trials (2G), 3 or more Go trials (+3G). The value of λ varies according to each trial history condition (see Fig. 3.4B). Each condition was simulated with four possible SSDs (160 ms, 260 ms, 370 ms and 470 ms). The parameters are the same as those defined in Methods section. λ_{stop} is set to 7.3 when the Stop signal appears, i.e. after the SSD. **(B)** Standard deviation (SD) of RT in the simulations (scaled to the range of the real data, correction factor of 8) shows the same trend as observed in the real data for different trial history conditions (Fig. 3.6C). The SD of RT was modulated by trial history consistent with the impact of VarCE itself. Error bars are SEM. **(C)** Mean firing rate along six different sequences observed in the simulations. As in Fig. 3.6E, the firing rate was averaged in the period between 80 ms and 410 ms after the onset of the Go-signal. As in the physiological data, the FR was not modulated by trial history. Error bars are SEM.

Chapter 4

Decision-making depends on an urgency signal modulated by context

How wonderful that we have met this paradox. Now we have some hope of making progress

Niels Bohr

The content of this chapter has not been published yet; it is part of a manuscript in preparation (Marcos et al., 2014a). The present study aims to provide a plausible mechanism that is able to explain two new experimental datasets that seem to lead to contradictory results (Huk and Shadlen, 2005; Thura et al., 2012). So far, decision making has been generally examined in tasks in which decisions are based on a perceptual property of a constant stimulus. In contrast, these new experimental paradigms intend to broaden the understanding of decision making into two different contexts where sensory evidence changes over time. However, while early evidence seems to have a stronger impact on decision making than late evidence in one context, the contrary occurs in the other. What is more, to date, it has been difficult to reconcile these two findings, and so that is what we intend to do in this study. We suggest that experimental results can be explained through the

same neural mechanism that is regulated by context. We use the theoretical decision-making model seen in the previous chapter (Chapter 3), and, besides, we extend it with a task-dependent urgency signal. Thus, decisions are the result of combined integration and urgency mechanisms which enable us to demonstrate that the apparently paradoxical results simply arise from differences in the strategy adopted by monkeys/subjects according to context, and at the same time- this strategy is regulated by the urgency signal. To validate our approach, we have designed a new experimental task where human subjects engaged in one context are confronted with trials from a different context. The experimental data confirms the predictions of our model.

The abstract has been published as a Society for Neuroscience abstract (Marcos et al., 2012a):

> In most real world situations we have to make decisions while information changes over time. Many models have successfully explained decision-making in tasks with constant information, but not in some tasks where the information varies over time. Here, we propose that behavior in any task depends on both the dynamics of the decision making network driven by perception and memory combined with a context-dependent regulation of urgency.
>
> Two recent experimental paradigms have been especially important for distinguishing between decision-making models: the variable motion discrimination (VMD) task and the pulse task. In both, subjects are required to detect the net direction of dots moving on a screen. In the VMD task the net motion increases or decreases every 225ms whereas in the pulse task this only occurs during 100ms pulses early or late in the trial. Subjects engaged in the VMD task do not show any difference in their performance between trials when a long positive or negative pulse, i.e. in favor or against the final net motion, is presented in the early epoch. In contrast, in the pulse task subjects do show a

change in their performance between trials with a short early positive and negative pulse, but not when the pulse is presented late in a trial. These results suggest that in the VMD task early information has weaker influence on the decision than late information, while the opposite occurs in the pulse task. Existing decision-making models make different predictions for each task: bounded integrator models predict that early information has more or same impact on the decision process whereas linear urgency models predict the opposite.

Using an attractor network for decision-making we show that a task-dependent urgency signal could explain the two paradoxical experimental results. The model is based on the Wilson and Cowan equations and comprises two populations of interconnected excitatory neurons that compete to select between two motion directions. It has an urgency signal that over time changes the energy needed to push the network to an attractor state. We show that to explain the behavior in each task, the urgency signal has to be regulated depending on the context. A signal with high starting value of urgency can explain the pulse task whereas a signal with low starting value of urgency explains data in the VMD task. We hypothesize that such a task-dependent change in urgency is related to a subjects experience of how accuracy improves over time, which is faster in the pulse task than in the VMD task. This has been tested using experiments in which trials of one task are interleaved with the other.

4.1 Introduction

Decision making has been an important focus of attention in the last decades (Roitman and Shadlen, 2002; Gold and Shadlen, 2007). It has been mainly studied in tasks in which monkeys or human subjects have to discriminate between two or more options based on a perceptual feature of a stimulus

(Roitman and Shadlen, 2002; Palmer et al., 2005; Churchland et al., 2008). Although these tasks have proven to be fundamental to understand the basic mechanisms of the decision-making process, they consider that the sensory evidence remains constant over time. However, this is not the case in most real-world situation where, not infrequently, information varies over time. In some of these cases, choices that might seem to be clearly supported by sensory evidence could be dismissed because of new upcoming evidence supporting an alternative one. Imagine a situation in which, for instance, you are planning to take the next exit (exit A) in the roadway to go home. Just when you are about to do it, you see a long queue of cars stopped in that exit. You might change your mind and continue until the next exit (exit B) so as to avoid a long retention and the subsequent delay. So, even if you were about to take exit A, this new information might make exit B become your final choice. Now, here is a critical question: do the mechanisms that have proven to explain decisions with constant information generalize in these cases?

Many models have successfully explained decision making in tasks with constant information (exit A being always better option than exit B). On the one hand, "bounded integrator" models (Stone, 1960; Laming, 1968; Ratcliff, 1978; Usher and McClelland, 2001; Wang, 2002; Mazurek et al., 2003; Bogacz and Gurney, 2007; Link and Heath, 1975) propose that evidence is sampled from the environment and accumulated over time until the accumulation reaches a threshold and the decision is made. In contrast, the urgency-gating model (Reddi and Carpenter, 2000; Cisek et al., 2009) proposes that the process of decision making is the result of a multiplication between sensory evidence and an urgency signal that is internal to each subject and grows linearly over the course of time. When the result of this multiplication is above a threshold, the decision is made. Even if both views propose a different mechanism by which decisions are formed, they can both explain the behavioral and neural data that has been observed in tasks with constant sensory information (Roitman and Shadlen, 2002; Churchland et al., 2008). For this reason, the mechanisms underlying decision making are not well-known yet. To shed some light on this issue, experiments

with information that varies over time have been proposed during the last years, but interestingly, in these cases both theoretical views have failed to completely explain the data (Huk and Shadlen, 2005; Cisek et al., 2009). That is why we propose a reconciled view in which decision making depends on the two of them: integration and urgency, i. e. the accumulation of sensory evidence is modulated by a context-dependent urgency signal.

One of the generally accepted standard tasks to study perceptual decision making is the random motion discrimination (RDM) task (Roitman and Shadlen, 2002; Hanks et al., 2006; Britten et al., 1992). In this task, dots moving in different directions are presented on a screen. In each trial, the experimenter controls the percentage of dots that coherently move towards the same direction (motion coherence), which is kept constant during the course of the trial. In the last years, to distinguish between decision-making models, this task has been extended by making the amount of sensory evidence dynamic over the course of a trial, i. e. the amount of motion coherence changes at different time onsets along the duration of a trial (Huk and Shadlen, 2005; Thura et al., 2012). Following this new line of research, two experimental tasks have been especially relevant because of their apparently paradoxical results: the RDM task with pulses (Huk and Shadlen, 2005) and the variable motion discrimination (VMD) task (Thura et al., 2012). In both tasks, subjects are required to detect the net direction of dots moving on a screen. In the RDM task with pulses, the net amount of motion coherence either increases or decreases throughout one short period of time. In contrast, in the VMD task, the net amount of motion coherence changes more than once within a single trial. Performance in the RDM task with pulses directly depends on the time onset of the pulse, thus being more influenced by an early appearance of a pulse than by a late one. Contrarily, performance of subjects in the VMD task is not affected by long pulses appearing early in a trial. These results suggest that, in the RDM task with pulses, late information has weaker influence on the decision than early information, while the opposite occurs in the VMD task. Likewise, existing decision-making models make different predictions for each task: bounded integrator models predict that early information has more or equal impact

than late information on the decision process (Huk and Shadlen, 2005), whereas linear urgency models predict the contrary (Cisek et al., 2009).

To explain the two paradoxical experimental results, we use an attractor network of binary decision making with a task-dependent urgency signal added. The model is based on the Wilson and Cowan equations (Wilson and Cowan, 1972) and comprises two populations of interconnected excitatory neurons that compete through mutual inhibition to choose one motion direction from two possibilities. It has an urgency signal which gradually adjusts the energy needed to push the network to an attractor state. By means of this, we corroborate that a regulation of the urgency signal according to context explains the observed behavior in each task. A signal with a high starting value of urgency (the network can be pushed into an attractor relatively easy from the start of the trial) explains the RDM task with pulses, whereas a signal with low urgency (the network needs high amount of energy to be pushed into an attractor state at the beginning of each trial) accounts for data in the VMD task. We presume that such a task-dependent change in urgency is related to the subjects experience during the task. They might try to increase the speed-accuracy trade-off by slowing down their decisions during the VMD task. One possible explanation for this would be that they need longer time to accumulate sensory evidence, since it continuously changes within a trial. In contrast, in the RDM task with pulses, the amount of motion coherence changes less often, and so it probably leads to a different decision strategy.

In order to test our theoretical approach, we designed an experimental task in which participants have to detect motion in VMD trials. During the execution of this task, a few number of pulse trials are interleaved. This way we proved that, in coherence with our hypothesis, the effect that pulses exert on performance reverses. In other words, late pulses have a stronger impact than early pulses, it being consistent with the idea that subjects modulate their urgency based on the context they experience.

4.2 Materials and Methods

Decision making with information that varies over time

Activity of neurons in areas such as the lateral intraparietal cortex (LIP) (Roitman and Shadlen, 2002; Churchland et al., 2008), the superior colliculus (Hanes and Schall, 1996b; Schall and Thompson, 1999), the frontal eye field (Thompson et al., 1997; Bichot et al., 2001) or prefrontal cortex (Kim and Shadlen, 1999) has shown to be correlated with the formation of a decision. However, it is not clear if the build-up activity of these neurons responds to an accumulation of evidence or to an urgency to commit to a decision as time progresses. With a view to further clarifying the mechanisms underlying decision-making, two new experiments have been proposed: the RDM task with pulses (Huk and Shadlen, 2005) and the VMD task (Thura et al., 2012). Both experimental paradigms deal with non-constant perceptual information, which allows us to investigate how performance and neural activity are biased according to the temporal dynamics of the perceptual input.

Huk and Shadlen (2005) analyzed the impact of modifying, for a short period of time, the total net motion coherence during a RDM task. Two monkeys performed the task while behavior and neural activity from neurons in LIP were recorded. A dynamic random-pixel texture appeared as a background to the moving dots presented on a screen (Fig. 4.1A). In all trials, the texture contained noisy pixels moving like dark television snow (*no-pulse* trials). In two thirds of the trials, the texture was moving during 100 ms either in the same direction (*positive-direction pulse* trials) as the moving dots or in the opposite direction (*negative-direction pulse* trials). The pulse in the texture occurred at one of five possible onsets: 100 ms, 150 ms, 211 ms, 287 ms and 392 ms. As it could be observed, the behavior of monkeys was indeed influenced by the motion direction of the texture: their responses became more accurate and the reaction time (RT) became shorter when the texture moved in the same direction as the total net motion of the moving dots, while, when the texture moved towards the opposite side, the

contrary occurred. It should be noted that an early pulse trial generated a stronger modulation of the activity of neurons in LIP than a late one.

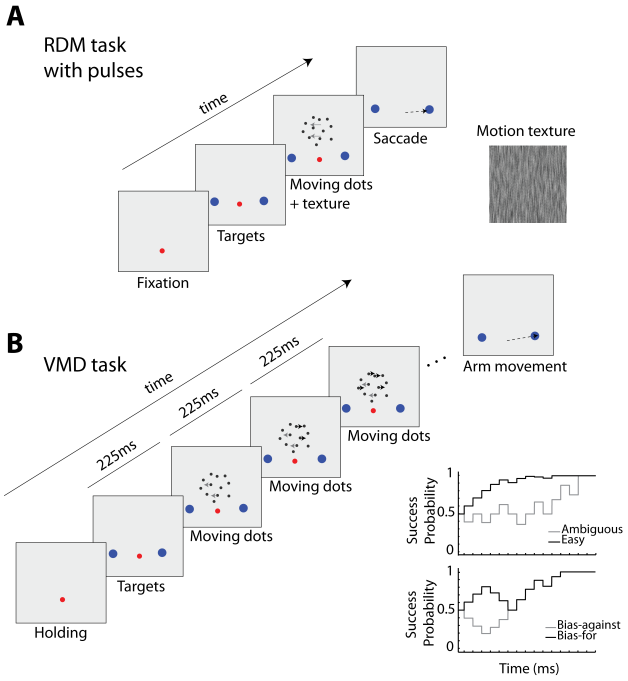

Figure 4.1: Experimental tasks. **(A)** RDM task with motion texture (Huk and Shadlen, 2005). All trials started with the appearance of a circular dot. After fixation, two targets appeared in opposing sides of the screen. After a random delay, the moving dots appeared in a 5 diameter aperture. Monkeys had to indicate, with a saccadic eye movement to one of the two targets, the net direction of the moving dots (RT task). In all trials, a texture with noisy pixels randomly moving (motion texture) was shown in the background. In one third of the trials, the pixels of the texture favored (for a short period) either the same direction of motion of the dots (positive-direction pulse trials) or the opposite direction (negative-direction pulse trials). The pulses had a duration of 100ms and occurred at one of five possible time onsets. **(B)** VMD task (Thura et al., 2012). Similar to what happened in the RDM task, a trial started with a circular central stimulus on the screen. Subjects were requested to place a cursor inside it. Then, two targets appeared in its perimeter. Immediately, random moving dots came into view on the center of the screen (within a 6-cm diameter circle). After 200ms, the dots started to increase their total net motion direction in steps of 3% of coherence that occurred every 225ms (with a maximum number of 15 steps). At that point,subjects were asked to make a prediction about towards which direction the dots would be moving at the end of the trial.

Thura et al. (2012) investigated the behavior of human subjects performing a VMD task. In this case, the amount of motion coherence was increased by 3% towards one of two possible directions of motion every 225 ms (Fig. 4.1B). Subjects were asked to predict the direction towards which most of the dots would be moving at the end of the trial. Moreover, trials with specific profiles where interleaved with trials where the motion coherence changed randomly. For instance, in easy trials the motion coherence incremented consistently towards one direction whereas, in an ambiguous trial, both directions of motion were similarly favored for almost the whole duration of the trial (small panel in Fig. 4.1B). Out of these special cases, two trial profiles were especially relevant because they leaded to contradictory behavioral results when compared to the ones observed in Huk and Shadlen (2005): *bias-for* and *bias-against* trials (small panel in Fig. 4.1B). In the bias-for trials, the amount of motion coherence increased towards the same direction in the first three steps; nevertheless, the motion coherence increased towards the opposite direction during the following three steps and, in the end, the motion coherence increased towards the same direction as it has done in the first three steps. The bias-against trials shared the same dynamics from the seventh step onwards although, in the first six steps, the behavior of the moving dots was reversed. Interestingly, humans did not show any bias in their behavior based on how the perceptual information was presented at the beginning of the trial. Neither RT nor accuracy were significantly different in these two kinds of trials. In Thura et al. (2012), subjects could be under time-pressure or non-time pressure conditions in which they had a maximum of 3s or 8s to respond, but we did not consider these two conditions in the development of our study, since its aim is merely to explain the main results, which do not depend on this distinction.

Decision-making models

The build-up activity observed in neurons recorded while monkeys perform a RDM task (Roitman and Shadlen, 2002) has been consistent with a variety of already existing decision-making models. Even though these models

give different explanations of the same observed phenomena, they can be described in general terms as:

$$\frac{dx_i(t)}{dt} = -Lx_i(t) + \lambda_i(t)u(t) + \sigma\xi(t) \tag{4.1}$$

where x_i is the decision variable that accumulates evidence in favor of one choice and against an opposite one, L is a leaky term that defines how quickly the decision variable can change its value, λ is the net sensory evidence, $u(t)$ is an urgency signal, ξ represents Gaussian noise with mean 0 and variance 1 and σ modulates the variance of the noise. The proposed models only differ from one another on the way they define each term of the equation. In all cases, the decision process is considered to end when a specific level of activity is reached (θ).

For instance, a well-known "bounded integrator" model is the drift-diffusion model (Stone, 1960; Laming, 1968; Ratcliff, 1978; Ratcliff et al., 2003; Smith and Ratcliff, 2004). In this model, $L = 0$ (infinite constant of integration) and $u(t) = C$ where C is a constant. In this case, x_i integrates sensory evidence with equal weight for the whole duration of a trial. An alternative "bounded integrator" model is called the leaky integrator model (Usher and McClelland, 2001) where $L \in]0, 1]$. Therefore, the decision variable in this model forgets about early sensory evidence in a trial as time progresses. Yet another alternative view is the urgency model (Reddi and Carpenter, 2000; Cisek et al., 2009), which is an extreme case of the leaky integrator. It defines that $L = 1$ and hence x_i does not depend on past sensory information but exclusively on current available information; thus, the urgency signal is described as $u(t) = gt$ where g defines the slope and t refers to time from the start of the decision process onwards.

In our simulations, we used the drift-diffusion model with parameters: $\theta = \pm 100$ and $\sigma = 2$; and the leaky integrator model with parameters: $L = 0.02$, $\theta = \pm 8$ and $\sigma = 0.45$. The trials simulated bias-for, bias-against and ambiguous cases, in accordance with the profile in the small panel in

Fig. 4.1B, using steps of 0.032 occurring every 225ms. The aforementioned simulations consisted of 1,000 trials per condition.

Attractor network with urgency

We propose an alternative mechanism to make decisions based on a combination of attractor dynamics and urgency. The model is an extension of the Wilxon and Cowan equations (Wilson and Cowan, 1972). It consists of two groups of excitatory neurons that have recurrent connections and that are connected to each other through an inhibitory group of neurons. Each excitatory group of neurons is sensitive to one of two possible motion directions, that is, left or right in our simulations (Fig. 4.2). An urgency signal changes the way the activity of the neural groups evolves over time, directly adjusting the energy needed to reach an attractor state. The activity of each of these neural populations is defined as:

$$\tau \frac{dU_{left}(t)}{dt} = -U_{left}(t) + f(\lambda_{left} + \omega_{+} U_{left} - \omega_{-} U_{right}) + \sigma\xi(t) \quad (4.2)$$

$$\tau \frac{dU_{right}(t)}{dt} = -U_{right}(t) + f(\lambda_{right} + \omega_{+} U_{right} - \omega_{-} U_{left}) + \sigma\xi(t) \quad (4.3)$$

where U defines the mean neural activity, τ is the time constant that determines how quickly the activity of the neural group changes its activity, λ_{left} and λ_{right} are the inputs associated with left and right respectively, ω_{+} is the recurrent connection, ω_{-} is the inhibitory connection and $f(.)$ is a sigmoid function defined as:

$$f(x) = \frac{F_{max}}{1 + e^{(-(x-b)u(t))}} \quad (4.4)$$

where F_{max} is the maximum value that the function can achieve, b is the center of the function and $u(t)$ is the urgency signal that defines the slope of the sigmoid function at time t. $u(t)$ changes its value according to a linear equation:

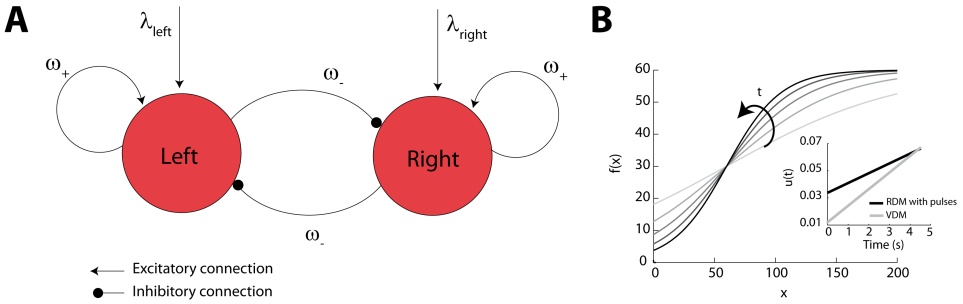

Figure 4.2: Attractor network with urgency. **(A)** Scheme of the architecture of the network. Two excitatory neural populations with recurrent excitatory connections that mutually inhibit each other. The neural populations are sensitive to either left or right motion. **(B)** Illustration of how the sigmoid slope changes as time progresses. *Small panel*, urgency signal in the RDM task with pulses (black) and in the VMD task (gray).

$$u(t) = u_0 + u_s t \tag{4.5}$$

where u_0 is the initial value of u, u_s is its slope and t relates to time. Fig. 4.2B illustrates the influence of the value of $u(t)$ in the sigmoid function. As u increases, the slope of the sigmoid also does, implying that the variation in the activity of the neural groups is more sensitive to changes in their input as time progresses. Fig. 4.13 shows the changes in the state of the network when u changes.

MT neurons have shown to be tuned to a specific motion direction of the dots, increasing their activity for a concrete direction and decreasing it for the opposite one. Area MT projects to LIP where the decision seems to be formed. The perceptual input to our decision-making model simulates the increase and decrease in activation observed in MT neurons:

$$\lambda = \lambda_0(1 \pm \frac{c}{100}) \tag{4.6}$$

where c is the total motion coherence, λ_0 is the baseline activity that becomes modulated, and the signs + and - refer to the preferred or non-

preferred direction of motion. When a positive-direction pulse is simulated, a specific amount of motion is added to the total net motion coherence input of the neural population sensitive to that direction and subtracted from the other one. In our simulations, we used pulses of 10% of motion coherence as in the modeling work of Huk and Shadlen (2005) and steps of motion of 3.2% for the VMD task.

The noise of the network was modeled with an additive Gaussian noise of mean 0, variance 1 defined by $\xi(t)$ and a scaled factor of $\sigma = 8$ spikes/s.

The decision process was considered terminated when the difference in activity between the two neural populations was above a predefined threshold (θ). In our simulations, we used the parameters: $\omega_+ = 1$, $\omega_- = 1.5$, $\tau = 20$ms, $F_{max} = 60$ spikes/s, $b = 60$ and $\theta = 57$ spikes/s. The initial value of the urgency signal u_0 and its slope u_s were different in each condition: $u_0 = 0.0337$ spikes/s and $u_s = 0.0072$ spikes/s in the RDM task with pulses and $u_0 = 0.012$ spikes/s and $u_s = 0.0122$ spikes/s in the VMD task. Small panel in Fig. 4.2B shows how the urgency signal changes over time for each context. Each condition was simulated using 5,000 trials.

Two different procedures were used in order to compare the results of our simulation with the data published in Huk and Shadlen (2005) and in Thura et al. (2012). To account for RTs in the RDM task with pulses, we added non-decision delays to the decision times (DTs) obtained from the model. To that purpose, we used the same procedure as in (Thura et al., 2012), adding 500ms to the results because this is the average time it took monkeys to make their decisions in the easiest trial (51% of motion coherence). This amount of motion coherence is thought to be nearly non-ambiguous and, therefore, the response is produced immediately. Any delay observed in these trials is considered to be due to non-decision latencies (Thura et al., 2012). In the case of the VMD task, we directly compare our results with the results published in Thura et al. (2012) since they report DTs in their study (the authors subtract non-decision time delays from their data).

Experimental paradigm

Together with the University of Montreal, we designed an experiment to validate the predictions obtained with our model. The experiment and data analyses were performed in Montreal. We show the results here to demonstrate that human subjects behave as predicted by our model (Carland et al., 2014).

n=32 subjects (17 females) participated in a RDM task with pulses and in a VMD task with interleaved pulse trials (positive-direction pulses in both cases). All participants gave their consent and the experimental procedure was approved by the University of Montral ethics committee. Each trial started when the subject placed the cursor in a small starting circle of 1cm diameter. This action was followed by the appearance in the center of the workspace of 200 dots randomly moving inside a circle of 6cm diameter together with 2 circle targets of 3.5 cm diameter placed 180 apart at a distance of 6cm from the center. In a VMD trial, after 200m the motion strength was increased by 3% towards one of two possible directions (top or bottom), 6 dots changed their direction of movement accordingly. The same procedure was applied every 225ms up to a maximum of 15 times. Only the resultant net motion was displayed; in other words, the dots always moved coherently towards one possible direction of motion. Subjects had to respond by choosing one of the two circle targets that were associated with the two potential directions of movement within 3s after the start of a trial. Subjects responded by doing an arm planar movement, using a cordless stylus pen embedded within a plastic cylinder on a digitizer tablet (CalComp). The tablet recorded the position of the pen with a temporal precision of 125Hz and a spatial accuracy of 0.013cm. All information from the task was shown using the reflection of a LCD monitor (60Hz refresh) in a half silvered mirror suspended 16cm above and parallel to the digitizer tablet so as to create the illusion that the image was floating on the plane of the tablet.

All subjects participated in 3 sessions: a session where they performed a RDM task with pulses ("blocked" condition"), and two sessions consisting

in a VMD task with interleaved pulse trials ("interleaved" condition). They had to achieve a total of 560 correct trials in the blocked condition and a total of 500 correct trials in the interleaved condition. The blocked condition consisted in trials with common baseline motion coherence of 3%. In 40% of trials the motion did not vary (no-pulse trials), but the remaining 60% of the trials contained a positive-direction pulse that doubled the motion coherence (to 6%) for 100ms, and that could occur at one of four possible onsets: 100ms, 200ms, 400 ms or 1600ms. In the VMD task, 20% of the trials were RDM trials, including no-pulse and pulse trials, that were identical (in percentage, duration and pulse onsets) to those in the RDM task with pulses. The remaining 80% of the trials comprised typical VMD trials.

The sequence of trials in each session was pseudorandomized (predefined sequence that was equal for all subjects). Subjects were given a maximum of 3000ms to report their decision, although they did not generally take more than 1800ms. All sessions started with 40 trials of 50% of motion coherence, that were used as baseline to compute decision times.

Data analyses

The psychometric functions are fits of the logistic equation:

$$P_{correct} = \frac{1}{1 + e^{(-(\beta_0 + \beta_1 c + \beta_2 I))}} \tag{4.7}$$

where β are fitted parameters, c is the total net motion of the dots and I is 0 for no-pulse trials, -1 for negative-direction pulse trials and 1 for positive-direction ones. The shift of the psychometric function was calculated by dividing the coefficient related to motion β_2 into the one related to the pulses effect β_1.

In the VMD task, the probability that the motion of the dots will favor right (R) at the end of the trial is given by (Cisek et al., 2009):

$$P(R|S_R, S_L, S_C) = \frac{S_C!^{min(S_C, 7-S_L)}}{2^{S_C}} \sum_{i=0} \frac{1}{i!(S_C - i)!} \tag{4.8}$$

where S_R is the number of steps that favored right, S_L is the number of steps that favored left and S_C is the number of steps that still remain.

To compare RTs in bias-for and bias-against trials, we consider RTs that are longer than 1125ms. This way, responses produced during the first six steps are not taken into account (Cisek et al., 2009; Thura et al., 2012).

4.3 Results

First, we briefly expound and explain the reasons why previous linear models of decision-making cannot account for the data obtained in these two experimental paradigms. Second, we show the results obtained with our proposed model. Our simulations are focused on the main results shown in Huk and Shadlen (2005) and Thura et al. (2012). On the one hand, for the RDM task with pulses, we prove that our model is able to explain the effect of pulses in RT and accuracy, as well as their influence on behavior according to its time onset (early pulses have greater influence than late ones). On the other hand, for the VMD task, we show that our model can explain the absence of differences between RT in bias-for and in bias-against trials. Consequently, our model predicts that RT and accuracy are greater affected by late pulses than by early ones in this context. Last, we provide experimental evidence supporting the validity of our model.

Linear models of decision making

The drift-diffusion model has been used in an attempt to explain the effect that pulses exert on behavior (Huk and Shadlen, 2005). In spite of the fact that the model could explain the overall effect of pulses, it could not account, however, for the greater effect on behavior and neural activity that early pulses had when compared to late pulses. Instead, the drift-diffusion model predicted that the effect of pulses is nearly the same regardless of their time onset. Consistent with this time-shift invariance, (Cisek et al., 2009) revealed that the drift-diffusion model fails to explain the lack of effect on RT due to bias-for and bias-against trials. Fig. 4.3A presents a

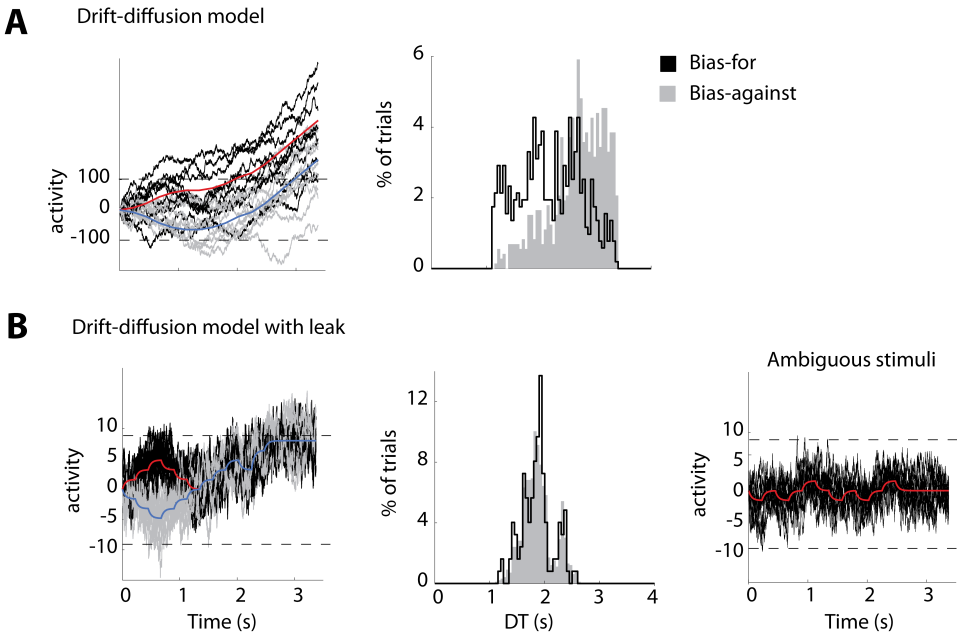

Figure 4.3: Simulations of two linear models. **(A)** Drift-diffusion model. *Left panel*, activity of drift-diffusion model during bias-for and bias against trials. Mean activity in bias-for trials (red) is higher than mean activity in bias-against trials (blue) in the 6th step (1125ms). At that point, the evidence towards each direction is the same, but it should be noted that the model inclines towards a different one in each type of trial. This inequality in activation results in different predicted behavior. *Right panel*, distributions of RT when only trials that finished after 1125ms are considered (K-S test, $p < 0.001$). **(B)** Drift-diffusion model with leak. *Left panel*, activity of the model during bias-for and bias against trials. A strong leak accounts for the same mean activity in bias-for (red) and bias-against (blue) trials. *Middle panel*, RT time distribution for decisions made after 1125ms in these two kind of trials (KS test, $p > 0.05$). *Right panel*, Simulation of the model in ambiguous trials. It is very unlikely that the accumulation reaches a decision threshold due to the high leak value needed to account for RTs in bias-for and bias-against trials.

simulation of this model using bias-for and bias-against trials. As observed, the model predicts shorter RTs for bias-for trials than for bias-against trials. The reason for this is that, when the same level of sensory evidence has been reached for both directions of motion (step 6, see Small panel in Fig. 4.1B), the accumulation of evidence favors the correct option in bias-for trials, but

the opposite option for bias-against trials. Hence this leads to a significant difference in RTs (KolmogorovSmirnov test, $p < 0.001$, in our simulations).

One could think that a leaky integrator, in which early information is forgotten, may account for the main result in the VMD task. Nevertheless, though this could be the case (Fig. 4.3B), the leak would need to be so strong that the accumulation of evidence would hardly reach the threshold when ambiguous trials were encountered (see small panel in Fig. 4.1B and right panel in Fig. 4.3B).

An urgency gating model has successfully explained the lack of effect of an early long pulse (bias-for and bias-against trials) on behavior (Cisek et al., 2009). However, in disagreement with Huk and Shadlen (2005), this model predicts that late information has greater impact on the decision-making process than late information, since the sensory evidence is directly multiplied by a growing signal.

So far, there is no consistency between the new experimental data with changing sensory evidence and decision-making models. That is why we propose a model which, by means of a combination of integration and urgency signal, is able to reconcile the experimental and theoretical views.

Task-dependent urgency

Our model was simulated in two conditions: the pulse task and the VMD task. All parameters excepting the initial value of urgency, u_0 and its slope, u_s, were kept constant in both tasks. We propose that these two values change due to the context/task in which subjects are involved. More specifically, we suggest that during the RDM task with pulses, the initial state of the urgency (u_0) should be higher than it is in the VMD task, as the amount of changes in the information provided is very low. On the contrary, the slope of this function (u_s) is steeper in the VMD task than in the RDM task with pulses. Both signals intersect in a high value of u towards the end of the trial, forcing the system to make a decision. The mechanism by which the urgency signal is formed over experience has not been addressed here,

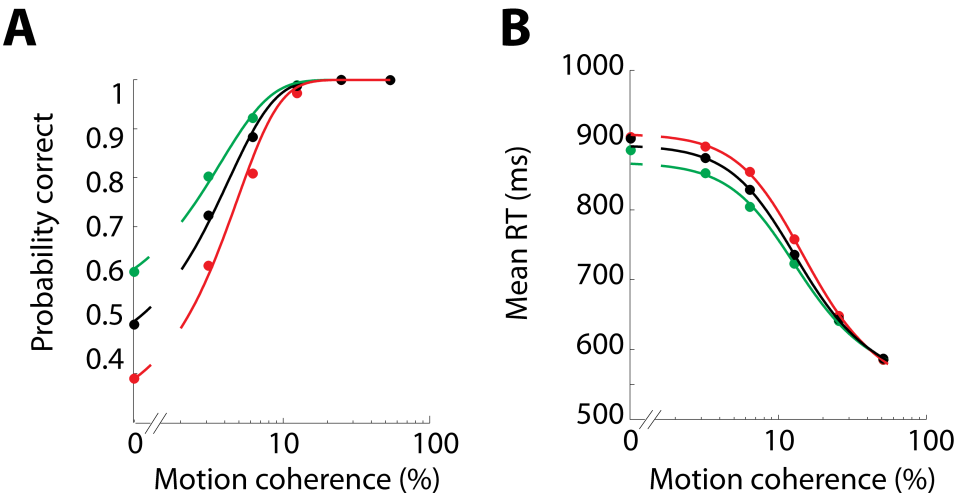

Figure 4.4: Pulse effects on behavior during the RDM task for positive-direction pulse (green), negative-direction pulse (red) and no-pulse trials (black). **(A)** Pyschometric functions. Probability of correct along different values of motion coherence. Positive-direction pulses shifted the pyschometric function by an amount equivalent to adding 1.41% or 1.54% of motion in the positive and negative-direction pulse trials respectively. **(B)** Chronometric functions. Mean reaction times along different values of motion coherence.

since it is out of the scope of the current study.

RDM task with pulses

To simulate the same conditions as in (Huk and Shadlen, 2005) we used six levels of motion coherence (0%, 3.2%, 6.4%, 12.8%,25.6% and 51.2%), pulses of 100ms at one of five possible time onsets (100ms, 150ms, 211ms, 287ms and 392ms) after the start of the trial. As it is the case with the real data (Huk and Shadlen, 2005), the probability of correct increases when the motion pulse has the same direction as the correct one (positive-direction pulse) and decreases when the opposite occurs (negative-direction pulse). This can be observed in the shift of the psychometric function towards the left, when a positive-direction pulse is simulated, and by a shift towards the right when a negative-direction pulse is simulated, in comparison with the no-pulse condition (Fig. 4.4A). In our simulations, the shift is equivalent to

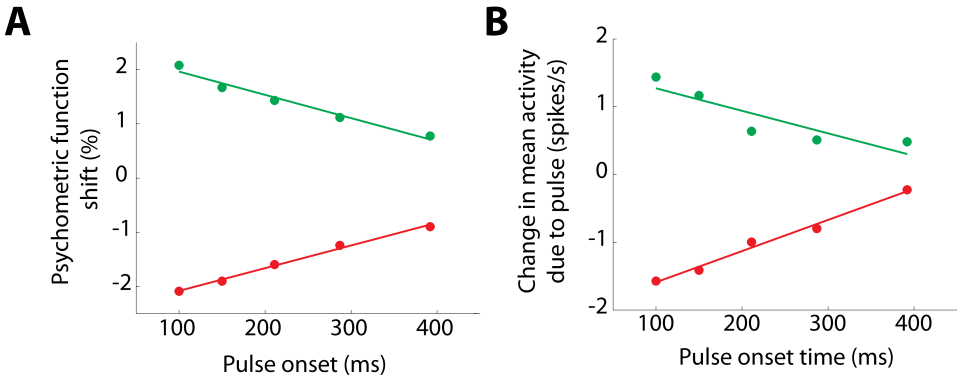

Figure 4.5: Pulse onset effect during the RDM task. **(A)** Increase or decrease in performance due to the time onset of a positive (green) or negative-direction (red) pulse when compared to no-pulse trials. Positive-direction pulses augment the probability of success while the opposite effect is observed with negative-direction pulses. **(B)** Changes in activity under the same conditions as in (A). Same effect as with performance observed on the mean activity of the neural population.

1.41% of motion coherence, added to the whole duration of the trial in the positive-direction pulse, and to an addition of 1.54% of motion coherence in the negative-duration pulse condition ($\pm$ 1.6% in (Huk and Shadlen, 2005)). Likewise, RTs are shorter in trials containing a positive-direction pulse and longer in trials with a negative-direction pulse in comparison with the no-pulse condition (Fig. 4.4B)

The effect that pulses had on the performance of subjects directly depended on the time onset of the pulse. Early pulses produced a greater shift of the psychometric function than late pulses (Fig. 4.5A). Similarly, Fig. 4.5B shows the same effect on the mean response of the neural population favoring the correct option. A similar trend is observed in the mean activity along different values of motion coherence (Fig. 4.10). In addition, the effect that pulses had on the mean activity of the neural population persisted for at least 600ms from its time onset (Fig. 4.11).

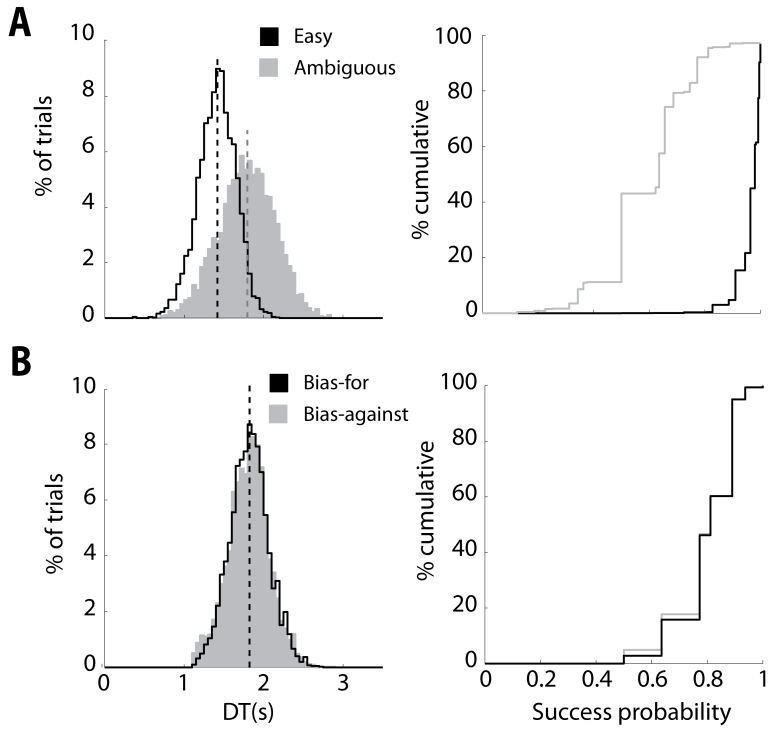

Figure 4.6: RT and success probability in VMD trials. **(A)** RTs and SPs in easy (black) and ambiguous (gray) trials. RTs and SPs are significantly different (KS test, $p < 0.001$). **(B)** RTs and SPs in bias-for and bias-against trials (K-S test, $p > 0.05$). Dashed lines are the means of the RT distributions.

VMD task

Our simulations were focused on trials with specific profiles: easy, ambiguous, bias-for and bias-against trials (see 4.2). As observed in the behavioral performance (Thura et al., 2012), in our simulations (Fig. 4.6A) RTs in easy trials are significantly shorter than in ambiguous trials (KS test, $p < 0.001$) whereas the success probability is significantly higher in easy trials than in ambiguous ones (K-S test, $p < 0.001$). On the other hand, RTs and probability of success in bias-for and bias-against trials were not significantly different (two-sample K-S test, $p > 0.05$). This lack of behavioral difference between the two cases is due to the low level from which the urgency function starts -i. e. the network does not easily falls into an attractor

state, and therefore most of the early information is "forgotten" after some time-. Yet, in 4% of the trials, the decision is made early in the trial ($RT < 1125$ms, data not shown). Fig. 4.12 shows performance accuracy based on the step in which the decision is made, consistent with experimental data (Thura et al., 2012). For bias-against trials, waiting is necessary to succeed, whereas in the ambiguous case longer RTs do not lead to more accurate responses.

VMD task with RDM trials with pulses

The results from the RDM task with pulses and the VMD task might seem contradictory. In the first task, early pulses have greater influence on behavioral performance and neural activity, while the opposite occurs in second. A question that rises from this paradox is: how can the same model reproduce both results? We hypothesize that, in order to increase the speed-accuracy ratio, subjects modulate their urgency signal. In the VMD task, the sensory evidence changes its strength much more often than in the RDM task with pulses. To increase the speed-accuracy ratio, subjects adopt different strategy policies: they become more conservative (slowing down their response time) in the VMD task than in the RDM task with pulses. This, on average, generates different urgency signals defined by context. Providing this is really the case, if RDM trials with pulses are interleaved when subjects perform a VMD task, the effect of pulses should be reversed. In other words, late pulses would have greater influence on behavior than early pulses when trials are interleaved in the VMD task. We simulated this condition using pulses of 100ms and 10% of motion coherence at five different pulse onsets (100ms, 200ms, 400ms, 800ms and 1600ms). As shown in Fig. 4.7, RTs became significantly shorter only when they occurred either at 800ms or 1600ms (2-way ANOVA, $p < 0.001$).

The effect that early and late pulses have on the performance accuracy and neural activity is also reversed when compared to the RDM task with pulses. Fig. 4.9A shows the change in probability of correct due to the pulse onset. As observed, late pulses have a greater effect on accuracy than

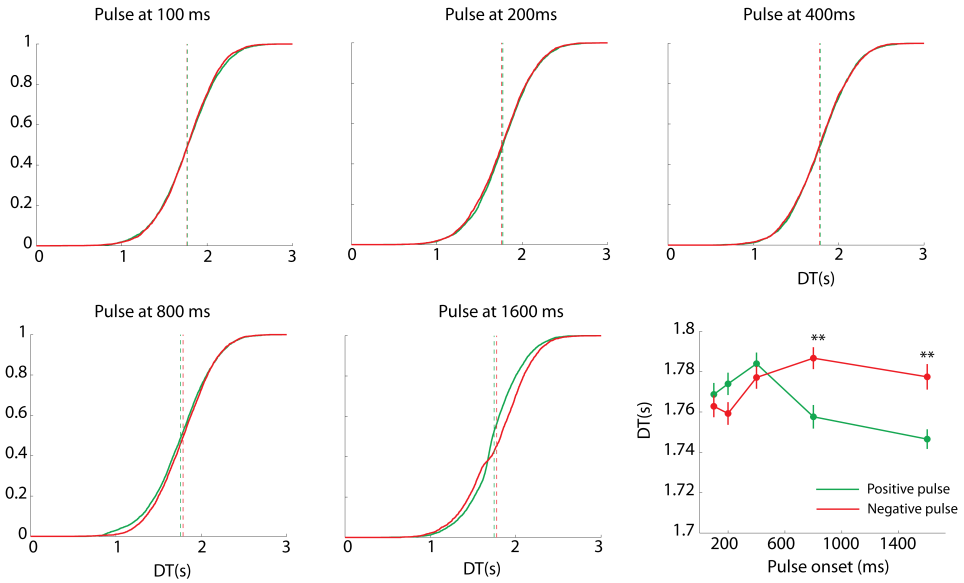

Figure 4.7: Pulse onset effect on RT during the VMD task. Cumulative distributions of RTs for five different pulse onsets: 100ms, 200ms, 400ms, 800ms and 1600ms. Positive-direction pulses are indicated in green and negative-direction pulses in red. The last panel shows the distributions in a single graph. Only pulses that appear late in the trial (800ms and 1600ms) have a significant effect on RT (2-way ANOVA, $* = p < 0.001$)

early pulses. Likewise, the change in the mean neural activity shows the same effect (Fig. 4.9B).

Experimental data

At first, we analyzed whether the contexts led to different adopted decision policies. To this end, we looked at RTs in trials with constant stimulus -i. e. RDM trials experienced under blocked or interleaved conditions-. Fig. 4.9A shows the mean RTs ($\pm$ SEM) of individual subjects for these two cases. All individual data points are above the diagonal line, indicating that subjects are slower when constant stimulus trials are encountered within the interleaved condition than when they appear in the blocked one ($p < 0.001$). Next, we assessed the influence that positive-direction pulses had on RTs depending on the time onset and context in which they appear. To this

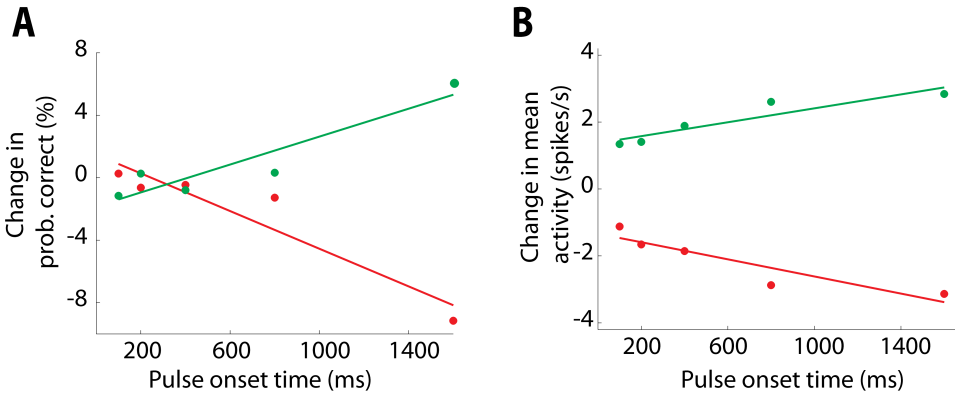

Figure 4.8: Pulse onset effect on performance and mean activity during the VMD task. **(A)** Change in the probability of correct due to pulse onset when a positive-direction (green) or a negative-direction (red) pulse appear in a trial in comparison with no-pulse trials. Late pulses have greater effect on accuracy than early pulses. **(B)** Change in mean neural activity of the model along different pulse onsets. Same trend as in (A) is observed in the mean activity.

purpose, we pooled together data from different subjects with similar RTs. Trials with pulses occurring at 1600ms were not considered in the analyses because most of the RTs were shorter than that. Cumulative distributions of RTs for the subset indicated in Fig. 4.9A are shown in Fig. 4.9B for blocked (Left panel) and interleaved conditions (Right panel). The 200ms pulse onset was the only effective one in both conditions ($p < 0.001$ in the block condition, $p < 0.002$ in the interleaved condition). The 100ms pulse onset influenced RTs only in the blocked condition ($p < 0.001$) whereas the 400ms pulse onset affected RTs in the blocked condition ($p < 0.001$). Hence, pulses that appeared early in the trial had an effect on the context in which RTs were faster (blocked condition), while they did not in the context of slower RTs (interleaved condition).

These experimental results validate the fact that, when subjects slow down their response times, early pulses do not influence decision-making any more, but late pulses do. Therefore, the effect of early and late pulses in behavior reverses, contrarily to contexts in which response times are faster. Thus, these results confirm the predictions of our model and strongly

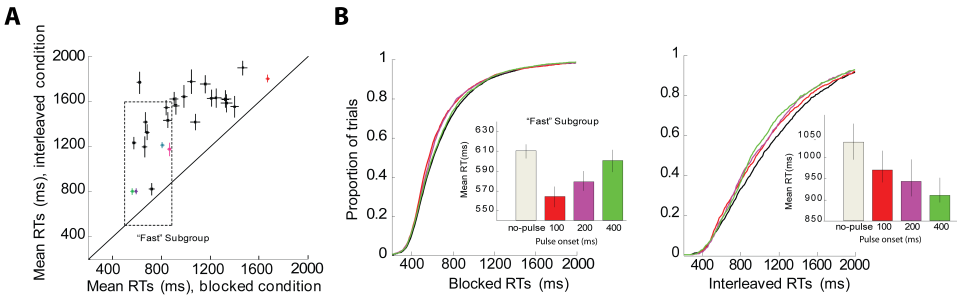

Figure 4.9: Experimental results in blocked and interleaved conditions. **(A)** RTs of individual subjects in the blocked (x-axis) and interleaved (y-axis) conditions. The dotted rectangle indicates the RTs from subjects that were pooled together. Errorbars are SEM. **(B)** Cumulative distributions of RTs and mean RTs for no-pulse (black), 100ms pulse onset (red), 200ms pulse onset (magenta), and 400ms pulse onset(green) for blocked (*Left panel*) and interleaved (*Right panel*) conditions.

support our proposal of the existence of a task-dependent urgency signal which modulates the integration of evidence during decision formation.

4.4 Discussion

Decision making in tasks with noisy but constant sensory evidence has been successfully explained by integrator and linear urgency models. The behavioral and neurophysiological data have been fitted following two different hypotheses: sensory information is accumulated over time until a decision bound is reached or, in contrast, it is not accumulated but directly multiplied by an urgency signal that grows over time. However, new experimental paradigms with noisy though not constant sensory evidence have critically challenged these models. In this study, we have proved that decision-making regarding both constant and non-constant information can be explained by an attractor model extended with an urgency signal. So, we suggest that the paradoxical results from Huk and Shadlen (2005) and Thura et al. (2012) respond to an integration of evidence modulated by a context-dependent urgency signal. To support it, we proceeded to perform an experiment with human subjects where we observed that, when RDM trials with pulses are interleaved in the VMD task, the effect that pulses have on behavior

reverses.

An attractor network with urgency had been previously proposed by (Standage et al., 2011) to control speed-accuracy trade-off in a way that is consistent with existent neural and behavioral data. The modulation of the gain of the urgency signal controlled the time constant of the integration during the course of a trial. Our study shows that, by using the same principles, we can reconcile previous experimental data with information that changes over time.

The possibility of the existence of an urgency signal modulating the decision-making process had been proposed previously (Reddi and Carpenter, 2000). It had been proposed that when we are confronted with decisions that have to be made as fast as possible, we implicitly encode an internal urgency signal. This signal causes a time-dependent increase in the neural activity associated with an imminent deadline to decide regardless of the sensory evidence itself (Churchland et al., 2008).

Although Wong et al. (2007) could explain the data of Huk and Shadlen (2005) with a standard attractor network (without urgency), the very same model would fail to explain the experimental data of Thura et al. (2012). The reason for this is that early sensory evidence has a stronger influence on the attractor dynamics than late one. This will inevitably result in a bias of the neural response between bias-for and bias-against trials, due to the difference in the six first time steps.

Consistent with our study, Tsetsos et al. (2012) showed that, in some cases, pulse effects in success probability were weaker or reversed when the time of observation of the stimulus became longer, in comparison with what was observed in Huk and Shadlen (2005). With a leaky competing accumulator (LCA) model (Usher and McClelland, 2001), they could account for primacy, recency and equal effects produced by pulses through modulation of the impact of inhibition or leakage in the decision process. However, as we have shown in Fig. 4.3B, the leakage has to be very strong to equalize the effect of long early pulses (bias-for and bias-against) during the VMD

task. This strong leakage would make the network to hardly reach an attractor state when trials with low motion coherence are interleaved. This way, although the LCA model proposes a plausible mechanism by which decision are made, it cannot account for situations in which the amount of sensory evidence changes within the same context.

In summary, our study proposes a mechanism by which decisions based on information that changes over time can be made. It advances the understanding of decision-making in two ways: first, we reconcile two apparently paradoxical results that have been segregated for the last years and, second, we propose a robust framework that can account for constant and dynamic information over time. Yet, the question that remains open is how the urgency signal is actually implemented by the brain. In the following chapter, we elaborate on an intrinsic neural mechanism based on the variance of the neural response that could answer this question.

4.5 Supplemental figures

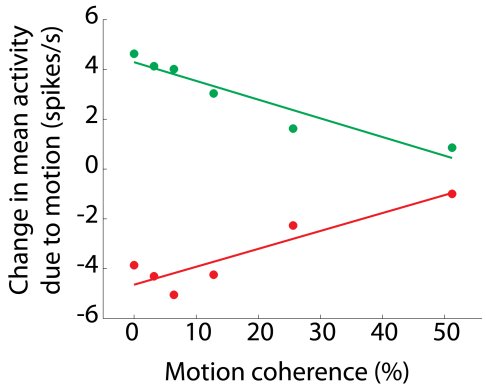

Figure 4.10: Effect of pulses along different amounts of motion coherence. Change in mean activity along different values of motion coherence for positive-direction (green) and negative-direction (red) pulses when compared to no-pulse trials. Pulses have a greater impact on the mean activity for low levels of coherence than when the stimulus is easy.

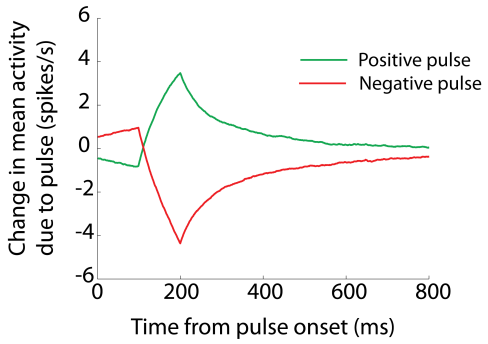

Figure 4.11: Persistent effect of pulses in mean activity. Mean effect of positive-direction (green) and negative-direction (red) pulses in the activity of the model when all conditions are averaged. The effect of pulses lasts for at least 600ms.

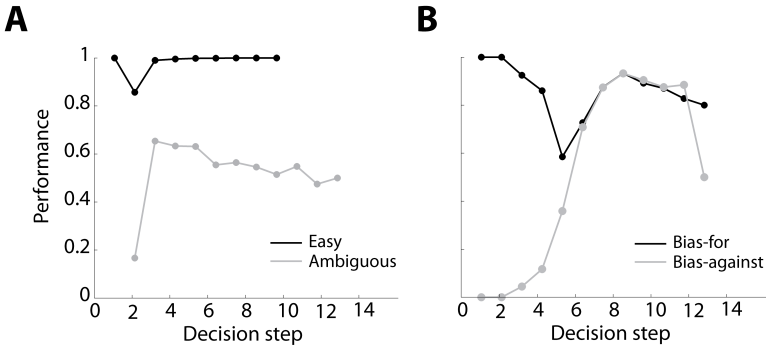

Figure 4.12: Accuracy at decision time. Probability of correct at the time of decision for easy/ambiguous trials **(A)** and for bias-for/bias-against trials ((**B**).

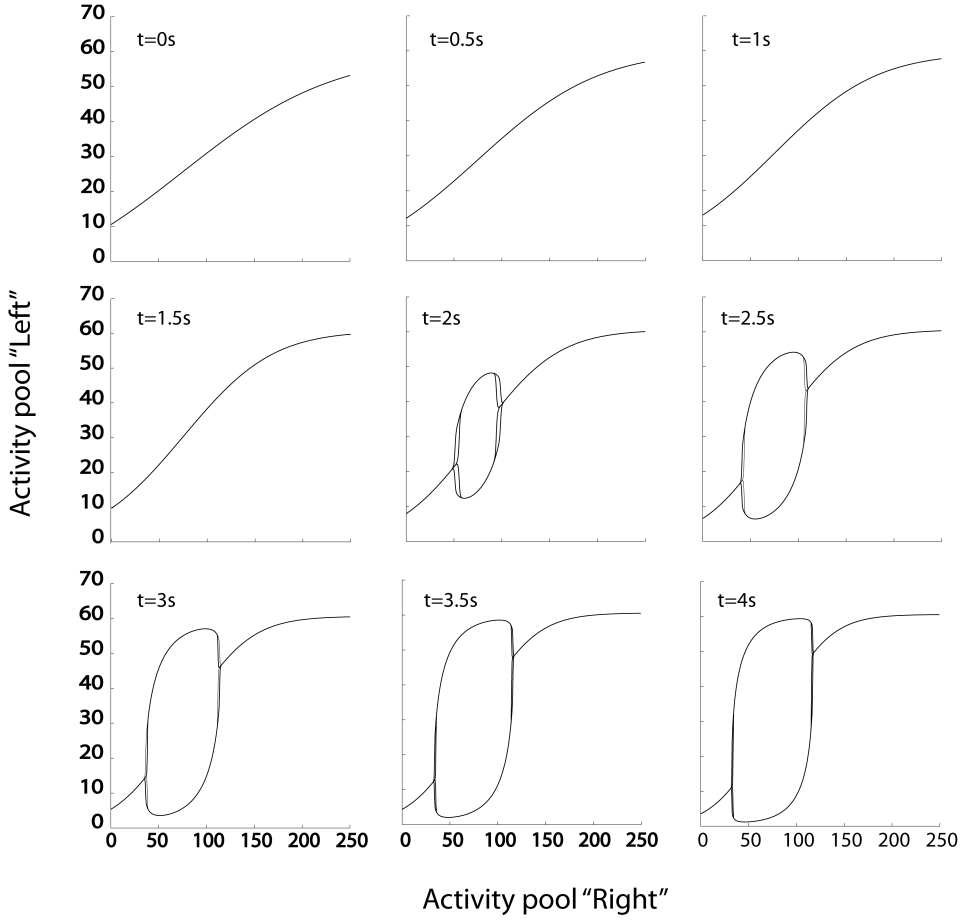

Figure 4.13: Urgency bifurcation diagram. Example of how the state of the network evolves within a trial for an arbitrary urgency signal. As observed, as time progresses, it is more likely that one of the two neural populations would fail into an attractor state.

Chapter 5

The cost of embodiment in decision making

Never let the body tell the mind what to do
George S. Patton

In the final contribution of this thesis we investigate the possible influence that the constraints of embodiment has in decision making. In other words, is there a cost to embodiment that biases perceptual decision making? If so, what is its neuronal substrate? In a combined experimental and theoretical study, we investigate this question demonstrating such an effect exists in humans performing a reaching task. We use a biologically constrained spiking neural model of binary decision making and show that it reproduces the human performance data we present. It does this by advancing two novel physiologically motivated principles. First, our model integrates perceptual evidence biased by motor cost through the lateral synaptic connections reflecting the learning history of the behaving system. Thus, our model shows that cost does not need to be computed locally and explicitly but rather can be expressed implicitly in the distributed neuronal dynamics. Second, in order to explain trials in which subjects give-up, the execution of an action is realized only when the variance across the neuronal pools is below a specific threshold. In this way, showing that the variance of

neuronal activity is functionally relevant, as opposed to recent suggestions to the opposite. Hence, we explain human performance data that reflects the cost of embodiment on the basis of implicit neuronal factors as opposed to depending on explicit computations of cost and or timing pointing to the key role of neuronal variability in the transformation of perception into action.

This study has been published in an abstract form in BMC Neuroscience (Marcos et al., 2013a) and is part of a manuscript in preparation (Marcos et al., 2014b). The abstract reads:

> Perceptual decision making has been widely studied using tasks in which subjects are asked to discriminate a visual stimulus and instructed to report their decision with an action. In these studies, performance is dependent on the accuracy of the choice and the action is merely a means of reporting. However, actions differ in their execution costs, and even subtle differences can influence how we choose between them. Can such execution costs influence decisions, even about purely perceptual discriminations? Here we show the results of a psychophysical experiment in which human subjects were presented with a random dot motion discrimination task and asked to report the perceived motion direction using movements that differed in biomechanical costs. We found that the pattern of decisions exhibited a significant bias towards the movement requiring the lower biomechanical cost, even when this bias reduced performance accuracy. These results can be reproduced with a realistic spiking neural model of binary decision making that accounts for motor and perceptual information in its synaptic connections and input respectively. With this assumption the model further predicts a link between neural response variance and the certainty/uncertainty of the decision being made.

5.1 Introduction

Many studies of perceptual decisions have shown that the timing and accuracy of choices are well predicted by a model in which neural activity builds up to a decision threshold at a rate related to stimulus strength (Gold and Shadlen, 2007; Heekeren et al., 2004; Smith and Ratcliff, 2004). For example, when subjects discriminate the direction of coherent motion in a random-dot motion (RDM) display, reaction time gets shorter and accuracy improves as the coherence of motion is increased. Furthermore, both the timing and accuracy of decisions can be biased by the prior probability that a given choice is correct or asymmetries in the expected payoff associated with a given choice (Hanks et al., 2011; Mulder et al., 2012). These effects are all consistent with the notion that during perceptual decision making the brain employs a strategy aimed at maximizing the expected value of the choices it makes (Gold and Shadlen, 2007).

In a previous theoretical study, we showed that the embodiment of an agent directly influences the flow and processing of perceptual information or behavioral feedback (Verschure et al., 2003). This raises the question whether embodiment and its constraints influences decision making in a similar way as prior probabilities and/or expected payoff. In other words, is there a cost to embodiment that biases decision making and perception? In a combined experimental and modeling study, we investigate whether the biomechanical costs of movement execution can influence perceptual decisions in a manner comparable to known factors such as prior probability or expected payoff. In addition, we use our theoretical study to answer the question whether such embodiment-derived costs are computed explicitly or can be accounted for in implicit terms and to asses which aspects of the neuronal dynamics are predictive of acting versus giving up.

Actions are executed at the expense of a variety of metabolic and temporal costs. The motor system is exquisitely sensitive to such costs, which influence not only the implementation of a movement (Dounskaia et al., 2011) but also the choice of movement itself. For example, when subjects are faced with a choice between two movements that differ in biomechanical

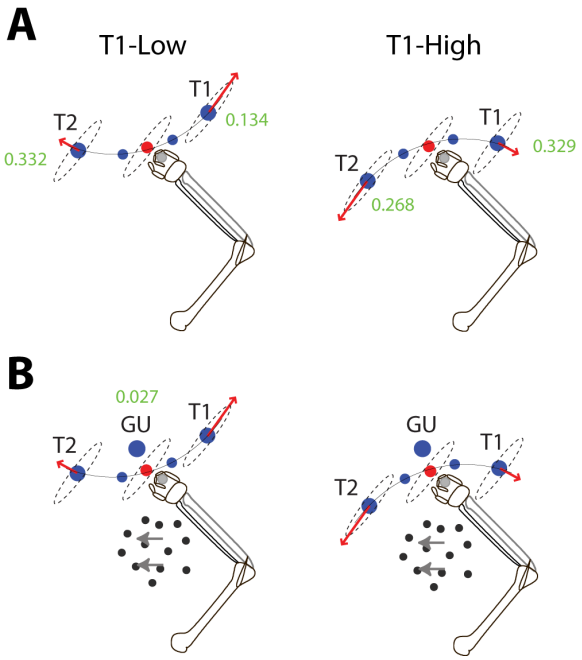

Figure 5.1: Experimental setup. **(A)** Biomechanics trials. The two possible configurations of targets are illustrated: Left, T1-Low configuration; Right, T1-High configuration. The arrows indicate the ideal trajectory that subjects are requested to perform to select either target T1 or target T2. Starting from the red central circle the movement must pass over the via-point (small blue circle) to reach T1 or T2. The ellipse that surrounds the red circle (not shown to subjects) indicates the arms mobility ellipse. In both configurations, the trajectory of the movements towards both targets is equally aligned with respect to the ellipse. However, in the T1-Low configuration the final trajectory of the movement to T1 is aligned with major axis of the ellipse whereas the movement to T2 ends aligned with the minor axis. The opposite occurs in the T1-High configuration. **(B)** RDM-biomechanics trials. Subjects are required to detect the net motion coherence of the moving dots that appear on the screen and to report it reaching T1 (right motion) or T2 (left motion) in one of the two possible configurations (T1-Low) or T1-High. An additional third target (give-up) placed in the same position in both configuration allowed the subjects to report that they were unable to detect the net motion of the stimulus. Numbers next to the targets indicate the total estimated muscle work needed to reach that target (in Joules).

costs (e.g. energetic demand), they tend to choose the one with a smaller cost (Cos et al., 2011, 2012). This occurs even if the candidate movements

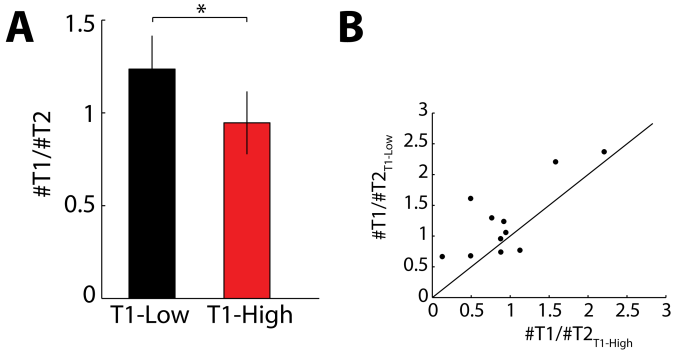

Figure 5.2: Index of asymmetry, ratio of T1 and T2 choices. **(A)** Number of trials where target T1 is selected divided by the number of trials where T2 is selected in T1-Low and T1-High configurations. Target T1 is selected significantly more often when the targets are placed in a T1-Low configuration (Wilcoxon paired signed rank test,$* = p < 0.05$).Error bars are SEM. **(B)** Comparison of the ratio T1 choices/T2 choices in T1-High (x-axis) and T1-Low (y-axis) for individual subjects. Consistent with (A) the majority of the dots are above the diagonal, meaning that most of the subjects (9 out of 11) select more often T1 when the trajectory follows a T1-Low configuration.

have identical launching costs and differ only at the end of the movement, implying that subtle biomechanical costs can be predicted well ahead of movement onset and influence action selection. We build our experiment on this the classic random-dot motion discrimination task (RDM; Britten et al. (1992)) in which subjects are asked to report their choices with a right- or leftward movement of opposite biomechanical cost (Cos et al., 2011). We allowed subjects a give-up option that they could select when they were not confident enough to make an informed guess.

Importantly, our task was explicitly defined as a purely perceptual discrimination task. That is, subjects were instructed to detect the direction of motion in a visual display and their movement was merely a means to report that decision. Therefore, if in a given trial a subject perceived motion to the right and made a leftward movement we consider this an error of perceptual decision making. In summary, if subjects in our task exhibit a biomechanical bias in their choices it means that rather than performing the perceptual discrimination task only, they were also allowing task irrelevant

biases to influence decisions that rendered their performance sub-optimal relative to the task instruction. To explain our experimental results, we used a spiking neural model of decision making (Wang, 2002; Marcos et al., 2013b). This model allowed us to investigate the neural dynamics of the potential mechanisms that integrate perceptual evidence and motor cost, as well as the mechanisms triggering decisions towards the give-up option. In particular, the model demonstrates that motor cost can be accounted for implicitly in PFC networks in terms of how the body schema is reflected in the lateral interaction among the elements of PFC decision-making circuits. In addition, our model demonstrates that it is the variance of neuronal activity that can be used as a criterion in deciding between action and inaction.

5.2 Results

Behavioral results

To quantify the effect of biomechanical cost in the selection of movement, we first assessed the subjects preference for either of two target in free choice trials (biomechanics trials). The two targets (T1 and T2) were presented in two possible configurations: T1-Low and T1-High (Fig. 5.1A; for details see Materials and Methods). Briefly, in the T1-Low configuration the biomechanical cost of moving to T1 was lower than the cost of moving to T2, and the opposite was true in the T1-High configuration. As expected, when allowed to choose freely subjects tended to select T1 more often in the T1-Low configuration compared to the T1-High configuration (mean value T1 choices/T2 choices of 1,24 and 0,95 in T1-Low and T1-High arrangements respectively; Fig. 5.2), consistent with previous studies (Cos et al., 2011, 2012).

We investigated whether this biomechanical cost dependent bias influences decision making during the RDM task, in which the perceived motion of the dots was reported by moving to either T1 or T2 in either the T1-Low or T1-High configurations (Fig. 5.1B). We observed that subjects exhibited

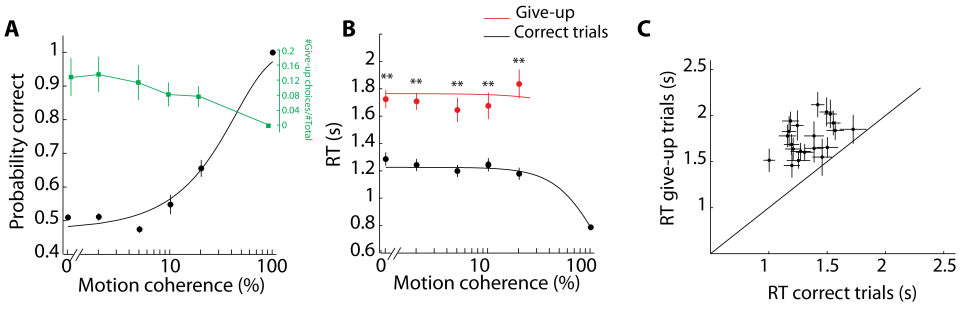

Figure 5.3: Behavioral performance during RMD-biomechanics trials. **(A)** Proportion of correct trials (black) and give-up (GU) choices (green) for different levels of motion coherence, data from all subjects. The performance starts at chance level (0.5) for 0% coherence and it reaches its maximum value of 1 for trials with 100% coherence. Experimental data (dots) were fitted with a logistic function (see Materials and Methods). In contrast, the number of GU choices decreases with motion coherence reflecting the decrease in the difficulty of the trial. **(B)** Mean reaction time for different levels of motion coherence for correct trials and trials where the give-up option was selected. RT was longer when subjects selected the give-up target than when they correctly detected the direction of motion and selected either T1 or T2 (repeated measures ANOVA with Bonferroni correction, $** = p < 0.001$). Experimental data (dots) were fitted with a hyperbolic tangent (see Materials and Methods). **(C)** Mean RTs for correct versus give-up trials. Each data point corresponds to one subject and one level of motion coherence (only the cases that contained more than four give-up trials are plotted). Error bars are SEM.

an increase of accuracy and decrease of their reaction rime (RT) as a function of increasing coherence of the motion cue (Fig 5.3A). Furthermore, consistent with previous studies they chose the give up option less frequently as the coherence increased (Fig. 5.3A), (Kiani and Shadlen, 2009), and these choices always had a longer mean RT (Fig. 5.3B and Fig. 5.3C). Individual RTs in Fig. 5.3C indicates a clear tendency for all subjects. These results confirm that subjects used the give-up target to report their impossibility to detect the dominant motion direction of the moving dots.

Most importantly, the difference in selecting T1 versus T2 targets to report the direction of motion also differed between the T1-Low versus T1-High configurations in the RMD-biomechanics. The probability of T1 choices for left- and rightward motion depended on the targets arrangement

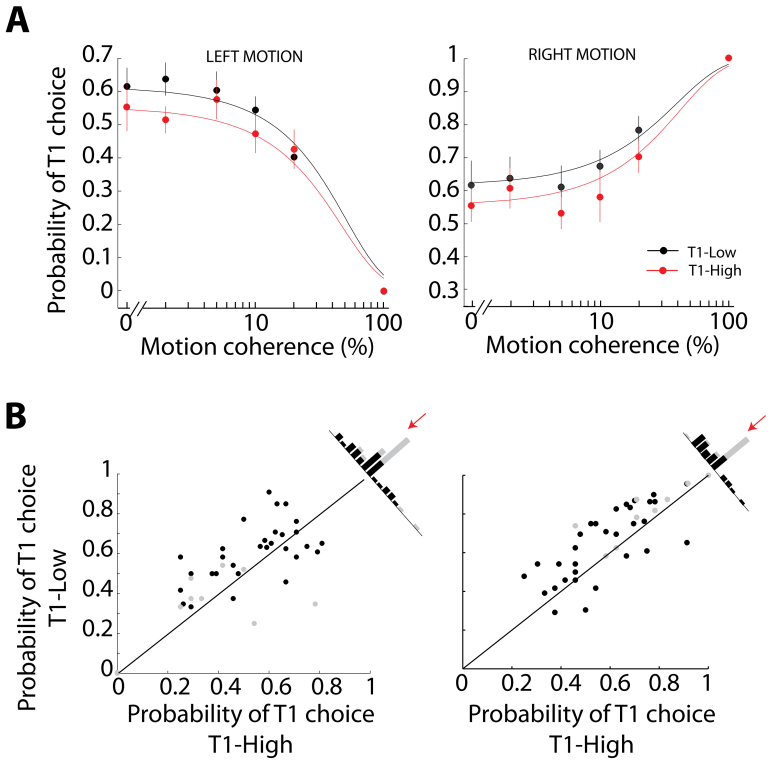

Figure 5.4: Behavioral influence of motor cost during RMD-biomechanics trials. **(A)** Probability of T1 choice for right and left motion along different values of motion coherence for the two biomechanical configurations (T1-Low and T1-High). The probability of selecting T1 is higher when the targets are placed in the T1-Low configuration compared with the T1-High configuration for both left (Left panel) and right (Right panel) motion direction (Wilcoxon signed rank test, $p < 0.05$; see Materials and Methods). In both biomechanical configurations the probability of selecting T1 is above 0,5 for 0% of motion coherence indicating a small perceptual bias towards right. **(B)** Comparison of individual performance for left (Left panel) and right (Right panel) direction of motion. Each dot corresponds to an individual probability of being correct for a specific level of motion coherence, solid black circles for 0%, 2%, 5% and 10% amount of motion coherence and solid gray circles for 20% and 100%. As observed, the majority of the dots are below the diagonal for left motion direction (Left panel), meaning that subjects are in general better at detecting left motion in the T1-High configuration, whereas the opposite occurs for right motion. These figures are consistent with the general behavior observed in (A). Red arrows indicate the mean of the population for all motion coherences.

(Fig. 5.4A). Subjects selected more often T1 for both directions of motion, left and right, when the targets were set in the T1-High configuration as opposed to T1-Low. We interpret this result as a tendency to detect more often right motion when the targets are placed on the T1-Low configuration in comparison to when they follow the T1-High configuration. This effect could be explained by fitting the probability of T1 selection with a psychometric function with three free parameters related to an initial bias, level of motion coherence and biomechanical configuration. The use of the third parameter was necessary to correctly explain the data (Wilcoxon signed rank test, $p < 0.05$, see Materials and Methods). The effect of the motor cost was equivalent to an addition of 6,6% of right motion to the stimulus in T1-Low with respect to T1-High. Fig 5.4B shows a consistent effect across individual subjects.

Is it possible that the influence of the biomechanical bias we observe in perceptual decision making (Fig. 5.4) is due to a strategy of simply choosing the easier movement by default during trials in which the subject is unsure of the direction of motion? This is unlikely because of two reasons. First, the presence of the give up option, whose biomechanical cost was lowest of all. That is, if a subject was unable to detect motion in a particular trial, they could just select the give up option and move on to the next trial. As expected, give up choices decreased in frequency as a function of motion coherence (Fig. 5.3A) and tended to be made with long reaction times (Fig. 5.3B). Second, we observe that the modulation of the selection modulation systematically varies with the motion coherence throughout the full range of coherence values (Fig. 5.4A and 5.4B). Hence, subjects decide to choose biased by biomechanical cost and thus not to give up for coherence values where the probability of giving up is very low (Fig. 5.3A). Thus, when subjects did choose T1 or T2, they presumably acted on the basis of a non negligible certainty on the motion coherence.

Modeling results

To investigate how the biomechanical cost of each movement may influence the perceptual decision-making process we used a realistic neural model of leaky integrate-and-fire neurons (Wang, 2002). The model consists of N_E excitatory neurons and N_I interneurons. The structure of the network is illustrated in Fig. 5.5A. Excitatory neurons are subdivided into two selective populations that encode the two possible directions of motion, left or right, and a population of non-selective neurons that simulates the activity of neighboring neurons that are not selective for any of the two motions. The three groups of excitatory neurons are connected to an inhibitory group that mediates the competition in the network (see Materials and Methods). The biomechanical preference of movements is reflected in the model by different weight in the recurrent connections of the selective pools (ω_{+T1}, ω_{+T2}). The decision-making process was considered to be terminated when the difference in activation between the two selective neural populations was above a predefined threshold (Roxin and Ledberg, 2008). We also propose that, in parallel, an additional system integrates the variance of the population response. Whenever the difference in activity between the two competing populations reaches the decision bound the value of the integrated variance is assessed: if the variance is above a predefined threshold the choice is considered to be "give-up" (an example of the model dynamics is illustrated in Fig. 5.9). As observed in the experimental data (Fig. 5.2A), the difference in the strength of the recurrent connections biases the selection of the targets in each biomechanical configuration (Fig. 5.5B, $p < 0.001$) when biomechanics trials (no motion coherence) are simulated. When the perceptual input is activated (RMD-biomechanics trials), the model shows the same asymmetry in the probability of T1 choices for right/left motion during T1-Low/T1-High configurations (Fig. 5.5C) as observed in Fig. 5.4A. This is due to the unbalanced weight of the recurrent connection of the neural populations that biases the competition between the neural populations favoring the one with the strongest recurrent connection.

We propose that the network decides to give-up when the accumulated

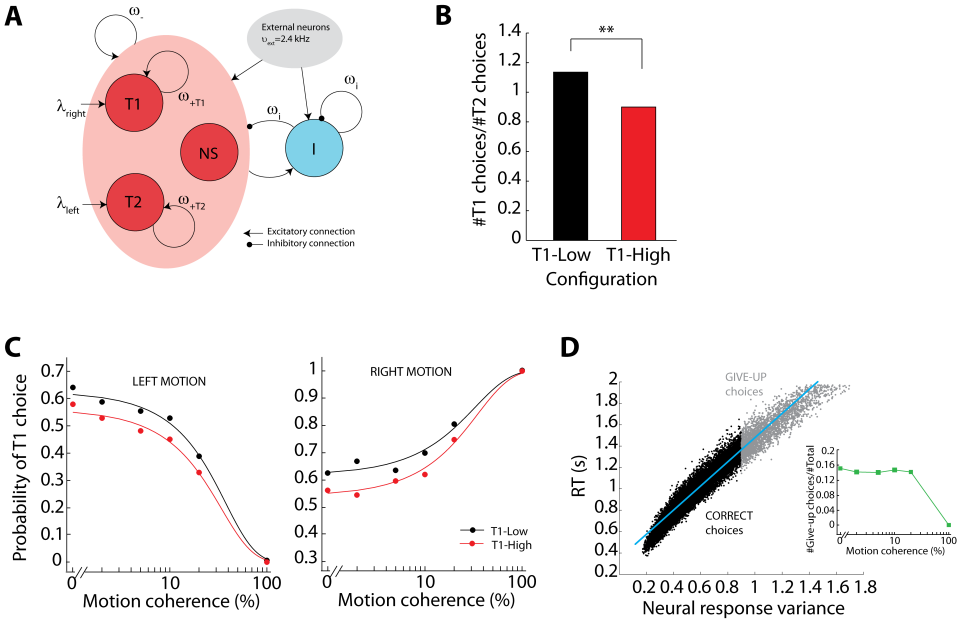

Figure 5.5: Decision-making neuronal model. **(A)** Layout of the spiking neural network. The network consists of one population of excitatory pyramidal neurons and one population of inhibitory interneurons. The excitatory population is subdivided into two selective populations (T1 and T2) and one non-selective population (NS). The 'T1' population is selective for right motion (λ_{right}) and generates actions towards target T1 whereas the 'T2' population is selective for leftward motion (λ_{left}) and is associated with a movement towards T2. Both selective populations compete through shared inhibition from the inhibitory population (I). The network comprises a total of 2,000 neurons that follow all to all connectivity. **(B)** T1/T2 ratio for both the high and low cost biomechanical condition, T1-Low and T1-High (binomial test, $** = p < 0.001$). **(C)** Probability of T1 choices for leftward (*Left panel*) and rightward (*Right panel*) motion. The results show a higher probability of selecting T1 for both, leftward and rightward, motion directions when the simulated targets are in the T1-Low configuration ($\omega_{+T1} = 1.6991$, $\omega_{+T2} = 1.7003$ for T1High and $\omega_{+T1} = 1.703$, $\omega_{+T2} = 1.699$ for T1-Low). **(D)** Relationship between RTs and variance of the neural response of the motion selective populations at the time of the decision. The data have been fitted using linear regression, with a slope of 1,15s and an offset of 0,32s ($r = 0.97$ and $p < 0.001$). Small figure shows GU choices obtained with the model for different levels of motion coherence.

variance at the time of the decision is high. This mechanism can correctly explain the longer RT in give-up trials compared to correct trials (Fig. 5.7A) and results in a number of give-up trials that falls to 0 for non-

ambiguous trials (100% of motion coherence) (small panel in Fig. 5.5D), following a similar trend as the real data (Fig. 5.3A). RT and variance of the response of the population are highly correlated and the later can be used as a predictor of the former. These results suggest that a measure of the variance of the population can be directly linked with uncertainty about the current decision to be made, i. e. the higher the variance the higher the uncertainty (Marcos et al., 2013b).

5.3 Discussion

Here, we investigated the influence of biomechanical factors in a perceptual decision-making task. In agreement with previous studies (Cos et al., 2011, 2012), our results showed that subjects tend to select more often the choice that requires a movement demanding less effort. Furthermore, here we show that this same tendency is present when subjects had to detect the direction of random dot motion. The bias in the perceptual decision making is equivalent to an additional amount of motion coherence favoring the direction that required a movement with lower motor cost (virtual increase motion towards right in the T1-Low case compared to the T1-High case). We interpret this bias as a result of a embodied decision-making process in which motor cost influences the flow of perceptual information, consistent with our previous theoretical work (Verschure et al., 2003). These results were reproduced by a spiking neural model of a binary decision-making task which showed that asymmetric recurrent connections related to each movement bias the selection of the movement that requires less cost. Our model also predicts that give-up options can be explained by the neural variance of the population response, suggesting a link between neural response variance and uncertainty.

Previous studies showed that biomechanical factors influence decisions between reaching movements (Cos et al., 2011, 2012). Other factors being equal, the target selected most frequently was the one with the lowest associated biomechanical cost. Here we have extended these results to show that biomechanical costs may also bias decisions in visual discrimination

tasks when decisions are reported with reaching movements. This shows that the motor system may unduly influence visual decisions even during tasks in which motor costs are completely unrelated to sensory evidence. Our simulated results were also consistent with this claim.

One of the critical points of this study was to distinguish whether the observed bias caused by the asymmetry in the biomechanical cost of movements in the perceptual decision-making process was due to an integration of sensory evidence and motor cost in the decision process or due to a selection of the default and less costly action when subjects are simply not able to detect the net motion of the dots. To overcome this issue, we added a give-up option that subjects were instructed to select if they were unable to distinguish the direction of motion. The presence of this option suggests that T1 or T2 selections occurred only when the subjects had detected rightward or leftward motion, respectively (Kiani and Shadlen, 2009).

The main result of the biomechanical influence in perceptual decision making also indicates that choices exhibit a mild lateral bias towards T1 in both T1-Low and T1-High, in a similar fashion to the bias reported by Resulaj et al. (2009). As a control for this, we performed an additional experiment in which subjects used the keyboard to report the motion direction of the moving dots. As in our main results they have a tendency to report that they detected right motion over left motion as indicated by a proportion of right choices above 0.5 (Fig. 5.8). Therefore, in our case, this bias may not be attributed to any conflict between motor and perceptual costs.

The spiking neural model allowed us to investigate the neural dynamics involved in the integration of biomechanical cost and sensory evidence during a perceptual decision-making task. Our results show that we can explain the observed behavioral bias by an asymmetry in the strength of the recurrent connection of the two selective populations. Moreover, our model suggests the existence of a variance monitoring system that assesses the certainty of the choice at the time of the decision and that can account for give-up choices and their associated long RTs. In Marcos et al.

(2013b), we showed that trial history and across-trial neural response variance were highly correlated and that the later could be used as a predictor of the former. Here, we have extended this study to show a possible link between variance of the neural response of the population with uncertainty (Churchland et al., 2011; Marcos et al., 2013b). In this study, to calculate the neural response variance within a trial, we have used the variance of the spike count observed in a time window, i.e., the deviation of a single neuron spike count from the mean spike count of the population. Although less biologically grounded than measures of variance for across-trial variability, this measure is equivalent to the variance of conditional expectation (VarCE) when the point process variance (PPV) is not considered in the calculation (theoretical value $\Phi = 0$). This assumption has shown to lead to equivalent results than the ones observed when PPV is not zero (Churchland et al., 2011; Marcos et al., 2013b) and has been widely used previously (Gur et al., 1997).

To explain the motor bias during a perceptual decision-making task observed in our experimental data, we asymmetrically changed in our model the recurrent connection of the two neural populations in a manner dependent on the simulated target configuration. Since the movements that are required to report the direction of the moving dots in our task are very common we suppose that subjects do not require any motor learning but instead they already tend to perform more often the less costly movement. We have modeled this by assuming that the strengths of the synaptic connections are asymmetric and constant during the task.

Kiani and Shadlen (2009) showed that an intermediate level of activity of neurons in the lateral intraparietal cortex (LIP) recorded from two monkeys was associated with their later selection of the give-up option in their study. To explain this mechanism, they used a bounded accumulation model and argued that at the time of the decision the mean firing rate activity is informative about the certainty of the monkeys, low activation meaning low certainty and vice versa. A similar explanation is suggested by the urgency-gating model (Cisek et al., 2009). For simplicity, here we

show that with a fixed threshold and using both firing rate and variance of the neural response we could also explain the results. This suggests a concrete role to the neural response variance and predicts that it reflects the certainty of the decision.

Neurons in parietal and frontal areas have shown to be involved in two-choice perceptual and motor decision-making tasks (Platt and Glimcher, 1999; Gold and Shadlen, 2000; Cisek and Kalaska, 2005; Gold and Shadlen, 2007; Mirabella et al., 2011). The mean firing rate activity of neurons in these areas have shown to be correlated with decision that is being made. An open question is now whether these same neurons show the modulation in their neural response variance with the uncertainty of the decision being made as predicted by our model and if this variance can be used as a predictor of the outcome of the decision, i. e. right/left choice or give-up.

5.4 Materials and Methods

Characterization of motor cost

To estimate the cost of movement (Fig. 5.1A and 5.1B) needed to reach T1 and T2 in the two biomechanical configurations we followed the same procedure as in previous studies (Hogan, 1985a,b,c). Motor cost was calculated using the alignment of the movement with the major or minor axis of the ellipse of movement and the structure of the arm and its distribution of mass (Fig. 5.6; see Supplemental Experimental Procedures).

Experimental design

Eleven subjects (7 females and 4 males, aged 24-34 years) performed the biomechanics trials of the experiment (Fig. 5.2A) and eight of these eleven subjects (5 females and 3 males, aged 24-34 years) also performed two more sessions that contained biomechanics and RMD-biomechanics trials (Fig. 5.2A and 5.2B). The experiment was approved by the Universitat Pompeu Fabra ethical committee.

The experimental setup consisted of a touchscreen computer (Sony VAIO L Series Touchscreen AIO PC 24", resolution of 1920x1080). Subjects made movements using a digitizing stylus whose position was sampled at 60 Hz. The control of the behavioural task, stimulus display, and synchronization of task events and signal recordings were performed by a custom written C/C++ program using the Openframeworks library [1]. The data from each session was transferred to a MySQL database (Oracle, Santa Clara, CA) for further analysis using custom designed Matlab scripts (Mathworks, Natick, MA).

The experiment consisted of two kinds of trials divided into three sessions: the biomechanics trials and the random motion discrimination-biomechanics trials (RMD-biomechanics trials). The first session consisted of a baseline experiment, aimed at characterizing the influence of biomechanics on the subjects decision, in the absence of any other sensory information than the potential directions of movement. To this end, we defined two kinds of biomechanicals configurations that we call: 'T1-Low' and 'T1-High'. To independently influence biomechanical factors associated with each of the two targets, we arranged them such that the path of the trajectory approaching a target was approximately aligned with either the major or minor axis of the arms biomechanical mobility ellipse calculated at the target (Fig. 5.2A). In the arrangement shown on the top of Fig. 5.2A, which we call the T1-Low condition, reaching movements to target 1 (T1) would arrive on a path aligned with the major axis of the mobility ellipse, while movements to target 2 (T2) would arrive on a path aligned with the minor axis. In contrast, in the T1-High arrangement shown on the bottom of Fig. 5.2A, movements to T1 would arrive along the minor axis and movements to T2 would arrive along the major axis. The session was divided into three blocks all of them exclusively containing biomechanics trials with a total of 100, 80 and 80 trials respectively. There were two types of trials: two-target trials (N=140) which were used to assess subject choices; and one-target trials (N=120), which were used to ensure that subjects had

[1] http://www.openframeworks.cc/

substantial experience performing all of the movements.

Each individual trial began when the central cue, or origin cue, (red central dot, Radius 0.47cm) was shown on the screen. Time started counting when the subject moved the stylus into the origin cue and held it there for a 500ms which we call Centre Hold Time (CHT). Next, the stimuli defining the potential trajectories were shown, either in the T1-Low or T1-High configurations. Each potential trajectory was defined by the central cue, a via-point (cyan dot, Radius 0.55cm), and a target (cyan dot, Radius 0.85cm). The via-point and target cues changed to a lighter cyan as the stylus moved over them.

The two remaining sessions consisted of both biomechanics trials and RMD-biomechanics trials. Each session was divided into four blocks of trials: the first block contained biomechanics trials (40 trials) whereas the other three blocks had exclusively RMD-biomechanics trials. In each experimental session, the first block was used as an exploratory task to allow subjects to experience the biomechanical cost of each motor action and contained 24 two-target trials and 16 one-target trials. The remaining three blocks had RMD-biomechanics trials. The dynamic random dots were displayed in a 5 circular aperture with a dot density of 16,7 dots per deg2 per s (Roitman and Shadlen, 2002; Hanks et al., 2006; Britten et al., 1992). Dots were placed in a random position or in a subsequent right or left position every three frames (50 ms). Coherently moving dots were displaced to produce $6^{\circ}s^{-1}$ motion. Six motion strengths were used: 0%, 2%, 5%, 10%, 20% and 100%. Each RMD-biomechanics trial started with a red central cue that subjects were required to reach with the stylus and hold during 500ms. After that period one of the two configurations of choices, T1-Low or T1-High, appeared on the screen during 200ms followed by dots moving on the screen with one of the predefined amounts of coherence towards right or left that randomly changed from trial to trial. Subjects were required to detect the direction of the moving dots and to report it as fast as possible performing one of the two possible movement trajectories that appeared in each trial or to select a 'give-up' option meaning that they were not able to

detect the dots direction of movement. This option ensured that subjects reported the direction of motion when they were convinced that they had detected it. Each of these three blocks contained a starting amount of 96 trials. When subjects selected the give-up option in a given trial that trial was placed at the end of the sequence of the block so that it appeared again. With this procedure we ensured that we had enough trials with responses that reflected the detection of motion coherence. To avoid that the duration of the session was dramatically increased by this procedure we limited the number of trials in each block to 120 trials. A trial had a fixed total duration of 4s, which means that if subjects selected a target before the end of the trial the screen appeared white for the remaining time. This was used to discourage random guessing. The inter-trial interval had a duration of 500ms.

Fits to behavioral data

Psychometric functions (Fig. 5.3A) were fitted by a logistic function (Wong et al., 2007):

$$P_{correct} = \frac{1}{1 + exp(-(\beta_0 + \beta_1 c'))} \tag{5.1}$$

where c' is the motion coherence level and β_0 and β_1 are free parameters.

Chronometric (Fig. 5.3B) curves were fitted by:

$$RT = \frac{\beta_2}{\beta_3 c'} tanh(\beta_2 \beta_3 c') + t_R \tag{5.2}$$

where β_2 and β_3 and t_R are free parameters.

Statistical test

To assess the bias in the probability of T1 choices for right/left motion direction between T1-Low and T1-High during the RMD-biomechanics trials we fitted the behavioral performance relative to the motion direction,

adding a third parameter to Eq. 5.1 related to the cost associated with the trajectory required to reach the specific target:

$$P_{T1} = \frac{1}{1 + exp(-(\beta_0 + \beta_1 c' + \beta_2 I))} \tag{5.3}$$

where I has a value of 1 when the targets are arranged in the T1-High configuration and 0 when they are placed in the T1-Low configuration. To investigate the bias in performance due to the configuration of the targets we evaluated the null hypothesis that the arrangement of the targets does not bias choices made by the subject (this occurs when the related parameter has a value of 0, $\beta_2 = 0$). The significance of the effect was calculated by comparing the distribution of the parameters β_2 to a distribution of median 0 (Wilcoxon signed rank test, $p < 0.05$). We estimated the strength of the effect in terms of motion coherence with the ratio β_2/β_1.

Spiking neural network

The spiking model comprise 2000 leaky integrate-and-fire neurons (Tuckwell, 1988). That represent pools of neurons in prefrontal cortex that implement decision-making mechanisms underlying the motor task we study (Wang, 2002). All parameters are identical to those presented in (Wang, 2002) except the weight of the recurrent connections of the excitatory populations. Briefly, it consists of $N_E = 1600$ excitatory pyramidal neurons (80% of total) and $N_I = 400$ interneurons (20% of total) (Braitenberg and Schtz, 1991). Of the excitatory neurons, 240 (15% of N_E) are activated by right motion, 240 (15% of N_E) by left motion, and the remaining 1120 neurons (70% of N_E) are not activated by any of the two motion directions. The network has all-to-all connectivity via three types of receptors: AMPA, NMDA and GABAA. The two motion selective neural populations compete with each other though shared recurrent inhibitory connections mediated by the interneurons. All neurons in the network receive an external input via excitatory connections (AMPA mediated) that simulates the background activity with a mean rate $\nu_{ext} = 2.4kHz$, according to a Pois-

son process independent for each neuron (Wang, 2002). The mean motion dependent input, λ, to the two motion selective excitatory populations was defined by modulating a baseline of spontaneous activity, μ_0, by adding or subtracting a motion dependent neuronal response, $\rho c'$, for preferred, μ_p, or non-preferred, $\mu_n p$, motion direction respectively:

$$\mu_p = \mu_0 + \rho c' \tag{5.4}$$

$$\mu_{np} = \mu_0 - \rho c'$$

with an added noise with standard deviation σ, following a Gaussian distribution. This modulation of activity simulates the effect of different levels of motion coherence on the activation of MT neurons that in turn project to the PFC decision-making network we model here (Britten et al., 1993, 1996). For simplification, we did not consider pooling or noise correlation between MT neurons (Zohary et al., 1994; Shadlen et al., 1996; Bair et al., 2001), since it does not affect the behavior of the model (Wang, 2002). In our simulations, $\mu_0 = 40Hz$, $\rho = \mu_0/4Hz$ and $\sigma = 10Hz$ and $\mu_{right} = \mu_p$ and $\mu_{left} = \mu_{np}$ when right motion is simulated and $\mu_{left} = \mu_p$ and $\mu_{right} = \mu_n p$ for left motion (Fig. 5.5A). At every 50ms the values of μ_p and μ_{np} are resampled from the Gaussian distribution. We added a value of 2,3 Hz to the T1 pool input to account for the observed bias towards right motion detection (probability of T1 choice greater than 0,5 in T1-Low and T1-High for 0 motion coherence; Fig. 5.4A) (Resulaj et al., 2009).

The decision generated by the model was considered to be realized when the difference in activation between the two competing pools reached a fixed threshold (Σ) of 35 Hz. In order to decide between giving up and executing the decision, a monitoring system integrated, in parallel, the total normalized spiking variance of the decision-making population over the course of the trial. The action to move towards T1 or T2 was executed only when the integrated value of variance was below a predefined threshold ($\phi = 0,9Hz^2$) at the time of the decision or "give-up" otherwise.

The mean firing rate and the variance of the populations were computed using a time window of 60ms and a sliding window of 5ms. Other window sizes yielded similar results.

Supplemental Experimental Procedures

Characterization of biomechanics and muscle work

There is a variety of biomechanical factors associated with any given movement, including passive inertial properties of the arm (Sabes and Jordan, 1997; Sabes et al., 1998), interaction torques (Gritsenko et al., 2011), and muscle visco-elastic properties (Goble et al., 2007; Dounskaia et al., 2011). Previous studies determined that the biomechanical costs associated to movements along different directions were predicted in anticipation of movement onset and biased the selection between movements (Cos et al., 2011, 2012). For consistency with these previous studies, we characterized biomechanics by using the alignment of the end-point trajectory with the major or minor axis of the ellipse of mobility and admittance, calculated for planar movements (Hogan, 1985a,b,c), as a metric of biomechanics. End-point mobility depends on joint configuration and captures the spatial anisotropies that result from the structure of the arm and its distribution of mass. For movements on the plane, mobility can be mathematically expressed as a 2x2 tensor matrix and can be visually represented as an ellipse whose major/minor axes indicate the directions of maximal/minimal sensitivity to perturbations. Eq. 5.5 describes the transformation from arm inertia to end-point mobility.

$$W(\theta) = J(\theta)I^{-1}(\theta)J'(\theta) \tag{5.5}$$

The mobility tensor $W(\theta)$ is the inverse of the inertia tensor $I(\theta)$. Eq. 5.5 transforms the mobility tensor from joint space into end-point space by using the Jacobian $J(\theta)$ of the arm. Since measures of arm inertia are not directly available, we have used a planar, simplified two-segment model of the arm, which describes each segment as a centre of mass, m1 and m2,

located at a fraction c1 or c2, along the respective segment length. Eq. 5.6 shows the resulting formulation the inertia tensor.

$$I(\theta) = \begin{pmatrix} m_1 c_1^2 l_1^2 + m_2 c_2^2 2_1^2 + 2 m_2 c_2 l_1 l_2 cos(\theta_2) & m_2 c_2 l_1 l_2 cos(\theta_2) + m_2 b_2^2 l_2^2 \\ m_2 c_2 l_1 l_2 cos(\theta_2) + c_2^2 l_2^2 & m_2 c_2^2 l_2^2 \end{pmatrix} \tag{5.6}$$

In Eq. 5.6, θ_1 and θ_2 are the shoulder and elbow angles as defined in Fig. 5.1. The remaining parameters are averaged mass and mass center distances, which have the following average values: $m_1 = 1.76Kg$, $m_2 = 1.65Kg$, $c_1 = 0.475$, $c_2 = 0.42$ (Sabes and Jordan, 1997). Arguably, there are more complex metrics that could be used to characterize the cost of biomechanics, e.g., a model of the arm that includes the main muscle groups. However, mobility provides an easily interpretable visual representation in the form of an ellipse, as the metric to capture the main features of the anisotropies of the arm in the plane. Our experimental paradigm was designed to vary the biomechanical cost (alignment of the trajectory with the major and minor axes of the mobility ellipse) to study their interactions on subjects response choices and perceptual decision making. We parameterized the movements from the origin to one of two targets and created two geometrical arrangements in which the movement to T1 and to T2 always implied opposite biomechanical costs. Specifically, in the first arrangement we aligned the arrival to T1 with the major axis of the ellipse and to T2 with its minor axis (T1-Low arrangement), and in the second we aligned the arrival to T1 with the minor axis and to T2 with the major axis (T1-High arrangement). Furthermore, we selected the movement initial directions to imply equal launching biomechanical costs by adding via-points along directions anti-symmetrical to the mobility axes (see Cos et al. (2011) for further detail).

Despite this qualitative characterization of biomechanical costs, biomechanics itself is an intrinsic cost of movement and can be best quantified from the consequence of moving along each specific direction. For example, moving along the major or minor axis of the mobility ellipse means

following a ridge or a valley of the biomechanical cost function, and consequently of the energetic demand of moving along those directions. Since we do not have direct access to a measure of the energy consumption, we used instead the muscle work over each entire movement to the target at each arrangement to quantify the cumulative biomechanical cost associated to a movement. To that end, τ_m predicted for the subjects trajectories (Eq. 5.7) necessary to perform each reaching movement (see Cos et al. (2011) for a detailed account of this calculation). Muscle work (W) was calculated by integrating the net torques along the trajectory from the origin to the target, as described by Eq. 5.7.

$$W = \int_{Origin}^{Target} \tau_m d\theta \tag{5.7}$$

The resulting estimates of muscle work for each movement and configuration are shown above the targets of Fig. 5.1B.

Spiking neural model

Our decision-making model follows the descriptions in Wang (2002). In this supplementary material we present its main characteristics. Except for the recurrent connections of the selective excitatory populations we use the same parameters as selected in Wang (2002).

The spiking neural model consists of 2,000 neurons from which 20% are inhibitory interneurons ($N_I = 400$) and 80% are excitatory pyramidal neurons ($N_E = 1600$). The excitatory neurons are subdivided into three populations: two selective populations each of them with fN_E neurons ($f = 0.15$) responsive for either right or left motion and a non-selective population formed by the remaining $(1-2f)N_E$ neurons that is not selective for any of the two stimuli. A neural population is formed by neurons sharing the same inputs and connectivity. Neurons are all-to-all connected.

Neurons

All the simulated neurons are modeled as leaky integrate-and-fire neurons (Tuckwell, 1988). The subthreshold activity of these kind of neurons can be described as:

$$C_m \frac{dV(t)}{dt} = -g_m \left(V\left(t\right) - V_L\right) - I_{syn}(t) \tag{5.8}$$

with a membrane capacitance $C_m = 0.5nF$ for pyramidal neurons and $C_m = 0.2nF$ for interneurons, a membrane leak conductance $g_m = 25nS$ for pyramidal neurons and $g_m = 20nS$ for interneurons and a resting potential $V_L = -70mV$. I_{syn} represents the total input current flowing into the neuron (see Eq. 5.9). When the membrane potential V reaches a membrane threshold of $V_{th} = -50mV$ a spike is emitted and the membrane potential drops to a reset potential $V_{reset} = -55mV$ with a refractory period $\tau_{ref} = 2ms$ for pyramidal neurons and $\tau_{ref} = 1ms$ for interneurons.

Synapses

The connections of the network are mediated by three types of receptors: AMPA, NMDA and GABA for the inhibitory connections. Recurrent excitatory post-synaptic currents (EPSCs) are mediated by AMPA ($_{IAMPA,rec}$) and NMDA ($I_{NMDA,rec}$) receptors. External inputs send to the network either from stimuli information or by background noise from neurons connected to the network are considered to be driven only by AMPA ($I_{AMPA,ext}$) receptors. All inhibitory post-synaptic currents (IPSCs) are exclusively driven by GABA receptors (I_{GABA}). Therefore, the total synaptic current can be described as:

$$I_{syn}\left(t\right) = I_{AMPA,rec}\left(t\right) + I_{NMDA,rec}\left(t\right) + I_{AMPA,ext}\left(t\right) + I_{GABA}\left(t\right) \tag{5.9}$$

where the currents follow the dynamics:

$$I_{AMPA,rec}(t) = g_{AMPA,rec}(t)(V(t) - V_E)\sum_j^{N_E} \omega_j s_j^{AMPA,rec}(t) \quad (5.10)$$

$$I_{NMDA,rec}(t) = \frac{g_{NMDA,rec}(V(t) - V_E)}{1 + [Mg^{2+}]\left(-\frac{0.062V(t)}{3.57}\right)}\sum_j^{N_E} \omega_j s_j^{NMDA,rec}(t) \quad (5.11)$$

$$I_{AMPA,ext}(t) = g_{AMPA,ext}(t)(V(t) - V_E)\sum_j^{N_{ext}} s^{AMPA,ext}(t) \quad (5.12)$$

$$I_{GABA}(t) = g_{GABA}(V(t) - V_I)\sum_j^{N_I} \omega_j s_j^{GABA}(t) \quad (5.13)$$

where $V_E = 0mV$ and $V_I = -70mV$ and ω_j is the strength of synaptic connection j. NMDA currents have a voltage dependency that is controlled by extracellular concentration of magnesium ([Mg2+=1mM]) (Jahr and Stevens, 1990). The synaptic conductances have values close to those experimentally measured (Destexhe et al., 1998): $g_{AMPA,rec} = 0.05nS$, $g_{NMDA,rec} = 0.165nS$, $g_{AMPA,ext} = 2.1nS$ and $g_{GABA} = 1.3nS$ for pyramidal neurons and $g_{AMPA,rec} = 0.04nS$, $g_{NMDA,rec} = 0.13nS$, $g_{AMPA,ext} = 1.62nS$ and $g_{GABA} = 1nS$ for interneurons. Each receptor has a fraction of open channels that is defined by s. For AMPA (external and recurrent) and GABA s_j changes over time as:

$$\frac{ds_j(t)}{dt} = -\frac{s_j(t)}{\tau_{decay}} + \sum_k \delta(t - t_j^k) \quad (5.14)$$

where the decay time of AMPA receptors for recurrent and external connections is $\tau_{AMPA,decay} = 2ms$ (Hestrin et al., 1990; Spruston et al., 1995) and for GABA receptors $\tau_{GABA,decay} = 5ms$ (Salin and Prince, 1996; Xiang et al., 1998). The sum over k represents the sum of spikes emitted by

presynaptic neuron j at time t_j^k. The fraction of open channels for NMDA receptor mediated currents are described by:

$$\frac{ds_j^{NMDA}(t)}{dt} = -\frac{s_j^{NMDA}(t)}{\tau_{NMDA,decay}} + \alpha x_j\left(t\right)\left(1 - s_j^{NMDA}\left(t\right)\right) \tag{5.15}$$

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_{NMDA,rise}} + \sum_k \delta(t - t_j^k) \tag{5.16}$$

where $\tau_{NMDA,decay} = 100ms$, $\alpha = 0.5ms^{-1}$ and $\tau_{NMDA,rise} = 2ms$ (Hestrin et al., 1990; Spruston et al., 1995).

Network connectivity

The network is fully connected and the weight of their connection follows a Hebbian rule: it is strong if the neurons are highly coupled in their response or weak otherwise. We assume that the weights were already learned before so we keep them constant over the simulations (only the weight of the recurrent connections change from conditions) Therefore, within each selective population the neurons have strong connection between them ($\omega_+ > 1$, $\omega_+ = 1.7$; $\omega_{+T1} = \omega_+ - 0.0009$, $\omega_{+T2} = \omega_+ + 0.0003$ for T1-High and $\omega_{+T1} = \omega_+ + 0.003$, $\omega_{+T2} = \omega_+ - 0.001$ for T1-Low). The connection weight between two selective populations and from the non-selective population to the selective ones have weaker values ($\omega_- < 1$; $1 - f(\omega_+ - 1)/(1 - f)$). Since the deviation from ω_+ of the recurrent connections is very low we neglect its effect on the calculation of ω_- (it does not affect our results). The rest of the connections have a value of 1.

Model inputs

To simulate spontaneous noise due to neighboring external neurons all neurons in the network receive an AMPA mediated external current that simulate uncorrelated Poisson spike trains of $\nu_{ext} = 2,4kHz$. The motion stimuli was simulated using the procedure explained in the main text Material and Methods.

5.5 Supplemental figures

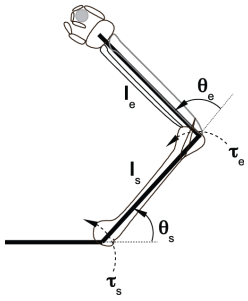

Figure 5.6: Model of the right arm. The modeled arm is a two-segment rigid body rotating around the shoulder and elbow joints. θ_s and θ_e are the shoulder and elbow angles, l_s and l_e are the upper arm and forearm lengths, and τ_s and τ_e are the shoulder and elbow torques, respectively.

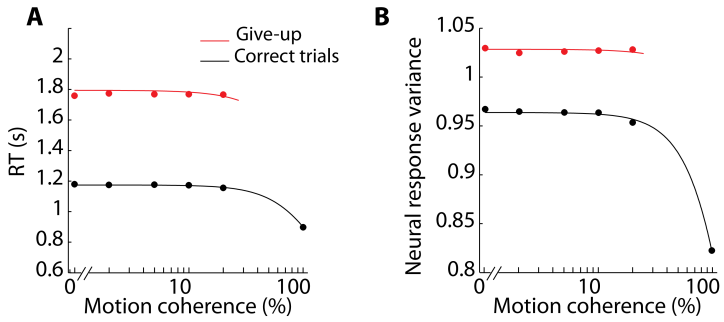

Figure 5.7: RT and neural response variance in simulations. Mean RT **(A)** and neural variance response **(B)** for correct (black) and give-up (red) trials. **(A)** Mean RTs are longer for give-up than for correct trials, as shown in the experimental data (Fig. 5.3B). **(A)** Neural response variance is higher in give-up trials compared to correct trials, following the same trend as in (A).

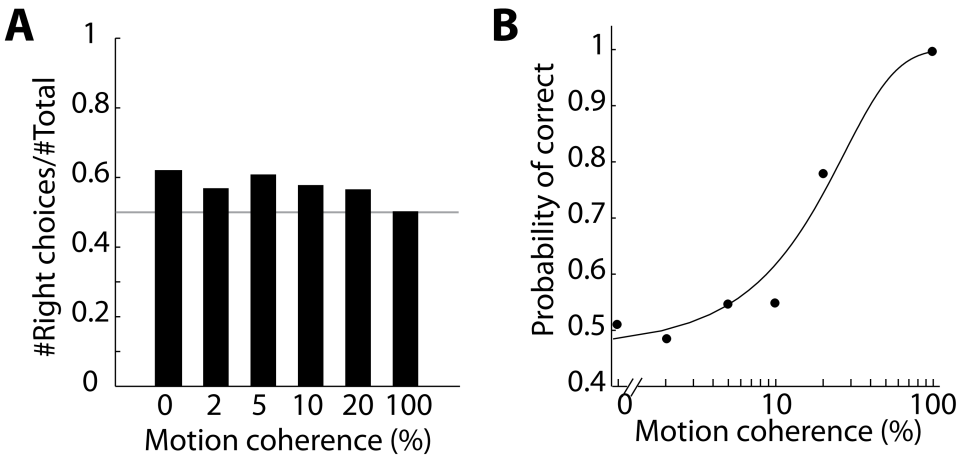

Figure 5.8: Behavioral performance in a control experiment. **(A)** Proportion of right motion detection in a control experiment. 3 of the 8 subjects from the main experiment performed a control experiment in which they were required to detect the direction of motion of the moving dots and to indicate it using the keyboard, where 'j' meant left direction, 'k' meant right direction and the 'space bar' was used as the 'give-up' option. The proportion of right motion detection was above 0,5 along all motion coherences showing a bias towards this direction. As in the main experiment, the number of times that subjects selected the 'give-up' option decreased as the difficulty of the trial also decreased: 25,14%, 23,43%, 15,73%, 18,18%, 3,37% and 0% of the trials for 0%, 2%, 5%, 10%, 20% and 100% of motion coherence respectively. **(B)** Probability of correct for different levels of motion coherence. As in the main experiment the probability of being correct is around chance level for 0% coherence and reaches its maximum value of 1 for 100% coherence.

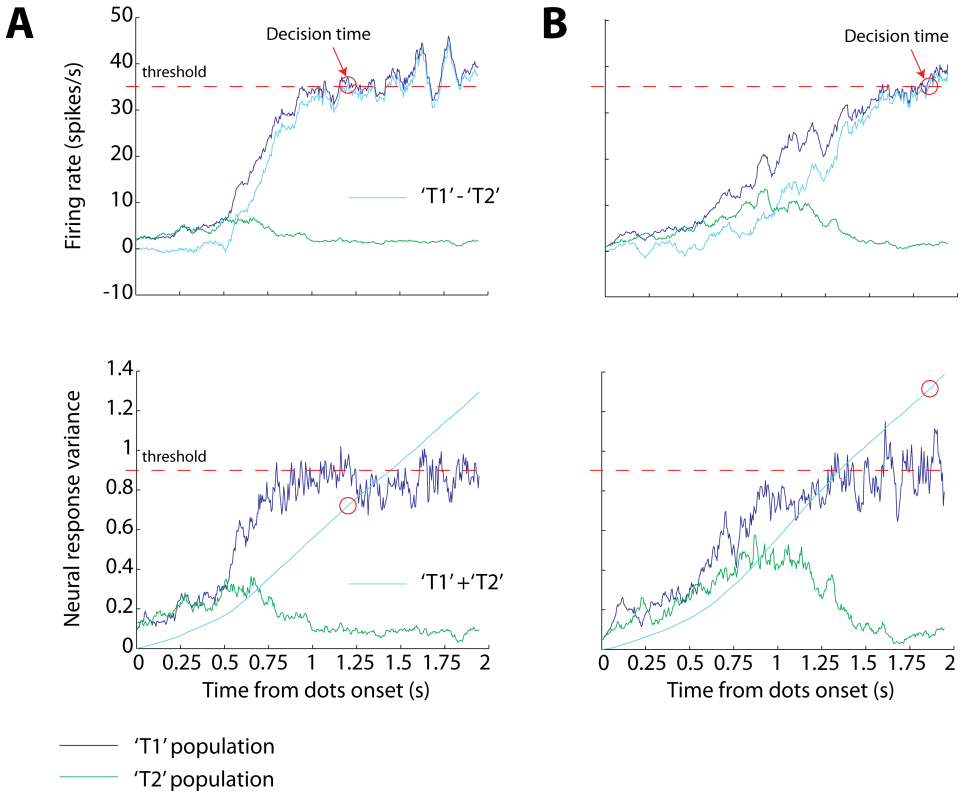

Figure 5.9: Example of the model dynamics in two trials resulting in different outcome: correct choice **(A)** and give-up choice **(B)**. Red dashed horizontal line indicates the decision threshold for mean firing rate (Top panel) and the threshold for the neural variance to select the give-up choice. Top panel shows the mean firing rate activity for each selective population (T1 and T2) and the difference in their activation (T1 mean firing rate minus T2 mean firing rate, cyan color). Bottom panel shows the neural variance of the two selective populations (T1 and T2) and the sum of these variances integrated over time and normalized by the number of neurons in the population (cyan color). When the difference in mean firing rate activity between T1 and T2 populations reaches the decision threshold the decision is considered to be terminated. At that same moment in time the value of the integrated variance is assessed: if it is above the variance threshold the choice is considered to be give-up, otherwise it is considered to be T1 or T2 depending on the population with the highest value of mean firing rate. **(A)** A trial that ended in a T1 choice. As observed, at the time of the decision the variance of the population response is below threshold and therefore T1 will be selected. **(B)** A trial that led to give-up choice is illustrated. In this case, at the time of the decision the integrated variance is above threshold resulting in the selection of the give-up option.

Chapter 6

Conclusions

The important thing is not to stop questioning.
Curiosity has its own reason for existing
Albert Einstein

In this thesis, we have used a combined experimental and theoretical approach to make progress towards a better understanding of embodied decision making. To this end, the studies presented here have focused on the influence that key elements related to embodiment -such as motor cost or experience- have in the neural correlates of decision making and the subsequent behavior. More specifically, we have addressed the following questions: how do existing decision-making models generalize to real-world situations in which an embodied agent is situated? (Chapter 2)What are the underlying mechanisms that cause a bias in behavior based on previous experience during a task? (Chapter 3) Does decision making depend on the context in which the embodied system is situated? (Chapter 4)Do the physical constraints of a body influence decision making even when it compromises accuracy? (Chapter 5). To shed some light on these issues, we have as stated before- analyzed them in a combined manner, by using experimental data and by working on the development of a decision-making model.

In the following sections, we revisit the specific contribution of each

study in the same order as they appear in this thesis and, furthermore, we propose new opportunities for future research they may originate. We finish this chapter by suggesting a model that integrates all the key elements related to embodiment as examined in this thesis.

6.1 Contributions

Generalization of integrator models to foraging

In a first stage, we have used a cognitive architecture to investigate the generalization of a well-known decision-making model to real-world situations (Chapter 2). To this purpose, we have extended the DAC architecture with an integrator model, called race model, and we have provided an artificial embodied agent along with it. The DAC architecture consists of three layers of different complexity that interact with one another and need to be orchestrated in order to produce proper behavioral patterns which lead to achieve specific goals. The added race model consisted of multiple processes that independently accumulated evidence in favor of each possible action proposed by each layer, and limited to the physical constraints of the embodied agent. Only when one of the accumulator processes reached a certain level of activation (threshold) was the decision made and, subsequently, the action triggered. The system was tested in foraging tasks with increased difficulty; in each task, the goal of the agent was to learn about a new environment and to use the acquired information to reach specific positions on it. Our study has revealed that the accumulation of evidence before committing to an action led to a highly robust system against sensory noise; additionally, it directly implied a compression of information stored in the memory. Interestingly, in comparison with a non-accumulator model of decision-making, this compression prompted a change in the way that information about a task was stored. In other words, it has opened a debate about the kind of information that is really relevant and therefore should be memorized, i. e., is it formed by the discrete components of each action? Or is it, in contrast, related to the goal itself and the action self-instantaneously generated? A further robot experiment suggested that the

latter might be the case.

In a second study included in the same chapter, we have gone one step further by enhancing the reactive layer of the DAC architecture with a homeostatic system, which regulated the internal drives of the artificial agent. The reactive control produced continuous actions triggered by direct stimulation of specific sensors associated with external events. The sequence of actions was converted into one unique goal by the contextual control. The goals were retrieved from memory and the actions required to reach them were self-generated. The study has shown an integrated manner in which both reactive and contextual controls could operate to control behavior in an optimal way. Reactive control was shown to be sufficient in tasks in which there was a direct path between the artificial agent and the next goal, but it proved to be insufficient when this was not the case. The self-generation of actions performed by the contextual control played a fundamental role in the success of tasks where obstacles were placed between goals and the agent.

Further research could be focused on a dynamical modulation of the accumulation of evidence for the race model. The mechanism could be based on the certainty/uncertainty that the embodied agent has about a specific location or goal. For instance, when a task is still being learned, actions proposed by the reactive or adaptive layer could have a greater impact on the decision-making process. This impact might systematically decrease as the agent learns about the task, in favor of the actions proposed by the contextual control. We predict that this modulation could lead to an optimal integration of information from the different layers where different levels of learning are considered.

Trial history modulates neural variability and performance

Once a task has been learned, perception and memory need to be integrated in order to properly make a decision. In the DAC architecture, this is achieved by biasing the behavior of the artificial agent according to the sequential order in which external events are encountered. In Chapter 3,

we have examined the neural mechanism causing this bias in behavior. To this end, we used a countermanding reaching task because it had previously been corroborated that the behavior in a current trial was modulated by the recent history of that trial (Emeric et al., 2007). Using experimental data from Mirabella et al. (2011), where two monkeys performed a countermanding reaching task, we have concluded that the same modulation of behavior as previously reported- was also found in our case. With the aim to find the neural substrate causing that bias in behavior, we have looked at the mean firing rate of single neurons recorded from PMd as well as at their across-trial variability. Our analyses have shown that the mean firing rate of the neuron did not show any modulation due to trial history, whereas their across-trial response variability (as measured by VarCE) did show a modulation. In addition, we have also demonstrated that VarCE could be used as a predictor of RTs, since both variables have proven to be strongly correlated.

Already existing decision-making models that are based on accumulation of sensory evidence propose that, to account for changes in RT, either baseline activity, decision threshold or the rate of the accumulation have to be modulated. However, we did not find any of these modulations in our neural data. For this reason, in order to explain the differences in RTs, we have used an attractor network for binary decision making in which two neural populations of excitatory neurons compete through mutual inhibition. We have enhanced this model with a monitoring system that modulated the strength of the input to the two competing populations, while the one related to perception was kept constant. Our simulations have shown that the addition of a memory-related signal could directly account for changes in the variability of the across-trial response, while the mean firing rate did not exhibit significant variation. Moreover, by using the difference in activation between neural populations, we could find an explanation for the variations in RT.

This study has contributed to the investigation of decision making with two important advances: first, it has demonstrated that neurons in

PMd encode trial history in their across-trial neural response variability and, second, it has predicted the existence of a trial history monitoring system that modulates neural activity and behavior. Further research could be focused on trying to confirm whether this monitoring system exists or not and -providing it does-, in which area of the brain it is originated. In addition, neural data simultaneously recorded from neurons in PMd could shed some light on the possible specific role of the neural response variance in a single trial. We have also intended to find an answer to the latter question, about the possible functional role of the neural response variability, in the theoretical study presented in Chapter 5.

Task-dependent modulation of decision-making

In Chapter 4, we have contributed to the study of decision making by reconciling two experimental results that seemed to be contradictory. The experiments were based on tasks in which the stimulus changed over time; in both tasks, monkeys or human subjects were requested to detect the direction of motion of the majority of dots presented on a screen. These new experimental paradigms resulted in sharply opposite conclusions. One of the tasks proved that short changes in the amount of coherent motion direction of the dots (pulses) altered the response time of monkeys performing the task, thus observing a higher impact on the responses when the pulses occurred early in the trial. On the contrary, the other task showed that the response time in human subjects did not vary when a long pulse occurred early in the trial. So far, none of the previously existing decision-making models had been able to explain the reason for this difference in the results.

We have enhanced the model from Chapter 3 with a task-dependent urgency signal and we have demonstrated that it could explain the apparently contradictory results, since the urgency signal changed the gain of the integration process during the course of a trial. Our study indicates that the only difference between both tasks was the starting level and the slope of that urgency signal, which depended on the given context. A high value of urgency made it more likely for the network to reach an attractor

state, whereas the opposite happened when the urgency value was low. In an extreme case, a low value of urgency could lead to a loss of information after some time. Hence, the high starting value of urgency could justify the observed results in the RDM task with pulses -i. e. high impact of early pulses on behavior- whereas low values of urgency could provide reasons for the absence of effect of long early pulses on RT (like the lack of impact of early information on behavior) during the VMD task. The way in which the urgency signal was acquired in each context still remains unknown, though future research could focus on a detailed analysis of the behavior of human subjects or monkeys during the first part of the tasks in order to ascertain it. The different manner in which they confronted the first trials could clarify how the decision-making strategy was adopted in each context, aside from provide further information about the adoption of different strategies in each specific context.

Our model made a strong prediction: if such a task-dependent urgency really exists, then when trials from one task are interleaved during the performance of the other, a reverse effect of pulses must be observed in behavior. In order to validate our model, we designed a new experimental paradigm which enables us to test this prediction. The experiment was run and the data analyzed in the University of Montreal as a collaborative work. The experimental results confirmed that human subjects had a bias caused by early pulses in their RT when they performed a RDM task with pulses. Additionally, the study revealed that, when subjects encountered the very same trials during the performance of a VMD task, they did not have a bias in their RT due to early pulses; however, they did have a bias when late pulses were presented. These experimental results confirmed the validity of our approach.

We suggest that further research could be focused on the specific neural modulation, implementing the task-dependent urgency signal that our model proposed. Making recordings from neurons in areas involved in the decision-making process, such as LIP (Roitman and Shadlen, 2002; Leon and Shadlen, 2003) or PMd (Mirabella et al., 2011), while monkeys

perform the RDM task with pulses and the VMD task could help to clarify this issue.

Motor cost in decision making

The final contribution of this thesis has highlighted the importance of motor cost in decision making. We performed an experiment with human subjects in which they were presented with RDM trials and had to report their decision with a planar movement towards one of two targets presented on the screen. The actions required to reach each distinct target differed in their biomechanical cost. Our experimental results revealed that subjects significantly tended to select the option that required less biomechanical cost more often, even if that tendency compromised their performance. An additional theoretical study, using a spiking neural model of binary decision making, proved that an asymmetric weight in the lateral connectivity of the neurons involved in the decision process could explain the observed bias. We therefore inferred an association between this asymmetry and the experience that a behaving system has during life. Particularly, we proposed that an embodied system generally tends to perform certain actions more frequently than others because of its physical constraints, and that this tendency is intrinsically represented at the neural level.

Our study has also revealed the potential functional role of neural response variability in behavior. We could explain trials in which subjects reported, by selecting a third target (give-up target), that they could not detect the motion coherence of the dots with the response variability of the neural network. The model generated actions towards one of the two targets connected to a specific motion direction only when the mean firing rate of the population related to it was above threshold and the accumulated response variability of the populations was bellow threshold. When the latter did not occur, the trial was classified as a give-up trial.

Simultaneous recordings from neurons in areas involved in decision making, such as the parietal and frontal areas (Platt and Glimcher, 1999; Gold and Shadlen, 2000; Cisek and Kalaska, 2005; Gold and Shadlen, 2007;

Mirabella et al., 2011), during tasks with a give-up option could verify the functional role of variability in this process. Following this line of research, the question whether a link between uncertainty and neural response variability exists, as predicted by our studies in Chapter 3 and Chapter 5), could be answered.

6.2 Concluding remarks

The studies conducted in this thesis have contributed to the advance in the understanding of embodied decision making by providing the field with new experimental results both at the behavioral and the neural level-, and by helping to develop a decision-making model that could explain each finding. Overall, the experimental results together with the theoretical approaches followed in this research have allowed for an improved knowledge of the behavioral and neural mechanisms implicated in embodied decision making.

In Fig. 6.1, we highlight the main contributions of our studies in a single integrated model. The decision-making process is represented by populations of excitatory and inhibitory neurons with recurrent connections. All excitatory connections inhibit one another through the inhibitory neural population, and each excitatory population is sensitive to a specific perceptual feature of a stimulus. The competition between the neural populations is biased in three ways: by a trial history monitoring system, by the physical constraints of an embodied system and by the context in which the behaving system is placed. On the one hand, the monitoring system modulates the input into the neural populations based on recent experience, causing a bias in behavior due to memory. On the other hand, the physical constraints of a body are represented in the lateral recurrent connections of the excitatory neurons. Neural populations associated with movements that have proven to be less costly during life experience have stronger recurrent connections. This way, their chances of winning the competition are higher than in the case of movements associated with high cost. And then again, the given context directly influences the gain of the integration of excitatory populations. This gain varies over the course of a trial differing in starting

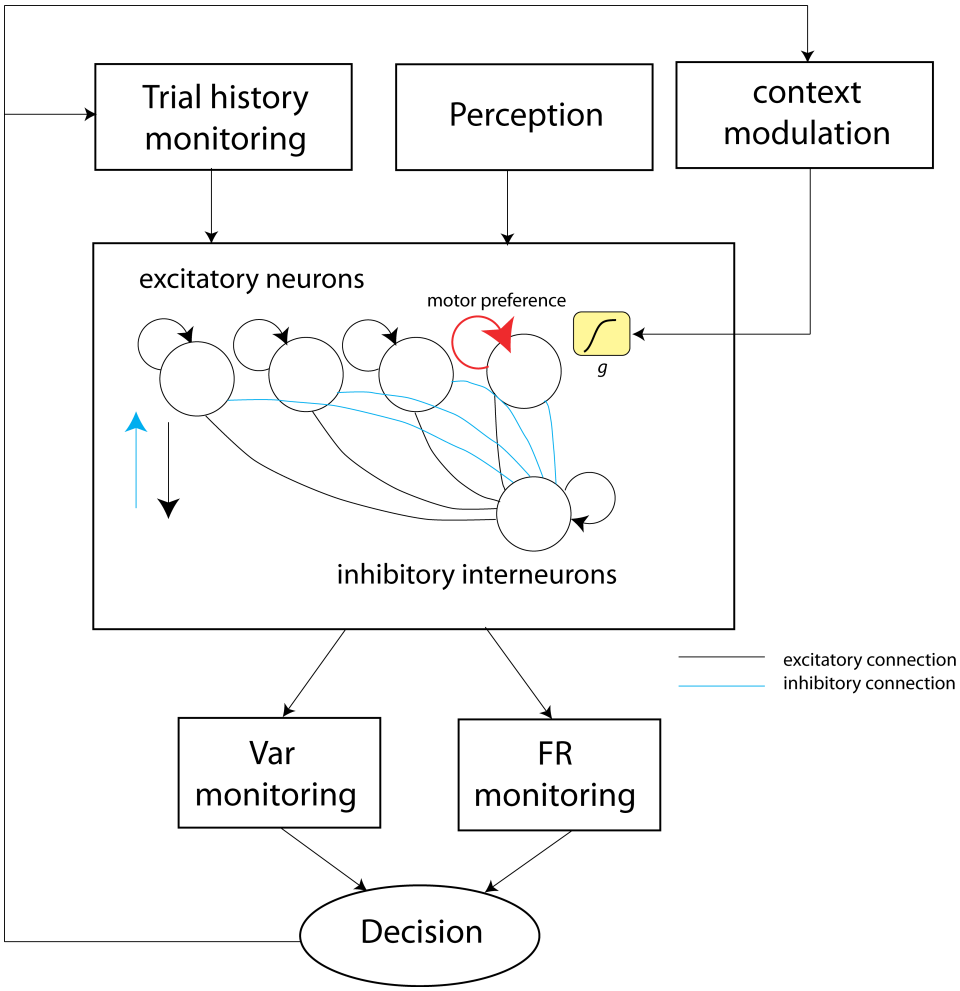

Figure 6.1: Integrated model. Summary of the main contributions of this thesis from a theoretical point of view. The core part of the model consist in a network of excitatory neurons that mutually inhibit each other through a group of inhibitory neurons. All the neural groups have recurrent lateral connections. The activity of the neural network is biased in three ways: through a monitoring system that modulates the perceptual input based on the recent trial history, through the motor cost of actions that are represented in the lateral recurrent connections and through the given context that varies the gain of the integration process within a trial and dependent on the current context. The read-out of the network consists of two processes: a variance (Var) and a firing rate (FR) monitoring processes. Monitoring system and context regulate their modulation based on the outcome of the decision.

value and slope according to context. The modulation produced by the monitoring system and by the context changes is based on the outcome of the decision-making process. In the former case, recent experience updates the value of the signal provided by the monitoring system; in the latter, the outcome of the decision modulates the starting and slope values of a task-dependent signal that, in turn, modifies the way in which the gain of the integration process changes within a trial. In order to make a decision, our integrated model looks at both mean firing rate and accumulated response variance. When the mean firing rate reaches a threshold, the level of the accumulated variance is checked and, if it is, besides, above a specific threshold, the model decides to give-up, or else the decision is made towards the action associated with the winning neural population.

The model clearly shows the contribution of many different processes, involving different brain areas, to decision making. LIP (Roitman and Shadlen, 2002; Leon and Shadlen, 2003), FEF (Gold and Shadlen, 2003), SC (Munoz et al., 2000; Ratcliff et al., 2007; Shen and Paré, 2007) or PMd (Mirabella et al., 2011) have shown to be some of these areas, in which the activity of neurons is correlated with the formation of decisions. In addition, activity of neurons in MT/V5 has shown to be tuned to visual motion (Ditterich et al., 2003; Newsome and Paré, 1988; Salzman et al., 1990, 1992) and to project into LIP, in this way, providing perceptual input to LIP. An open question is where the processes that bias the activity in decision-making related areas, such as the trial history monitoring system, might take place. Further research, conducted with areas recorded simultaneously, could shed some light on this issue.

The integrated model summarizes our contributions to the understanding of the neural mechanisms underlying embodied decision making. What is more, it also provides a new framework for future theoretical research that may lead to the design of new experimental paradigms, and so the loop of experimental-theoretical approaches is restarted. It equally demonstrates the implications of our work not only in biological science and neuroscience but also in applied sciences such as robotics. This denotes the

multidisciplinary character of the studies conducted here, also /visible in the diversity of conferences and workshops where they have been presented (see Publications).

To conclude, the behavioral and neural findings and the theoretical models included in this thesis, as well as the future lines of research proposed, complement and improve our understanding of embodied decision making. Indeed, we believe that these studies may constitute an important impetus for an increased shift of main focus, in which embodiment will be considered a highly relevant factor for decision making, to such an extent that the process cannot be completely understood without it.

Bibliography

Abbott, L. F. Realistic synaptic inputs for model neural networks. *Network*, 2:245–258, 1991. 90

Afraz, S-R., Kiani, R., and Esteky, H. Microstimulation of inferotemporal cortex influences face categorization. *Nature*, 442:692–695, 2006. 11

Albantakis, L. and Deco, G. The encoding of alternatives in multiple choice decision making. *Proceedings of the National Academy of Sciences USA*, 106(25):10308–10313, 2009. 11, 71, 86

Allred, S., Liu, Y., and Jagadeesh, B. Selectivity of inferior temporal neurons for realistic pictures predicted by algorithms for image database navigation. *Journal of Neurophysiology*, 94:4068–81, 2005. 11

Amit, D. J. and Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex*, 7(3):237–52, 1997. 89

Andersen, R. A., Asanuma, C., Essick, G., and Siegel, R. M. Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *Journal of Comparative Neurology*, 296:65–113, 1990. 13

Andersen, R. A., Brotchie, P. R., and Mazzoni, P. Evidence for the lateral intraparietal area as the parietal eye field. *Current Opinion in Neurobiology*, 2:840–46, 1992. 13

Asaad, W. F., Rainer, G., and Miller, E. K. Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, 21:1399–1407, 1998. 48

Asanuma, C., Andersen, R. A., and Cowan, W. M. The thalamic relations of the caudal inferior parietal lobule and the lateral prefrontal cortex in monkeys: divergent cortical projections from cell clusters in the medial pulvinar nucleus. *Journal of Comparative Neurology*, 241:357–81, 1985. 13

Bair, W., Zohary, E., and Newsome, W.T. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *Journal of Neuroscience*, 21:1676–1697, 2001. 146

Barca, L. and Pezzulo, G. Unfolding visual lexical decision in time. *PLoS ONE*, 7(4):e35932, 2012. 29

Bayes, T. Towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370–418, 1763. 36

Baylis, V., Salter, L., and Locke, R. Pathways for continence care: an audit to assess how they are used. *British Journal of Nursing*, 12:857–63, 2003. 11

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., and Pouget, A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron*, 74:30–39, 2012. 28

Bichot, N. P., Chenchal Rao, S., and Schall, J. D. Continuous processing in macaque frontal cortex during visual search. *Neuropsychologia*, 39: 972–982, 2001. 103

Blatt, G. J., Andersen, R. A., and Stoner, G. R. Visual receptive field organization and cortico- cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Comparative Neurology*, 299: 421–45, 1990. 13

Bogacz, R. and Gurney, K. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation*, 19: 442–477, 2007. 10, 21, 100

Boucher, L., Palmeri, T. J., Logan, G. D., and Schall, J. D. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychological Review*, 114(2):376–97, 2007. 23

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Biophysical Journal*, 12:4745–4765, 1992. 11, 12, 101, 131, 143

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, 10:1157–1169, 1993. 12, 146

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, 13:87–100, 1996. 146

Brunel, N. and Wang, X. J. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, 11(1):63–85, 2001. 89

Butler, A. and Hodos, W. *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation*. Wiley, 1996. 16

Carello, C. D. and Krauzlis, R. J. Manipulating intent: evidence for a causal role of the superior colliculus in target selection. *Neuron*, 43(4):575–83, 2004. 29

Carland, M., Marcos, E., Thura, D., and Cisek, P. Perceptual decisions are better explained by urgency-gating than by sensory accumulation. *In prep.*, 2014. 110

Chafee, M. V. and Goldman-Rakic, P. S. Inactivation of parietal and prefrontal cortex re- veals interdependence of neural activity during memory-guided saccades. *Journal of Neurophysiology*, 83:1550–66, 2000. 13

Chapuis, A. and Gélis, E. *Le monde des automates; étude historique et technique*, volume 2. E. Gélis, 1928. 7

Chen, X., Scangos, K. W., and Stuphorn, V. Supplementary motor area exerts proactive and reactive control of arm movements. *Journal of Neuroscience*, 30(44):14657–75, 2010. 17

Churchland, A.K., Kiani, R., and Shadlen, M.N. Decision-making with multiple alternatives. *Nature Neuroscience*, 11:693–702, 2008. 14, 23, 71, 100, 103, 122

Churchland, A.K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., and Shadlen, M.N. Variance as a signature of neural computations during decision making. *Neuron*, 69(4):818–831, 2011. 26, 27, 73, 81, 85, 88, 89, 94, 140

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Santhanam, G., Sahani, M., and Shenoy, K. V. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, 13:369–378, 2010. 26, 71

Churchland, M.M., Yu, B.M., Ryu, S.I., Santhanam, G., and Shenoy, K.V. Neural variability in premotor cortex provides a signature of motor preparation. *Journal of Neuroscience*, 26(14):3697–3712, 2006. 71

Cisek, P. and Kalaska, J.F. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron*, 45(5):801–814, 2005. 29, 71, 75, 141, 163

Cisek, P., Puskas, G. A., and El-Murr, S. Decisions in changing conditions: the urgency-gating model. *Journal of Neuroscience*, 29(37):11560–71, 2009. 23, 100, 101, 102, 106, 111, 112, 114, 140

Clarke, D. M. *Descartes' philosophy of science (Studies in intellectual history)*. Pennsylvania State Univ. Pr. (Txt), 982. 7

Coe, B., Tomihara, K., Matsuzawa, M., and Hikosaka, O. Visual and anticipatory bias in three cortical eye fields of the monkey during an adaptive decision-making task. *Journal of Neuroscience*, 22(12):5081–90, 2002. 29

Cooper, J.M. and Hutchinson, D.S. *Plato: complete works*. Hackett Pub., 1997. 6

Cos, I., Bélanger, N., and Cisek, P. The influence of predicted arm biomechanics on decision-making. *Journal of Neurophysiology*, 105(6):3022–

3033, 2011. 10, 19, 20, 130, 131, 132, 138, 147, 148, 149

Cos, I., Medleg, F., and Cisek, P. The modulatory influence of end-point controllability on decision-making of motor actions. *Journal of Neurophysiology*, 105(6):1764–1780, 2012. 10, 19, 130, 132, 138, 147

de Almeida, L., Idiart, M., and Lisman, J.E. A second function of gamma frequency oscillations: An E%-max winner-take-all mechanism selects which cells fire. *Journal of Neuroscience*, 29(23):7497–7503, 2009. 38, 66

De Rosa, R. Cartesian sensations. *Philosophy Compass*, 4(5):780–792, 2009. 8

Deco, G. and Hugues, E. Neural network mechanisms underlying stimulus driven variability reduction. *PLoS Computational Biology*, 8(3):e1002395, 2012. 79

Del Giudice, P., Fusi, S., and Mattia, M. Modelling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. *Journal of Physiology*, 97(4-6):659–81, 2003. 89

Destexhe, A., Mainen, Z.F., and Sejnowski, T.J. Kinetic models of synaptic transmission. *Methods in Neuronal Modeling*, page 126, 1998. 151

Ditterich, J. Stochastic models of decisions about motion direction: behavior and physiology. *Neural Networks*, 19:981–1012, 2006. 22, 23

Ditterich, J., Mazurek, M., and Shadlen, M. N. Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature Neuroscience*, 6: 891–98, 2003. 13, 166

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., and Friston, K. J. How the brain learns to see objects and faces in an impoverished context. *Nature*, 389:596–99, 1997. 11

Domenech, P. and Dreher, J.C. Decision threshold modulation in the human brain. *Journal of Neuroscience*, 30(43):14305–14317, 2010. 83

Dorris, M. C. and Glimcher, P. W. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*, 44 (2):365–78, 2004. 29

Dounskaia, N., J., Goble, and W., Wang. The role of intrinsic factors in control of arm movement direction: Implications from directional preferences. *Journal of Neurophysiology*, 105:999–1010, 2011. 129, 147

Drai, D., Benjamini, Y., and Golani, I. Statistical discrimination of natural modes of motion in rat exploratory behavior. *Journal of Neuroscience Methods*, 96:119–131, 2000. 50

Dröscher, A. The history of the golgi apparatus in neurones from its discovery in 1898 to electron microscopy. *Brain Research Bulletin*, 47(3): 199–203, 1998. 8

Duff, A. and Verschure, P. F. Unifying perceptual and behavioral learning with a correlative subspace learning rule. *Neurocomputing*, 73(10-12): 1818–1830, 2010. 36, 37, 63

Duff, A., Rennó-Costa, C., Marcos, E., Luvizotto, A.L., Giovannucci, A., Sánchez Fibla, M., Bernardet, U., and Verschure, P.F.M.J. Distributed adaptive control: A proposal on the neuronal organization of adaptive goal oriented behavior. In *From Motor Learning to Interaction Learning in Robots*, pages 15–41. In: Sigaud, O., Peters, J. (eds.) From Motor Learning to Interaction Learning in Robots, 2010. ISBN 978-3-642-05180-7. 30, 51

Emeric, E.E., Brown, J.W., Boucher, L., Carpenter, R.H., Hanes, D.P., Harris, R., Logan, G.D., Mashru, R.N., Par, M., Pouget, P., Stuphorn, V., Taylor, T. L., and Schall, J. D. Influence of history on saccade countermanding performance in humans and macaque monkeys. *Vision Research*, 47(1):35–49, 2007. 10, 18, 72, 73, 80, 81, 160

Ergil, K. V. Greek & roman medicine. In *Ancient Healing: Unlocking the Mysteries of Health and Healing Through the Ages*, pages 68–98. Publications International Ltd., 1997. 5

Faisal, A. A., Selen, L. P., and Wolpert, D. M. Noise in the nervous system. *Nature Reviews Neuroscience*, 32:292–303, 2008. 28

Fodor, J. A. *Modularity of Mind*. MIT Press, Cambridge, MA, USA, 1983. 29

Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. Visual

categorization and the primate prefrontal cortex: Neurophysiology and behavior. *Journal of Neurophysiology*, 88:929–41, 2002. 11

Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, 23:5235–46, 2003. 11

Friedman, H. R. and Goldman-Rakic, P. S. Coactivation of prefrontal cortex and inferior parietal cortex in working memory tasks revealed by 2DG functional mapping in the rhesus monkey. *Journal of Neuroscience*, 14: 2775–88, 1994. 13

Fries, W. Cortical projections to the superior colliculus in the macaque monkey: a retrograde study using horseradish peroxidase. *Journal of Comparative Neurology*, 230:55–76, 1984. 13

Geisler, W. S. and Albrecht, D. G. Bayesian analysis of identification performance in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. *Vision Research*, 35:2723–2730, 1995. 88

Goble, J. A., Zhang, Y., Shimansky, Y., Sharma, S., and Dounskaia, N. V. Directional biases reveal utilization of arms biomechanical properties for optimization of motor behavior. *Journal of Neurophysiology*, 98:1240–1252, 2007. 147

Gold, J.I. and Shadlen, M.N. Representation of a perceptual decision in developing oculomotor commands. *Nature*, 404:390–394, 2000. 71, 141, 163

Gold, J.I. and Shadlen, M.N. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience*, 23:632–651, 2003. 10, 49, 166

Gold, J.I. and Shadlen, M.N. The neural basis of decision making. *Annual Review of Neuroscience*, 30:535–574, 2007. 15, 19, 29, 34, 71, 99, 129, 141, 163

Golgi, C. Sulla struttura della sostanza grigia dell cervello. *Gazzetta Medica Italiana - Lombardia*, 33:244–246, 1873. 8

Grammaticos, P. C. and Diamantis, A. Useful known and unknown views of the father of modern medicine, hippocrates and his teacher democritus.

Hellenic Journal of Nuclear Medicine, 11(1):2–4, 2008. 6

Green, D. M. and Swets, J. A. *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966. 21

Grill-Spector, K., Kushnir, T., Hendler, T., and Malach, R. The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3:837–4, 2000. 11

Gritsenko, V., Kalaska, J. F., and Cisek, P. Descending corticospinal control of intersegmental dynamics. *Journal of Neuroscience*, 31:11968–11979, 2011. 147

Gur, M., Beylin, A., and Max Snodderly, D. M. Response variability of neurons in primary visual cortex (V1) of alert monkeys. *Journal of Neuroscience*, 17(8):2914–2920, 1997. 140

Hanes, D. P. and Schall, J. D. Neural control of voluntary movement initiation. *Science*, 274:427–430, 1996a. 48

Hanes, D. P., Patterson, W. F., and Schall, J. D. Role of frontal eye fields in countermanding saccades: visual, movement, and fixation activity. *Journal of Neurophysiology*, 79:817–834, 1998. 17, 18

Hanes, D.P. and Schall, J.D. Neural control of voluntary movement initiation. *Science*, 274(5286):427–430, 1996b. 72, 103

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., and Shadlen, M. N. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *Journal of Neuroscience*, 31(17):6339–6352, 2011. 129

Hanks, T.D., J., Ditterich, and Shadlen, M.N. Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature Neuroscience*, 9:682–689, 2006. 101, 143

Hartland, C., Bredechem, N., and Sebag, M. Memory-enhanced evolutionary robotics: the echo state network approach. In *Proceedings of the Eleventh conference on Congress on Evolutionary Computation*, pages 2788–2795, 2009. 56

Heekeren, H. R., Marrett, S., Bandettini, P. A., and Ungerleider, L. G. A

general mechanism for perceptual decision-making in the human brain. *Nature*, 431:859–862, 2004. 129

Hernandez, A., Zainos, A., and Romo, R. Temporal evolution of a decision-making process inmedial premotor cortex. *Neuron*, 33(6):959–72, 2002. 29

Hestrin, S., Sah, P., and Nicoll, R. Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron*, 5:247–253, 1990. 151, 152

Hogan, N. Impedance control: an approach to manipulation: part I-theory. *Journal of Dynamic Systems, Measurement, and Control*, 107:1–7, 1985a. 141, 147

Hogan, N. Impedance control: an approach to manipulation: part II-implementation. *Journal of Dynamic Systems, Measurement, and Control*, 107:8–16, 1985b. 141, 147

Hogan, N. Impedance control: an approach to manipulation: part III-applications. *Journal of Dynamic Systems, Measurement, and Control*, 107:17–24, 1985c. 141, 147

Horwitz, G. D., Batista, A. P., and Newsome, W. T. Representation of an abstract perceptual decision in macaque superior colliculus. *Journal of Neurophysiology*, 91(5):2281–96, 2004. 29

Huettel, S.A., Song, A.W., and McCarthy, G. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *Journal of Neuroscience*, 25(13):3304–3311, 2005. 81

Huk, A. and Shadlen, M. N. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience*, 25(45):10420–10436, 2005. 14, 16, 23, 97, 101, 102, 103, 104, 105, 109, 112, 114, 115, 116, 121, 122

Iizuka, Hiroyuki and Di Paolo, E. A. Extended homeostatic adaptation: Improving the link between internal and behavioural stability. In Asada, M., Hallam, J. C. T., Meyer, J.-A., and Tani, J., editors, *From Animals to Animats 10*, volume 5040 of *Lecture Notes in Computer Science*, pages 1–11. Springer Berlin Heidelberg, 2008. 53

Ito, S., Stuphorn, V., Brown, J.W., and Schall, J.D. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, 302(5642):120–122, 2003. 80

Jahr, C.E. and Stevens, C.F. Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *Journal of Neuroscience*, 10:3178–3182, 1990. 151

Janssen, P. and Shadlen, M. N. A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience*, 8:234–41, 2005. 13

Johnson, P.B. and Ferraina, S. Cortical networks for visual reaching: intrinsic frontal lobe connectivity. *European Journal of Neuroscience*, 8(7): 1358–1362, 1996. 82, 83

Johnson, P.B., Ferraina, S., Bianchi, L., and Caminiti, R. Cortical networks for visual reaching: physiological and anatomical organization of frontal and parietal lobe arm regions. *Cerebral Cortex*, 6(2):102–119, 1996. 82

Jung, M. W. and McNaughton, B. L. Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus*, 3:165–182, 1993. 48

Kiani, R. and Shadlen, M. N. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, 324(5928):759–64, 2009. 133, 139, 140

Kim, J. N. and Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, 2:176–185, 1999. 10, 49, 103

Kirk, G. S. *The presocratic philosophers: a critical history with a selections of texts*. Cambridge University Press, 1983. 6

Kuipers, B. The spatial semantic hierarchy. *Artificial Intelligence*, pages 191–233, 2000. 61

Laming, D. R. J. Information theory of choice-reaction times. *Wiley, New York*, 1968. 10, 21, 100, 106

Leon, M. I. and Shadlen, M. N. Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, 38:317–327, 2003. 10,

13, 162, 166

Lewis, J. W. and Van Essen, D. C. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *Journal of Comparative Neurology*, 428:112–37, 2000. 13

Link, S. W. and Heath, R.A. A sequential theory of psychological discrimination. *Psychometrika*, 40:77–105, 1975. 10, 21, 100

Logan, G. D. and Cowan, W. B. On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91:295–327, 1984. 16, 23, 35, 38, 71, 72, 93

López, L., Vouloutsi, V., Escuredo Chimeno, A., Marcos, E., Bermúdez, S., Ziyatdinov, A., Mathews, Z., Perera Lluna, Alexandre, and Verschure, P. Moth-like chemo-source localization and classification on an indoor autonomous robot. In Pramatarova, Assoc. Lilyana D., editor, *On Biomimetics*, pages 453–466. InTech, 2011. ISBN SBN 978-953-307-271-5.

Luppino, G., Rozzi, S., Calzavara, R., and Matelli, M. Prefrontal and agranular cingulate projections to the dorsal premotor areas F2 and F7 in the macaque monkey. *European Journal of Neuroscience*, 17(3):559–578, 2003. 82, 83

Mackintosh, N. J. *Conditioning and associative learning*, volume 3. Oxford psychology series Reprint. Clarendon Press, 1990. 31

Mallot, Hanspeter A. and Basten, Kai. Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing*, 27(11): 1658–1670, 2009. Cognitive Systems: Perception, Action, Learning. 61

Marcos, E. and Verschure, P. F. M. J. Optimal memory acquisition and recall in real-world foraging. In *FENS Forum 2010 (FENS 2010)*, Amsterdam, Netherlands, 2010.

Marcos, E., Duff, A., Sánchez-Fibla, Martí, and Verschure, Paul F. M. J. The neuronal substrate underlying order and interval representations in sequential tasks: a biologically based robot study. In *International Joint Conference on Neural Networks (IJCNN)*, Barcelona, Spain, 2010a. 36,

44, 46, 47

Marcos, E., Sánchez-Fibla, M., and Verschure, P. F. M. J. The complementary roles of allostatic and contextual control systems in foraging tasks. In Doncieux, S., Girard, B., Guillot, A., Hallam, J., Meyer, J-A., and Mouret, J-B., editors, *From Animals to Animats 11*, volume 6226 of *Lecture Notes in Computer Science*, pages 370–379. Springer Berlin Heidelberg, 2010b. ISBN 978-3-642-15192-7. 49

Marcos, E., Sánchez-Fibla, M., and Verschure, P. F. M. J. From continuous analog to discrete symbolic representations of the world in optimal foraging: a robot study. *4th International Conference on Cognitive Systems (CogSys)*, 2010c.

Marcos, E., Pani, P., Brunamonti, E., Ferraina, S., and Verschure, P. F. M. J. Behavioral and neural modulation during motor decision making in a countermanding task. In *2011 Society for Neuroscience meeting (SfN 2011)*, Washington, US, 2011.

Marcos, E., Carland, M., Thura, D., Cisek, P., and Verschure, P. F. M. J. Decision-making depends on an urgency signal modulated by context. In *2012 Society for Neuroscience meeting (SfN 2012)*, New Orleans, US, 2012a. 98

Marcos, E., Duff, A., Sánchez-Fibla, M., and Verschure, P. F. M. J. Generalization of integrator models to foraging: a robot study using the DAC9 model. In Prescott, T. J., Lepora, N. F., Mura, A., and Verschure, P. F. M. J., editors, *Biomimetic and Biohybrid Systems*, volume 7375 of *Lecture Notes in Computer Science*, pages 156–167. Springer Berlin Heidelberg, 2012b. ISBN 978-3-642-31524-4. 34

Marcos, E., Cos, I., Cisek, P., Girard, B., and Verschure, P. F. M. J. Biomechanical costs of reaching movements bias perceptual decisions. *BMC Neuroscience*, 14(1):P408, 2013a. 128

Marcos, E., Pani, P., Brunamonti, E., Deco, G., Ferraina, S., and Verschure, P. Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. *Neuron*, 78(2):249–255, 2013b. 70, 132, 138, 139, 140

Marcos, E., Ringwald, M., Duff, A., Sánchez-Fibla, M., and Verschure, P. F. M. J. The hierarchical accumulation of knowledge in the distributed adaptive control architecture. In Baldassarre, Gianluca and Mirolli, Marco, editors, *Computational and Robotic Models of the Hierarchical Organization of Behavior*, pages 213–234. Springer Berlin Heidelberg, 2013c. ISBN 978-3-642-39874-2. 30

Marcos, E., Carland, M., Thura, D., Cisek, P., and Verschure, P. F. M. J. Decision-making depends on a task-dependet urgency. *In prep.*, 2014a. 97

Marcos, E., Cos, I., Cisek, P., Benot, G., and Verschure, P. F. M. J. The cost of embodiment in decision-making. *In prep.*, 2014b. 128

Mathews, Z., Lechón, M., Blanco-Calvo, J. M., Dhir, A., Duff, A., Bermúdez i Badia, S., and Verschure, P. F. M. J. Insect-like mapless navigation based on head direction cells and contextual learning using chemo-visual sensors. In *Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, IROS'09, pages 2243–2250. IEEE Press, 2009. 48

Mathews, Z., Bermúdez i Badia, S., and Verschure, P. F. M. J. PASAR: An integrated model of prediction, anticipation, sensation, attention and response for artificial sensorimotor systems. *Information Sciences*, 186 (1):1–19, 2012. 36

Maunsell, J. H. R. and Van Essen, D. C. Functional properties of neurons in the middle temporal visual area (MT) of the macaque monkey: II. binocular interactions and the sensitivity to binocular disparit. *Journal of Neurophysiology*, 49:1148–67, 1983. 11

Mazurek, M.E., Roitman, J. D., J., Ditterich, and Shadlen, M. N. A role for neural integrators in perceptual decision making. *Cerebral Cortex*, 13: 1257–1269, 2003. 10, 21, 100

McEwena, B. S. and Wingfield, J. C. The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, 2:2–15, 2003. 50

Meyer, J.-A., Guillot, A., Girard, B., Khamassi, M., Pirim, P., and Berthoz, A. The psikharpax project: towards building an artificial rat. *Robotics*

and Autonomous Systems, 50(4):211–223, 2005. 51

Mirabella, G., Pani, P., Par, M., and Ferraina, S. Inhibitory control of reaching movements in humans. *Experimental Brain Research*, 174(2): 240–255, 2006. 18, 71, 72, 77, 80

Mirabella, G., Pani, P., and Ferraina, S. Neural correlates of cognitive control of reaching movements in the dorsal premotor cortex of rhesus monkeys. *Journal of Neurophysiology*, 106(3):1454–1466, 2011. 10, 17, 18, 71, 72, 82, 83, 84, 87, 93, 141, 160, 162, 164, 166

Mountcastle, V. B., Steinmetz, M. A., and Romo, R. Frequency discrimination in the sense of flutter: psychophysical measurements correlated with postcentral events in behaving monkeys. *Journal of Neuroscience*, 10:3032–44, 1990. 11

Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., and Forstmann, B. U. Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, 32(7):2335–2343, 2012. 129

Munoz, D. P. and Schall, J. D. Concurrent, distributed control of saccade initiation in the frontal eye field and superior colliculus. In Hall, W. T. and Moschovakis, A., editors, *The superior colliculus: New approaches for studying sensorimotor integration*, pages 55–82. New York: CRC Press, 2003. 23

Munoz, D. P., Dorris, M. C., Paré, M., and Everling, S. On your mark, get set: brainstem circuitry underlying saccadic initiation. *Canadian Journal of Physiology and Pharmacology*, 78:934–944, 2000. 10, 166

Nawrot, M.P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, 169(2):374–390, 2008. 80, 88

Nelson, M.J., Boucher, L., Logan, G.D., Palmeri, T.J., and Schall, J.D. Nonindependent and nonstationary response times in stopping and stepping saccade tasks. *Attention, Perception and Psychophysics*, 72(7):1913–1929, 2010. 18, 72, 80

Newsome, W. T. and Paré, E. B. A selective impairment of motion percep-

tion following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, 8:2201–11, 1988. 13, 166

Niwa, M. and Ditterich, J. Perceptual decisions between multiple directions of visual motion. *Journal of Neuroscience*, 28(17):4435–4445, 2008. 14

O'Malley, C. D. *Andreas Vesalius of Brussels, 1514-1564*. Berkeley: University of California Press, 1964. 7

Op de Beeck, H., J., Wagemans, and Vogels, R. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nature Neuroscience*, 4:1244–52, 2001. 11

Osborne, L. C., Lisberger, S. G., and Bialek, W. A sensory source of motor variation. *Nature*, 437:412–416, 2005. 28

Palmer, J., Huk, A. C., and Shadlen, M. N. The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5: 376–404, 2005. 100

Paré, M. and Hanes, D.P. Controlled movement processing: superior colliculus activity associated with countermanded saccades. *Journal of Neuroscience*, 23(16):6480–6489, 2003. 17, 18, 72

Pavlov, I. P. *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. London: Oxford University Press, 1927. 7, 30, 37

Pesaran, B., Nelson, M. J., and Andersen, R. A. Free choice activates a decision circuit between frontal and parietal cortex. *Nature*, 453(7193): 406–9, 2008. 29

Pezzulo, G., Barsalou, L., Cangelosi, A., Fischer, M., McRae, K., and Spivey, M. J. The mechanics of embodiment: a dialogue on embodiment and computational modeling. *Frontiers in Cognition*, 2:1–21, 2011. 29

Platt, M.L. and Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature*, 400:233–238, 1999. 71, 141, 163

Pouget, P., Logan, G.D., Palmeri, T.J., Boucher, L., Par, M., and Schall, J.D. Neural basis of adaptive response time adjustment during saccade

countermanding. *Journal of Neuroscience*, 31(35):12604–12612, 2011. 73, 82

Pylyshyn, Z. W. *Computation and Cognition: Toward a Foundation for Cognitive Science*. MIT Press, Cambridge, MA, USA, 1984. 29

Rainer, G., Lee, H., and Logothetis, N. K. The effect of learning on the function of monkey extrastriate visual cortex. *PLoS Biology*, 2:e44, 2004. 11

Ramon y Cajal, S. A new concept of the concept of the histology of the central nervous system (1892). In A., Rottenberg D. and Hochberg, F. H., editors, *Neurological Classics in Modern Translation*, page 729. New York, Hafner, 1977. 8

Rao, R. P. N. An optimal estimation approach to visual perception and learning. *Vision Research*, 39:1963–89, 1999. 19

Ratcliff, R. A theory of memory retrieval. *Psychological Review*, 83:59–108, 1978. 10, 21, 100, 106

Ratcliff, R. and Rouder, J. N. Modeling response times for two-choice decisions. *Psychological Science*, 9:347–356, 1998. 35

Ratcliff, R., Cherian, A., and Segraves, M. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *Journal of Neurophysiology*, 90:1392–1407, 2003. 21, 106

Ratcliff, R., Hasegawa, Y. T., Hasegawa, R. P., Smith, P. L., and Segraves, M. A. Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *Journal of Neurophysiology*, 97:1756–1774, 2007. 10, 49, 166

Reddi, B. A. J. and Carpenter, R. H. S. The influence of urgency on decision time. *Nature Neuroscience*, 3(8):827–30, 2000. 100, 106, 122

Redgrave, P., Prescott, T. J., and Gurney, K. The basal ganglia: a vertebrate solution to the selection problem. *Neuroscience*, 89(4):1009–23, 1999. 16

Redish, A. David, editor. *Beyond the Cognitive Map: From Place Cells to*

Episodic Memory. Cambridge, MA, MIT Press, 1999. 61

Renart, A., Brunel, N., and Wang, X. J. Mean-field theory of recurrent cortical networks: from irregularly spiking neurons to working memory. In Feng, Jianfeng, editor, *Computational Neuroscience: A Comprehensive Approach*. A CRC Press Company. Boca Raton, 2003. 89

Rennó-Costa, C., Luvizotto, A., Marcos, E., Duff, A., Saáchez-Fibla, M., and Verschure, P. F. M. J. Integrating neuroscience-based models towards an autonomous biomimetic Synthetic Forager. In *2011 IEEE International Conference on Robotics and Biomimetics*, pages 210–215, Phuket, Thailand, December 2011. IEEE. ISBN 978-1-4577-2138-0.

Rescorla, R. and Wagner, A. A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In (Eds.), A. H. Black & W.F. Prokasy, editor, *Classical Conditioning II: Current Research and Theory. Appleton Century Crofts, New York*, pages 64–99. 1972. 63

Resulaj, A., Kiani, R., Wolpert, D. M., and Shadlen, M. N. Changes of mind in decision-making. *Nature*, 461(7261):263–6, 2009. 139, 146

Rieger, M. and Gauggel, S. Inhibitory after-effects in the stop signal paradigm. *British Journal of Psychology*, 90:509–518, 1999. 18, 72, 80

Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Engineering*, 10:137–145, 1963. 87

Rocca, J. Galen and the ventricular system. *Journal of the History of the Neurosciences*, 6(3):227–39, 1997. 6

Roitman, J .D. and Shadlen, M.N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22:9475–9489, 2002. 10, 12, 13, 49, 99, 100, 101, 103, 105, 143, 162, 166

Romo, R., Hernández, A., Zainos, A., Lemus, L., and Brody, C. D. Neuronal correlates of decision-making in secondary somatosensory cortex. *Nature Neuroscience*, 5(11):1217–25, 2002. 29

Romo, R., Hernández, A., and Zainos, A. Neuronal correlates of a per-

ceptual decision in ventral premotor cortex. *Neuron*, 41(1):165–73, 2004. 29

Roxin, A. and Ledberg, A. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Computational Biology*, 4(3):e1000046, 2008. 78, 79

Sabes, P. N. and Jordan, M. I. Obstacle avoidance and a perturbation sensitivity model for motor planning. *Journal of Neuroscience*, 17:7119–7128, 1997. 19, 147, 148

Sabes, P. N., I., Jordan M., and Wolpert, D. M. The role of inertial sensitivity in motor planning. *Journal of Neuroscience*, 18:5948–5957, 1998. 10, 19, 147

Said, H. M. *Traditional Greco-Arabic and Modern Western Medicine: Conflict Or Symbiosis*. Hamdard National Foundation, Pakistan, 1975. 6

Salin, P.A. and Prince, D.A. Spontaneous GABAA receptor mediated inhibitory currents in adult rat somatosensory cortex. *Journal of Neurophysiology*, 75:1573–1588, 1996. 151

Salzman, C. D., Britten, K. H., and Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346: 174–77, 1990. 13, 166

Salzman, C. D., Murasugi, C. M., Britten, K. H., and Newsome, W. T. Microstimulation in visual area MT: effects on direction discrimination performance. *Journal of Neuroscience*, 12:2331–55, 1992. 13, 166

Sanchez-Fibla, M., Bernardet, U., Wasserman, E., Pelc, T., Mintz, M., Jackson, J. C., Lansink, C., Pennartz, C., and Verschure, P. F. M. J. Allostatic control for robot behavior regulation: a comparative rodent-robot study. *Advances in Complex Systems*, 13(3):377–403, 2010. 52

Scangos, K.W. and Stuphorn, V. Medial frontal cortex motivates but does not control movement initiation in the countermanding task. *Journal of Neuroscience*, 30(5):1968–1982, 2010. 10, 17, 72

Schall, J. D. and Thompson, K. G. Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience*, 22:241–259, 1999. 103

Shadlen, M. N. and Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, 86:1916–1936, 2001. 12, 13, 35, 48

Shadlen, M.N. and Newsome, W.T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18(10):3870–3896, 1998. 80

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience*, 15:3870–3896, 1996. 12, 13, 146

Shen, K. and Paré, M. Neuronal activity in superior colliculus signals both stimulus identity and saccade goals during visual conjunction search. *Journal of Vision*, 7:15.1–13, 2007. 10, 166

Sheynikhovich, D., Chavarriaga, R., Strosslin, T., Arleo, A., and Gerstner, W. Is there a geometric module for spatial orientation? insights from a rodent navigation model. *Psychological Review*, 116(3):540–566, 2009. 51

Smith, P.L. and Ratcliff, R. Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27:161–168, 2004. 21, 34, 71, 106, 129

Spruston, N., Jonas, P., and Sakmann, B. Dendritic glutamate receptor channel in rat hippocampal CA3 and CA1 pyramidal neurons. *Journal of Physiology*, 482:325–352, 1995. 151, 152

Standage, D., You, H., Wang, D.-H., and Dorris, M. C. Gain modulation by an urgency signal controls the speedaccuracy trade-off in a network model of a cortical decision circuit. *Frontiers in Computational Neuroscience*, 5:1–14, 2011. 122

Stone, M. Models for choice reaction time. *Psychometrika*, 25:251–260, 1960. 10, 21, 100, 106

Tenenbaum, J. B. and Griffiths, T. L. Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24:629–40, 2001. 19

Thevarajah, D., Mikulic, A., and Dorris, M. C. Role of the superior colliculus in choosing mixed-strategy saccades. *Journal of Neuroscience*, 29(7): 1998–2008, 2009. 29

Thompson, K. G., Bichot, N. P., and Schall, J. D. Dissociation of visual discrimination from saccade programming in macaque frontal eye field. *Journal of Neurophysiology*, 77:1046–1050, 1997. 103

Thorndike, E. Animal intelligence. *Macmillan, New York*, 1911. 31, 37

Thura, D. and Cisek, P. Deliberation and commitment in the premotor and primary motor cortex during dynamic decision-making. *Neuron*, 81(6): 1401–1416, 2014. 16

Thura, D., Beauregard-Racine, J., Fradet, C.-W., and Cisek, P. Decision-making by urgency-gating: theory and experimental support. *Journal of Neurophysiology*, 108(11):2912–30, 2012. 15, 23, 97, 101, 103, 104, 105, 109, 112, 117, 118, 121, 122

Tolman, E. and Honzik, C. Insights in rats. *University of California Publications in Psychology*, 4(14):215–232, 1930. 56

Tsetsos, K., Gao, J., McClelland, J. L., and Usher, M. Using time-varying evidence to test models of decision dynamics: Bounded diffusion vs. the leaky competing accumulator model. *Frontiers in Neuroscience*, 6:79, 2012. 122

Tuckwell, H. C. Introduction to theoretical neurobiology. *Cambridge University Press, Cambridge, UK*, 1988. 145, 150

Uchida, N. and Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nature Neuroscience*, 6:1224–29, 2003. 11

Uka, T. and DeAngelis, G. C. Linking neural representation to function in stereoscopic depth perception: roles of the middle temporal area in coarse versus fine disparity discrimination. *Journal of Neuroscience*, 26: 6791–802, 2006. 11

Usher, M. and McClelland, J. L. The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, 108:550–592, 2001. 10, 21, 22, 100, 122

Verbruggen, F. and Logan, G.D. Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, 12(11):418–424, 2008. 16, 18, 71, 72, 80

Verschure, P. and Althaus, P. A real-world rational agent: unifying old and new AI. *Cognitive Science*, 27:561–590, 2003. 30, 36, 46, 51, 52

Verschure, P. and Pfeifer, R. Categorization, representations, and the dynamics of system-environment interaction: a case study in autonomous systems. In Meyer, J.-A., L., Roitblat H., and W., Wilson S., editors, *From Animals to Animats: Proceedings of the Second International Conference on Simulation of Adaptive behavior*, pages 210–217. MIT Press, 1992. 63

Verschure, P., Krose, B., and Pfeifer, R. Distributed adaptive control: The self-organization of structured behavior. *Robotics and Autonomous Systems*, 9:181–196, 1993. 30

Verschure, P. F. M. J., Pennartz, C. M. A., and Pezzulo, G. The why, what, where, when and how of goal-directed choice: neuronal and computational principles. *Philosophical Transactions of the Royal Society B: Biological Sciences. In Press*, 2014. 30

Verschure, Paul F. M. J. Distributed adaptive control: A theory of the mind, brain, body nexus. *BICA (In Press.)*, 2012. 36

Verschure, P.F.M.J. and Coolen, A.C.C. Adaptive fields: Distributed representations of classically conditioned associations. *Network*, 17(2):189–206, 1991. 30

Verschure, P.F.M.J., Voegtlin, T., and Douglas, R.J. Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature*, 425(6958):620–624, 2003. 30, 35, 36, 51, 54, 71, 83, 129, 138

Wald, A. *Sequential analysis*, volume 24. New York: Wiley, 1947. 21

Wallis, J. D. and Miller, E. K. From rule to response: neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, 90(3): 1790–806, 2003. 29

Walton, M.E., Devlin, J.T., and Rushworth, M.F.S. Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, 7(11):1259–1265, 2004. 80

Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5):955–968, 2002. 11, 24, 25, 86, 100, 132, 136, 145,

146, 149

Wilson, H.R. and Cowan, J.D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24, 1972. 25, 73, 75, 85, 90, 102, 107

Wong, K.-F. and Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, 26(4):1314–1328, 2006. 25

Wong, K-F., Huk, A. C., Shadlen, M. N., and Wang, X-J. Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision-making. *Frontiers in Computational Neuroscience*, 1:6, 2007. 122, 144

Wyss, R. *Sensory and motor coding in the organization of behavior*. PhD thesis, ETHZ, 2003. 39, 40

Wyss, R., König, P., and Verschure, Paul F. M. J. A model of the ventral visual system based on temporal stability and local memory. *PLoS Biology*, 4(5):e120, 2006. 40, 55

Xiang, Z., Huguenard, J.R., and Prince, D.A. GABAA receptor mediated currents in interneurons and pyramidal cells of rat visual cortex. *Journal of Physiology*, 506:715–730, 1998. 151

Yang, T. and Shadlen, M. N. Probabilistic reasoning by neurons. *Nature*, 447(7148):1075–80, 2007. 29

Yoshida, W. and Ishii, S. Resolution of uncertainty in prefrontal cortex. *Neuron*, 50(5):781–789, 2006. 81

Zandbelt, B.B. and Vink, M. On the role of the striatum in response inhibition. *PLoS ONE*, 5(11):e13848, 2010. 83

Zohary, E., Shadlen, M.N., and Newsome, W.T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370:140–143, 1994. 146