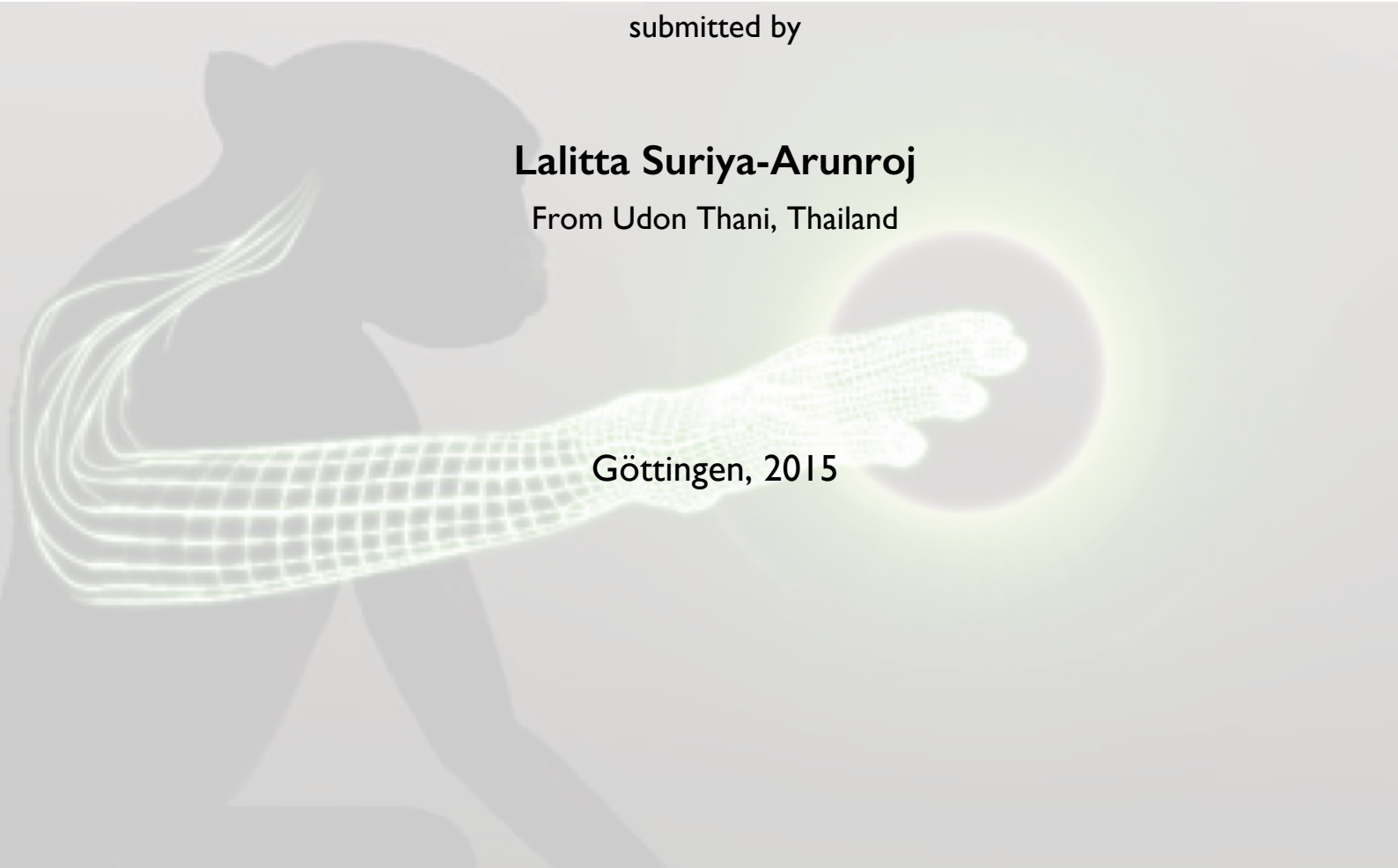
Neural basis of rule-based decisions with graded choice biases

Dissertation
for the award of the degree

"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

within the doctoral program **Sensory and Motor Neuroscience**
of the Georg-August University School of Science (GAUSS)



submitted by

Lalitta Suriya-Arunroj
From Udon Thani, Thailand

Göttingen, 2015

Thesis Committee:

Prof. Dr. Alexander Gail, *Sensorimotor Group, German Primate Center*

Prof. Dr. Stefan Treue, *Cognitive Neuroscience Laboratory, German Primate Center*

Prof. Dr. Florentin Wörgötter, *Institute of Physics III - Biophysics, Georg-August Universität*

Members of the Examination Board

Referee: Prof. Dr. Alexander Gail
Sensorimotor Group, German Primate Center

2nd Referee: Prof. Dr. Stefan Treue
Cognitive Neuroscience Laboratory, German Primate Center

Further members of the Examination Board

Prof. Dr. Florentin Wörgötter
Institute of Physics III - Biophysics, Georg-August Universität

Prof. Dr. med. Hansjörg Scherberger
Primate Neurobiology, German Primate Center

Prof. Dr. Andreas Glöckner
Georg-Elias-Müller-Institute of Psychology, Georg-August Universität

Dr. Igor Kagan
Decision and Awareness Group, German Primate Center

Date of oral examination: *24th July 2015*

(cover picture: guardian.co.uk)

I hereby declare that this thesis has been written independently and with no other sources and aids than quoted.

_____

Göttingen, June 2015 Lalitta Suriya-Arunroj

Acknowledgements

First and foremost, I am deeply thankful to Prof. Dr. Alexander Gail for giving me the opportunity to work on an exciting project under excellent conditions in his laboratory. Though I started at rock bottom, his expert advice and constant support helped me in no small degree to develop technical skills and research philosophy in the complex world of neurophysiology. His guidance also involved giving me space to develop my own ideas, which we refined through discussion. I am grateful to Prof. Dr. Stefan Treue and Prof. Dr. Florentin Wörgötter for helpful and supportive suggestions during advisory committee meetings throughout the past four years. It is also my pleasure to thank Dr. Igor Kagan, Prof. Dr. Andreas Glöckner and Prof. Dr. med. Hansjörg Scherberger for kindly agreeing to evaluate my work.

I am indebted to Stephanie Westendorff, Christian Klaes, Shubo Chakrabarti, and Bahareh Taghizadeh-Sarshouri, who gave me considerable support, especially in the initial phase of task programming and data analysis, as well as providing their codes which I could modify in order to analyze my data. And I am also grateful to Pierre Morel, who introduced me to use of a generalized linear mixed model. I would like to thank, in particular, Sina Plümer who was involved in almost everything from providing lab shoes, support on monkey works and setup devices, to cheering me up during my low times. My thanks also go to Leonore Burchardt, Dirk Prüße, and all animal caretakers (Janine, Debbie, Carina and Andrea) who supported on monkey works and taught me innumerable tricks, Klaus Heisig who constructed all parts needed for the setups; Ralf Brockhausen, and Kevin Windolph during my early PhD time, for software, backup system, and active support on troubleshooting; and Beatrix Glaser, for all administrative help. I would like to thank all the lab members for not only providing me with feedback and supportive suggestions on my work but also for making my life in this place so enjoyable, with crazy jokes and puns about monkeys in the preparation room and animal area, great time during the Christmas party and retreat, barbecue lunch in summer, amusing German expressions, and so on. Thanks are also due to the Weekend Club (Enrico, Cliodhna, and again Sina) and the Midnight Club (Suresh) who never abandoned me alone in the lab.

I am very appreciative of support from practical students; Johanna Müller and Laura Geurtz collected data for control experiments in the psychophysics project, Raffaella Bancone

helped me with eye data analysis in the monkey physiology project, and Nicola Popp supported me with the ROC analysis. Their contribution gave me not only the opportunity to develop teaching and supervising skills but also the opportunity to reflect on my thoughts through their interesting questions and discussions. I would like also to thank all 41 subjects who participated in psychophysics experiments and two non-human subjects, Humphrey and Kaspar, who trained me (in a good way) at the same time as I trained them.

I could not say thank you enough to Nicole and Jim, for kind support with my English since I was at the University of Strasbourg, as well as Cliodhna, Pierre, and Enrico who gave me some feedback and proofread parts of my thesis.

I am indebted to long-term support from the Office of The Civil Service Commission (OCSC) who provided me with the prestigious scholarship and the opportunity to come and pursue my studies in Europe since 2003 as well as the Office of educational affairs (OEA) for help and support during my time abroad. Without this opportunity, I would not have ended up writing this PhD thesis.

Being so far from home, I must not forget my gratitude to the various Thai student communities throughout Europe, especially ones in Strasbourg and Göttingen, where friendship was built and homesickness was nursed. The tastes of the food at our cooking parties brought back a sense of home and our multidisciplinary discussions on various topics such as politics, law, culture, society, and sciences, helped to expand my awareness beyond the scope of neurosciences and decision-making. Special thanks are also due to Manussawee, Monthol, Chadapa, Atchara, and a few more Thai friends who took turns to pack my lunches and dinners during the last month before my thesis submission and did not accept any cent from me. I am particularly thankful to Pananya for her presence, mental support, and food supply during my whole PhD time.

Last, but not least, a special thank is dedicated to my family. Since I was a child my parents have always allowed me to decide my own route through life and have supported me in the important decisions I have had to make. My three brothers created for me a chaotic, unpredictable and colorful environment which enabled me to ease my way into working with monkeys without too traumatic transition.

The most courageous **decision** you make each day is to be in a **good mood** :)

La décision la plus courageuse que vous prenez chaque jour est d'être de bonne humeur.

~Voltaire

Table of contents

General introduction

In every hour of every day we are faced with options; which route to take to work, where to go for a meal, which stocks to invest in. However, with limitations of movement, time and budget, our options are constrained. We are obliged to *cut off* (decide: from Latin dēcīdere, to cut off) and reduce the available options until we are left with a single course to pursue.

Since the 1970s, the decision-making process has been described by the centralized view of a *good-based* model in which options are *cut off* within the executive brain areas in the frontal lobe (Norman & Shallice 1980; Padoa-Schioppa 2011), with the resulting choice outcome fed into sensorimotor areas in order to plan the *chosen* movement. More recently, an increasing body of evidence has revealed a decision-making network spanning multiple brain regions, including sensorimotor cortices which have been found to represent potential motor plans before the decision is made (Cisek 2012). Representation of potential actions has been proposed to play a crucial role in decision-making: options are first identified so that values can be assigned to each of them, then the decision-making process will weigh these options and, finally, produce a choice (Rangel et al. 2008). This emerging view regards *action-based* decision-making, at least for decision involving movement, as a competition between movement plans (Cisek 2007). Neural activity in the sensorimotor system has been shown to represent, before the decision is made, available options (Klaes et al. 2011; Thura & Cisek 2014; Yang & Shadlen 2007) and reward expectation (Sugrue et al. 2004; Platt & Glimcher 1999). Furthermore, perturbation of the sensorimotor systems, via micro-stimulation or inactivation of the neural population representing one of the potential action plans, has been shown to bias choice probability (McPeek & Keller 2004; Carello & Krauzlis 2004; Schieber 2000; Oliveira et al. 2010; Cisek 2012). However, to our knowledge, no study has so far shown that the effect of imbalanced planning alone, without reward contrast between options, could also induce selection bias.

This thesis deals with the influence of planning on decision-making and comprises two main experiments, upon which the structure of this thesis will be organized. First, a psychophysical experiment was conducted with human participants, in which

behavioral responses were collected and analyzed to provide first insights into how the decision-making process deals with different task manipulations. The second electrophysiological experiment in monkeys allowed us to record activities at the level of single neurons while rhesus monkeys (*macaca mulatta*) solved choice tasks, yielding data with high temporal and spatial resolution. As the behavioral task involved reach movements, our sensorimotor areas of interest were the dorsal premotor cortex (PMd) and the parietal reach region (PRR), which are known to form the frontoparietal network of reach planning.

This first chapter will start with brief literature reviews of the different perspectives on decision-making mechanisms in the brain, then it will introduce the current state of knowledge of sensorimotor transformations in the frontoparietal network, followed by an overview of known properties of neurons in the areas of interest. It will introduce some terminology and experimental paradigms frequently employed in the field, and will end with the summarized aim of this PhD project. The second chapter will describe and discuss the psychophysical experiment, entitled "I plan therefore I choose: Free-choice bias due to prior action-probability but not action-value". The third chapter will deal with the monkey physiology experiment under the title "Biased action selection due to imbalanced action preparation". Finally, the last chapter will comprise a brief summary and general discussion of the project.

Neuroscience of decision-making

Decision-making is one of the most intensively studied cognitive processes, thanks to rich inputs from various fields such as computer science, economics, neuroscience and psychology, to name a few. Decision neuroscientists – or, using the more fashionable term, neuroeconomists – combine knowledge and theories on decision-making from perspectives of economics, psychology, and neuroscience (Glimcher & Rustichini 2004) and seek to understand the neural encoding of economical values, such as expected value, utility (subjective values), risk, ambiguity, etc. (Dolan & Sharot 2012). The main aim is to locate the brain areas involved in decision-making, to characterize which steps of the process are involved, and to explain and predict different choice behaviors in response to various factors.

Decision-making in the hedonic brain

As choice is in general regarded to be driven by preference (Dolan & Sharot 2012; Samuelson 1938) or utility (Glimcher & Fehr 2014; Neumann & Morgenstern 1944) and a rational decider is considered to maximize benefits and minimize costs, the first candidate brain areas for value weighing were searched for and found in the dopaminergic pleasure centers (see Olds & P. Milner 1954). Neural activities correlated to reward expectation were discovered in dopaminergic neurons (Tobler et al. 2005) as well as neurons receiving dopaminergic projections in the basal ganglia (Kawagoe et al. 1998; Knutson et al. 2001) and the prefrontal cortex (PFC) (Pearson et al. 2014; Padoa-Schioppa 2011) (Figure 1).

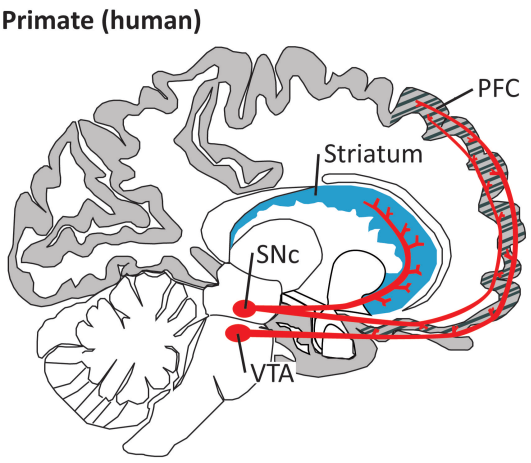


FIGURE 1 | Dopaminergic projections from the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc) to the prefrontal cortex (PFC; hatched area) and striatum (blue) in primate brain (figure from Puig et al. 2014)

Decision-making in the executive brain

Problems with decision-making are a typical characteristic seen in patients with frontal lobe damage (Bechara et al. 1997; Fellows & Farah 2007). Specifically, patients with lesions in the ventromedial prefrontal cortex (vmPFC) show poor risk assessment, impaired stimulus-reward associative learning, inconsistent preference and an unusual way of organizing information during multi-attribute decision-making (Bechara et al. 1997; Fellows & Farah 2005a; Fellows & Farah 2005b; Fellows 2006). The vmPFC is therefore believed to play an important role in combining values from different attributes of available options (e.g. color, taste, benefits, etc.) into a *single common currency* for comparing options during decision-making (Levy & Glimcher 2012; Pearson et al. 2014). This view has been supported by electrophysiological studies in monkeys showing that available juice options and the selected juice were represented by neurons in the vmPFC and the orbitofrontal cortex (OFC; an area within vmPFC) (Padoa-Schioppa 2011; Strait et al. 2014; Padoa-Schioppa & Assad 2006) (Figure 2). In the framework of the good-based model, all values are combined and compared within vmPFC/OFC, then the chosen good guides an action plan (Padoa-Schioppa 2011).

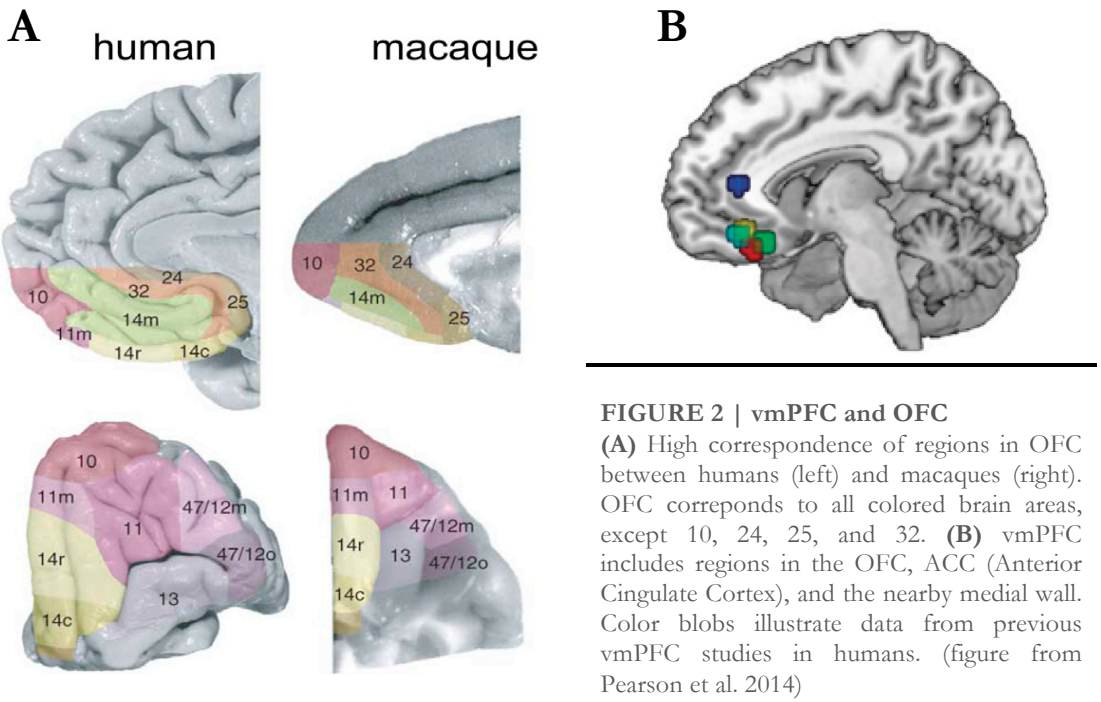


FIGURE 2 | vmPFC and OFC
(A) High correspondence of regions in OFC between humans (left) and macaques (right). OFC correponds to all colored brain areas, except 10, 24, 25, and 32. **(B)** vmPFC includes regions in the OFC, ACC (Anterior Cingulate Cortex), and the nearby medial wall. Color blobs illustrate data from previous vmPFC studies in humans. (figure from Pearson et al. 2014)

Decision-making in the action-planning brain

The traditional view of decision-making as a serial process – decision-making then action plan – has been challenged by recent findings showing that the representation of alternative action plans emerges before the decision is made and, importantly, these representations also reflect decision values, based on which we could predict subsequent choice behaviors (Platt & Glimcher 1999; Sugrue et al. 2004; Klaes et al. 2011; Thura & Cisek 2014; Yang & Shadlen 2007). According to the action-based view, as choice is generally expressed by actions, the sensorimotor network could be one of the brain loci where choice occurs as the outcome of competition between action plans (Cisek 2007). Information that is crucial for weighing the options, such as sensory evidence, values, and others, are provided by other input areas, e.g. the basal ganglia, OFC, the anterior cingulate cortex (ACC) (Cisek 2012) (Figure 3).

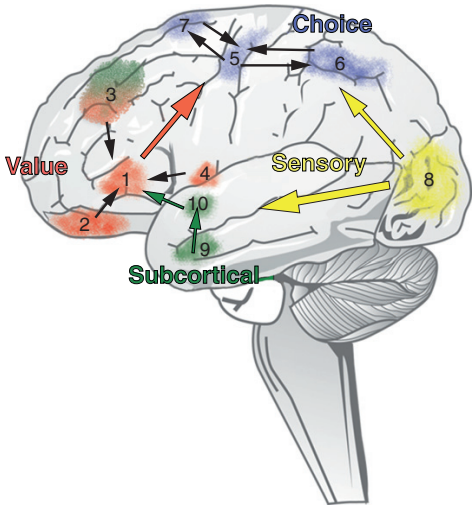


FIGURE 3 | One possible schema depicting the decision network.
Value information from cortical and sub-cortical structures converges into a single common value representation within the **valuation circuitry** before passing on to the **choice-related motor control circuitry**. Sensory signals can arise from various modalities but only visual signals are illustrated here. (1) vmPFC, (2) OFC, (3) dorsolateral prefrontal cortex (DLPFC), (4) Insula, (5) Primary motor cortex (M1), (6) Posterior Parietal Cortex (PPC), (7) Frontal eye fields, (8) Visual cortex, (9) Amygdala, and (10) Striatum. (figure from Levy & Glimcher 2012)

In a broader framework beyond the stage of value comparison, decision-making has been segmented into five computational processes: identifying and representing potential courses of action (representation), assigning the predicted costs and benefits associated with each option (valuation), comparing values and making the choice (action selection), evaluating the desirability of the actual outcome (outcome evaluation), and finally updating the systems according to this feedback (learning) in order to improve future decisions (Rangel et al. 2008; Rangel & Hare 2010). On the basis of this framework, action plans first have to be established so that value can be associated with each of the action plans, and a decision finally made.

Models of decision-making

Another important approach towards understanding of the decision-making mechanism is to implement computational models in order to simulate the decision-making processes and to predict decision outcomes, given various input parameters. One of the most influential concepts in the modeling of a selection mechanism is the rise-to-threshold process, conceptualizing decision-making as a gradual accumulation of evidence until it reaches a threshold, at which moment the choice is determined (Smith & Ratcliff 2004). A model can consist of two (or more) accumulators, representing the available options, which race to the threshold with the winner determining the choice (race model, Figure 4b). Alternatively, the model can consist of a single accumulator representing the contrast between two options (decision variables: DV) that drifts to alternative bounds, with the bound reached first determining the final choice (drift diffusion model, Figure 4a) (Ratcliff 1981; Gold & Shadlen 2007; Summerfield & Tsetsos 2012). In both cases, the number of thresholds matches the number of options to be considered.

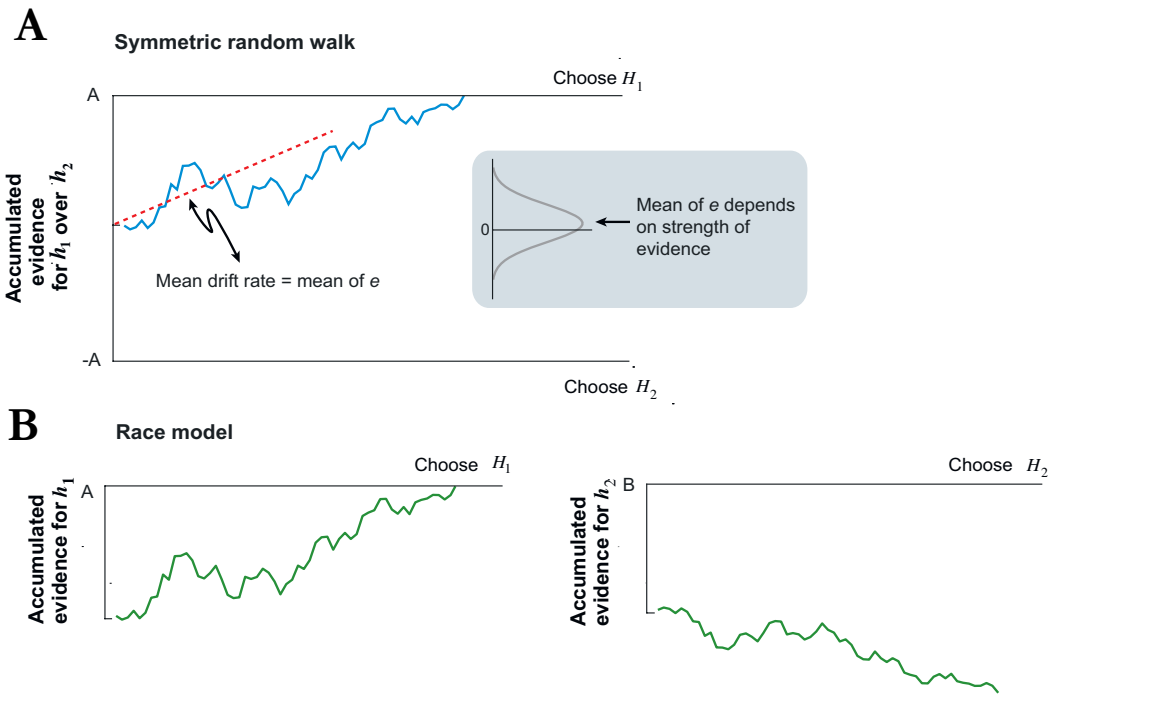


FIGURE 4 | Random walk model or drift-diffusion model (A) and race model (B).
Noisy evidence is accumulated until it reaches a bound. In the bounded diffusion model (A), the difference between the evidence for H1 and H2 is accumulated, whereas in the race model (B), each accumulator accumulates evidence for each alternative. If competing accumulators are inversely correlated, the race model is nearly identical to a symmetric random walk. (Gold & Shadlen 2007; Shadlen & Kiani 2013)

On the basis of this concept, biasing factors can intervene in the decision-making process in a few ways: shortening the migration distance of the DV in favor of one option, either by raising the starting point or lowering the threshold (e.g. Ratcliff 1978; Ratcliff 1985; Bogacz et al. 2006; Mulder et al. 2012); or accelerating migration speed (increasing drift rate) towards one option (e.g. Ratcliff 1981; Diederich & Busemeyer 2006; Roitman & Shadlen 2002). These mechanisms allow the DV to reach one bound faster and, as the bound that is reached first determines the choice, results in a higher choice probability towards that bound. Which process is involved in which situation and which brain areas they occur in are currently topics of ongoing research (Summerfield & Tsetsos 2012; Shadlen & Kiani 2013).

The next question then is what information is taken into account by the DV and how this information drives our decision. The influence of reward on behavior has long been in the spotlight of various research fields. Reward is well known to attract attention, motivate, reinforce, and thus shape our behavior. Robust *motivational* effect of reward can improve performance and speed up actions (Dayan & Balleine 2002; Franchina & Brown 1971; Hassani et al. 2001; Hollerman et al. 1998) and *directing* effect of reward has been shown to guide choice preference towards the most highly reward option, in various species, known as the matching law (Bonem & Crossman 1988; Edwards 1956; Herrnstein 1961; Mir et al. 2011).

The decision-making process is generally perceived as a value comparison process. However, the *directing* effect of reward is not the only factor on which we weigh the options. According to statistical decision theory, there are at least three basic computational elements of decision: evidence – information supporting different options, value – costs and benefits associated to each option, and *prior* – the predicted probability of a particular event to occur, that will be integrated into a decision variable and interpreted by a decision rule (threshold) in order to produce a choice (Gold & Shadlen 2007).

Many efforts have then been spent to investigate each of these decision elements within the established computational framework. Previous studies have shown that *migration distance* of the DV can be shortened by (1) *prior* probability, that induces

anticipation of an event to occur and/or facilitates preparation of own movement to react, or (2) speed-accuracy tradeoff – whether we solve the task with the focus on being fast (shorter distance) or being accurate (longer distance) (Bogacz et al. 2006; Simen et al. 2009; Mulder et al. 2012). The strength of sensory evidence can steer *drift rate* (Roitman & Shadlen 2002; Churchland et al. 2008; Shadlen & Kiani 2013). However, the influence of reward differences on decision-making process is less clear, due to evidences for both baseline shift (Maddox 2002; Bogacz et al. 2006; Mulder et al. 2012) and drift rate change (Diederich & Busemeyer 2006).

In my opinion, the influence of *prior*, which is the predicted probability of an event to occur, is of particular interest for two reasons. First, *prior* can imply expected probability of receiving reward (Gold & Shadlen 2007), in cases where the expected event is known to be rewarding. The probabilities of event outcome and reward overlap most evidently in the context of risky choice or the two-alternative forced choice paradigm that has been extensively employed to investigate valuation during decision-making. Facing a risky choice situation, a decider has to choose between options associated with different reward probabilities and, after having committed to a choice, will get feedback about whether the choice was right or wrong. In this setting, the predicted probability of one target being valid is identical to the expected probability of receiving reward on that target, which, according to the matching law (Herrnstein 1961; Pierce & Epling 1983; Sugrue et al. 2004; Lau & Glimcher 2005), matches the probability of the decision-maker choosing that option. With the overlap of three different probabilities – probability of each option being valid, expected probability of each option delivering reward (reward expectancy), and subjects' choice probability towards each option – it is challenging to infer whether choice probability followed reward expectancy or target predictability.

Second, *prior* probability is known to allow anticipation. This implies the possibility to plan an associated action in advance of expressing choice. As we were interested in the influence of action plan on decision, we needed a new paradigm in which two situations provided comparable expected values, which are the product of reward probability and amount (Neumann & Morgenstern 1944), but only one encouraged

movement planning. The distinction between target probability and reward probability would allow us to compare whether choice probability matched degrees of movement planning in response to target predictability or whether it reflected pure preference due to expected reward contrast without the involvement of movement planning.

Sensorimotor transformation

In order to properly interact with our environment, our brain has complex tasks to accomplish, ranging from *integrating information* from different sensory modalities (e.g. light, sound, taste, smell, etc.), *planning a desired action* by transforming information from sensory coordinates (e.g. image on the retina) into movement parameters (e.g. muscle activation), and *executing the movement*. The central stage, sensorimotor transformation, functions in a distributed way over the frontoparietal network both in a bottom-up manner when the movement is spontaneously guided by visual stimuli (visual-guided movement), and a top-down manner whenever the movement involves higher cognitive processes, e.g. context dependency, transformation rules (further details below), motivational factors, etc. In the latter case, the sensorimotor network acts in concert with higher cognitive area, e.g. the prefrontal cortex (PFC), in order to succeed in a given desired movement (Mountcastle et al. 1975; Andersen & Buneo 2002; Wise & Murray 2000; Miller & Cohen 2001).

Brain mechanisms of visual-guided movement planning

Studies from different neuroscience disciplines have revealed strong reciprocal connections between parietal and frontal areas, more precisely the posterior parietal cortex (PPC) and the premotor cortex (PMC), which function together for many aspects of action planning (Andersen & Cui 2009; Pandya & Kuypers 1969; Kurata 1991). Involved in visually guided movement planning, this circuitry receives inputs from the extrastriate visual cortex (Felleman & Van Essen 1991; Marconi et al. 2001; Wise et al. 1997; Tanné et al. 1995), which encompasses visual areas other than primary visual cortex and is responsible for specialized, higher-order processing of visual signals (Orban 2007; Maunsell & Newsome 1987).

The extrastriate visual cortex comprises two systems; one extends ventrally towards the temporal lobe, characterized as the "perception for recognition" pathway, and another dorsally towards the parietal lobe, the "perception for action" pathway (Ungerleider & Mishkin 1982; Goodale & A. D. Milner 1992b; Fagg & Arbib 1998; O'Reilly 2010). Furthermore, while the ventral stream extracts features of the object to identify what we see via the inferotemporal (IT) cortex, the dorsal stream sends information essential for movement planning to the following parietal areas: the inferior parietal lobule (IPL: 7a and 7b) as well as the medial and lateral intraparietal area (MIP: area 5 and LIP: area 7ip), the medial dorsoparietal area (MDP) and area 7m of the superior parietal lobule (SPL)(Felleman & Van Essen 1991; Wise et al. 1997; Cavada & Goldman-Rakic 1989).

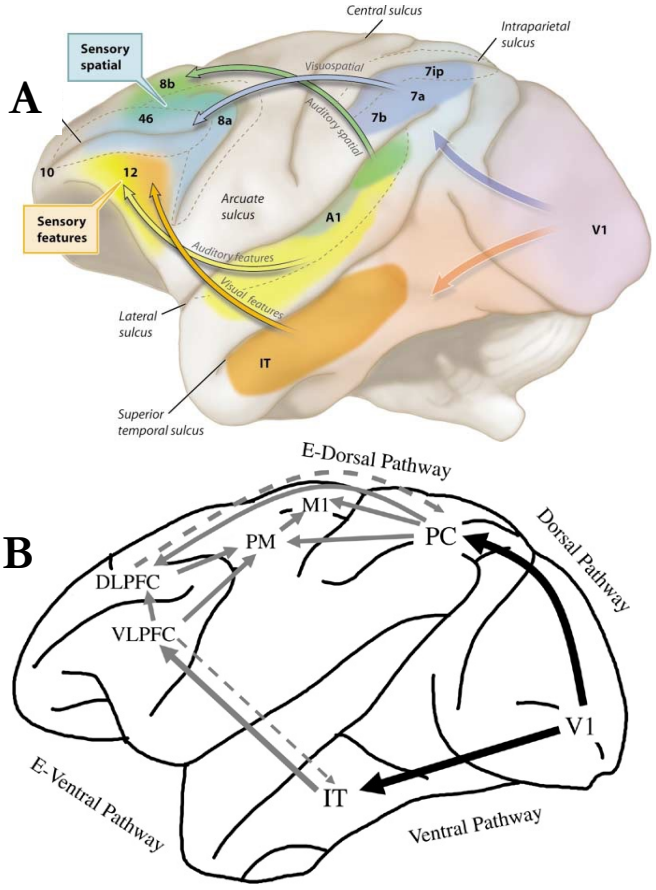


FIGURE 5 | extended dorsal and ventral pathways in the macaque brain **(A)** An artistic view shows the separation between dorsal and ventral visual streams (auditory pathway also shown is beyond the scope of this Introduction). Spatial information from the visual cortex (purple) reaches the PPC (areas 7a/7b/7ip - blue) then reaches further to the dlPFC (areas 46/8a - light blue). Feature information is relayed at the IT (orange) then further transmitted into the vlPFC (12 - orange) (figure from Arnsten 2003) **(B)** A schematic view shows further connections beyond the extended- dorsal and ventral pathways. V1 - primary visual cortex, PC - parieral cortex, PM - premotor cortex, M1 - primary motor cortex (figure from Sakagami et al. 2006)

Further differences between the pathways are seen through to the frontal areas. IT has strong neuronal projections into the ventrolateral prefrontal cortex (vlPFC) whereas the IPL projects further into the dorsolateral prefrontal cortex (dlPFC) and the SPL projects into the dorsomedial prefrontal cortex (dmPFC), known as the extended

dorsal and extended ventral pathways, respectively (Wise et al. 1997; Cavada & Goldman-Rakic 2005; Petrides & Pandya 1999; Petrides 2002; Sakagami et al. 2006; Goldman-Rakic 1996). Finally, both pathways converge in the PMC and the motor cortex. (Figure 5)

The PPC, as previously mentioned, has reciprocal connections with the PMC and both areas share similar information coding (Sakagami et al. 2006; Johnson et al. 1996; Wise et al. 1997). As well as receiving input from the PPC, the PMC also has interconnections with the dlPFC and the vlPFC, which are believed to be the loci where information from the ventral and the dorsal streams are integrated (Sakagami et al. 2006) (Figure 5b).

In cases where a movement is directly signaled by sensory cues without invocation of high-level cognitive process, e.g. touching a spot on a touch screen, this automatic sensorimotor transformation can be processed rapidly in a bottom-up fashion through the parieto-premotor circuit (Desmurget & Sirigu 2009; Pisella et al. 2000; Gail et al. 2009; Sakagami et al. 2006). However, when the movement goal has to be inferred based on arbitrary transformation rules, e.g. touching not at the spot on the screen but next to it, the abstract rules are believed to join the sensorimotor circuit at the premotor counterpart via the PFC (Wise & Murray 2000; Toni et al. 2001; Gail et al. 2009; di Pellegrino & Wise 1991; Crammond & Kalaska 1994; Westendorff et al. 2010).

The combinatorial processing within the frontal and parietal network computes a movement-direction vector, but no details of the muscular contraction command (Mountcastle et al. 1975). The motor command is then transferred to the effectors via the primary motor area (M1), the brain stem and the spinal cord. However, there is no clear-cut functional boundary between PMC and M1, rather a functional gradient. Motor output to the spinal cord is not limited to M1, as premotor-spinal projections also exist (Luppino et al. 2005; Johnson et al. 1996). Finally, the hierarchical distinction between the functions of PMC and M1 is more pronounced during sophisticated behaviors such as spatial transformation (Wise et al. 1997).

Frontoparietal network of reach movements

The spatial input from the dorsal visual stream progressively diverges into parallel subsystems in the PPC (Goodale & A. D. Milner 1992a), each specialized towards the demands of different sensorimotor functions and effectors, such as grasping in the anterior intraparietal area (AIP) and the ventral premotor cortex (PMv), eye movements in the lateral intraparietal area (LIP) and the frontal eye field (FEF), and reaching in the PRR and PMd (Figure 6) (Andersen & Cui 2009; Mountcastle et al. 1975; Batista et al. 1999; Snyder et al. 1997; Rizzolatti & Luppino 2001; Baumann et al. 2009). Each of these subsystems specifies the spatial parameters of different kinds of potential actions (e.g. Snyder, Batista & Andersen 2000a; Mazzoni et al. 1996; Andersen & Buneo 2002; Baumann et al. 2009; Colby et al. 1996) and plays a direct role in guiding their execution during movement (Resulaj et al. 2009; Cisek & Kalaska 2010; Pastor-Bernier et al. 2012).

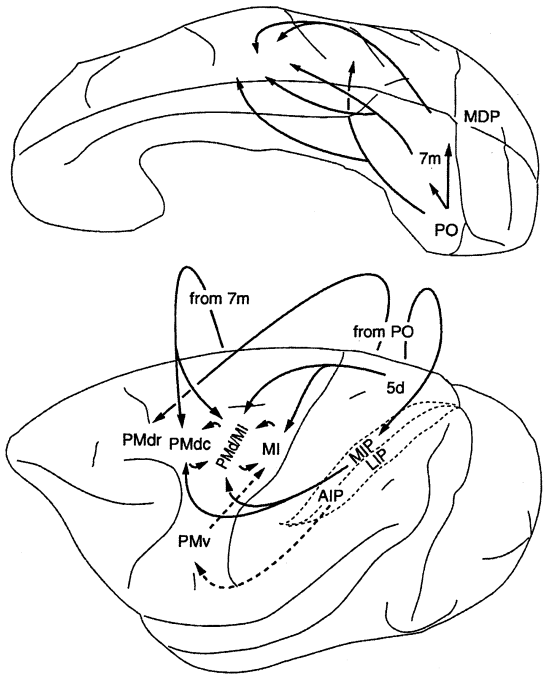


FIGURE 6 | Network of reach areas
Output from PO projects to PRR (MIP and MDP) and the rostral aspect of PMd. PRR and PMd are reciprocally connected while AIP and PMv are also connected. The rostral aspect of PMd also receives a projection from the prefrontal cortex (not illustrated here). The caudal PMd and M1 are also reciprocally linked. (figure from Snyder, Batista & Andersen 2000a)

Parietal reach region (PRR) is a functional definition for the brain region located in the PPC which shows activities related to reach planning. PRR is located in the medial bank of the intraparietal sulcus, hence the name MIP (medial intraparietal area) which overlaps to some degree with PRR. PRR also includes dorsal aspects of areas parieto-

occipital area (PO) (MDP or V6a) (Snyder, Batista & Andersen 2000a). PRR receives inputs from visual extrastriate areas and somatosensory areas and projects to PMd (Johnson et al. 1996; Snyder, Batista & Andersen 2000a).

The dorsal premotor cortex (PMd) is located between the PFC and M1. Most parietal input to PMd originate from the PPC and, in parallel, PMd receives additional inputs from the PFC as well as the PO that have direct connections with the visual areas (Colby et al. 1988; Johnson et al. 1996; Wise et al. 1997). PMd projects motor output to both M1 and the spinal cord (Luppino et al. 2005; Johnson et al. 1996).

Motor goal tuning and rule-based reaching tasks

Directional selectivity, tuning function, and planning activity

Electrophysiological experiments in monkeys have shown that PMd and PRR neurons are selectively active depending on the direction of the upcoming reach. A **directional selective** neuron will exhibit maximal firing activity when the future reach direction is aligned with the neuron's coding direction, so called **preferred direction**, with lower firing rates when reach is aimed towards other directions, and minimal responses in cases of a reach towards the opposite direction. The directional selectivity of a neuron can be visualized by plotting a **tuning function**, depicting a neuron's activity as a function of reach direction (Figure 7).

The directional selectivity in **motor goal tuned neurons** emerges (up to several seconds) before the movement initiation and is referred to as **planning activity** (Gnadt & Andersen 1988; Batista & Andersen 2001; Quian Quiroga et al. 2006). As well as demonstrating selectivity to upcoming movement direction, PMd and PRR neurons also show spatial selectivity to visual stimuli (visually tuned neurons), when stimuli are presented at different locations in space (spatial cues), and can also be selective to the movement direction at the time of movement (peri-movement neurons). A single neuron can display different tuning properties evolving throughout

cue presentation, movement planning, and movement execution e.g. visuomotor neurons (Crammond & Kalaska 1996; Gail & Andersen 2006).

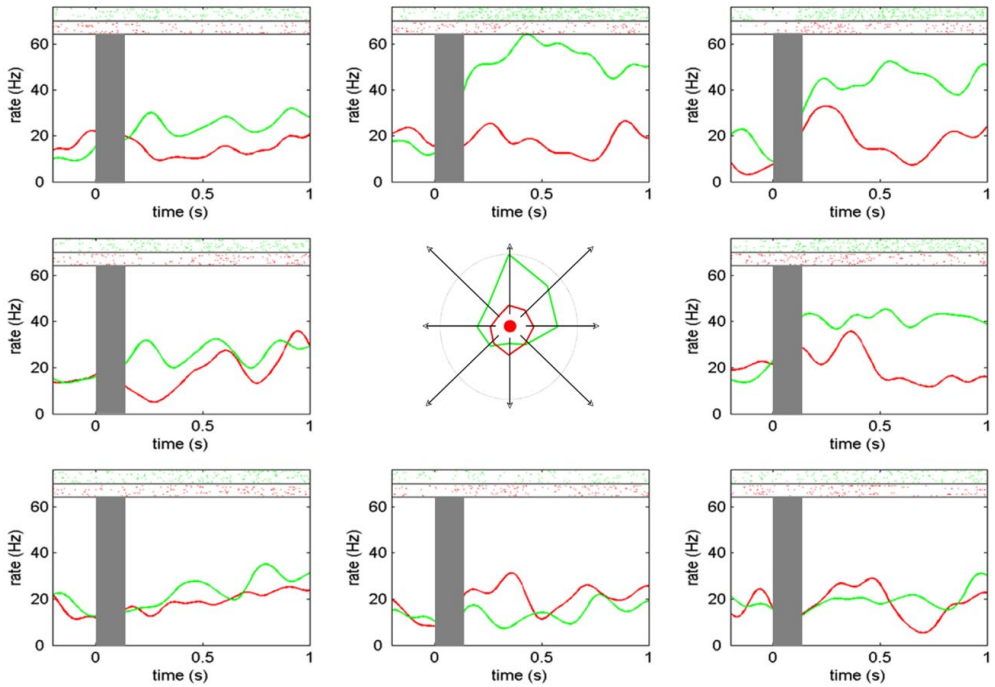


FIGURE 7 | Response of a PRR cell to delayed reaches (green) and delayed saccades (red)
Each plot shows the response to a different direction (shown in the center plot with arrows), with the raster plots displayed on top and mean firing rates on the bottom. Gray rectangles mark the time of target presentation. The center inset show the spatial tunings of the cells in the time period 150 –750 ms. Reach movements start after 1s. This PRR cell is directionally tuned for reaches (green) to the upper direction. (figure from Quian Quiroga et al. 2006)

Memory-guided reach task and transformation rules

In order to experimentally probe the motor planning activities of the neurons, a memory-reach task is typically used. In the first stage of a simple memory-reach task, a spatial cue flashes on the screen. The cue indicates the position of the target to be memorized but does not yet prompt the subject to move. As soon as the subject is aware of the target location, directionally selective neurons start to show sustained activities depending on their preference and the future reach direction, with high activity if the reach direction matches the neurons' preferred direction. However, as the locations of the flashed spatial cue and the upcoming reach overlap in this task design, and given the knowledge that neurons can be tuned for both memorized cue location and future reach location, we cannot distinguish whether the neural tuning

reflects a visual memory trace or motor planning. Such dissociation is made possible by integrating into the task **transformation rules**, which allow the separation of reach goals from the cue position. Depending on the rules, the reach goal can arbitrarily be located anywhere relative to the cue position.

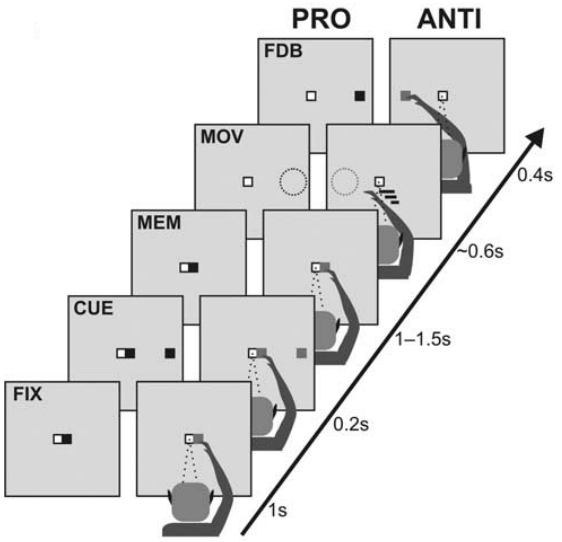


FIGURE 8 | Anti-reach task. The color of a centrally presented stimulus (green/blue) instructs pro- or anti-reaches. The movement goal (dotted circles; not visible to the monkey) is, in case of a pro-reach, at the location (left, right, up, or down) of the previously flashed cue (CUE) or, in case of an anti-reach, opposite of it. (figure from Gail & Andersen 2006)

One classic paradigm is the anti-reach task, in which the color of the spatial cue instructs the subject to either directly reach towards the cue direction itself or to reach in the opposite direction (Figure 8). This task design successfully distinguished pure motor goal tuned neurons – exhibiting spatial selectivity depending on reach direction and not on cue location – from visually tuned neurons, with selectivity depending on cue locations (Crammond & Kalaska 1994; Gail & Andersen 2006; Gail et al. 2009; Westendorff et al. 2010; Klaes et al. 2011).

It is noteworthy that as motor goal representations are known to be modulated by eye positions and eye movements (Buneo et al. 2002; Snyder, Batista & Andersen 2000b), studies on movement intention thus typically require subjects to maintain gaze at one position (ocular fixation) throughout the experimental trial.

Rule-based reach selection tasks

Since decision-making in the sensorimotor system has started to be in the focus, various tasks have been designed in order to reveal various decision parameters in the sensorimotor brain areas, such as reward value and probability (Sugrue et al. 2004; Platt

& Glimcher 1999; Lau & Glimcher 2005; Pastor-Bernier & Cisek 2011), evidence accumulation (Roitman & Shadlen 2002; Coallier et al. 2015), and decision commitment (Thura & Cisek 2014). However, the reward-related signals detected in the brain might not directly encode the reward itself but could turn out to represent something else, such as motivation, salience, prediction of sensory stimuli or sensory outcome (e.g. taste of the juice reward, hand touching money), as well as motor responses, which could be both anticipation of movements to receive reward (e.g. approaching a prey, reaching for a bottle of milk in the fridge) and consummatory responses (e.g. chewing or sucking). The encoding of value and its consequence are then worth being properly categorized for a better understanding of value encoding as well as learning and decision mechanisms in the brain. (O'Doherty 2014)

Following the same reasoning, when a decider has to choose between touching two icons, each of which is associated with a certain value, we cannot rule out whether the neural signals reflect options by encoding the icons associated to different values (stimulus value) or to the reach plans towards each of the icon (action value) because the sensorimotor system encodes both visual stimuli and future movements.

Here again, we could employ transformation rules to distinguish stimulus and movement encoding. Indeed, a recent study from our laboratory (Klaes et al. 2011) showed that the sensorimotor cortices encode both alternative motor goals even when the reach goals had to be inferred based on an abstract rule. Furthermore, when the monkeys showed biased choice behavior, PMd and PRR neurons also showed biased motor goal representation. However, the balanced and biased scenarios were recorded in two separate datasets. It is of high interest therefore to investigate whether one and the same population of neurons could encode alternative motor goals in a graded fashion, reflecting degrees of subsequent choice bias. To achieve this aim, we induced graded degrees of movement planning using a *prior* indicating the probability of two potential reach goals, and, importantly, a rule-based reach goals in order to access pure motor-goal encoding.

Of particular interest here is the findings that neural activity in the sensorimotor system has been shown to be modulated depending on reward expectation (Sugrue et

al. 2004; Platt & Glimcher 1999). In the work of this thesis, experimental manipulations of *prior* probability, in contrast to planning-independent expected reward of options at the decision moment, together with the rule-based motor goals, instead of motor goals physically marked by visual cue, allowed us to probe whether reward-independent choice bias would be reflected by degrees of movement plans in sensorimotor system.

Aim of the project

Since the effect of action planning on decision-making has never been clearly distinguished from the effect of reward expectation, the human psychophysical study was initiated with the aim of tackling this problem. We designed experimental paradigms in which we could directly test whether preliminary action plans, in contrast to preliminary reward expectations, can influence choice between equal-valued options.

As stated at the end of the section *Neuroscience of decision-making*, effects of *prior* probability and expected value could not be dissociated in the classic risky choice paradigm due to an overlap of reward probability, target probability, and choice probability. In addition, due to the fact that subjects typically have to immediately make a choice once all potential options are offered, this paradigm provides no way of preventing subjects from immediately planning towards the higher-value option (Figure 9: Risky choice). In order to manipulate action planning in a graded fashion, we designed **pre-cued instruction tasks**, in which a pre-cue indicated two potential targets and their assigned probabilities of occurrence as target in a trial. A subsequent rule cue validated these assigned probabilities. Subjects had to follow the imperative rule cue; if they did not, the trial was considered unsuccessful (Figure 9: Pre-cued instruction). The *prior* knowledge provided by the pre-cue thus helps subjects to predict the valid target and encourages them to plan their later movements.

While increasing target probability, the pre-cue also increases reward expectation for that target. We then compared subjects' response to another task in which we matched the expected value (reward probability × amount) while discouraging subjects from

planning by not providing information about probability of occurrence. In this second task, the pre-cue indicated the amount of reward that subjects would receive in case each target would be instructed while we kept the probability of each target to be instructed always at 50%. This probability-neutral pre-cue carried no information that would be useful to predict the valid target and to start any movement plan. This results in two tasks: PROB task – in which the probability was varied and the reward amount was fixed – and AMNT task – in which the probability was fixed and the reward amount was varied. On top of this, we included **choice trials**, which were randomly interleaved and which always provided equal reward, in order to probe the choice bias induced by *prior* probability and potential payoff.

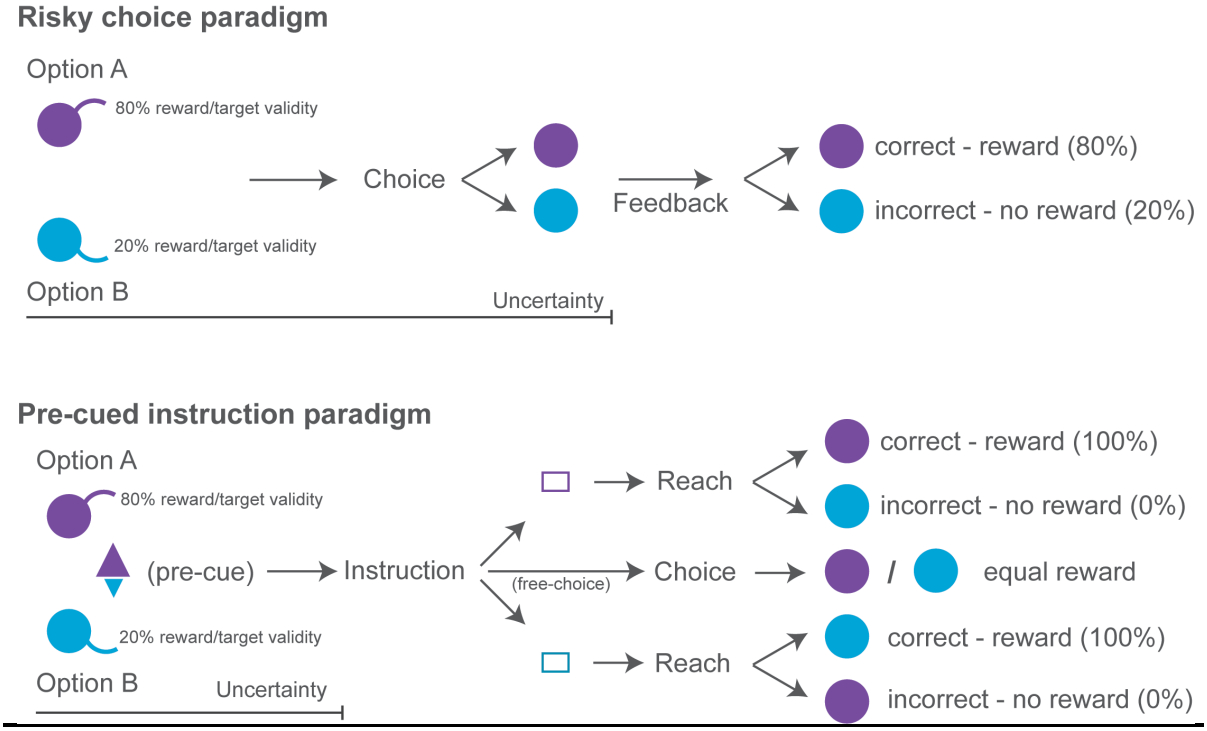


FIGURE 9 | risky-choice and pre-cued instruction tasks. Risky choice paradigm: a trial starts with presentation of two options with different probabilities of occurrence, which matches with the probability of reward in this paradigm. Subjects can make up their mind as soon as the options are revealed. After choosing, a feedback signal depending on the *prior* probability, indicates to the subject whether their choice was right or wrong (i.e. choice rewarded or not). **Pre-cued instruction paradigm:** a trial starts with two alternatives with different assigned probabilities of occurrence. Subjects cannot choose as soon as they are aware of options but have to wait for a subsequent instruction. The rule-cue indicates which cued target is valid for this trial, on the basis of the probability previously announced by the pre-cue. The choice (white) rule cue allows subjects to choose between both options with equal reward. In this case, subjects are under uncertainty only until the instruction appears, afterwards subjects are completely risk-free.

Next, since the sensorimotor decision-making mechanism has been proposed as a competition between action plans, it is then essential to investigate whether motor-

planning activities in our paradigm show a graded pattern reflecting the graded level of *prior* manipulation as well as graded subsequent choice bias.

In the pro-anti reach task, one spatial cue generates two movement options with unequal difficulties: pro-reach – reaching towards the spatial cue – is easier than anti-reach – reaching to the opposite direction to the spatial cue. Klaes et al. (2011) unexpectedly detected biased choice behavior, probably due to the anti-reach which was more difficult for the monkeys and required more intensive training than the pro-reach (Figure 10: Pro-Anti task). The monkeys were subsequently trained to perform balanced choice behavior for the balanced dataset.

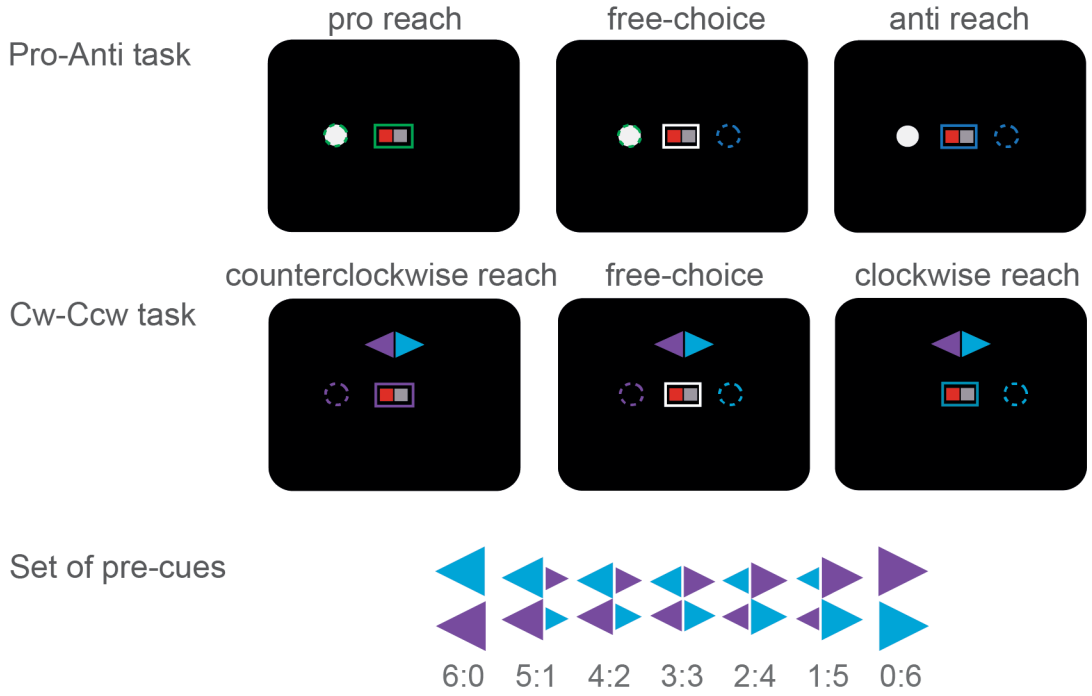


FIGURE 10| pro-anti reaches and Clockwise (*cw*) – counterclockwise (*ccw*) reaches. Pro-anti reaches (top row): the green rule-cue (box framing the centered gaze and hand fixation points) indicates the pro-rule, instructing a reach towards the spatial cue (white circle); the blue rule-cue indicates the anti-rule, instructing a reach in the opposite direction of the spatial cue; and the white rule-cue announces a choice trials in which subject can freely choose between pro or anti reaches. **Cw-ccw reaches (middle row):** pre-cue triangles indicate two potential motor goals, located 90° degree *cw* and *ccw* to the pre-cue. Colored (magenta or cyan) rule-cue instructs a reach to the *cw* or *ccw* reach goal pointed by the pre-cue triangle of the matched color (colors can be swapped; no fixed association of colors and rules). The white rule-cue allows subject to freely choose between *cw* and *ccw* rules. Dashed circles indicate correct reach targets for each cue configuration; these were not shown on the screen during the experiment. **Pre-cue set (bottom):** the pre-cue triangles can be of different relative sizes, indicating different probabilities of each possible reach goal which is subsequently instructed by the rule-cue.

In our case, as we were directly dealing with bias manipulation, we designed a behavioral task in which both reach options were rule-based to avoid any initial bias due to unequally difficult rules as options. The clockwise and counterclockwise rules

relative to the cue position were introduced with symbolic cues (double triangles), guiding reach directions ±90° relative to the cue position (Figure 10: Cw-ccw task). Additionally, we applied different degrees of bias manipulation by assigning different relative probabilities of each rule, mediated to subjects by the size of the triangles.

It is worth mentioning that studies of responses that depend on context or arbitrary rules are of interest not only because we could obtain neural signals representing motor goals during movement planning but, in a broader context, they could also lead to a better understanding of other challenging rule-based behaviors such as symbol-guided responses and language (Wise & Murray 2000).

We used rule-based reach selection tasks with both human and monkey subjects for the compatibility of behavioral results.

To summarize, in this PhD work, we aimed to study the influence of planning on decision-making between rule-based reach goals. The ultimate aim of the work is to gain further understanding of sensorimotor roles on both sensorimotor transformation and decision-making. It is hoped that this work will provide another piece of the puzzle contributing to a better understanding of neural decision-making mechanism.

Original manuscripts

This chapter contains the following research manuscripts:

I plan therefore I choose: Free-choice bias due to prior action-probability but not action-value

Lalitta Suriya-Arunroj and Alexander Gail (*Front Behav Neurosci*. 2015 Nov 25;9:315)

Author contribution: LS and AG designed the experiment. LS implemented the experiment. LS collected and analyzed the data. LS wrote the manuscript; AG edited the manuscript. Authors discussed the results and commented on the manuscript at all stages.

Biased action selection due to imbalanced action preparation.

Lalitta Suriya-Arunroj and Alexander Gail (in preparation)

Author contribution: LS and AG designed the experiment. LS implemented the experiment and trained the animals. LS collected and analyzed the data. LS wrote the manuscript; AG edited the manuscript. Authors discussed the results and commented on the manuscript at all stages.

I plan therefore I choose: Free-choice bias due to prior action-probability but not action-value

In this chapter, we investigated the influence of a preliminary action plan on decision-making. We induced movement planning using the predictability of target locations provided by probabilistic pre-cues (PROB task: varied probability, fixed reward amount) and compared subjects' responses with the case when pre-cues indicated the potential payoff (AMNT task: varied reward amount, fixed probability). This design then provided seven bias levels with compatible reward expectancies (Probability × Amount) across tasks, with preliminary action planning encouraged only in the PROB task. We tested whether the preliminary plan towards one option, induced by predictability in PROB task, would differentially bias subsequent choice between equally valued options, compared to pure preferability, *prior* reward expectation without preliminary plan in AMNT task.

The results of this study showed that the target predictability induced strong and graded choice bias, depending on the bias degree induced. In addition to choice bias, subjects also showed faster reaction times in a graded fashion when choosing the biased option. However, the preferability induced only a slight choice bias on average. Indeed, the majority of subjects did not show choice bias at all. Importantly, the choice bias, in a subset of subjects, resulting from our payoff manipulation was not accompanied by a reaction time benefit.

We interpreted these results in light of the drift-diffusion model (DDM). Among other results, the fact that the biasing effect of the probabilistic pre-cue persisted in choice trials, where there was no evidence provided and both options offered equal reward, supported the baseline shift process. However, such baseline shift process fails to describe the limited biasing effect of *prior* reward expectation, when movement planning was discouraged, which then favored the drift rate change explanation.

This study provided evidence for the influence of action planning on decision-making and thereby supports the idea that the underlying mechanisms overlap.

I plan therefore I choose: Free-choice bias due to prior action-probability but not action-value

Lalitta Suriya-Arunroj[1] and Alexander Gail[1,2,3]

[1] Sensorimotor Group, German Primate Center, Göttingen, Germany
[2] Bernstein Center for Computational Neuroscience, Göttingen, Germany
[3] Georg August University, Göttingen, Germany

Abstract

According to an emerging view, decision-making and motor planning are tightly entangled at the level of neural processing. Choice is influenced not only by the values associated with different options, but also biased by other factors. Here we test the hypothesis that preliminary action planning can induce choice biases gradually and independently of objective value when planning overlaps with one of the potential action alternatives.

Subjects performed center-out reaches obeying either a clockwise or counterclockwise cue-response rule in two tasks. In the probabilistic task, a pre-cue indicated the probability of each of the two potential rules to become valid. When the subsequent rule-cue unambiguously indicated which of the pre-cued rules was actually valid (instructed trials), subjects responded faster to rules pre-cued with higher probability. When subjects were allowed to choose freely between two equally rewarded rules (choice trials) they chose the originally more likely rule more often and faster, despite the lack of an objective advantage in selecting this target.

In the amount task, the pre-cue indicated the amount of potential reward associated with each rule. Subjects responded faster to rules pre-cued with higher reward amount in instructed trials of the amount task, equivalent to the more likely rule in the probabilistic task. Yet, in contrast, subjects showed hardly any choice bias and no increase in response speed in favor of the original high-reward target in the choice trials of the amount task.

We conclude that free-choice behavior is robustly biased when predictability encourages the planning of one of the potential responses, while *prior* reward expectations without

action planning do not induce such strong bias. Our results provide behavioral evidence for distinct contributions of expected value and action planning in decision-making and a tight interdependence of motor planning and action selection, supporting the idea that the underlying neural mechanisms overlap.

We conclude that free-choice behavior is particularly biased when pre-cues allow the planning of one response over another alternative. Our results provide behavioral evidence for distinct contributions of expected value and action planning in decision-making process. The results also provided evidence for the tight interdependence of decision behavior and motor planning, thereby supporting the idea that the underlying neural mechanisms overlap.

Keyword: reach movement, decision-making, action selection, motor planning, bias, *prior* probability, expected value

Introduction

During economic choice, we weigh potential options. In general, the most beneficial or least costly option directs our decision. When deciding between equally valued options, an economic decision results in choice of either option with equal probability, known as matching behavior (Herrnstein, 1961; Lau and Glimcher, 2005; Pierce and Epling, 1983; Sugrue et al., 2004). However, factors other than value-based preference can also influence our choices. In statistical decision theory, value is one building block, which together with *prior* probability and evidence, forms the three main computational elements of decision-making processes (Gold and Shadlen, 2007). Furthermore, choices normally require action selection. Independent of the expected reward, differences in physical effort between response options can be an obvious motor-related decision factor (Burk et al., 2014; Kurniawan et al., 2011; Rigoux and Guigon, 2012). But more than that, covert motor-related factors, like *prior* action planning, could also be expected to influence later overt responses, and hence decision outcomes (Cisek, 2007; Gallivan et al., 2015). Here we compare the effect of *prior* motor planning with *prior* value-based preference on choice probabilities and choice reaction times in situations with balanced reward (neutral free choice) at the time of the actual decision.

Imagine you are late for dinner. Your two alternative routes home are on average equally fast but depend on current traffic. In scenario 1, the slow truck in front of you indicates a turn to the right. This will likely make you plan a left turn at the next intersection, because with the truck going right, the left route now is faster, hence has a higher value. If the truck then unexpectedly pulls into a rest area you are free to choose between equal-valued options (neutral choice) once you arrive at the intersection (time of commitment). Will you stick with your *prior* plan, even though both alternatives would now be equally attractive? In scenario 2, the truck's indicators are so dirty that you cannot identify an indicated direction, but you know that the left route would allow you to overtake the truck more easily than the right route. This means the left route has a higher (truck-conditional) value, as the truck would slow you down less. If the truck then unexpectedly pulls into a rest area, both options are equal-valued again. Will you decide based on your *prior* value-based preference? Both scenarios have in common that the value is initially higher for the left route, in scenario 1 due to an imbalance in *prior* probability, in scenario 2 due to an imbalance in *prior* value of a decision-critical event that is yet to happen (the truck's actual turn = evidence). Also, in both scenarios, you will be provided with evidence immediately before your commitment, and most times (truck turns left or right) this will directly instruct your choice (go the other way). Finally, in both scenarios, the expected value becomes neutralized when the truck pulls into the rest area, rendering both of your alternatives equally valued. The scenarios differ in their intuitive effect on action planning. In scenario 1 your initial preference based on the probability of the later event encourages you to preliminarily plan your action. Instead, in scenario 2, even though you also have an initial value-based preference, the situation discourages specific action planning until you know the truck's actual turn, since each turn is equally likely. Will an *a priori* value-based preference in which action planning is discouraged affect your later neutral choice differently to an *a priori* preference that is associated with an action plan? This is the question we address in this study.

To investigate the effect of motor planning on choice, we need to dissociate planning from value-based preference. It is known that movement planning is encouraged by motor-goal predictability. For example, tasks with probabilistic pre-cues, as used in early attention studies (Posner et al., 1980), have been adopted in sensorimotor studies to test

the ability to plan movements in Parkinson's disease patients (Jahanshahi et al., 1992; Leis et al., 2005; Praamstra et al., 1996; Stelmach et al., 1986). In such tasks, the pre-cue correctly indicates the location of an upcoming target with a typical probability of 80% (cue validity) while in the remaining 20% of trials the non-cued target will be instructed. A subsequent imperative cue instructs the subject when and towards which target location to act. The *prior* information contained in the pre-cue encourages subjects to plan the movement towards the pre-cued target, confirmed by effects of cue-validity on reaction times and, in some cases, movement times (Leis et al., 2005). Following the rationale of cue validity, a neutral pre-cue indicating equal probability of occurrence for each target should not evoke imbalanced preliminary planning. We used this rationale for testing the biasing effect of motor planning on reward-balanced choices. For this we manipulated the degree of motor planning by different degrees of *motor-goal predictability*.

However, probabilistic pre-cues can confound predictability with *preferability of a motor goal*. Among multiple targets, if the validity of one target becomes larger, the probability of receiving reward at that target also increases, and hence the expected value, defined as the product of probability and amount of reward (Gold and Shadlen, 2007; Levy and Glimcher, 2012; Neumann and Morgenstern, 1944), will also increase for the higher-validity target. In order to disentangle the effect of planning from the effect of reward expectation, we designed two tasks with matched expected rewards but only one of which encouraged preliminary action planning.

In the context of probabilistic choice behavior, a *prior* should have more impact when evidence is weaker (Körding and Wolpert, 2004; Vilares et al., 2012). Therefore, the effect of the *prior* is typically investigated in situations when decisions are based on ambiguous evidence. For example, using random dot motion stimuli, *priors* were shown to influence the interpretation of ambiguous visual sensory evidence (e.g. Mulder et al., 2012) and affect the latency of action initiation (e.g. Carpenter and Williams, 1995). Yet, if the perceptual interpretation of the evidence is one-to-one associated with a behavioral response then a *prior* is also likely to invoke preliminary action planning. It can therefore be difficult to disentangle whether the effect of the *prior* on choice is mediated via an effect on sensory processing or on action planning. Here we test the effect of *prior*

probabilities without any perceptual uncertainty, emphasizing the effect of action planning on choice.

Even though our experiment did not utilize ambiguous cumulative sensory evidence towards a perceptual decision, but rather immediate unambiguous evidence (instructed trials) or rule-neutral evidence (choice trials), we find it helpful to conceptualize our study in the context of drift-diffusion models (DDM). On the basis of the DDM, conceptualizing decision processes as a gradual accumulation of evidence towards one of two alternative boundaries (e.g. Ratcliff, 1978; Ratcliff et al., 1999), bias can be explained by different computational mechanisms: (1) a shorter migration distance, either due to a baseline shift (Ratcliff, 1985) or a bound shift (Ratcliff, 1978), or (2) a change in drift rate (Ratcliff, 1981). These mechanisms allow accumulated evidence to reach one bound with smaller reaction times (RT) and higher choice probabilities (CP). Which mechanism is responsible for RT reduction and CP increase in which behavioral context is a topic of ongoing research (Summerfield and Tsetsos, 2012). In the context of perceptual decision-making, previous studies showed that *prior* probability adapts migration distance (Bogacz et al., 2006; Mulder et al., 2012; Simen et al., 2009) while strength of evidence steers drift rate (Coallier and Kalaska, 2014; Coallier et al., 2015; Hanks et al., 2015; Roitman and Shadlen, 2002). However, the effect of expected value as been accounted for by different explanations: baseline shift (Bogacz et al., 2006; Maddox, 2002; Mulder et al., 2012) or drift rate change (Diederich and Busemeyer, 2006). The diversity of explanations for the effect of expected value could be due to insufficient systematic dissociation of other factors from the effects of movement planning as previous studies usually involved probabilistic choice tasks, which we avoid here. Additionally, we test the specific hypothesis that preference leads to the same biasing effects as planning, except for a downscaling factor that reduces effect strength (Bogacz, 2007; Maddox and Bohil, 1998; Mulder et al., 2012). Such downscaling should be particularly obvious when testing multiple levels of bias, in which case it should be possible to estimate the value-based bias from the probabilistic bias by applying a fixed gain factor. Instead of using only a single level of bias manipulation, we therefore probed for a graded effect of graded *prior* probability as opposed to graded value.

To our knowledge, no study has directly tested the proportional effect of action planning on choices in which there is no difference in expected value between options. According to emerging evidence, neural mechanisms overlap between decision-making and movement planning (Cisek, 2007; Coallier et al., 2015; Gallivan et al., 2015; Hanks et al., 2015; Klaes et al., 2011; Lindner et al., 2010; Scherberger and Andersen, 2007). We therefore hypothesize that previously planned actions should bias later neutral choices in favor of these actions, independently of reward expectation.

Methods

Participants

43 subjects (30 females, age (mean ± SD): 27.45 ± 4.89) participated in the study as paid volunteers. Among the 43 subjects, 31 participated in both AMNT and PROB tasks (on two separate days; 19 did PROB task first and 12 did AMNT task first), 10 in only PROB task, and 2 in only AMNT task. All subjects were healthy, right-handed, and had self-reported normal or corrected-to-normal vision. Detailed written instructions were given to the subjects before the experiment. Prior to each recording session, subjects were familiarized with the set-up and practiced the task. All subjects gave written informed consent for participation. Experiments were in accordance with institutional guidelines for experiments with humans and adhered to the principles of the Declaration of Helsinki. The experimental protocol was approved by the ethics committee of the Georg-Elias-Mueller-Institute for Psychology, University of Goettingen.

We included all 43 subjects in the analyses. When comparing among conditions within each task with *post-hoc* tests, we included the 33 subjects who participated in the AMNT task and 41 subjects who participated in the PROB task. When comparing between tasks with *post-hoc* tests, we included the 31 subjects who participated in both tasks.

Rule-selection task with sequential cueing

The idea of the study was to investigate the influence of rule predictability and pure preference on choice behavior. The task implements the idea of the traffic example in the

introduction. We designed a center-out reach task with sequential cueing. A first pre-cue raised expectations on either the probability or the value of a later rule instruction (Figure 1). The rule instruction (rule-cue) provided final information on the actual rule and hence the action(s) to be rewarded. Two potential reach goals had to be inferred from the single pre-cued location based on clockwise (*cw*) and counterclockwise (*ccw*) transformation rules. We implemented two variants of this rule selection task, one in which the *prior* probability of either rule to be instructed was announced in advance by the pre-cue (PROB task), another in which the reward of either rule, in case it would be instructed, was announced (AMNT task). The choice experiment was risk-free, since at the time of the required behavioral response (decision), there was no uncertainty about the outcome; subjects were either instructed about the correct response immediately before the decision (instructed trials = truck turns left or right in the example from the *Introduction*), or they were free to choose among both options with 100% reward probability and equal reward amount for each option (choice trials = truck pulls into rest area). Note that the pre-cue was only informative about the reward structure of the instructed trials, while free-choice options were safe and equal-valued. As a consequence, subjects could achieve 100% reward probability with proper performance in all task conditions. The reward delivered in each trial was accumulated and translated into the compensation that participants received at the end of the session (see below).

Due to the temporal separation between pre-cue and rule-cue, each choice was preceded by a brief planning period (approaching the intersection in the truck example). During the planning period subjects were uncertain about the type of trial (instructed or choice), and uncertain about what the instruction will be. For optimal performance, subjects in response to the final instruction had to either follow their initial expectation or countermand it (instructed trials), or freely choose (choice trials). For example, a pre-cue in the AMNT task might raise the initial expectation that a left-side reach would be preferable, but the rule cue at the end of the planning period could still indicate a left-side or right-side single correct option (instructed trials) or two correct options (choice trial). The pre-cue could not be ignored, though, since the rule-cue only was meaningful in relation to the pre-cue.

The idea of this task design was that with a majority of instructed trials the pre-cue would induce a trial-by-trial behavioral bias, either based on predictability or preferability. In a first step, we had to confirm that both manipulations were effective in by analyzing subjects' responses in the instructed trials. In instructed trials the manipulations had an actual effect on the reward outcome, hence an effect on behavior had to be expected. As the main research question, we test if the variable *predictability* and *preferability* had the same effects on free-choice behavior. We probed this with randomly interspersed choice trials. Since the choice trials were value-balanced and risk-free, any choice probabilities and other behavioral biases should reflect the subject's *a priori* bias induced by the predictability or preferability resulting from the pre-cue, since no further immediate evidence supporting either rule is provided during the remainder of a choice trial.

Importantly, in the AMNT task, we aimed to induce a preference in the subjects without encouraging planning of the according action since either rule was equally likely to be instructed. On the other hand, the PROB task instead would encourage preliminary planning of the action that was associated with the most likely correct option since, in the likely case of an instructed trial, the instruction would match the rule expectation. We tested the assumption that predictability, as opposed to preferability, would lead to target-specific motor planning by additionally analyzing movement execution parameters in instructed trials. If such test yielded evidence for motor planning in predictable but not in preferable trials then any effects of predictability on choice behavior could be explained by an effect of motor planning.

We then compared the effect on risk-free choice behavior of either a preliminary value-based target preference (henceforth referred to as "preferability", AMNT experiment) or a preliminary motor plan towards a probable target (henceforth referred to as "predictability", PROB experiment). We will first describe in detail the elements of the task that are common to both experiments and then the differences in the reward schedule of the two experiments.

Graded bias manipulation in the rule-selection tasks

Subjects in both experiments had to choose between either a clockwise (*cw*) or counter-clockwise (*ccw*) spatial mapping rule to infer the reach target relative to the position of the pre-cue (rule-selection task). Subjects were requested to perform reaches from the center to one of the four cardinal (0°, 90°, 180°, and 270°) targets in the periphery (center-out reach) on a touch screen (eccentricity of 9 cm with ~40 cm of screen distance, depending on subjects' reaching range), while maintaining gaze at the screen center (eye fixation) throughout the trial. The target locations had to be inferred from an incongruent cue location and were not marked by visual stimuli (rule-based movement).

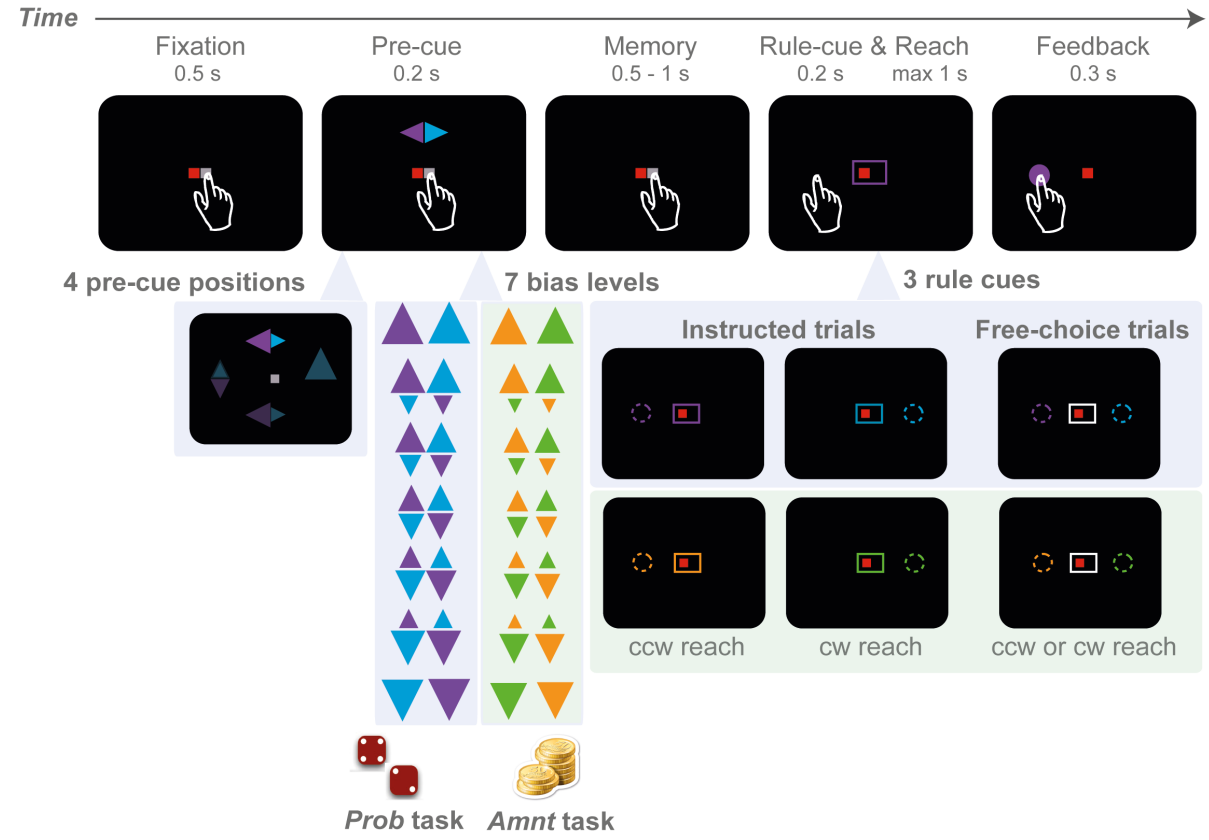


FIGURE 1 | Rule-selection tasks with probabilistic (PROB) and amount (AMNT) bias manipulations. A trial starts with eye-hand fixation. Next, a pre-cue appears at one of the four cardinal locations (top, down, left, right), indicating two potential goals and the probability of each goal in the PROB task (big triangle - high probability), or the reward amount associated with each goal in the AMNT task (big triangle - more reward). In the PROB task, the reward amount is kept constant in all conditions, whereas in the AMNT task the probability is kept equal in all conditions. After a memory period, the rule-cue appears and either indicates one valid target or leaves the options open so that subjects can freely choose. Both options in choice trials always provide equal value.

Each trial started with a fixation period. Small red and white squares were presented at the center of the screen as eye and hand fixation points, respectively (Figure 1). Subjects initiated the trial by directing gaze to the eye fixation spot and, at the same time, touching

the hand fixation spot (tolerance window: 3 cm radius). After a random period of 500-1000 ms fixation, a spatial pre-cue flashed briefly (200 ms) at one of the four cardinal target positions. The pre-cue consisted of two differently colored triangles, one pointing to the *cw*, the other to the *ccw* direction. The two triangles indicated the only two possible reach targets in a given trial, one at 90° *cw* and one at 90° *ccw* from the pre-cue, at identical eccentricity.

There was no fixed association between rules (*cw* or *ccw*) and colors. Subjects had to remember the pre-cue colors and match these with the subsequent color information of the rule-cue to complete the trial correctly. At the end of the following random-length memory period (500-1000 ms), the rule-cue appeared. The rule-cue consisted of a box framing the fixation points and was either colored to match one of the pre-cue triangles (instructed trials), or color-neutral (white; choice trials). In the instructed trials the colored rule-cue narrowed down the two potential targets to only one single correct target (*cw* or *ccw*). The reward probability for the instructed target was always 100%, for the alternative target 0%. In the choice trials, both potential targets indicated by the pre-cue were rendered valid with 100% reward probability by the color-neutral rule-cue. In choice trials the reward amount was always fixed and independent of the size of the pre-cue or any previous choice responses (reward-all schedule). Simultaneously with the onset of the rule-cue, the hand fixation point disappeared ("go" signal) and subjects had to reach toward the instructed or chosen target within a maximum of 1000 ms. In each block of six trials, two trials were randomly set to be choice trials, and the remaining four were instructed trials. Each unsuccessful trial was reinserted into the trial sequence.

In case of successful acquisition of a rewarded target, subjects received positive feedback in the form of a circular patch at the target position with an encouraging high-pitched tone (coin sound). If subjects failed to reach the correct target, the trial was aborted and a demotivating low-pitched tone was played. Failure trials included aborted trials due to ocular fixation breaks, incorrect reaches to locations on the screen outside the tolerance window (3 cm radius) around the valid target(s), and reaches later than the maximal response time. Subjects were explicitly requested to respond as accurately and as rapidly as possible.

Subjects had to perform the rule-selection task in two variants, which differed only in the instructed trials, not the choice trials. The idea was to induce a graded level of predictability (PROB task) or preferability (AMNT task) between the two possible rules. For this, the relative size of the two triangles in the pre-cue was varied in seven steps corresponding to seven instructed expectation levels. Note, for convenience we refer to these levels jointly as 'bias conditions' even though the behavioral bias that will be potentially induced by this task parameter is our tested variable. In the zero-bias trials the triangles had equal size. In the 100% bias level conditions either only the *cw* or the *ccw* triangle was visible and larger than in the zero-bias condition. In the intermediate bias conditions the two triangles had intermediate sizes. The seven instructed-bias conditions used the following combinations of pre-cue triangles (base lengths of the triangle): {3.5:0.0, 3.0:0.7, 2.5:1.4, 2.0:2.0, 1.4:2.5, 0.7:3.0, 0.0:3.5}. The bias level was kept constant within each block of six trials as will be described below.

Probabilistic rule-selection task (PROB)

The idea of the PROB task was to induce a graded level of rule predictability across trials, without a difference in the final value of the two motor-goal options. In the PROB task, we assigned the colors magenta and cyan to the two pre-cue triangles. The size of the pre-cue triangles indicated the likelihood with which the *cw* and *ccw* rules would be instructed by the rule-cue later in the trial. The reward for the targets associated with either rule was identical. Seven bias levels corresponded to likelihoods of {6:0, 3:1, 2:1, 1:1, 1:2, 1:3, 0:6} for an instruction of the *ccw* or *cw* rule, respectively. In the case of a bias level 2:1 towards *ccw*, subjects were offered 4 *ccw* trials and 2 *cw* trials. Subjects received 3 reward units for each correctly performed instructed trial and 1.5 reward units for either choice in the choice trials. Thus, at the 100% bias levels (6:0 or 0:6), the low probability rule had 0% chance of getting 3 units in instructed trials and 100% of getting 1.5 units in choice trials, resulting in an expected value of $0\times3\times2/3 + 1\times1.5\times1/3 = 0.5$ reward units. The high probability rule had a 100% chance of getting 3 units in instructed trials and 100% chance of getting 1.5 units in choice trials, resulting in an expected value of $1\times3\times2/3 + 1\times1.5\times1/3 = 2.5$ reward units. The ratios of initially expected values (EV) associated with the two rules at the seven bias levels were then {2.5:0.5, 2:1, 1.83:1.17, 1.5:1.5, 1.17:1.83,

1:2, 0.5:2.5}. Note that these initial EVs were only valid for the time between the pre-cue and the rule-cue. After the rule-cue, the final value for instructed trials was 3 for the instructed rule, zero for the non-instructed rule, and 1.5 for both rules in the choice trials.

Reward-amount based rule-selection task (AMNT)

The idea of the AMNT task was to induce a graded preference for the different options without being encouraged to plan an according movement. In the AMNT task we assigned colors orange and green to the two pre-cue triangles. The size of the pre-cue triangles indicated the amount of reward that would be associated with each rule in case it was instructed later in the trial. The probability of each rule being instructed was kept 50:50. The reward units at the seven bias levels corresponded to {6:0, 5:1, 4:2, 3:3, 2:4, 1:5, 0:6}. For example, in the case of a bias level 4:2 towards *ccw*, subjects got 4 reward units for an instructed *ccw* reach and 2 units for an instructed *cw* reach, whereas in the case of a bias level 6:0 towards *cw*, subjects received 6 reward units for *cw* reach but nothing for *ccw* reach (but still have to reach to the unrewarded target to complete that experimental block and proceed to the next block). The length of the feedback sound at the end of successful trials matched the amount of reward subjects received in that given trial. The ratios of initial EV associated with the two rules at the seven bias levels were then {2.5:0.5, 2.17:0.83, 1.83:1.17, 1.5:1.5, 1.17:1.83, 0.83:2.17, 0.5:2.5}. Again, these EVs were only valid for the time between pre-cue and rule-cue. After the rule-cue, the final EV for instructed trials was equal to the reward units assigned to the instructed rule (6:0, 5:1, etc.), always zero for the non-instructed rule, and 1.5 for both rules in the choice trials.

Importantly, we matched the preliminary EVs in the PROB and AMNT tasks as closely as possible, given the block structure of trials. In five out of seven conditions the EVs matched exactly. In two conditions the EV ratios matched approximately (PROB 2:1, AMNT 2.17:0.83). Choice trials and zero-bias trials were identical between both tasks in all other respects. Thus, prior to the rule-cue (when subjects did not know yet if a given trial will be instructed or choice) and after the rule-cue, the EVs for choice trials in both tasks matched. Hence our task design ensures that any observed differences in the free-choice behavior between the two tasks should be attributable to biases that were

introduced by the purposeful manipulation of the expectation for the instructed trials, and not to differences in the choice trials.

We also ran a control experiment for the AMNT task in which we doubled the reward contrast between high- and low-valued options. The reward units at the seven bias levels in this AMNT-double task corresponded to {12:0, 8.5:0.5, 5:1, 3:3, 1:5, 0.5:8.5, 0:6}. 16 subjects who had previously participated to the AMNT experiment were invited to perform the AMNT-double task. The subgroup selection depended only on subject availability and was independent of previous performance on the AMNT task.

Subject compensation and bonus

Recording sessions terminated when subjects reached 600 successful trials. By design, the same amount of reward was reached in both types of task. In the PROB task, in one block of six trials, the four instructed trials (4 × 3 units) and two choice trials (2 × 1.5 units) led to 15 reward units. To reach 600 trials, subjects needed to complete 100 blocks, i.e. a total of 15 × 100 = 1500 units. In the AMNT task, in each biased block, two out of four instructed trials delivered high reward and another two delivered low reward, e.g. in a 1:5 condition block, subjects received (5 × 2) + (1 × 2) = 12 reward units. Two choice trials (2 × 1.5 units) added to the same total of 15 units per block as in the PROB task. We converted 3 reward units to 2 Euro cent, which finally made 2/3 × 1500 units = 1000 cent, thus €10 per session. As there was no penalty for aborted trials and subjects had to reach the same number of successful trials, the total compensation per session was identical between tasks.

Additional to the baseline compensation of €10 for each accomplished session, a bonus of up to €6 for good performance could be achieved: performance under 50%: no bonus; 50%: bonus of €1; then each step of 5% will add €0.5 until reaching maximal bonus of €6 at 100%). Alternatively, subjects received a compensation of €6 per hour, if this yielded the higher compensation. For example, subjects with very high performance typically spent about one hour and received €15-16 whereas subjects who made many error trials and/or multiple pauses (self-paced task design) spent about two hours in the setup and received €12-13).

Pre-recording procedure and balancing

Subjects were required to maintain gaze at the center of the screen. For this, a calibration of the eye-tracking system was first carried out. Then a short (5-10 minutes) training session was run to accustom subjects to the task and setup condition. Since our experiment aimed at quantifying biasing effects, we wanted subjects to explore the range of possible free-choice responses before the start of the experiment. For this, we ran an initial balancing session for each subject to discourage subjects from repeating the same default reach choices through the rest of the experimental session. The balancing task contained only trials with a zero-bias condition (equal triangle sizes) and differed from the rule-selection task described above only in the reward schedule that we applied on the choice trials. Instead of rewarding both options with 100% probability, we used a bias-minimizing reward schedule (BMRS). In the BMRS the reward probabilities for free-choice targets were calculated based on the individual subject's choice history. The less often a target was freely chosen in the previous two choice trials, the higher the reward probability in favor of this target was (Klaes et al., 2011):

$$p(R_{cw})=F(n_{ccw}-n_{cw})$$

$$p(R_{ccw})=F(n_{cw}-n_{ccw})$$

where n_{cw} is the total number of rewarded *cw* reaches and n_{ccw} is the total number of rewarded *ccw* reaches. F was defined as:

$$F(x)=\begin{cases}1, & x>1\\ 2/3, & x=1\\ 1/2, & x=0\\ 1/3, & x=-1\\ 0, & x<-1\end{cases}.$$

Subjects were explicitly told that chosen targets would successively stop being rewarded and they needed to explore all possible reaches to complete this task. The balancing task was run until the subject made at least two *cw* and two *ccw* reaches at each pre-cue position, which means at least 16 choice trials. As choice trials made up 33% of all trials, the balancing task then comprised a minimum of 48 trials.

Apparatus and data acquisition

Subjects were seated in a dimly lit room facing an LCD screen (19" ViewSonic VX922) mounted behind a transparent touch sensitive screen (IntelliTouch, ELO Systems, CA, USA), with a chinrest and forehead band used to stabilize head position. The screen was mounted with a tilt of 33° from the vertical for subject's comfort, with the lower edge on the table at ~40 cm distance from the chinrest base and the top edge at eye level. The luminance of all colored stimuli was in the range 12-13 cd/m^2 (luminance meter LS-100, Minolta, Japan). Luminance was measured at eye level when positioning the color cues at the top of the four positions used in the experiment, i.e. at a direction of 90° from the screen center and with an eccentricity of 9 cm. Throughout the trial, the gaze direction of the subjects was constrained at the central fixation point (red square) within a tolerance window of 3 cm (~4.3° VA radius). Eye positions were monitored by a camera placed in front of the screen's lower edge (Eyelink 1000, Kanata, Canada). A real-time LabView program running on a PXI computer (National Instruments) was used to control the tasks and to register relevant stimulus properties, event timings, and subject's behavioral responses in each trial.

Behavioral data analysis

The main goal of this study was to quantify the biasing effect of predictability and preferability on choice behavior. Since preliminary analysis revealed symmetric effects of *cw/ccw* rule in our data (effects of interaction between *Bias* and rule types (*cw-ccw*) on reaction times in instructed *follow* trials: PROB task: t-statistic = -1.59, $p > 0.05$; AMNT task: t-statistic = 0.11, $p > 0.05$; interaction between *Bias* and rule types (*cw-ccw*) on choice probabilities: PROB task: t-statistic = -0.07, $p > 0.05$; AMNT task: t-statistic = -0.44, $p > 0.05$, GLMM; see details on GLMM below), we chose the absolute value of the pre-cued bias level as the independent variable and merged all trials with different pre-cue positions. In other words, we grouped the data into four bias conditions, one zero-bias condition plus three non-zero bias conditions. Bias degrees were quantified by the contrast in preliminary EV associated with each pair of reward amount or probability: higher EV - lower EV / higher EV + lower EV. Bias degrees were {0 0.22 0.33 0.67} for the PROB task and {0 0.22 0.45 0.67} for the AMNT task.

Additionally, we sorted the data according to rule-congruency, i.e. according to whether the reach was conducted to the same (*follow*) or the opposite (*against*) direction as the direction indicated by the bigger pre-cue triangle (bias direction). Note that *follow* and *against* responses could occur by instruction in instructed trials and by subjects' choice in choice trials and the instructed trials of the probabilistic task are equivalent to typical cue-validity tasks in which *follow* trials would correspond to *valid* trials, and *against* trials to *invalid* trials.

We analyzed error rate (ER), reaction time (RT) in both error and correct trials, and movement time (MT) in instructed trials. ERs were defined as the fraction of trials not leading to a successful target acquisition within the reach period, either due to mis-reaching or fixation breaks (often occurring together). As errors other than miss-reaches were in general very rare and both targets were considered valid in choice trials, we report ERs and error RTs only in instructed trials. RTs were defined as the time between the go-signal and the subject's release of the touch screen from the fixation position and MTs as the time between the subject's releases of the fixation point to the time that the subject's finger arrived at the target position. Both RTs and MTs were corrected for display and touch screen delays. Trials with invalid RTs (0.5% of total number of trials) were excluded from the RT analysis as subjects might have prematurely released the screen before the rule-cue was perceived. As rejection threshold we used 2.5 interquartile ranges below Q1 (25% quartile) or 100 ms, whichever value was higher.

We analyzed RTs and choice probability (CP) in choice trials. CP was defined as the fraction of correct choices following the bias introduced by the pre-cue. For the zero-bias condition we show the fraction of *cw* choices.

We tested for biasing effects in all aforementioned dependent variables. For this, a generalized linear mixed model ('`fitglme`'; MATLAB R2014b) was fitted to assess influences of bias degrees on ER, RT, MT, and CP, as well as differential effects between PROB and AMNT tasks. Full models included the factors bias degree (*Bias*: continuous variable), rule congruency (*Congruency*: categorical responses *follow* vs. *against* biased direction), and task type (*Task*: categorical variable: AMNT vs. PROB) and all interaction terms, as fixed effects. Note that we considered ER, RT, and MT at zero-bias degree in

both *follow* and *against* categories to keep the zero-bias level included in both *follow* and *against* fittings. Subjects were included as random effect (uncorrelated random intercepts and slopes for bias levels, congruency, and tasks) to account for the variance across subjects. The likelihood of the models including or excluding different fixed and random effects were compared using the Matlab function '`compare(model1, model2)`'.

We used the following model to test overall differential effects of bias between *follow-against* responses on ER (binomial response), RTs, and MTs with interaction term between Bias and Congruency in each task separately:

$$X \sim Bias * Congruency + (Bias * Congruency \mid Subjects), \quad (M1)$$

and to test the differential effects on ER, RTs, MTs between tasks:

$$X \sim Bias * Congruency * Tasks + (Bias * Congruency * Tasks \mid Subjects). \quad (M2)$$

Next, only for RTs, we additionally tested the differential effects on error RTs between tasks using the model:

$$RT \sim Bias * Error\ (success/error) * Tasks + (Bias * Error * Tasks \mid Subjects). \quad (M3)$$

When instructed to go against the bias, DDMs predict short error RTs in case of a bias mechanism mediated by a baseline shift whereas long error RTs in case of a drift rate change (Simen et al. 2009; Leite & Ratcliff 2011). As errors were rare in instructed *follow* trials and choice trials, we inspected error RTs only in instructed *against* trials.

When there was a biasing effect, we asked further whether (a) the biasing effect was symmetric for costs (*against*) and benefits (*follow*) and whether (b) the biasing effect was graded, i.e. scaled with the strength of the bias signal. With the model M1 we computed the slopes for *follow* and *against* and compared their confidence intervals to test whether the absolute values of the slope differed (asymmetry) or overlapped (symmetry) between both conditions. Note that the obviously different slopes (Figure 2B) between *follow* and *against* conditions were the reasons why we introduced congruency as a factor in the model. By modeling the data separately for each 'branch' of the bias factor separately, we got better linear fits than when treating the seven bias levels as a single factor (data not

shown). We tested for graded biasing effects versus a single step-like effect of bias, with *post-hoc* tests (paired t-tests with *Bonferroni* corrections for multiple comparisons) comparing each pair of successive bias conditions.

We tested biasing effects on CP using a separate full model without the *Congruency* term as there was no *follow-against* distinction in this case:

*CP (binomial response) ~ Bias * Tasks + (Bias * Tasks | Subjects).* *(M4)*

With additional *post-hoc* tests corrected for multiple comparisons we tested the graded effect as introduced above.

Results

We probed the biasing effect of motor-goal preferability and predictability on later risk-free choices in two steps. First, we wanted to confirm that both our value and probability manipulations were effective in affecting subject behavior. To do so, we analyzed error rates (ER), reaction times (RT), and movement times (MT) in instructed trials. If the pre-cue had a biasing effect then there should be costs and benefits involved with having to go *against* or being allowed to *follow* one's internal bias as a result of an instruction. Specifically, *follow* trials should lead to faster RTs and lower ER than *against* trials. Second, we wanted to test if preferability and predictability have similar or differential biasing effects on later choice behavior. For this, we analyzed RTs and choice probabilities (CP) in choice trials. If preferability and predictability induce the same biasing mechanism, we should observe the same response pattern across tasks. If instead preferability and predictability induce different biasing mechanisms, we should observe different response patterns. The former hypothesis might still be supported if the behavioral response patterns in instructed and choice trials are the same in the PROB and AMNT task, except for a potentially reduced effect in the AMNT task compared to the PROB task (Maddox and Bohil, 1998; Mulder et al., 2012).

Performance comparison PROB vs. AMNT task

We first tested if the AMNT and the PROB experiment were different in task difficulty. For this we did not compare only the total performance, but particularly the zero-bias trials. Zero-bias trials were identical in both experiments, except for the task context.

Average performance was 86.84 ± 2.1% in AMNT task (N = 33) and 91.53 ± 0.8% (mean ± SEM) in PROB task (N = 41). Subjects who participated in both tasks showed slightly better overall performance in PROB task than AMNT task (AMNT: 88.13 ± 1.7%, PROB: 91.91 ± 0.8%, mean ± SEM, N = 31, $p < 0.05$, paired t-test). However, importantly, when comparing performance only in the zero-bias condition of the instructed trials, which served as our reference condition, subjects performed equally well (AMNT: 87.53 ± 1.75%, PROB: 89.67 ± 1.37%, mean ± SEM, N = 31, $p > 0.05$, paired t-test).

Subjects performed both tasks with shorter RT in instructed compared to choice trials (AMNT [instructed-choice]: -24.48 ± 2.47 ms, $p < 10^{-9}$, PROB [instructed-choice]: -38.00 ± 4.24 ms, $p < 10^{-9}$, mean ± SEM, N = 31, paired t-test) and overall responded faster in PROB task than in AMNT task (AMNT-PROB: 21.03 ± 6.62 ms, $p < 0.01$, mean ± SEM, N = 31, paired t-test). However, when comparing only RTs from the zero-bias condition, subjects showed no RT difference between tasks (instructed [AMNT-PROB]: 7.95 ± 7.87 ms, $p > 0.05$, choice [AMNT-PROB]: -0.31 ± 8.07 ms, $p > 0.05$, mean ± SEM, N = 31, paired t-test).

In summary, the AMNT and the PROB task contexts did not lead to performance differences in the zero-bias trials, i.e. in the trials which do not differ between the tasks. In trials with non-zero bias levels, subjects responded on average more quickly and made fewer errors in the PROB task than in the AMNT task.

Effects of a priori preferability and predictability in instructed behavior

We first needed to establish whether our manipulations of probability and amount were strong enough to be effective. We compared ERs, RTs, and MTs in instructed trials between the PROB and AMNT task to test if they depended on the bias degree. If so, we

further tested if bias degree led to symmetry in costs and benefits, and if the effect was gradually increased with increasing bias degree. In both tasks we found significant effects of bias degree on instructed behavior, but with differences in individual aspects, as detailed in the following.

Error rates (ER) depended on the instruction to follow or go against the bias in both tasks (Figure 2A). A generalized linear mixed effect model (GLMM) showed higher ERs when subjects were instructed to reach against the bias than follow the bias in both tasks (interaction between *Bias* and *Congruency*: AMNT: t-statistic = 6.07, $p < 10^{-8}$; PROB: t-statistic = 10.05, $p < 10^{-9}$, N = 43). The effect on ERs differed between the PROB and AMNT task quantitatively. The ER full model confirmed that the ER cost was significantly higher in *against*-instructed trials in the PROB task than in the AMNT task (interaction between *Bias* and *Tasks* on *against* trials: t-statistic = -3.47, $p < 0.001$). Further there were overall significant difference in ER patterns between tasks (interaction between *Bias*, *Congruency*, and *Tasks*: t-statistic = -3.96, $p < 10^{-4}$).

The ER dependency on the bias degree was symmetric in strength in instructed trials of both tasks. Decrease in response to *follow* instructions was not significantly steeper than the ER increase in response to *against* instructions. (AMNT-*against*: 95% confidence interval CI = [-0.01, 1.03] (increase in percentage of error trials per one bias degree); AMNT-*follow*: CI = [-2.89, -0.97]; PROB-*against*: CI = [1.41, 3.21]; PROB-*follow*: CI = [-3.61, -1.92]).

ER in neither tasks showed a gradual effect as function of bias degree (Figure 2A).

Reaction times (RT): depended on the bias degree in both tasks similarly to ERs, but partially in a more gradual fashion (Figure 2B). GLMM showed an overall differential effect of biasing degree on RT between *follow* and *against* instructed reaches in both tasks (interaction between *Bias* and *Congruency*: AMNT: t-statistic = 9.72, $p < 10^{-9}$; PROB: t-statistic = 12.28, $p < 10^{-9}$). This means the bias degree induced systematic costs and benefits for RTs in instructed reaches. Also, the RT full model showed significantly different patterns of RTs between tasks (interaction between *Bias*, *Congruency*, and *Tasks*: t-statistic = 6.23, $p < 10^{-9}$), and a smaller RT benefit in the AMNT task as compared to the PROB task (interaction between *Bias* and *Tasks*: t-statistic = -5.65, $p < 10^{-9}$).

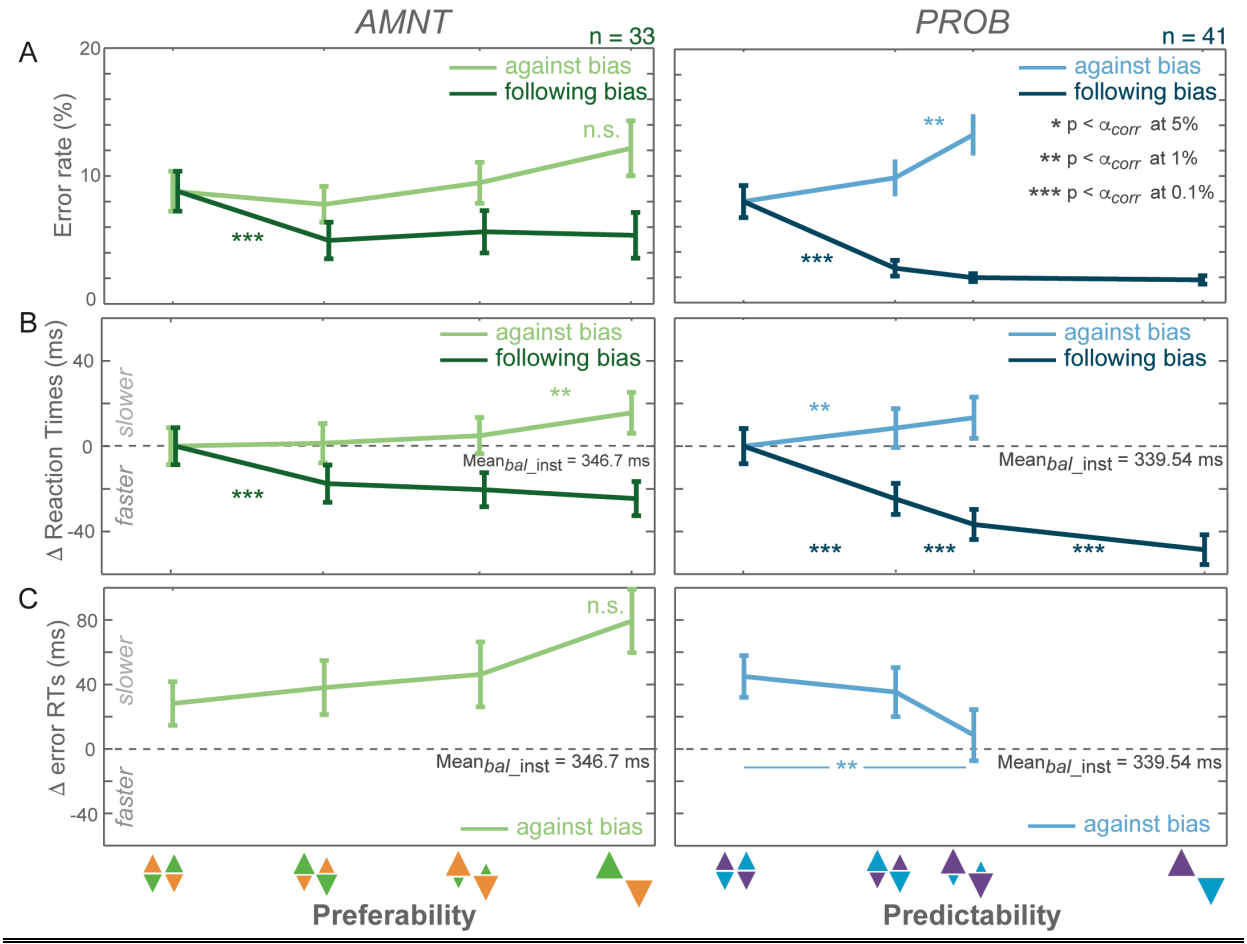


FIGURE 2 | Biasing effects on instructed responses. (A) Average error rates (ER): proportion of error trials committed during the reach period, **(B)** average reaction time (RT) difference from average RTs in instructed trials of zero-bias condition of each task in all bias levels in AMNT (left) and PROB (right) tasks, **(C)** average error RT difference (zero line indicates average correct RTs in instructed trials of zero-bias condition of each task). Dark and light colors represent *follow* and *against* responses, respectively. Error bars depict standard errors. (* $p < \alpha_{corr}$ at 5%, ** $p < \alpha_{corr}$ at 1%, *** $p < \alpha_{corr}$ at 0.1%, paired t-test with *Bonferroni* correction).

RT costs and benefits in the instructed trials of each task were not symmetric. The PROB task showed a significantly larger RT benefit of following than cost for going against the bias (PROB-*against*: CI = [16.99, 46.79] (increase in RTs (ms) per one bias degree); PROB-*follow*: CI = [-85.31, -63.71]). Asymmetry was not as strong in the AMNT task, since the absolute values of the confidence limits partly overlapped (AMNT-*against*: CI = [12.94, 27.50]; AMNT-*follow*: CI = [-44.41, -27.39]).

RTs in instructed trials partially showed a gradual increase in effect strength with increasing bias degree (Figure 2B). In the AMNT task, RTs decreased in one significant step when subjects were instructed to *follow* the bias. In the PROB task RTs decreased with each step in response to *follow* instructions with increasing bias level.

Analysis of error RTs. Depending on which mechanism explains the decision process, RTs in unsuccessful trials can be expected to be either faster or slower than successful trials (e.g. Bogacz, 2007; Brown and Heathcote, 2008; Heitz, 2014; Smith and Ratcliff, 2004). We analyzed error RTs when an instruction to go *against* a bias was violated; since error rates in *follow* trials were too rare to be analyzed properly. Error RTs depended on the bias condition in opposite ways in the AMNT and PROB task. Error RTs showed significant differential effects in response to *follow* or *against* instructions in the PROB task, but not in the AMNT task (interaction between *Bias* and *Congruency*: AMNT: t-statistic = -1.07, $p > 0.05$; PROB: t-statistic = 2.59, $p < 0.01$). Due to limited number of errors in *follow* trials of both tasks, slope analyses detected no significant slopes in follow trials (AMNT-*follow*: CI = [-36.90, 67.87]; PROB-*follow*: CI = [-64.10, 7.51]; not shown). Importantly, in *against* trials, slope analyses revealed significant increase in error RTs with bias levels in AMNT task, as opposed to significant decrease in PROB task (AMNT-*against*: CI = [8.64, 80.66]; PROB-*against*: CI = [-127.78, -50.155]; Figure 2C). RT differences between individual consecutive bias levels did not reach significance in either task, likely due to the limited number of error trials.

Overall, analysis of the instructed trials showed that manipulation of both preferability and predictability were effective and had consequences on RTs and ERs, yet, with indications from the error trials analysis that underlying mechanisms might differ.

Effects of a priori preferability and predictability on free-choice behavior

The main goal of our study was to investigate the effects of *prior* probability and expected amount on risk-free, reward-balanced choice behavior. We quantified choice probabilities (CP), and RTs for the randomly interspersed choice trials.

Choice probabilities (CP) depended on the bias degree in both tasks (AMNT: t-statistic = 2.14, $p < 0.05$; PROB: t-statistic = 7.90, $p < 10^{-9}$), however with a significantly smaller biasing effect in AMNT task than in PROB task (interaction between *Bias* and *Tasks*: t-statistic = -6.88, $p < 10^{-11}$). *Bonferroni*-corrected *post-hoc* t-tests revealed a graded biasing effect in the PROB task, in which choice bias became stronger for each step of bias degrees. In contrast, choice bias in the AMNT task showed only a single significant step

between balanced and [1/3] bias degree, and no difference between different non-zero bias degrees (Figure 3A).

Choice reaction times were only affected by bias degree in the PROB task. GLMM showed a differential effect between follow-against reaches in PROB task but not in the AMNT task (interaction between *Bias* and *Congruency*: AMNT: t-statistic = -1.91, $p > 0.05$; PROB: t-statistic = 7.02, $p < 10^{-9}$). The RT full model confirmed significantly different patterns of RTs between tasks (interaction between *Bias*, *Congruency*, and *Tasks*: t-statistic = 7.16, $p < 10^{-9}$) with the PROB task showing a significantly larger effect on RT benefits when following the bias (interaction between *Bias* and *Tasks*: t-statistic = 8.49, $p < 10^{-9}$). In addition, a clear asymmetry of choice RTs was revealed in the PROB task (PROB-*against*: CI = [-30.16, -4.44]; PROB-*follow*: CI = [-92.78, -61.25]).

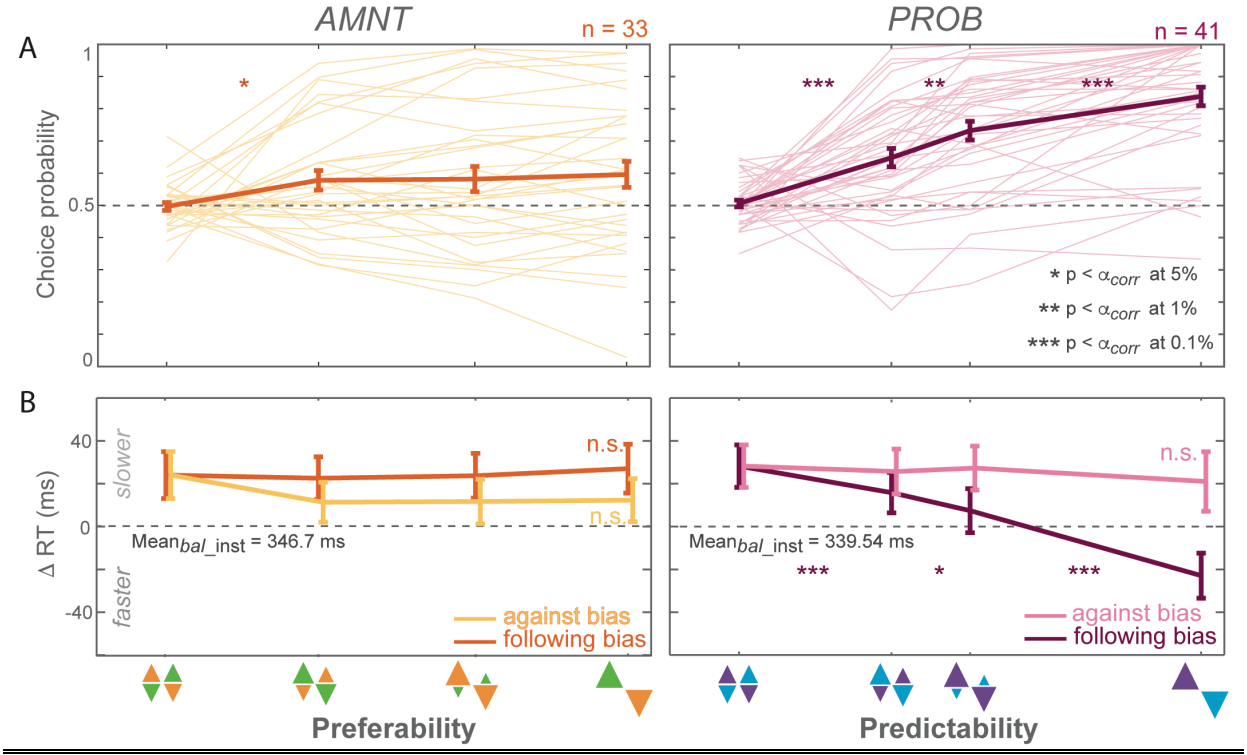


FIGURE 3 | Biasing effects on free-choice responses. (A) Average choice probability (CP): fractions of choice toward the biased direction (at the zero-bias level, as there is no biased direction, fractions of clockwise choice are presented), and **(B)** average reaction time (RT) difference (zero line indicates average RTs in instructed trials of the zero-bias condition of each task) in all bias levels in AMNT (left) and PROB (right) tasks. Dark and light colors represent *follow* and *against* responses, respectively. The error bars represent standard errors. (* $p < \alpha_{corr}$ at 5%, ** $p < \alpha_{corr}$ at 1%, *** $p < \alpha_{corr}$ at 0.1%, paired t-test with *Bonferroni* correction).

RTs in the choice trials showed a gradual increase in effect strength with increasing bias degree only for benefits in the PROB task. *Post-hoc* tests showed no significant RT

differences between individual neighboring bias degrees in the AMNT task but RTs that increased gradually with the bias degree in *follow*-choices in the PROB task (Figure 3B).

Notably, the RT benefit in *follow* reaches was comparable between instructed and choice trials in the PROB task, while in the AMNT task there was only a benefit for *follow* instructions, not for *follow* choices (interaction between *Bias* and trial types: AMNT: t-statistic = -1.21, $p < 10^{-9}$; PROB: t-statistic = 7.02, $p > 0.05$). This means, while predictability affected later choice behavior in the same way as instructed behavior, in contrast, preferability showed clear effects in instructed behavior but did not generalize to the choice behavior.

No effect of doubling the reward in the AMNT task

The observed limited effects of preferability as compared to predictability might be due to lack of EV contrast. EVs cannot be linearly translated into subjective expected utility which is more directly linked to choice (Neumann and Morgenstern, 1944; Savage, 1954). To rule out this possibility, we conducted a control experiment in which we doubled the reward amount ratio between high- and low-value options.

Error rates (ER) depended on the instruction to follow or go against the bias in the AMNT-double task. A GLMM showed higher ERs when subjects were instructed to reach against the bias than follow the bias (interaction between *Bias* and *Congruency*: t-statistic = 3.86, $p = 0.0001$; $N = 16$). ER decrease in response to *follow* instructions was significantly steeper than the insignificant increase in response to *against* instructions. (*against*: CI = [-0.40, 1.29]; *follow*: CI = [-5.00, -1.14]). Similarly to the standard AMNT task, ERs decreased when subjects were instructed to follow the bias (Figure 4A).

Reaction times (RT) depended on the bias degree in the AMNT-double task similarly to the standard AMNT task. GLMM showed an overall differential effect of biasing degree on RT between *follow* and *against* instructed reaches in both tasks (interaction between *Bias*

and *Congruency*: t-statistic = 7.10, $p < 10^{-9}$). As in the standard AMNT task, asymmetry was weak or absent in the AMNT-double task, since the absolute values of the confidence limits partly overlapped (*against*: CI = [14.43, 39.95]; *follow*: CI = [-60.03, -30.71]). RT decrease in *follow* trials and increase in *against* trials showed similar pattern as in the standard AMNT task (Figure 4B).

Choice probabilities (CP) showed insignificant increase with bias degrees in the AMNT-double task (t-statistic = 1.62, $p > 0.05$), and no CP difference was detected between the standard AMNT and the AMNT-double tasks (interaction between *Bias* and *Tasks* (AMNT vs. AMNT-double): t-statistic = 0.1114, $p > 0.05$).

Choice RTs were not affected by bias degree in the AMNT-double task. GLMM showed no differential effect between *follow-against* reaches in the AMNT-double task (interaction between *Bias* and *Congruency*: t-statistic = -1.91, $p > 0.05$).

CPs and choice RTs showed no difference between neighboring bias degrees (Figure 4C-D).

In summary, doubling of the reward amount contrast between high- and low-value options did not change the strength of biasing effects in the AMNT task.

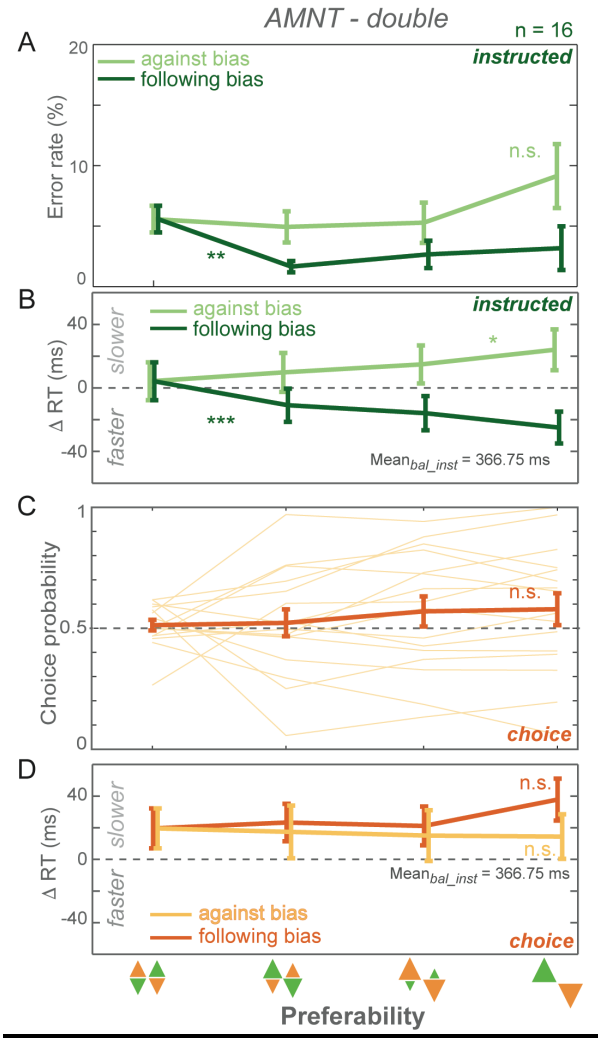


FIGURE 4 | Results of AMNT task with double reward size (AMNT-double). (A) Average error rates, **(B)** average instructed RT difference, **(C)** Choice probabilities, **(D)** average choice RT difference (zero line indicates average RTs in instructed trials of zero-bias condition of each task) in all bias levels. Dark and light colors represent *follow* and *against* responses, respectively. Error bars represent standard errors. (* $p < \alpha_{corr}$ at 5%, ** $p < \alpha_{corr}$ at 1%, *** $p < \alpha_{corr}$ at 0.1%).

Subject sub-grouping based on choice bias

CPs showed a weaker bias in the AMNT task than in the PROB task on average across all subjects. We asked if this was because some subjects showed no bias in the AMNT task, while others might show a bias of the same magnitude as in the PROB task. If so, would the pattern of RT results for subjects with a strong CP bias in the AMNT task look similar to the pattern of RT results in the PROB task? CPs across subjects varied, especially in the AMNT task. We subdivided subjects depending on their bias in the choice probabilities to test, first, if individual subjects showed a choice pattern contrary to the average pattern described above; and second, whether subjects' choice behavior in choice trials would predict the RT patterns in instructed and in choice trials. Using a GLMM on each subject's CP, we distinguished three classes of subjects: CP-biased (significant positive slope), CP-unbiased (slope not significantly different from zero), or CP-counter-biased (significant negative slope).

Most subjects in PROB task were CP-biased (33 biased, 8 unbiased, and no counter-biased). In contrast, the majority of subjects in the AMNT task were CP-unbiased (10 biased, 20 unbiased, and 3 counter-biased). Out of the 31 subjects who participated in both experiments, 27 were biased and 4 were unbiased in the PROB task. Out of these 27 biased subjects in the PROB task, 9 were biased, 17 were unbiased, and one was counter-biased in the AMNT task. And out of four unbiased subjects in the PROB task, two were also unbiased and two were counter-biased in AMNT task. No subject showed CP-bias in the AMNT task and was unbiased or counter-biased in the PROB task. In summary, no individual subject showed a reversed pattern of CPs in the AMNT task to the average pattern of CPs.

In the PROB task, RT benefits of instructed *follow* reaches matched RT benefits when choosing *follow* reaches in choice trials in both biased and unbiased subjects. GLMM showed no interaction between *Bias* and trial types (instructed vs. choice) in either subgroup [PROB (biased): t-statistic = 0.82, $p > 0.05$; PROB (unbiased): t-statistic = 1.49, $p > 0.05$]. The biasing effect differed between subject subgroups on instructed RTs but not in choice RTs [interaction between *Bias* and subgroup: PROB (instructed): t-statistic = 2.83, $p < 0.01$; PROB (choice): t-statistic = 1.06, $p > 0.05$]. Notably, shorter RTs when

choosing against bias seemed to come from unbiased subjects while biased subjects showed flat RTs in *against* trials of both tasks [AMNT-*against* (biased): t-statistic = -1.96, $p > 0.05$; AMNT-*against* (unbiased): t-statistic = -2.78, $p < 0.01$; PROB-*against* (biased): t- statistic = -1.07, $p > 0.05$; PROB-*against* (unbiased): t-statistic = -2.86, $p < 0.01$].

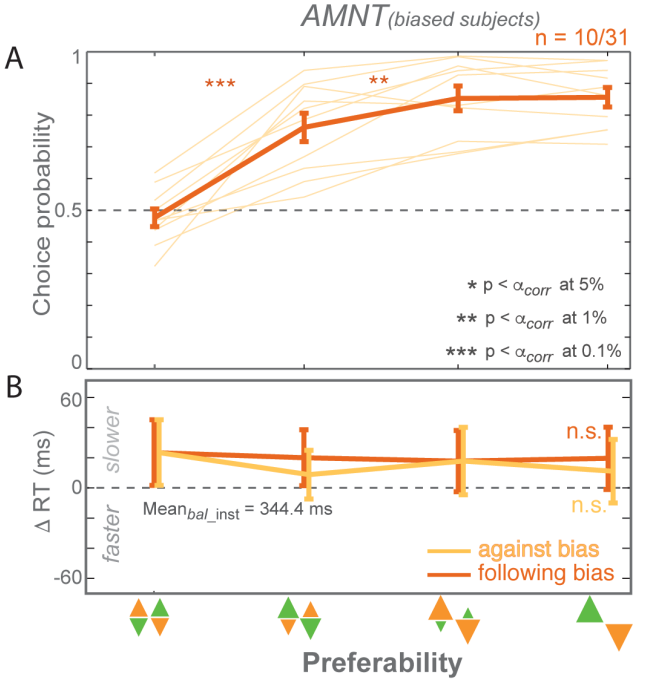


FIGURE 5 | Responses of CP-biased subjects in AMNT task. (A) Choice probabilities, and **(B)** average choice RT difference (zero line indicates average RTs in instructed trials of zero-bias condition) in all bias levels. Dark and light colors represent *follow* and *against* responses, respectively. The error bars depict standard errors. (* $p < \alpha_{corr}$ at 5%, ** $p < \alpha_{corr}$ at 1%, *** $p < \alpha_{corr}$ at 0.1%).

In summary, dividing subjects into CP-biased and CP-unbiased subgroups showed that even the minority of subjects who were biased in their choice behavior in the AMNT task lack a biasing effect on choice RTs (Figure 5). This also means they showed significantly different RT behavior between the AMNT and PROB tasks and further supports the idea that the underlying decision processes are different in both tasks

Movement times

Studies on motor planning had previously shown that *invalid* pre-cueing can affect not only RT but also MT (e.g. Leis et al., 2005). While classical DDMs to not consider MT, in the context of affordance or motor-oriented models of decision-making, motor kinematics can reveal additional insights (Gallivan et al., 2015). The prediction would be that having to go against planned movement should require any preliminary motor plan to be suppressed and lead to slower movement execution (Cisek, 2012).

Movement time (MT) analysis. GLMM showed a differential effect on MT between *follow* and *against* reaches in both tasks, yet the effect in the AMNT task was minimal (interaction between *Bias* and *Congruency*: AMNT: t-statistic = 2.01, $p < 0.05$; PROB: t-statistic = 5.76, $p < 10^{-8}$). Correspondingly, the MT full model confirmed significantly

different patterns of MTs between tasks (interaction between *Bias*, *Congruency*, and *Tasks*: t-statistic = -5.03, $p < 10^{-6}$), and a substantially higher MT of *against* reaches in the PROB task as compared to the AMNT task (interaction between *Bias* and *Tasks*: t-statistic = 5.70, $p < 10^{-7}$).

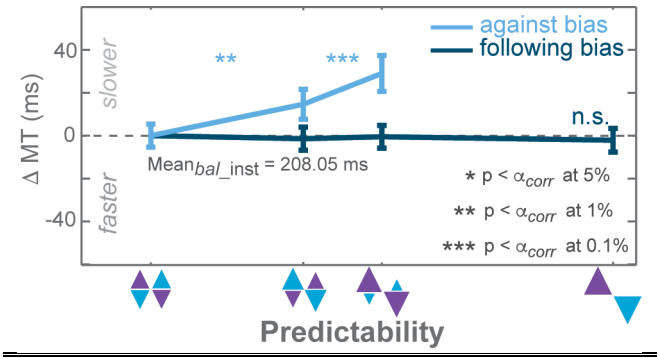


FIGURE 6 | Biasing effects on movement times (MT) instructed against trials of the PROB task. Zero line indicates average MTs in instructed trials of zero-bias condition. Dark and light colors represent *follow* and *against* responses, respectively. Error bars depict standard errors. (* $p < \alpha_{corr}$ at 5%, ** $p < \alpha_{corr}$ at 1%, *** $p < \alpha_{corr}$ at 0.1%, paired t-test with *Bonferroni* correction)

In the AMNT task, *follow* and *against* slopes did not significantly deviate from zero (AMNT-*against*: CI = [-3.14, 11.55]; AMNT-*follow*: CI = [-12.82, 0.93]) whereas costs and benefits of MTs in the PROB task showed clear asymmetry (PROB-*against*: CI = [55.92, 107.2]; PROB-*follow*: CI = [-17.82, -1.03]).

Notably, only MT cost in *against* trials of the PROB task showed gradual biasing effect whereas MTs remained unchanged in *follow* instructed trials in the PROB task (Figure 6), choice trials in the PROB task, and all types of trials in the AMNT task (not shown).

As previous studies showed that motor planning reduces motor variability (e.g. Churchland et al., 2006; Harris and Wolpert, 1998; Todorov, 2004), we tested as an additional confirmation of the MT result, if there was a biasing effect on endpoint variability (EPV), defined as the average distance of reach endpoints to the mean reach endpoint for each target location. In congruence with MT results, GLMM showed significantly higher EPVs of *against* reaches in the PROB task as compared to the AMNT task (interaction between *Bias* and *Tasks*: t-statistic = 2.42, p = 0.01).

Also consistent with the MT results, EPV only in *against* but not *follow* trials of the PROB task showed significant deviation from zero (PROB-*against*: t-statistic = 2.69, $p < 0.01$; PROB-*follow*: t-statistic = 1.87, $p > 0.05$) while only *follow* slope marginally deviated from zero in the AMNT task (AMNT-*against*: t-statistic = -0.15, $p > 0.05$; AMNT-*follow*: t-statistic = -2.29, $p < 0.05$).

In summary, only when being instructed against a bias in the PROB task, subjects showed effects on motor execution, visible by elevated MTs and EPVs with increasing bias degree. Neither instructions to go against a bias in the AMNT task nor choosing freely in any task had an effect on MTs and EPVs.

Discussion

In our rule-based selection task with sequential cueing, we informed participants about the probability (PROB task) or reward amount (AMNT task) of a subsequent instruction. Using interspersed, equal-valued and risk-free choice trials, we probed to what extent the *a priori* predictability or preferability of the upcoming rule induced a behavioral bias in subjects. Our results showed multiple biasing effects of movement planning due to predictability when compared to conditions that dealt with preferability of a goal but without planning being encouraged. (1) Subjects' responses were faster and less prone to errors in instructed trials when the final instruction matched the more likely or higher-valued rule (*follow* trials). (2) Responses were slower and more error-prone when the instruction matched the less likely or lower-valued rule (*against* trials). The strength of *follow* and *against* effects was in general not symmetric. (3) In the absence of an instruction, without any objective advantage, subjects more frequently chose the originally more likely rule, while this was true to a much lesser degree for the originally higher-valued rule. Subjects gained a reaction time advantage only in case of choosing the originally more likely rule, not the originally higher-valued rule. (4) Having to go against the more likely rule (but not against the higher-valued rule) slowed movement times and raised endpoint variability, while freely chosen movement execution was unaffected by *prior* expectations. Our results indicate a structural difference between decision biases resulting from predictability or preferability. These results are not consistent with the idea that the value-based decision process is a graded version of the probability-based decision process, as suggested earlier. Instead our results suggest that preliminary action planning is the major driving force for pending choice behavior and acts via different mechanisms than preference.

Probing bias with neutral choice trials

Inducing bias with instructed trials and probing bias with choice trials was an important feature of our task design that revealed differences between the consequences of predictability and preferability. Previous studies have compared the effect of *prior* probability with the effect of reward amount ("potential pay-off") on choice using partly ambiguous sensory evidence (Leite and Ratcliff, 2011; Maddox and Bohil, 1998; Mulder et al., 2012; Simen et al., 2009). In one study, subjects had to decide between two alternatives of a random dot motion stimulus, the probability or the reward amount of which was announced at the beginning of each trial and – different to here – guaranteed until the end of the trial (Mulder et al., 2012). Choices were risky due to the uncertainty about the evidence provided by the partly ambiguous stimulus. Results showed a weaker effect of potential payoff compared to *prior* probability. However, the fact that the difference in value between options was known from the start of each trial might have encouraged subjects to preliminarily plan the corresponding action, since in risky choices this would on average be advantageous. The fact that both types of *prior* expectations led to RT differences was taken as an indication for a shift of the DDM baseline in both cases. The reduced strength of effect for *prior* value expectations compared to probability expectations was accounted for by assuming an intermediate baseline shift but otherwise equivalent underlying mechanisms (Bogacz, 2007; Maddox and Bohil, 1998; Mulder et al., 2012). The competition-between-reward-and-accuracy-maximization (COBRA) hypothesis (Maddox, 2002; Maddox and Bohil, 1998) was proposed to explain the intermediate baseline shift in these perceptual decision-making experiments. In COBRA the reduced biasing effect of payoff manipulation is due to a conflict between (biased) reward and (unbiased) accuracy maximizing criteria, while both criteria show (biased) synergistic effects in case of probabilistic manipulation (Ashby et al., 1998; Bogacz et al., 2006; Maddox, 2002; Mulder et al., 2012). While the intermediate-baseline-shift model could account for the behavior observed in our instructed trials it does not predict the observed patter of CP and RT in the choice trials, as discussed in the following paragraphs.

Two features of our experimental design helped to decide whether biases induced by predictability and preferability can be accounted for by the same mechanism. First, two

sequential cues provided subjects with the necessary information, with the second cue either instructing a specific rule (100% evidence) or allowing subjects to choose freely between both rules (rule-neutral evidence). In the DDM concept, the onset of the second cue (rule-cue), which also signals the subjects to immediately make a reach, marks the initiation of the drift process. Obeying causality, this implies that the rule-cue can only have an influence on the drift period whereas the baseline could only be set by the *prior* knowledge provided by the pre-cue. Importantly, as the free-choice cue provides no additional evidence supporting either rule, an effect of the *prior* information – either probability or reward amount – on the baseline should persist in the choice trials. In contrast, any effect of expectation that does not affect instructed and choice trials in the same way cannot be mediated by baseline changes but must indicate changes during the drift process following the instructive rule-cue. As a second important feature, we varied *prior* expectations gradually. If *prior* expectations on value are based on the same mechanism as *prior* expectations on probability (except for a scaling factor) then graded effects in one case should also result in graded effects in the other case.

By introducing reward-neutral interspersed choice trials we also avoided confounds between planning and motivation. If *a priori* expected values and later actual values are always identical, then motivational effects of the actually different rewards (Dayan and Balleine, 2002; Franchina and Brown, 1971; Hassani et al., 2001; Hollerman et al., 1998; Mir et al., 2011) cannot be disentangled from the effect of *a priori* expectations on the reward. This makes it difficult to account for potential effects of action planning and motivation when assessing the effect of *a priori* expected value compared to *prior* probabilities. Here we compare *a priori* preferability and predictability in their effect on equal-valued choice trials, thereby avoiding motivational confounds.

Predictability leads to reduced migration distance

Our results from the PROB task support the view that *prior* probability affects migration distance in DDM. Equivalent effects of probability bias were observed in instructed and choice trials, allowing for a mechanism that starts prior to the rule-cue, i.e. during the DDM baseline period. In choice trials, in particular, as subjects always received the same reward for each possible choice, the reason that both options were not chosen equally

often cannot be due to a reward difference but must result from the biasing effect of the *prior* probability. A reduction of the migration distance to the boundary associated with the more likely rule can well explain the ER and RT benefits observed for following the biased rule by instruction and by choice, as well as the CP shift in choice.

Further support for a reduced migration distance towards the threshold of the predicted target (biased threshold) is provided by a higher frequency of errors with fast RTs in case of instruction to go against the bias. A short migration distance implies that the threshold can be reached with small fluctuations towards the biased threshold, which comes at the cost of wrongly choosing the option associated with the closer boundary (Bogacz et al., 2010; Heitz, 2014) thus occur at shorter RTs than correctly instructed responses (Brown and Heathcote, 2008; Smith and Ratcliff, 2004).

However, the reduced migration difference towards the biased threshold cannot have been achieved by a pure baseline shift. With a pure baseline shift, one would expect symmetric costs and benefits, since the migration distance towards one boundary is reduced by the same amount as it is increased for the other boundary (Figure 7A). Instead, we found that the ER and RT benefits for following the bias were larger than the costs for going against the bias.

Because of the direction of the asymmetry in our data, we rule out that the biasing effects in the PROB task are explained by a bias-proportional anti-symmetric change in drift rates. By this, we mean an increased drift rate towards the biased option and decreased drift rate towards the counter-biased option, each proportional to the bias degree. In this case one would expect a cost-benefit asymmetry opposite to the asymmetry observed, i.e. larger RT costs than RT benefits (Figure 7B). This is because in the DDM the RTs are reciprocally proportional to the drift rate. Increasing drift rate hence leads to an RT benefit that is smaller in absolute value than the RT costs associated with a same-amount decrease in drift rate. This is the same logic that explains how a drift-rate which symmetrically varies around an average drift rate creates the typically observed left-skewed RT distributions. While a pure baseline shift is not sufficient to explain the behavioral results, the ER and RT data in the PROB task also cannot be explained by a pure or an additional bias-proportional anti-symmetric change in drift rate.

Rather, the asymmetry in gradual costs and benefits can be explained with models that, in addition to the baseline shift, allow (1) a bias-proportional gradually reduced migration distance by a lowered boundary for the biased option or the counter-biased option or both (Figure 7C) or (2) a bias-proportional symmetric increase in drift rate towards both options (Figure 7D).

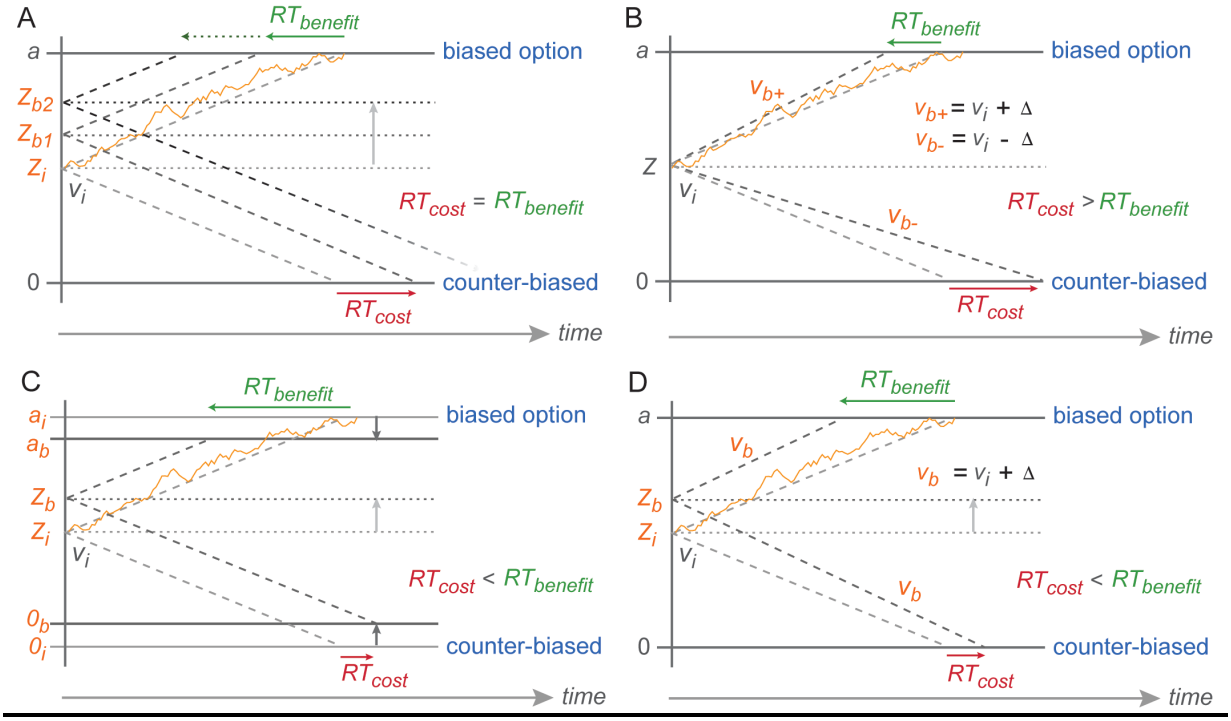


FIGURE 7 | Predicted symmetric effect of baseline shift and asymmetric effect of drift rate change. (A) Baseline shift of the same magnitude should result in symmetric RT cost-benefit whereas **(B)** RT shortening toward biased reach goal due to drift rate increase (v_{b+}) (from the neutral drift rate: v_i) should be smaller than RT cost toward counter-biased reach goal due to drift rate decrease (v_{b-}) of the same magnitude (Δ). Partial boundary lowering **(C)** and drift rate increase **(D)** explanations for asymmetric cost-benefit RTs in instructed reaches. *0*, *a*, *v*, *Z* represent counter-biased boundary, biased boundary, drift rate, and baseline, respectively. The subscript *i* indicates *initial* parameters in absence of any bias (non-biased trials) whereas subscript b represents parameters in case of bias with higher number depicting higher degree of bias and + vs. − depicting parameters toward biased vs. counter-biased boundaries, respectively.

While the behavioral consequences of predictability were largely equivalent between instructed and choice trials, there were also two differences. First, RTs were overall 20-30ms slower in choice trials, and, second, no RT costs were imposed when subjects chose freely against the biased rule. Neither difference contradicts the idea of a reduced migration distance for the biased option.

The average RT offset of 20-30 ms is not surprising since instructed trials always provided instantaneous and unambiguous evidence whereas choice trials provided rule-neutral evidence. Higher RTs for choice compared to instruction could be due to: (1) a

slower drift rate due to absence of clear evidence in the choice case (Hanks et al., 2015; Roitman and Shadlen, 2002); (2) higher decision thresholds in the case of symmetric reward choice (Cavanagh et al., 2011; Cavanaugh et al., 2006; Summerfield and Tsetsos, 2012); (3) an increased duration of non-decision time, which delays migration initiation, due to unclear stimuli (Coallier and Kalaska, 2014; Mulder and van Maanen, 2013); or some combination of these possibilities. Only processes that occur with or after the rule-cue can account for differences between instructed and choice trials since subjects were unaware of the trial type prior to the rule-cue. Unless thresholds became adapted with presentation of the rule-cue, this would argue in favor of different drift rates for instructed and choice trials. As subjects had to respond within a fixed time limit and waiting longer would not have provided additional evidence, the exact amount of RT offset in choice trials is probably determined by an internal urgency signal (Cisek et al., 2009).

The fact that we did not find an increasing RT cost for choices against increasing bias is also consistent with the idea of a reduced migration distance and can be explained by the stochastic nature of the diffusion process in DDM (Brown and Heathcote, 2008; Heathcote and Love, 2012). With neutral evidence provided by the rule-cue, subjects will chose against the bias only when the decision variable due to random fluctuations is coincidentally around the level of the original baseline for counter-biased trials, or even closer to the counter-biased boundary at the time of the commitment to a choice. With an increasing shift in the baseline away from the initial baseline level (towards the biased boundary), *against* choices become less and less likely (explaining the CPs), but *against* choice RTs are still independent of the degree of choice bias since they always start with the same distance from the counter-biased boundary (Figure 8).

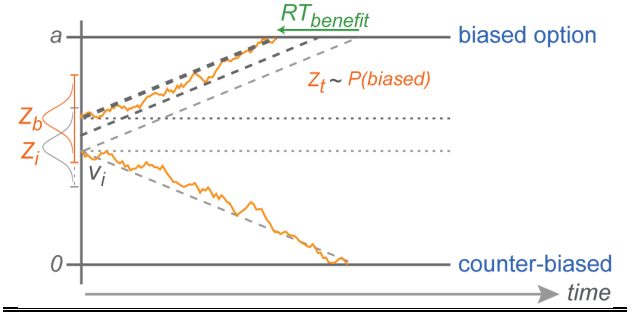


FIGURE 8 | Explanation of counter-biased choice. *Z* depicts the range in which the stochastic baseline shift operates. Counter-biased choice is possible only when the baseline shift in a given trial is still close to the initial baseline (lightest gray horizontal dotted line in the middle). This explains the pattern of choice RTs against the bias that are similar to choice RTs in the zero-bias condition.

In summary, our results from the PROB task confirm the hypothesis of a reduced migration distance due to predictability. This reduction is well explained by the combination of a bias-proportional baseline shift and either a bias-proportional reduction in threshold separation or a symmetric drift-rate increase.

Preferability has a different effect on choice than predictability

A main question of our study was whether planning of an optional action *per se* is responsible for later choice biases. We therefore tested for differences in bias between conditions in which an action is more likely to be requested later, compared to when the same action leads to a higher potential reward in the unpredictable case that it will be requested. Our results from the instructed trials of the AMNT task fit the predictions of the intermediate-baseline-shift hypothesis discussed above, while the results from the choice trials do not.

Predictability and preferability cause structurally, not just gradually, different behavioral bias. Consistent with earlier findings (Leite and Ratcliff, 2011; Maddox and Bohil, 1998; Mulder et al., 2012; Simen et al., 2009) and the intermediate-baseline-shift hypothesis (Bogacz et al., 2006), we observed a stronger biasing effect on ER and RTs in instructed PROB trials than in instructed AMNT trials. Note that the bias degrees in our experiment were carefully matched in terms of *a priori* expected value between both tasks (see Methods). One could still expect different subjective utilities between corresponding bias degrees of both tasks, depending on the subject's level of risk aversion, e.g. devaluing the higher rewards of the AMNT task which have only a 50% chance of becoming available in the end. Yet, we do not think that differences in expected utility explain the quantitative difference in the strength of biasing effects between both task types for two reasons. First, doubling of the contrast between high and low reward {3:3 1:5 0.5:8.5 0:12} in our control experiment did not alter the behavioral findings. Second, our maximal RT benefit of approximately 20 ms in the AMNT task matched the magnitude observed in a previous study with an even higher reward contrast of 20:1 (Staudinger et al., 2011). Rather, the results suggest that however high the reward ratio is, the RT costs and benefits driven by reward amount manipulation do not reach the level of RT costs and benefits observed with probability manipulation.

Our results suggest that preferability leads to drift rate changes, not to baseline changes. First, we should have observed similar patterns of biasing effects in the choice trials and the instructed trials of the AMNT task if biasing effects in the AMNT task were mediated by baseline shifts. Yet CPs were much weaker and RT differences were absent in choice trials of the AMNT task, suggesting an effect that occurs at the earliest at the appearance of the rule-cue, i.e., an effect that is independent of baseline shifts. Our results are therefore more compatible with the idea that in the AMNT task the bias-dependent RT costs and benefits are explained by drift rate adaptations that reflect the final expected reward after integration of the rule-cue. Therefore, in reward-balanced choice trials the subject's initial expectations are neutralized, leading to a lack of bias-dependent costs or benefits. The additional bias-independent fixed offset of RT between choice trials and instructed trials is of the same amount (20-30 ms) as in the PROB task, hence is likely to have the same mechanistic explanation.

Second, in the PROB task we found that RT benefits in instructed and choice trials and CPs in choice trials gradually increased with increasing levels of predictability. None of the three gradual effects was observed with increasing preference in the AMNT task. RT benefits in instructed trials and CPs in choice trials increased more or less in a single step as soon as the preferability was unbalanced, without a further increase with increasing bias degree. These observations contradict the idea that the effect of preferability is just an attenuated form of the effect of predictability.

Taken together, biasing effects due to pure motor-goal preferability are limited compared to predictability and likely restricted to processes following the final rule instruction. Once the final reward value is known after the rule instruction, adaptation of drift rate could reflect motivational effects for the immediately pending action, including shallow drift rates corresponding to demotivation when subjects were instructed to reach low- or non-rewarded targets.

Predictability and movement planning

As we illustrated with the real-world example in the introduction, it is plausible to believe that an above-chance likelihood of later being instructed to aim for a specific goal

encourages movement planning to achieve that goal. This should be the case in the biased trials of our PROB task. In contrast, chance likelihood of either goal renders preliminary movement planning toward one of the two remaining options pointless, even with varying preference for the two alternatives as in our AMNT task. This assumption was supported by our observed movement kinematics. In contrast to experiments requesting button presses (Maddox and Bohil, 1998; e.g. Mulder et al., 2012), our subjects performed extended reaches, allowing for such analysis. It was only in the PROB task, and not the AMNT task, that we found significant increase of MTs and EPVs when subjects had to go against an increasing bias. This suggests countermanding of a preliminary movement plan only in the PROB task. Based on neuronal evidence from motor planning areas, it has been proposed that when monkeys face multiple movement alternatives, the multiple candidate actions are simultaneously reflected in the movement planning activity of sensorimotor cortex preceding choice (Cisek, 2007; Klaes et al., 2011; Lindner et al., 2010; Scherberger and Andersen, 2007); In particular, a non-preferred or unselected action might not be completely suppressed before the chosen action is initiated (Cisek, 2012). Our MT and EPV results showed an effect on motor execution consistent with the idea of subjects having to disengage from a predominating motor plan in favor of a less predominating alternative plan in the PROB task only.

In summary, by dissociating preference-independent action planning (biased PROB trials) from action-independent preference (biased AMNT trials), we were able to link the processes underlying predictability with action planning within the DDM framework. According to this view, our results provide evidence that action planning modulates the migration distance in DDM, while preference modulates drift rate.

Conclusion

Our results suggest different mechanisms underlying biasing effects of *prior* predictability and preferability in decision-making. This finding supports the affordance competition hypothesis (Cisek, 2007); preliminary competitive movement planning in favor of one of two potential equal-valued movement options can induce a graded choice bias and reaction time advantage, while value-based preferences alone do not.

Acknowledgements

We thank Johanna Müller for help in the collection of the AMNT-double dataset, Cliodhna Quigley for feedback and proofreading, Sina Plümer and Klaus Heisig for technical support. This work was supported by the Federal Ministry for Education and Research (BMBF, Germany, grant 01GQ1005C), the German Research Foundation (DFG, grants SFB 889 and Research Unit GA1475-B2).

References

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., and Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review* 105, 442–481.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends in Cognitive Sciences* 11, 118–125. doi:10.1016/j.tics.2006.12.006.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review* 113, 700–765. doi:10.1037/0033-295X.113.4.700.

Bogacz, R., Hu, P. T., Holmes, P. J., and Cohen, J. D. (2010). Do humans produce the speed–accuracy trade-off that maximizes reward rate? *The Quarterly Journal of Experimental Psychology* 63, 863–891. doi:10.1080/17470210903091643.

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cogn Psychol* 57, 153–178. doi:10.1016/j.cogpsych.2007.12.002.

Burk, D., Ingram, J. N., Franklin, D. W., Shadlen, M. N., and Wolpert, D. M. (2014). Motor effort alters changes of mind in sensorimotor decision making. *PLoS ONE* 9, e92681. doi:10.1371/journal.pone.0092681.

Carpenter, R. H., and Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377, 59–62. doi:10.1038/377059a0.

Cavanagh, J. F., Wiecki, T. V., Cohen, M. X., Figueroa, C. M., Samanta, J., Sherman, S. J., et al. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nature Neuroscience* 14, 1462–1467. doi:10.1038/nn.2925.

Cavanaugh, J., Alvarez, B. D., and Wurtz, R. H. (2006). Enhanced performance with brain stimulation: attentional shift or visual cue? *Journal of Neuroscience* 26, 11347–11358. doi:10.1523/JNEUROSCI.2376-06.2006.

Churchland, M. M., Afshar, A., and Shenoy, K. V. (2006). A Central Source of Movement Variability. *Neuron* 52, 1085–1096. doi:10.1016/j.neuron.2006.10.034.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362, 1585–1599. doi:10.1146/annurev.neuro.20.1.25.

Cisek, P. (2012). Making decisions through a distributed consensus. *Current Opinion in Neurobiology* 22, 927–936. doi:10.1016/j.conb.2012.05.007.

Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in Changing Conditions: The Urgency-Gating Model. *Journal of Neuroscience* 29, 11560–11571. doi:10.1523/JNEUROSCI.1844-09.2009.

Coallier, É., and Kalaska, J. F. (2014). Reach target selection in humans using ambiguous decision cues containing variable amounts of conflicting sensory evidence supporting each target choice. *Journal of Neurophysiology* 112, 2916–2938. doi:10.1152/jn.00145.2014.

Coallier, É., Michelet, T., and Kalaska, J. F. (2015). Dorsal premotor cortex: neural correlates of reach target decisions based on a color-location matching rule and conflicting sensory evidence. *Journal of Neurophysiology* 113, 3543–73. doi:10.1152/jn.00166.2014.

Dayan, P., and Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron* 36, 285–298.

Diederich, A., and Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: bound-change, drift-rate-change, or two-stage-processing hypothesis. *Percept Psychophys* 68, 194–207.

Franchina, J. J., and Brown, T. S. (1971). Reward magnitude shift effects in rats with hippocampal lesions. *J Comp Physiol Psychol* 76, 365–370.

Gallivan, J. P., Barton, K. S., Chapman, C. S., Wolpert, D. M., and Randall Flanagan, J. (2015). Action plan co-optimization reveals the parallel encoding of competing reach movements. *Nature Communications* 6, 7428. doi:10.1038/ncomms8428.

Gold, J. I., and Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annu. Rev. Neurosci.* 30, 535–574. doi:10.1146/annurev.neuro.29.051605.113038.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* 520, 220–223. doi:10.1038/nature14066.

Harris, C. M., and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature* 394, 780–784. doi:10.1038/29528.

Hassani, O. K., Cromwell, H. C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *Journal of Neurophysiology* 85, 2477–2489.

Heathcote, A., and Love, J. (2012). Linear deterministic accumulator models of simple choice. *Front Psychol* 3, 292. doi:10.3389/fpsyg.2012.00292.

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front Neurosci* 8, 1–19. doi:10.3389/fnins.2014.00150/abstract.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav* 4, 267–272. doi:10.1901/jeab.1961.4-267.

Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of Neurophysiology* 80, 947–963.

Jahanshahi, M., Brown, R. G., and Marsden, C. D. (1992). Simple and choice reaction time and the use of advance information for motor preparation in Parkinson's disease. *Brain* 115 (Pt 2), 539–564.

Klaes, C., Westendorff, S., Chakrabarti, S., and Gail, A. (2011). Choosing Goals, Not Rules: Deciding among Rule-Based Action Plans. *Neuron* 70, 536–548. doi:10.1016/j.neuron.2011.02.053.

Körding, K. P., and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*

427, 244–247. doi:10.1038/nature02169.

Kurniawan, I. T., Guitart-Masip, M., and Dolan, R. J. (2011). Dopamine and effort-based decision making. *Front Neurosci* 5, 81. doi:10.3389/fnins.2011.00081.

Lau, B., and Glimcher, P. W. (2005). Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys. *J Exp Anal Behav* 84, 555–579. doi:10.1901/jeab.2005.110-04.

Leis, B. C., Rand, M. K., Van Gemmert, A. W. A., Longstaff, M. G., Lou, J. S., and Stelmach, G. E. (2005). Movement precues in planning and execution of aiming movements in Parkinson's disease. *Exp. Neurol.* 194, 393–409. doi:10.1016/j.expneurol.2005.02.014.

Leite, F. P., and Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making* 6, 651–687.

Levy, D. J., and Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology* 22, 1027–1038. doi:10.1016/j.conb.2012.06.001.

Lindner, A., Iyer, A., Kagan, I., and Andersen, R. A. (2010). Human posterior parietal cortex plans where to reach and what to avoid. *Journal of Neuroscience* 30, 11715–11725. doi:10.1523/JNEUROSCI.2849-09.2010.

Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *J Exp Anal Behav* 78, 567–595. doi:10.1901/jeab.2002.78-567.

Maddox, W. T., and Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *J Exp Psychol Learn Mem Cogn* 24, 1459–1482.

Mir, P., Trender-Gerhard, I., Edwards, M. J., Schneider, S. A., Bhatia, K. P., and Jahanshahi, M. (2011). Motivation and movement: the effect of monetary incentive on performance speed. *Exp Brain Res* 209, 551–559. doi:10.1007/s00221-011-2583-5.

Mulder, M. J., and van Maanen, L. (2013). Are accuracy and reaction time affected via different processes? *PLoS ONE* 8, e80222. doi:10.1371/journal.pone.0080222.g001.

Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., and Forstmann, B. U. (2012). Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. *Journal of Neuroscience* 32, 2335–2343. doi:10.1523/JNEUROSCI.4156-11.2012.

Neumann, Von, J., and Morgenstern, O. (1944). Theory of games and economic behavior. Princeton University Press Available at: https://archive.org/stream/theoryofgamesand030098mbp/theoryofgamesand030098mbp_djvu.txt.

Pierce, W. D., and Epling, W. F. (1983). Choice, matching, and human behavior: A review of the literature. *Behav Anal* 6, 57–76.

Posner, M. I., Snyder, C. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J Exp Psychol* 109, 160–174.

Praamstra, P., Meyer, A. S., Cools, A. R., Horstink, M., and Stegeman, D. F. (1996). Movement preparation in Parkinson's disease. Time course and distribution of movement-related potentials in a movement precueing task. *Brain* 119, 1689–1704.

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

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review* 88, 552–572.

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review* 92, 212–225.

Ratcliff, R., Van Zandt, T., and McKoon, G. (1999). Connectionist and diffusion models of

reaction time. *Psychological Review* 106, 261–300.

Rigoux, L., and Guigon, E. (2012). A model of reward- and effort-based optimal decision making and motor control. *PLoS Comput Biol* 8, e1002716. doi:10.1371/journal.pcbi.1002716.

Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience* 22, 9475–9489.

Savage, L. J. (1954). The Foundations of Statistics. Wiley Publications in Statistics Available at: http://books.google.de/books?id=zSv6dBWneMEC&printsec=frontcover&hl=de&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

Scherberger, H., and Andersen, R. A. (2007). Target Selection Signals for Arm Reaching in the Posterior Parietal Cortex. *The Journal of Neuroscience* 27, 2001–2012. doi:10.1523/JNEUROSCI.4274-06.2007.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: Empirical tests of theoretical predictions. *J Exp Psychol Hum Percept Perform* 35, 1865–1897. doi:10.1037/a0016926.

Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences* 27, 161–168. doi:10.1016/j.tins.2004.01.006.

Staudinger, M. R., Erk, S., and Walter, H. (2011). Dorsolateral prefrontal cortex modulates striatal reward encoding during reappraisal of reward anticipation. *Cerebral Cortex* 21, 2578–2588. doi:10.1093/cercor/bhr041.

Stelmach, G. E., Worringham, C. J., and Strand, E. A. (1986). Movement preparation in Parkinson's disease. The use of advance information. *Brain* 109 (Pt 6), 1179–1194.

Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787. doi:10.1126/science.1094765.

Summerfield, C., and Tsetsos, K. (2012). Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms. *Front Neurosci* 6. doi:10.3389/fnins.2012.00070.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience* 7, 907–915. doi:10.1038/nn1309.

Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A., and Körding, K. P. (2012). Differential representations of prior and likelihood uncertainty in the human brain. *Curr. Biol.* 22, 1641–1648. doi:10.1016/j.cub.2012.07.010.

Biased action selection due to imbalanced action preparation.

In this chapter, we investigated the effect of inducing movement planning in a gradual fashion on neural activity in sensorimotor areas, using the *PROB* task in monkeys. While the monkeys were solving the task with reach movements, we recorded neural activities in PMd and PRR to monitor reach-movement planning activities in response to different levels of predictability. As the reach options were rule-based, 90 ° clockwise or counterclockwise to the pre-cue position, we could ensure that the measured neural activities were purely motor-plan related. Specifically, we tested whether a preliminary action plan mediated by the probabilistic pre-cue would also induce choice bias in monkeys, and whether PMd and PRR neurons showed graded representation of both alternative rule-based reach goals reflecting graded levels of subsequent choice bias.

The behavioral results showed that manipulation of the monkeys' *prior* using the biasing pre-cues induced gradual choice bias despite balanced expected values at the time of action selection. The neural data analyses also revealed that the spatial selectivity of individual PMd and PRR neurons during motor planning was modulated according to the degree of later choice bias, i.e. weak spatial selectivity in balanced trials and increasingly stronger selectivity with increasing bias in the animals' behavior.

Furthermore, analyses of relative choice signal latencies around the reaction times during which action plans turned into chosen action showed choice signals emerged earlier in PMd than in PRR. This analysis provided additional results regarding the roles of PMd in decision commitment.

Our finding corroborates the relevance of sensorimotor areas in rule-based motor goal planning and action selection.

Biased action selection due to imbalanced action preparation.

Lalitta Suriya-Arunroj[1] and Alexander Gail[1,2]

[1]Sensorimotor Group, German Primate Center, Göttingen, Germany
[2]Bernstein Center for Computational Neuroscience, Göttingen, Germany

Abstract

According to an emerging view of action-based decision, at least for decisions involving movement responses, decision is regarded as competition between action plans. Neural activity in the sensorimotor system has been shown to represent, before the decision is made, alternative movement goals with the activity levels that reflect subsequent choice preference. However, to our knowledge, no study has so far shown that the effect of imbalanced planning alone, without reward contrast between options, could also induce selection bias.

In the present study, we employed *rule-based* reach task, in which reach goals were not physically marked by visual cues but had be inferred based on abstract rules, e.g. in case of avoiding an object, in order to dissociate the neural encoding of reach goals from that of visual stimuli. Additionally, we induced graded levels of *rule predictability* using probabilistic pre-cues, indicating the likelihood of each alternative rule to be valid at the end of the trial (instructed trials) while we randomly interspersed with choice trials, in which monkeys could freely choose both available rules, delivering reward with equal probability. While the monkeys performed the task, we recorded single unit activities in dorsal premotor cortex (PMd) and parietal reach region (PRR), the frontoparietal network of reach planning. Here, we test whether graded rule predictability induces graded representation of reach goals in PMd and PRR neurons and whether these graded reach planning activities would reflect graded selection bias between two reach goals.

Behavioral results show that graded levels of rule predictability successfully induced graded reaction time benefit and choice bias in monkeys despite balanced expected values at the time of action selection. Neural results show that the spatial selectivity of PMd and

PRR neurons during movement preparation was modulated according to the degree of rule predictability and mirrored the degree of subsequent choice bias. Furthermore, movement-related activities of both areas, PMd leading PRR in most cases, reflected reaction time pattern in response to different predictability levels. The latency result supports the role of PMd in the commitment process of sensorimotor decisions. The effect of imbalanced reach plans on reach selection provides evidence against serial process of decision-making, terminating before the chosen action is planned, while supporting the distributed process of decision-making involving action plans once they are available. Our findings emphasize the relevance of sensorimotor areas in rule-based action planning and decision-making.

Keyword: reach movement, motor planning, bias, *prior* probability, PMd, PRR, sensorimotor areas, rule-based decision-making

Introduction

In everyday situations, we often direct our movement towards targets guided by visual stimuli, i.e. reaching towards a cup, hitting a baseball, or chasing prey (direct movements). However, sometimes the brain has to face a more challenging situation in which motor goals are not physically marked by visual stimuli but have to be inferred from visual input based on abstract rules (inferred/rule-based movements). For example, we are on a ride and suddenly a friendly-looking dog appears on the street. To keep going straight and run over the dog (direct movement) will of course not be our option. Intuitively, we would rather pass either to the right or left (rule-based motor goals) relative to the dog (the visual stimulus). Here, we are interested in the roles of sensorimotor circuits in representing rule-based reach movements as well as in choosing between them.

The frontoparietal circuit for reach movements, comprising reciprocally connected areas in the superior lobe of the posterior parietal cortex (PPC) and the premotor cortex (PMC), is known to contribute to sensorimotor transformation processes (Andersen & Cui 2009; Kurata 1991). Specifically, the parietal reach region (PRR) in the medial intraparietal sulcus of the PPC, and the dorsal premotor cortex (PMd) show sustained motor-goal tuning not only when the impending reach goal is directly cued by visual

stimuli but also when the reach goal has to be inferred from a spatial cue based on spatial transformation rules (Gail & Andersen 2006; Gail et al. 2009; Westendorff et al. 2010; Crammond & Kalaska 1994). Both PMd and PRR neurons are therefore believed to be involved in sensorimotor transformations for reach movements including integration of abstract rules.

When monkeys are faced with two alternatives, neurons within the sensorimotor system also reflect both options during decision-making, with their activities modulated by the value of decision outcomes. After then the multiple options are reduced to the final choice (Coallier et al. 2015; Cisek & Kalaska 2005; Sugrue et al. 2004; Platt & Glimcher 1999; Dorris & Glimcher 2004). This evidence ignited the idea that decision-making processes parallel the sensorimotor transformation in the sensorimotor system (Cisek 2007; Gold & Shadlen 2007). Indeed, PMd and PRR neurons have been shown to simultaneously encode alternative reach goals depending on subjects' preference, even when selection is based on competitive transformation rules (Klaes et al. 2011). The use of abstract transformation rules, which deviate movement directions away from the visual stimulus position, provide privileged access to pure motor planning activities, without the confound of visual cue representation. This latter finding was thus clear evidence confirming that action plans could emerge before the decision termination and reflect subsequent choice preference.

According to previous studies using single-neuron recordings in monkeys, neural correlates of reward expectation, reflecting subsequent choice preference, have been found in sensorimotor areas (Sugrue et al. 2004; Platt & Glimcher 1999; Pastor-Bernier & Cisek 2011). Much less is known about the direct effect of movement planning on choice: whether a preliminary movement plan that overlaps with one of the alternatives would have an influence on subsequent choice. If the decision process implies a competition between motor plans (Cisek 2007; Cisek 2012), imbalanced action planning should influence choice behavior. We therefore investigated whether graded degrees of action plans, without reward asymmetry between options at the decision moment, would induce gradual choice bias in monkeys.

To achieve this, we designed a rule-based reach selection task in which instructed and choice trials were randomly intermixed (Suriya-Arunroj & Gail 2015). We induced *predictability* by using probabilistic pre-cues, announcing the *prior* probability of two alternative reach goals to be later instructed and rewarded. The pre-cues were meaningful only in instructed trials; in choice trials, both options were rewarded with equal probability. In contrast to previous experiments on *priors* (Platt & Glimcher 1999; Yang & Shadlen 2007; Sugrue et al. 2004) and Posner's task (Posner et al. 1980), our instructed trials allowed us to induce graded predictability of subsequent rule instructions, hence graded levels of movement preparation, and with choice trials we could probe for *prior*-dependent choice biases, irrespective of reward expectancy differences at the time of the actual decision (Klaes et al. 2011). In our previous study (Suriya-Arunroj & Gail 2015), human subjects showed graded behavioral bias in terms of both choice and reaction times. By contrasting behavioral bias with another task, providing comparable expected values between options but imbalanced movement plans were discouraged, subjects showed hardly choice bias and no benefit or deficit in choice reaction times.

Here, while monkey subjects performed the reach selection task with probabilistic pre-cueing, we measured activity of PMd and PRR neurons to investigate how rule predictability affects the encoding of reach goals, as reflectance of reach plans, in the sensorimotor cortices and whether this encoding of reach goals reflect monkeys' subsequent choice behavior.

Methods

Rule-based reach selection task

Two male rhesus monkeys (*Macaca mulatta*; Monkeys H and K; aged 9 and 11; weighing 10.34 kg and 9.80 kg) were trained to perform rule-based center-out reaches with sequential cueing (Figure 1). In each trial, a peripheral reach goal was located at one of the four cardinal locations (0°, 90°, 180°, and 270°; eccentricity 8 cm, subtending a visual angle of 12-12.5° at 36.5-38 cm screen distance; technical details below). These locations were not marked by visual stimuli, but had to be inferred from a cue by spatial

transformation rules (rule-based center-out reach). A pre-cue before and a rule-cue after an instructed delay required the monkeys to accumulate information that was sequentially provided (sequential cueing), then to reach to the inferred reach goal location on a blank space (memory guided reach) (as previously implemented in Suriya-Arunroj & Gail 2015).

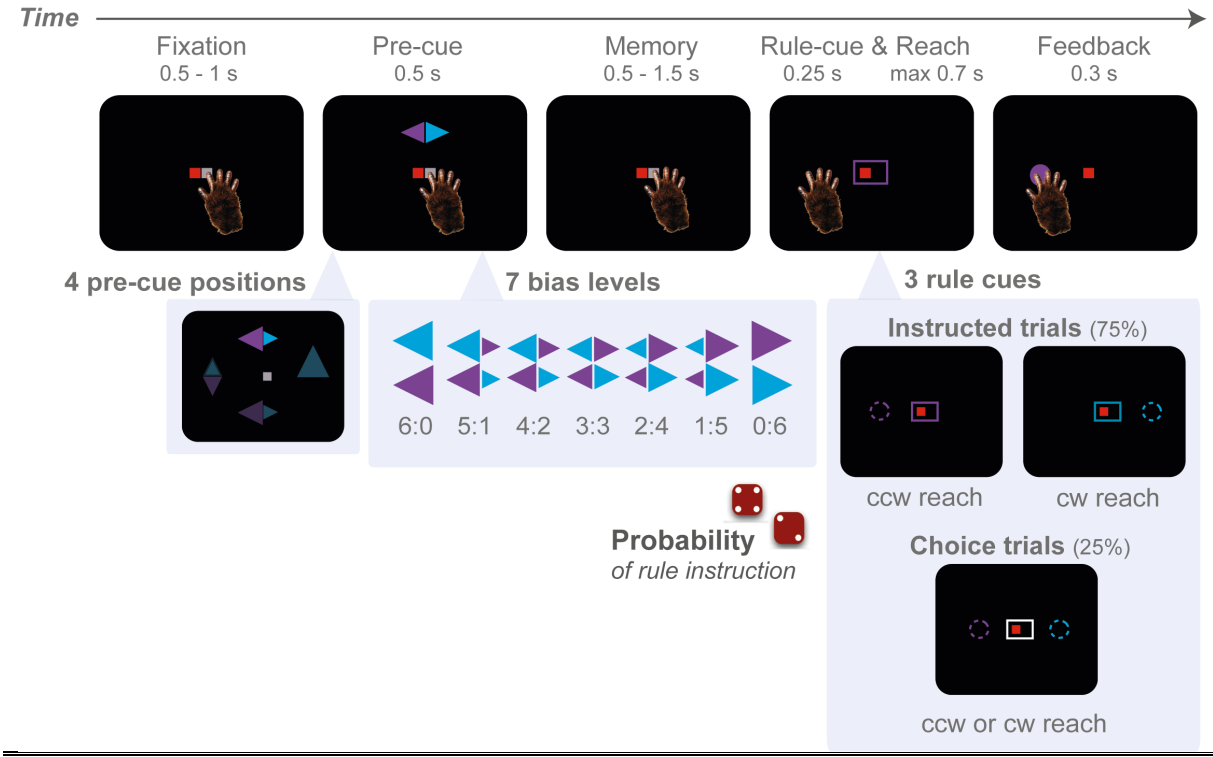


Figure 1: Rule-selection task with probabilistic bias manipulation: a trial starts with eye-hand fixation. Then, a pre-cue appears at one of the four cardinal locations (top, down, left, right), indicating two potential goals and the probability of each goal (big triangle – high probability). After a memory period, the rule-cue appears and either indicates one valid target or leaves the options open so that the monkey can freely choose. Both options in choice trials provide reward with equal probability.

Each trial started with a fixation period. The monkey initiated the trial by directing the gaze to the eye fixation spot (small red square) and touching the hand fixation spot (adjacent white square) at the center of the screen. After a random period of 500-1000 ms fixation (uniform distribution), a spatial pre-cue flashed briefly (500 ms) at one of the four cardinal locations with an eccentricity of 8 cm from the screen center. The pre-cue consisted of two differently colored triangles (magenta and cyan), one pointing to the clockwise (*cw*), the other to the counterclockwise (*ccw*) direction. The two triangles indicated the two possible spatial transformation rules (rule) and hence the only two possible reach targets in a given trial, one at 90° *cw* and one at 90° *ccw* from the pre-cue, at identical eccentricity. The association between rule and color was not constant. Instead, the monkey had to remember the two colored triangles and match them with the later

single-color rule-cue in order to complete the trial correctly. In addition, the two pre-cue triangles could be of different sizes, indicating the likelihood with which the *cw* and *ccw* rules would be instructed later in the rule-cue (further details below). After a random-length memory period (500-1500 ms; uniform distribution), the rule-cue appeared. It consisted of a box framing the fixation points and was either colored (magenta or cyan; instructed trials), or color-neutral (white; choice trials). In the instructed trials, the colored rule-cue narrowed down the two potential reach goals to a single correct goal (*cw* or *ccw*), which was indicated by the triangle in the pre-cue which had the same color as the rule-cue. The reward probability for the instructed target was always 100%, for the invalid target 0%. In the choice trials (25% of the trials), a white rule-cue offered two potential reach goals (*cw* and *ccw*). Both alternatives were valid choices and were rewarded with equal probabilities, independent of the size of the pre-cue and, different to Klaes et al. (2011), independent of the previous choice statistics. Simultaneously with the presentation of the rule-cue, the hand fixation square disappeared ("go" signal) and the monkey had to reach toward the instructed goal (instructed trials) or chosen goal (choice trials) within a maximum of 800 ms and hold the goal position for 300 ms (feedback period).

In case of successful target acquisition and reward, the animal received positive feedback in the form of a circular patch of the same color as the rule-cue at the target position with an encouraging high-pitch tone as well as juice reward. However if he failed, the trial was aborted, a circular feedback patch was displayed briefly at the correct reach position and a demotivating low-pitch tone was played, followed by 5 s timeout. Aborted trials could be due to ocular fixation breaks, incorrect reaches to locations on the screen outside the tolerance window (3 cm) around the valid (invisible) target(s), and reaches later than the response time limit. Each unsuccessful trial was reinserted into the trial sequence randomly.

Monkeys were required to keep ocular fixation throughout the trial. Monkey K managed to perform the task as requested with good performance (87% correct). In contrast, monkey H's performance was never greater than 50% while fixation was imposed. We therefore relaxed the ocular fixation requirement and allowed monkey H to briefly look at the pre-cue during the pre-cue period without the trial being aborted (yielding 88%

performance). However, fixation of the fixation point had to be re-acquired before the pre-cue disappeared (500 ms of pre-cue presentation) and maintained throughout the rest of the trial, particularly during motor planning and execution.

Manipulating the rule predictability

We aimed to induce a graded level of rule predictability in the monkey without a difference in the final value of the two motor-goal alternatives. For this we manipulated the probability with which the *cw* or *ccw* rule would be instructed by using the relative sizes of the pre-cue triangles to announce this probability to the monkey. We chose seven predictability levels corresponding to likelihoods of {6:0, 5:1, 4:2, 3:3, 2:4, 1:5, 0:6} for an instruction of the *ccw* or *cw* rule, respectively. These relative likelihoods were varied in a block-wise fashion: eight successful trials of a given predictability level were required before the next predictability level was randomly chosen. Two out of these eight trials were randomly set to be choice trials and the remaining six trials were instructed with different positions of the pre-cue and different assignments between pre-cue color and rotation rule. For example, in case of predictability level 0:6 towards *cw*, the monkey received the cue instructing *cw* reach in all six instructed trials whereas, in case of a bias level 5:1 towards *ccw*, the monkey was offered five instructed *ccw* trials and one instructed *cw* trial. The monkey always received one reward unit for each correctly performed instructed trial. In choice trials, we assigned a random (50%) chance of receiving more or less ([1.5 vs 0.5] or [2 vs 0] reward units) to both potential reaches. We did not use 100% reward probability for each alternative goal to discourage the monkey from repeating a default choice throughout the session and instead maintaining explorative behavior of both response alternatives. The expected value (reward probability × reward amount) was thus $0.5 \times (1.5/2 + 0.5/2)$ in choice trials with [1.5, 0.5] reward units and $0.5 \times (2/2 + 0/2)$ for [2, 0] reward units, yielding 0.5 units in all choice trials. Thus, at the 100% predictability levels (6:0 or 0:6), the low probability rule had 0% chance of being instructed, hence this rule would offer 0% of getting reward in instructed trials and 50% chance of receiving high or low reward in choice trials, this resulted in an initial (before the rule-cue was shown) expected value of $0 \times 6/8 + 0.5 \times 2/8 = 0.125$ reward units for the low probability rule. The high probability rule provided a 100% chance of delivering one

unit in instructed trials and, again, 50% of getting high or low reward in choice trials, resulting in an initial expected value of $1\times6/8 + 0.5\times2/8 = 0.875$ units. Calculating for the full range of bias conditions, the ratios of initial expected values (EV) associated with the two rules at each predictability level were then {0.875:0.125, 0.75:0.25, 0.625:0.375, 0.5:0.5, 0.375:0.625, 0.25:0.75, 0.125:0.875}. Note that these initial EVs were only valid for the time between the pre-cue and the rule-cue. After the rule-cue, the EV for instructed trials always turned to be one unit for the instructed rule, zero for the non-instructed rule, and 0.5 for both rules in the choice trials.

Pre-recording balancing procedure

Since our experiment aimed at quantifying the biasing effects of *prior* expectations, we wanted the monkeys to explore the range of possible choice responses before the start of the experiment. For this, we ran an initial balancing session in order to discourage the monkeys from repeating the same default reach choices throughout the experimental session. The balancing task contained only trials with the zero-bias condition and differed from the rule-selection task described above only in the reward schedule that we applied on the choice trials. Instead of rewarding both options with random probability, we used the bias-minimizing reward schedule (BMRS). In the BMRS the reward probabilities for choice targets were calculated based on the monkeys' choice history. The less often a target was freely chosen in the previous two choice trials, the higher the reward probability in favor of this target was (Klaes et al. 2011):

$$p(R_{cw})=F(n_{ccw}-n_{cw})$$

$$p(R_{ccw})=F(n_{cw}-n_{ccw})$$

where n_{cw} is the total number of rewarded cw reaches and n_{ccw} is the total number of rewarded ccw reaches. F was defined as:

$$F(x)=\begin{cases}1, & x>1\\ 2/3, & x=1\\ 1/2, & x=0\\ 1/3, & x=-1\\ 0, & x<-1\end{cases}.$$

The balancing task was run until the monkey made at least two *cw* and two *ccw* reaches at each pre-cue position, which means at least 16 choice trials. As choice trials accounted for 33% of all trials in the balancing task, the session then comprised a minimum of 48 successful trials.

Setup and data acquisition

The monkey was seated in a chair facing an LCD screen (19 inch ViewSonic LCD VX922; 5 ms off–on–off response time; screen distance of 36.5-38 cm). The monkey chair had a front plate at the level of the monkey's chest that could be opened in front of his dominant hand to allow reach movements onto the transparent touch sensitive screen (IntelliTouch, ELO Systems, CA, USA), which was directly mounted in front of the LCD screen. The luminance of all instructive stimuli was in the range between 12-13 cd/m^2, measured using a luminance meter (Minolta, LS-100, Japan; measured at eye level with 40-cm distance from the screen, when positioning color cues at top position; 90° from screen center with eccentricity of 9 cm). Throughout each trial, gaze direction was always constrained as long as the central fixation point (red square) was on the screen. For this purpose, the eye positions were recorded by an infrared camera placed above the screen (224 Hz CCD camera; ET-49B; Thomas Recording). Hand position had to be maintained on the central fixation point (white square) before the go-signal, or at the reach target after the go-signal. Both hand fixation and reach target allowed a tolerance window of 3 cm (4.5-4.7° VA) radius. A real-time LabView program running on a PXI computer (National Instruments) was used to control the tasks as well as to register all stimulus properties, event timing, and behavioral responses in each trial.

Behavioral data analysis

We analyzed whether the different levels of rule predictability induced a bias in error rate (ER), reaction time (RT), and choice probability (CP). With bias we mean a difference of any of these behavioral parameters in choice trials depending on whether the actually conducted behavior (the chosen action) matched or did not match the action which initially (by means of the pre-cue) had to be expected to be requested more likely in case of an instructed trial. Note that 'bias' is most meaningfully defined in choice trials, since

only in choice trials the final expected values are equal between both response options, and hence any behavioral differences cannot be accounted for by objective differences between both options. Nevertheless, we apply the 'bias' terminology also to differences of behavioral parameters as a function of rule predictability in instructed trials, since these behavioral differences must be the consequence of differences in the *a priori* cognitive state before the go-cue.

To analyze bias, we merged all the trials with different reach directions (i.e. different cue positions) and resorted the data such that ERs, RTs, and CPs were shown as a function of the absolute value of the pre-cued bias level, independent of the *cw* and *ccw* rules but dependent of whether the reach was conducted towards the same (*follow*) or the opposite (*against*) direction as indicated by the bigger pre-cue triangle (bias direction). This grouping resulted in one neutral 'zero-bias' condition plus three different bias conditions. We quantified bias conditions by the absolute normalized difference between the two initial EVs associated with each bias level [higher EV - lower EV / higher EV + lower EV]. We refer to this difference as bias degree and the four levels were {0 0.25 0.5 0.75}. Note that *follow* and *against* responses were defined only in non-zero bias conditions. The *cw* reaches were arbitrarily considered as *follow* reaches in zero-bias condition for convenience of presentation but without relevance for the conclusions.

RTs were defined as the time between the onset of the go-signal and the release of the hand from the touch screen. RTs were corrected for monitor display and touch screen latencies. Trials with very short RTs were excluded from RT analysis as the monkey might have prematurely released the screen before the rule-cue was perceived but those trials were not aborted online during task execution. As off-line rejection threshold we used the maximum of the following two values: 2.5 times the interquartile ranges below the 25% quartile of the RT distribution within a recording session or 100 ms.

In the error rate (ER) analysis, we counted the percentage of *wrong* selection trials in which the monkey violated the instruction by reaching to the non-instructed alternative of the two potential targets. ERs in instructed trials were split into *follow* and *against* in the same way as in the RT analysis. Wrong trials were not defined in choice trials, as any choice was considered correct. Other types of errors, like errors in response timing,

fixation breaks, or selection of orthogonal goals, were rare and were not analyzed separately.

In the choice probability (CP) analysis, in the three non-zero bias degrees, CPs were defined as the fraction of correct choices following the bias direction.

Additionally, on the basis of the drift diffusion model (DDM), a decision-making model conceptualizing decision process as a gradual accumulation of evidence until the decision variable reaches one of the decision bounds (e.g. Ratcliff 1978; Ratcliff et al. 1999), bias can be explained by a shorter migration distance, either due to a baseline shift (Ratcliff 1985) or a bound shift (Ratcliff 1978), or a change in drift rate (Ratcliff 1981). In case of bias due to baseline shifts, DDM predicts short wrong RTs, i.e. erroneously reach the closer bound when having to reach the correct farther one, whereas it predicts long wrong RTs when the bias is mediated by drift rate changes (Simen et al. 2009; Leite & Ratcliff 2011). We also analyzed the wrong RTs in order to reveal the bias mechanism induced by the predictability manipulation.

With a generalized linear mixed model (GLMM; the '`fitglme`' function implemented in MATLAB R2014b) we quantified influences of bias degrees on ERs and RTs. Full models included the bias degree (*Bias*: continuous variable), the response congruency (*Congruency*: categorical responses; *follow* vs *against* bias reach direction), and the interaction term as fixed effects. Monkey subjects were included as random effects (random slopes for *Bias* and *Congruency*) to account for variance of biasing effects across monkeys. The likelihood of the models including or excluding different variables was compared using the Matlab function '`compare(model1, model2)`'. Our final model was:

$$X \sim Bias * Congruency - Congruency + (Bias{:}Congruency \mid Monkeys), \qquad (M1)$$

to test for differential effects of bias between *follow-against* responses on ERs (binomial response) and RTs with possible interactions between *Bias* and *congruency*.

When there was a significant biasing effect, *post-hoc* tests in the form of t-tests with *Bonferroni* corrections for multiple comparisons, were performed to compare each pair of successive bias degrees.

We tested for a biasing effect on CP using a separate full model without the *Congruency* term as there was no *follow-against* distinction:

CP (binomial response) ~ Bias + (Bias | Monkeys), *(M2)*

followed by *post-hoc* tests for multiple comparisons.

Finally, we analyzed wrong RTs using the model:

*RT ~ Bias*Congruency*Success - Congruency + (Bias:Congruency:Success | Monkeys), (M3)*

followed by post-hoc tests for comparison between all neighboring bias degrees.

Animal preparation and recording procedure

All experiments complied with institutional guidelines on Animal Care and Use of the German Primate Center and with European (Directive 2010/63/EU) and German national law and regulations, and were approved by regional authorities where necessary.

Both monkeys were first implanted with a titanium head holder that was custom-fit to the head based on computer-tomographical surface reconstruction of the skull (3di GmbH, Jena, Germany). After being trained with the task, the monkeys were implanted with two magnetic resonance imaging (MRI) compatible recording chambers above PMd and PRR, in the left hemisphere contralateral to the monkeys' right dominant hand. One chamber was placed above PRR (Horsley Clarke coordinates: -12.50/-10.00 mm (monkey H/monkey K) lateral; -13.50/-18.50 mm anterior), the other chamber above PMd (-19.00/-13.00 mm lateral; 22.00/20.00 mm anterior). Chamber placement was guided by pre-surgical structural MRI and confirmed by postsurgical MRI, showing the implanted chambers relative to the brain and guiding the placement of recording electrodes (Supp. Figure 1). Sustained and direction-selective neural responses during memory-guided center-out reach planning were used as a physiological signature in both areas to confirm the imaging-based positioning (further details on *direction selectivity profiles* below). All surgical and imaging procedures were conducted under general anesthesia and proper analgesia.

Extracellular neural recordings were conducted from up to five microelectrodes simultaneously in each cortical area using a five-channel Microdrive ("mini-matrix"; Thomas recording, Giessen, Germany). In most sessions, simultaneous recordings were conducted in both areas. The raw signal from each electrode was pre-amplified (20×; Thomas recording), bandpass filtered, and amplified (154 Hz to 8.8 kHz; 400-800×; Plexon, Dallas, TX), while being subjected to on-line spike-sorting (Sort Client; Plexon). In addition to spike times, all spikes waveforms were digitized (40 kHz), recorded, and subjected to off-line control of sorting quality and stationarity (Offline Sorter; Plexon).

Neural data selection and directional selectivity profiles

All recorded units with sufficiently good isolation, stability, and activity (firing rates > 5 Hz in any task epochs; see details on analyzed epochs below), regardless of their task-relatedness or specific tuning properties, were included in the neural data analyses unless explicitly stated otherwise. Neuronal responses from successful trials were analyzed.

As known from previous studies, spatiotemporal selectivity profiles ("tuning") of individual neurons in PMd and PRR can change over time from cue related to motor related (Westendorff et al. 2010; Gail et al. 2009; Crammond & Kalaska 1994). Neural selectivity was computed to reveal cue-related, planning-related, and movement-related responses of individual neurons. We computed the selectivity of single-cell spike rates for the direction of the pre-cue/motor goal at different times during the trials. The analysis time windows were: 300 ms before pre-cue onset (baseline), 300 ms after pre-cue onset (pre-cue), the last 300 ms before the rule/go-signal (motor-goal), the time from the rule/go-signal until the reach offset (movement). This basic selectivity analysis included only the data from the extreme bias trials where the definite motor-goal was announced as soon as the pre-cue appeared, independent of cw and ccw rules (Supp. Figure 2).

Directional selectivity was quantified with a directional tuning vector (DTV). The DTV for visual-related (for baseline and pre-cue epochs) and movement-related (for motor-goal and movement epochs) properties was calculated relative to the cue location and reach direction, respectively. The DTV is defined as the vector average across all center-

out cue/reach directions $\vec{u}_i$ (unit vectors) weighted with the corresponding mean spike rates r_{ij} of neuron *j* as follows:

$$DTV_j = \sum_{i=1}^{4} r_{ij}\vec{u}_i.$$

The direction of the DTV defines the preferred direction (PD) of a neuron. The PD can take any value between 0° and 360°, i.e., 0° corresponds to a rightward cue/reach position. Significance of directional tuning was tested with a nonparametric one-way ANOVA (Kruskal–Wallis) with the four different cue/reach directions as groups and sample sizes defined by the number of identical trial repetitions ($p < \alpha = 0.01$).

We characterized the neurons according to their selectivity over the course of the trial. We labeled neurons with significantly tuned cue-related response during the pre-cue presentation as visually tuned neurons (Vis), with tuned movement-planning activity during the memory period, prior to the movement, as motor-goal neurons (Mot), and with tuned activity during the movement as peri-movement neurons (Mov). Each neuron could have multiple labels. We provide the frequency of occurrence of the different neuron types in Figure 4. Any further analyses were dependent on motor-planning activity. Therefore, only motor-goal tuned neurons were used in most of the following analyses, unless indicated otherwise.

Neural population tuning

In order to test whether planning activities on average depended on bias degrees, we computed tuning functions relative to the bias direction independent of pre-cue position and reach direction (Figure 2A). When visualizing the data during the pre-cue and the memory period, we sorted the trials into four conditions, according to four bias degrees. For visualization of the data during the reach period, we split the data into 15 conditions: four instructed *follow* reaches, three instructed *against* reaches (absence of *against* instruction in the extreme bias condition), four chosen *follow* reaches, and four chosen *against* reaches.

We visualized neural population tuning in two ways. First, we aligned the interpolated tuning profiles for each motor-goal neuron relative to the PD in the late memory period

(last 300 ms), normalized to the maximal activity of full-bias trials at the neuron's PD, then averaged the aligned and normalized tuning function across all motor-goal tuned neurons (Klaes et al. 2011) (Figure 6A). Second, we plotted non-interpolated normalized activities of each motor-goal tuned neuron at different bias degrees against rotated bias directions, bias directions relative to each neuron PD (Figure 6A - insets). Note that the population tuning plots were used only for illustrative purposes and that we refer to interpolated firing rates and real PDs, which can take any value between 0° and 360°, only for illustrative purposes. The quantitative analyses were based on the real neural activities restricted to the four discrete directions (up, down, left, right) which we had sampled in our task. The preferred direction of a neuron was then defined by the direction toward which the motor-goal evoked the maximum response, denoted PD_{max}.

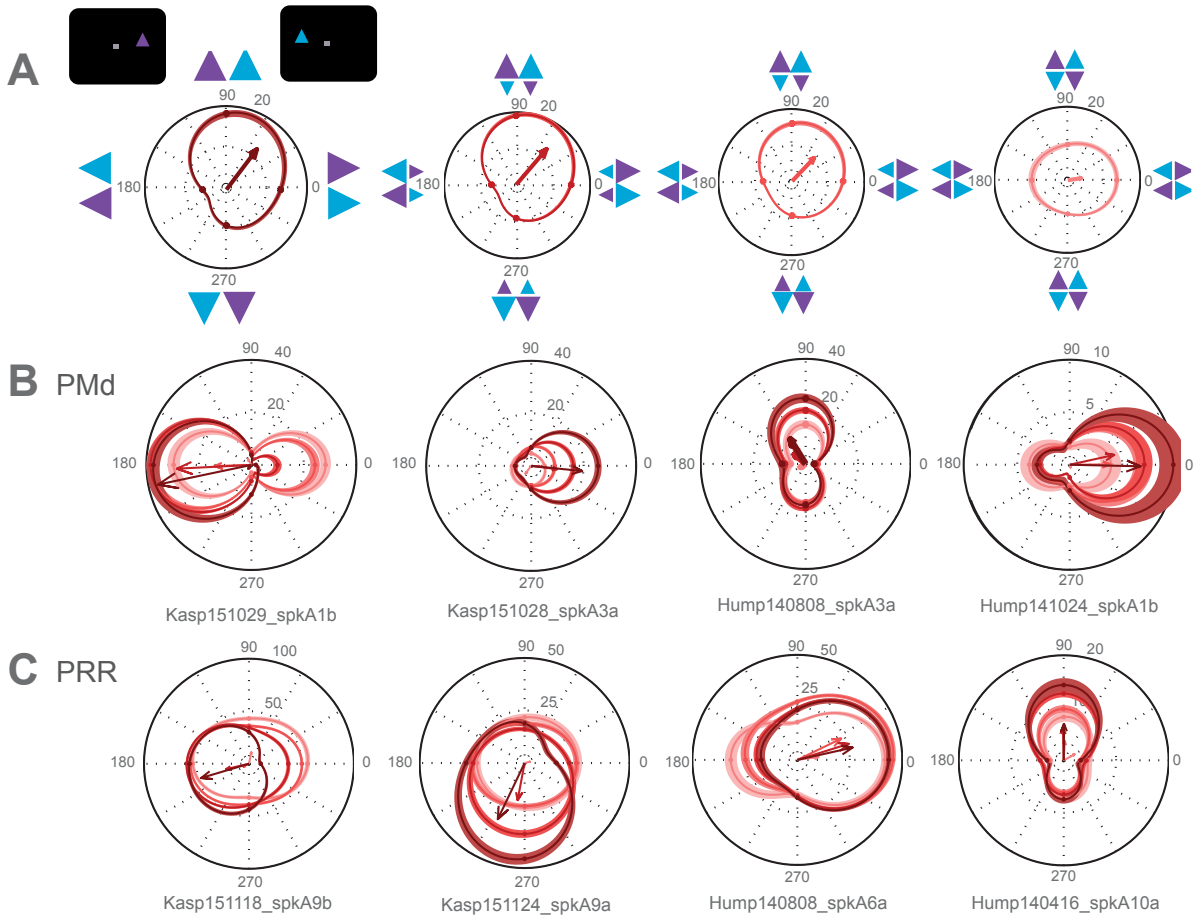


FIGURE 2 | Single neuron examples. (A) Tuning functions relative to biased directions. The biased direction is defined as the direction pointed to by the bigger pre-cue triangle independent of pre-cue location, e.g. for the upward biased direction, the pre-cues can be either on the left (*cw*) or the right (*ccw*) side. In the zero-bias condition, *cw* direction was set to be biased direction. Activities from late memory period of each neuron in trials with the same biased direction were averaged in order to compute the tuning function for all bias conditions. **(B)** PMd and **(C)** PRR example neurons represented as explained in (A).

Time-resolved spiking activities

To illustrate the temporal dynamics of population tuning, we computed spike densities by convolving each spike with a causal kernel defined as:

$$R(t) = \frac{\tau_g - \tau_d}{\tau_d^2} * \left(1 - e^{\frac{-t}{\tau_g}}\right) * e^{\frac{-t}{\tau_d}},$$

where $R(t)$ is the spike density at the time point *t*. The rise time constant τ_g was set to 2 ms, and the decay time constant τ_d was set to 50 ms.

Average spike densities across trials with identical conditions were sampled at 1 ms resolution and aligned to the onset of either pre-cue, go-cue, or reach onset (adapted from Westendorff et al. 2010).

Analysis of graded modulation of motor planning activity

We characterized the graded modulation of motor planning activity in two ways. First, at the population level, we computed the non-normalized mean firing rate of all motor-goal tuned neurons during the last 300 ms of the memory period at PD_{max}, OD, and orthogonal trials. Among the orthogonal trials, we dissociated trials in which the bias direction was towards PD_{max} + $\pi/2$ (plus, Op) and PD_{max} - $\pi/2$ (minus, Om), respectively. In order to assess graded planning activity, we tested whether bias degrees influenced planning activity in a graded fashion by using, first, a GLMM, including the effect of bias degrees (*Bias*), bias directions (*BiasDir*: PD_{max} vs. OD), as well as the interaction term as fixed effects and individual units as random effect as follows,

*Activity ~ Bias * BiasDir + (Bias * BiasDir | Units),* *(M4)*

then t-tests with *Bonferroni* correction for *post-hoc* comparisons.

Second, at the single cell level, we assessed the 'gradedness' of each individual neuron by extracting trial-by-trial normalized firing rates at the PD_{max} and OD. Then we used a linear model,

*Activity ~ Bias * BiasDir,* *(M5)*

to test whether neuronal activity was significantly modulated at the PD_{max} and/or at the OD with bias degrees. Significant positive slope at the PD_{max} indicates neuron's response that increases with increasing bias towards its PD_{max} and significant negative slope at the OD means a neuron decreases its response with increasing bias towards its OD. Next, we quantified the proportion of individual neurons that showed a significant effect in both PD_{max} and OD, only in PD_{max}, only in OD, or no significant modulation and plotted all neurons according to their slope at the PD_{max} and OD. In the plot, neurons that show only significant effect at the PD_{max} would gather along y-axis whereas neurons that show only significant effect at the OD along x-axis. We analyzed the distribution of all neural categories by transforming Cartesian coordinates of each neuron into angular coordinates (0°: OD-, 90°: PD_{max}+) then test for unimodality using Hartigans' dip test (J. A. Hartigan & P. M. Hartigan 1985).

Analysis of neural coactivation

Between each pair of motor-goal tuned neurons recorded in the same session, we computed Pearson correlation coefficients of the planning activities (late memory) across all trials (signal correlations) as a function of the distance between the neurons' PDs. The dependency of signal correlation from PD distance was computed separately for each bias degree. We were interested particularly in the effect of predictability on the signal correlation of neuron pairs having opposite PD_{max}, since these are most informative about co-encoding of two potential motor goals (see Results). We used a GLMM, including *Bias* as a fixed effect and neuron pairs as a random effect as follow,

$$CorrCoef \sim Bias \mid pairs, \qquad (M6)$$

then t-tests (Bonferroni-corrected) for *post-hoc* comparisons.

Analysis of choice predictive activities & receiver operating characteristic (ROC)

We choice trials, we investigated to which extend planning activities during the memory period could predict monkeys' subsequent choices. We gathered all choice trials and, for each single unit, sorted the trials into PD_{max}, Op, OD, and Om trials (identical to *Analysis*

of graded modulation). In addition, we split the conditions into *Reach-in* and *Reach-out* choices, for reaching into versus out of the neurons' PD_{max}, respectively.

First, we tested the difference between *Reach-in* and *Reach-out* choices (*ReachDir*) using GLMM,

*Activity ~ Bias * BiasDir * ReachDir + (Bias * BiasDir * ReachDir | Units),* *(M7)*

followed by *post-hoc* comparisons using *Bonferroni*-corrected t-tests.

Second, we computed an ROC for each neuron at each bias degree. The area under the ROC determines levels of choice predictability, where 0.5 corresponds to chance level, 1.0 to perfect predictability of choice based on neural response differences. We computed the *within-condition* ROC, i.e. *Reach-in* vs. *Reach-out* choices in *Bias-in* and *Bias-out* trials separately, to test how well we could predict *follow* vs. *against* choices at different bias degrees. Also, we computed the *between-condition* ROC, i.e. *Bias-in* vs. *Bias-out* conditions, to test how reliable we could predict *Reach-in* choice in *Bias-in* trials vs. *Reach-out* choice in *Bias-out* trials.

Analysis of neural response latencies and reaction times

Previous studies (Churchland & Shenoy 2007; Johnson et al. 1999; Suminski et al. 2015) that fractions of premotor neurons exhibit complex changes in their activities at the transition from planning to movement epochs. We employed a multivariate Euclidean distance analysis, which is more suited for analysis of complex activities during reach movements. For this analysis, we could include all single units we recorded in each cortical area without selection based on their tuning properties. First, we computed spike densities by convolving each spike train with a Gaussian kernel, for stronger smoothing, as compared to causal kernel, in order to work with neural velocity, of width $\sigma = 50$ ms (Westendorff et al. 2010), as follows

$$R(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \times e^{\frac{-t^2}{2\sigma^2}}.$$

Then, we computed Euclidean distance, without dimensionality reduction (adapted from Ames et al. 2014), between trials in which monkeys reached two opposing reach goals at

each pre-cue location (up, down, left, right) and each condition (all bias degrees, instructed vs. choice, *follow* vs. *against*). We then averaged the neural distances across cue locations to get one distance per condition. We estimated the variability of the neural distance by bootstrapping (N = 10,000 resampled distances per condition). Next, we quantified the neural response latencies by calculating the time bin, after the go-cue onset, in which the distance reached the maximal velocity (MVT: maximal velocity time). The difference in latencies between two cortical areas of interest for each condition was tested by permutation tests; thereby we randomly reassigned each neuron to one of the brain areas such that the number of units in both sets matched the original sample size (N = 10,000 samples). The percentage of random permutations leading to a latency difference larger or equal to the original sample served as the p-value and when the p-value was less than 5%, we considered the latency difference significant.

Similarly, we quantified the peak latencies by calculating the time bin in which the slope became lower than 0.3 (typical slope when the neural distance started to reach its peak or plateau; PT: peak time). We did not use the maximum neural distance because the PRR population showed no peak but a plateau in neural distance during reach movements. The peak latencies between two cortical areas were compared against RTs and the difference was tested by permutation tests as described above.

Results

The main goal of the study was to test whether graded predictability would induce graded reach plans of two rule-based motor goals, reflected by graded neural encoding of these potential motor goals in the frontoparietal sensorimotor network and whether the graded reach goal encoding can explain subsequent monkeys' choices. We first tested whether the pre-cued motor-goal predictability led to behavioral bias based on error rates (ER), reaction times (RT), and choice probabilities (CP), performed by two monkeys in 89 sessions (monkey H, 50 sessions; monkey K, 39 sessions). Then, we tested for predictability-dependent neural modulations in directional reach goal selectivity in PRR and PMd.

Biasing effects of motor-goal predictability on monkeys' behavior

Predictability of the more likely motor-goal instruction enhanced instructed as well as freely chosen reaches towards the biased direction, and weakened opposite reaches.

Error rates: Generalized linear mixed models (GLMM; M1) showed, with increasing bias degrees, increasing difference in ERs between the trials in which the rule-cue instructed to reach against the initial bias (*against* trials) and the trials in which the rule-cue agreed with the initial bias (*follow* trials) (interaction between *Bias* and *Congruency* (*against* - *follow*): t-statistic = 29.59, $p < 0.001$). Additionally, ERs showed a gradual increase in *against* and decrease in *follow* trials along the whole range of bias degrees (*against*: t-statistic = 53.92, $p < 0.001$; *follow*: t-statistic = -21.26, $p < 0.001$; pairwise comparisons shown in Figure 3A).

Reaction times: mixed model analyses (M2) showed a significant differential effect on RTs between *against* and *follow* trials (interaction between *Bias* and *Congruency*: t-statistic = 17.21, $p < 0.001$) and, consistent with ER results, slower RTs when the monkeys were instructed to reach against the bias and faster RTs when the instruction followed the initial bias (increase in RTs (ms) per one bias degree: *against*: estimate ± SE = 30.12 ± 1.25, t-statistic = 24.02, $p < 0.001$; *follow*: estimate ± SE = -46.63 ± 0.61, t-statistic = -76.31, $p < 0.001$; *Bonferroni*-corrected t-tests shown in Figure 3B).

Wrong RTs: GLMM (M3) showed a significant decrease in wrong RTs response to *against* instructions when the monkeys erroneously followed the bias, and an increase in response to *follow* instruction when he erroneously went *against* the bias (interaction between *Bias* and *Congruency* (*against* - *follow*): estimate ± SE = -57.11 ± 1.64, t-statistic = -34.87, $p < 0.001$). The GLMM also confirmed short error RTs ("fast errors") in *against* trials and "slow errors" in *follow* trials (*against* (*correct* - *wrong*): estimate ± SE = 86.34 ± 2.45, t-statistic = 35.17, $p < 0.001$; *follow*: estimate ± SE = -47.97 ± 1.97, t-statistic = -24.40, $p < 0.001$) (Figure 3C).

Choice probabilities (CP): GLMM analysis also indicated a biasing effect on CPs in the choice trials; the probability of choosing *follow* increased with the bias degree (t-statistic = 8.17, $p < 0.001$) and *Bonferroni*-corrected t-tests confirmed graded biasing effects; the choice bias became gradually stronger as a function of bias degree (Figure 3D). With the

strongest bias level monkeys almost exclusively (89 %) chose the biased reach direction, even though there was no objective advantage over the opposite direction in the choice trials.

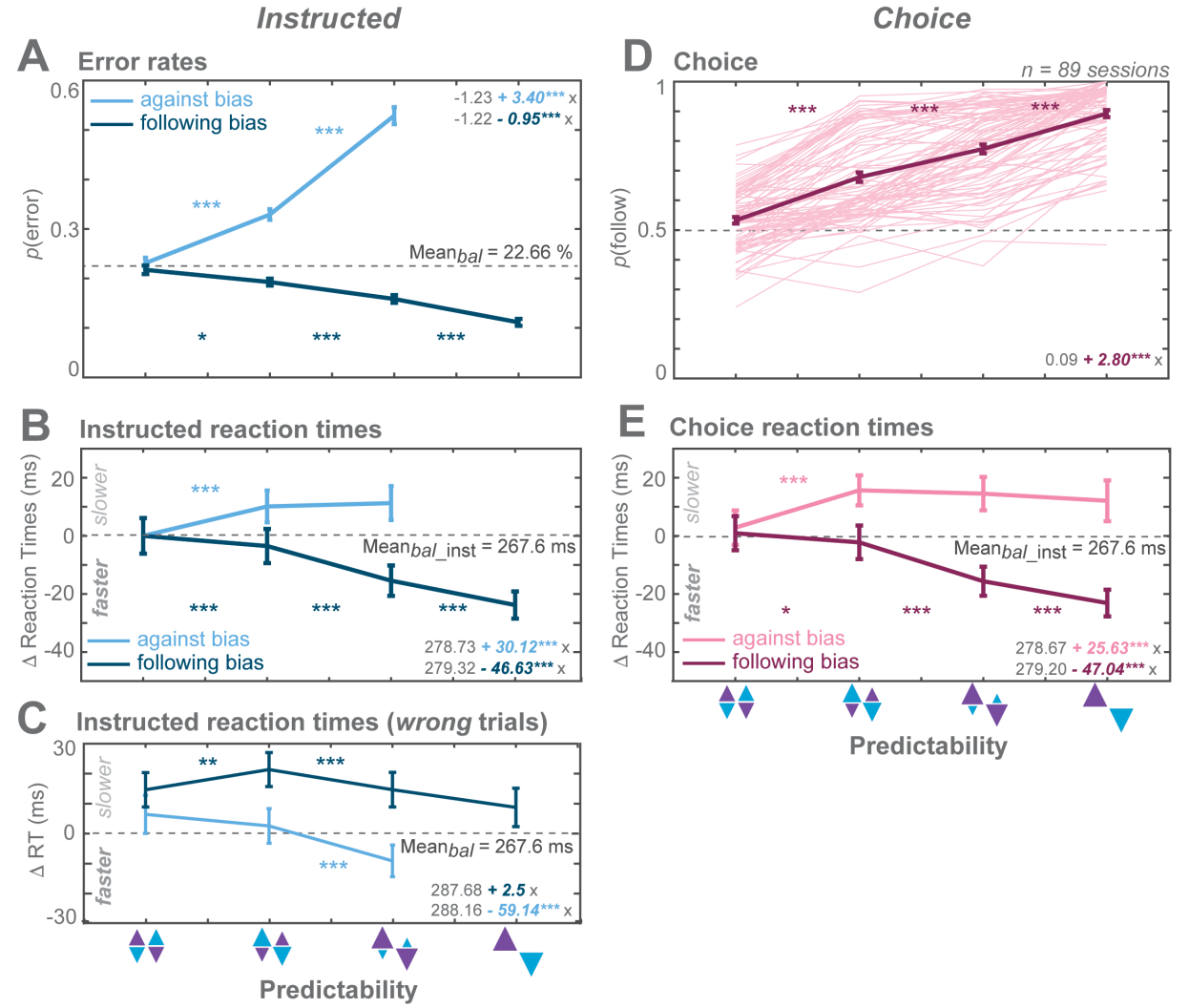


FIGURE 3 | Biasing effects on the monkeys' behavior. (A) Average error rates (ER), **(B)** average reaction times (RT) difference (zero line indicates average RTs in instructed trials of zero-bias condition), **(C)** average wrong RT difference of instructed trial at all bias levels are shown. **(D)** Average choice probabilities (CP) and **(E)** average choice RTs at all bias levels are shown. Dark colors represent *follow* reaches and light colors represent *against* responses. Error bars depict standard errors. Each panel displays fitted equation using GLMM for each trace. (*p < α_{corr} at 5%, **p < α_{corr} at 1% , ***p < α_{corr} at 0.1%; *Bonferroni*-corrected t-test).

Choice RTs also showed differential biasing effect between *follow* and *against* trials (interaction between *Bias* and *Congruency*: estimate ± SE = 70.93 ± 8.55, t-statistic = 8.30, $p < 0.001$) with slight increase in choice RTs when the monkeys chose against the bias and gradual RT advantages with *follow* choices increasing with bias degrees (*follow*: estimate ± SE = -43.12 ± 14.76, t-statistic = -2.92, $p < 0.01$; *against*: estimate ± SE = 25.63 ± 1.75, t-statistic = 14.61, $p < 0.001$). Figure 3E shows *post-hoc* multiple comparisons between neighboring conditions.

In summary, the monkeys performed faster and made fewer errors in instructed *follow* trials compared to instructed *against* trials, mostly in a graded manner with increasing bias degree. They also chose the biased option more frequently, with shorter RTs, when allowed to freely choose. The behavioral results therefore indicate that the pre-cues were effective in biasing the monkeys' behavior in a two-fold way: First, instructions that match the *prior* expectations create a behavioral benefit, and non-matching instructions create costs. The behavioral bias confirms graded levels of preparedness in response to graded predictability, induced by the probabilistic pre-cueing. Second, the bias shows in choice trials despite balanced reward at the moment of decision, i.e., despite lacking objective advantage, thereby justifying the term 'bias'.

Neural motor goal selectivity

We recorded extracellular single-neuron spiking activities from 561 cells in dorsal premotor cortex (PMd) (monkey H, 238; monkey K, 323) and 517 cells in parietal reach region (PRR) (monkey H, 248; monkey K, 269) while the monkeys performed the rule-based reach selection task. We distinguished between (a) visually encoding neurons spatially selective during the pre-cue presentation, (b) motor-goal encoding neurons selective to directions for pending reaches during the late memory period, and (c) movement encoding neurons selective to reach directions during the movement, where a neuron can contribute to multiple categories (see Methods and Figure 4). 53% of PMd (297/561) and 50% of PRR (261/517) neurons fulfilled the criteria for motor-goal encoding and were subjected to all following analyses. All neurons contributed to the latency analysis (see below).

A subset of neurons showed selectivity profiles that evolved over different task epochs. Neural categorization analysis revealed multiple differences among neurons in both areas. For example, while ~50% of neurons in both areas were selective in the motor-goal period, a two-fold larger fraction in PRR (49%) than PMd (26%) also showed selectivity during visual cue presentation. Interestingly, in PMd ~20% of the motor-goal plus movement neurons changed their preferred direction (with a minimal angle difference of 90°) from planning to movement period (Supp. Figure 2). In PRR large changes in PD between motor-goal and movement only occurred in 4% of neurons. More importantly

for the purpose of this study, the majority (~70%) of neurons in both areas that showed selectivity during the visual cue period and the motor-goal period kept their PD constant between both periods. As the visual cue and the motor goal locations were never overlapped in our task, this means that most neurons which were active during pre-cue presentation became inactive during movement planning, and vice versa. Therefore, the visual-to-motor-goal PD consistency confirms that during the motor-goal period, neural activity is motor-goal related, and does not reflect visual memory of the spatial pre-cue (Gail & Andersen 2006; Gail et al. 2009; Westendorff et al. 2010; Klaes et al. 2011).

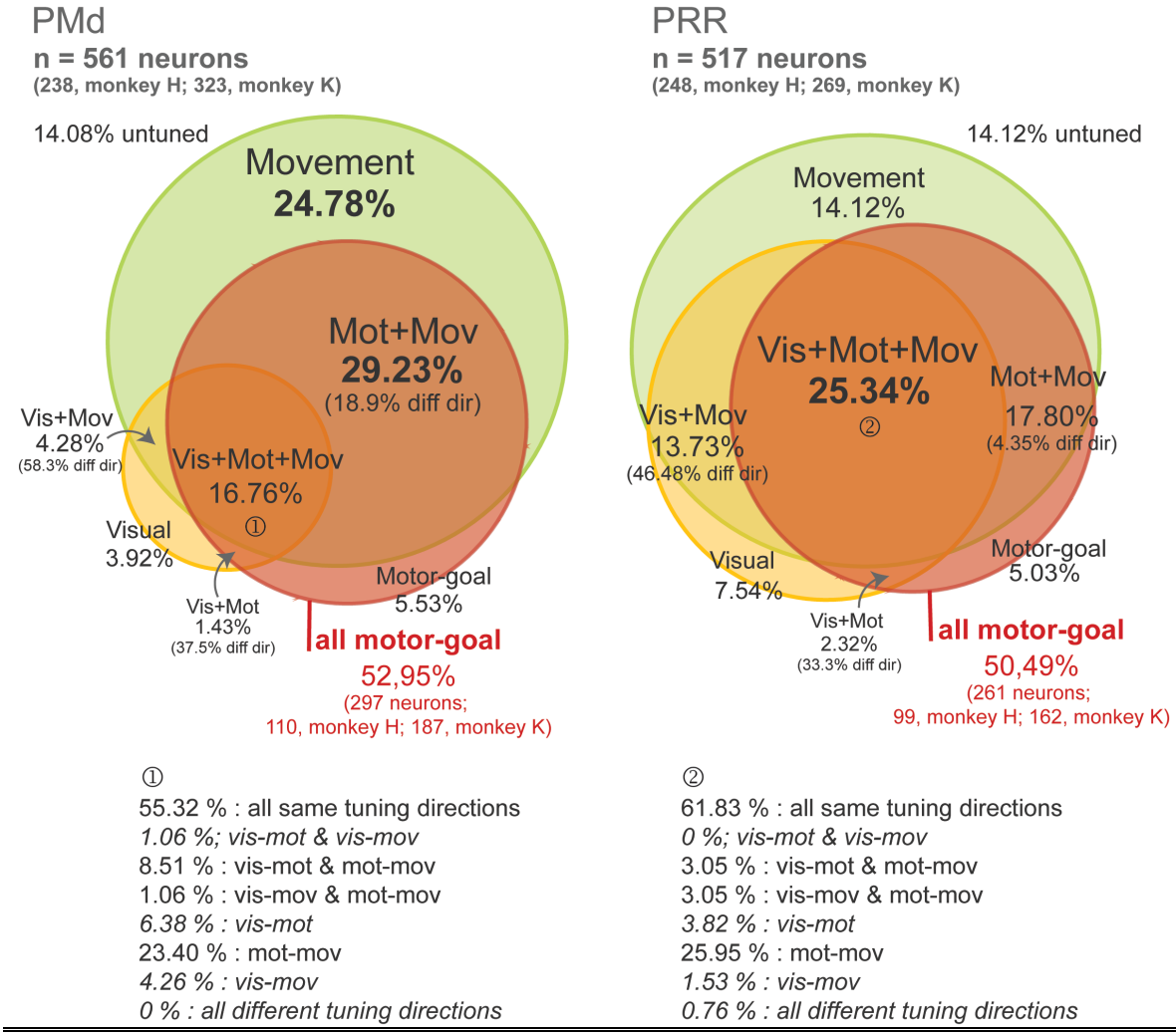


FIGURE 4 | Neural categorization in PMd and PRR. Venn diagrams show percentage of neurons of different types (Vis: visually tuned neurons, Mot: motor-goal tuned neurons, Mov: peri-movement neurons). For neurons that showed multiple tuning properties, fractions of neurons that keep tuning direction, change tuning directions, and some combinations of the possibilities, are shown.

Graded encoding of potential motor goals in PMd and PRR

As has been previously shown (Cisek 2007; Coallier et al. 2015; Klaes et al. 2011), neural selectivity in PMd and PRR is consistent with the idea of encoding potential movement options in ambiguous choice situations. Here we confirm that bimodal choice signals are in fact motor-goal related (characterized by sustained activation despite the absence of visual targets at the motor-goal locations) and provide evidence that potential motor goal encoding is modulated by *prior* probabilities consistent with an induced behavioral bias, and thereby is suited to support the idea of graded encoding of two potential plans simultaneously.

The time-resolved population activities in the current experiment showed initial cue-related activities during the pre-cue presentation, sustained activities during the memory period, and movement-related peaks around the movement time. Especially the average activities prior to movement onset were modulated by the bias level in both areas (Figure 2B-C & Figure 5A). We compared the population response in all bias levels over the course of the trials with several important first observations to be made (for detailed quantitative analyses and further results see *Analyses of graded planning activity* and below).

First, average responses were highest when a pre-cue with highest bias level indicated a likely reach goal towards the neurons' PD (dark red traces; *Bias-in*). Neural responses were intermediate when two competitive reach goal alternatives were equally valid (light-color traces), and lowest when the extreme biasing pre-cue pointed away from the PD (dark blue traces; *Bias-out*). Second, in the zero-bias condition (light-color traces), the average population activity represented two separate rule-based potential motor goals at two opposite directions. This is seen by the fact that during zero-bias motor planning the activity for motor goals towards the PD_{max} and OD were higher than for motor goals towards the orthogonal directions (grey dotted traces) and higher than during baseline. These two patterns (gradation and bimodal encoding at intermediate gradation level) occurred in the same population of neurons with an experimentally controlled trial-by-trial bias. More than our previous results in which two similar scenarios were shown in two separate datasets, in one of which the monkeys showed an accidental bias (Klaes et

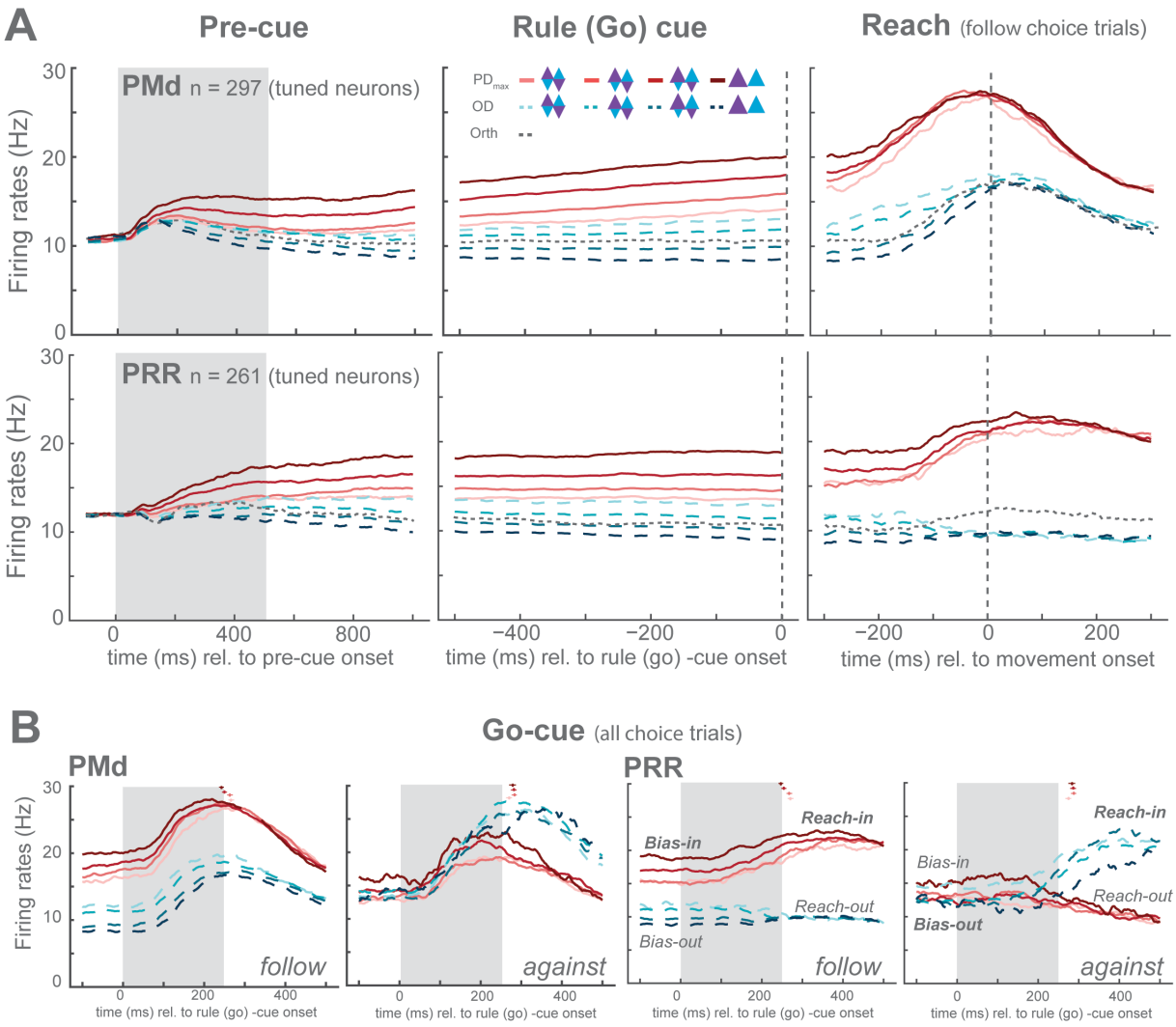


FIGURE 5 | | Population average of spike density functions. (A) Population average of spike density functions throughout trial. The average data at maximum (PD_{max}: red traces), opposite (OD: dashed blue traces) and orthogonal (Orth: dotted grey traces) directions from different bias conditions are presented. Due to multiple traces, we omitted error bars in these plots for visibility. Grey shaded areas correspond to the duration of pre-cue presentation, when the data are aligned to pre-cue onset (left column), and dashed vertical lines correspond to the moment when the rule cue appears (middle column), and when the reach movement is initiated (rigth column). Before the rule-cue appears, there is no physical distinction between *follow* and *against* reaches or instructed and choice trials; the data are plotted jointly before the rule-cue but split into instructed vs. choice and *follow* vs. *against* reaches after the rule-cue. Only the *follow* choice trials are shown. **(B)** Population average of spike density functions of all choice trials aligned to go-cue onset, with average reaction times. The average data at PD_{max} (red traces) and OD (dashed blue traces) from different bias conditions are presented. Small dots above the curves represent averaged reaction times in *follow* and *against* trials of each bias condition and horizontal bars represent standard errors. Grey shaded areas correspond to the duration of go-cue presentation.

al. 2011), these results support encoding of graded motor goal preference in individual neurons in PRR and PMd. Third, during the reach period, the PRR neural population showed elevated activities selectively at the PD_{max} whereas PMd showed high activities in all reach directions, which was probably due to the fact that many PMd neurons changed their tuning directions from planning to reach epoch (Figure 4 and Supp. Figure 2). Forth, in both brain areas, *Reach-in* activities reached similar peak levels around the time when

the movement was initiated despite starting from different levels of planning activities in different bias conditions (Figure 5B). In the context of drift-diffusion model, this finding supports a baseline shift hypothesis, leading to higher choice probabilities and faster reaction times towards the biased option as observed in the behavioral results.

Analyses of graded planning activity

After showing how motor-goal predictability modulates responses during the whole motor planning period, we next analyzed how modulation of planning activity depended on the bias direction and bias degrees, first at the population level, then at the level of individual neurons.

We compared neural activity levels during the last 300 ms of the memory period, for all bias directions and all bias degrees, to assess directional modulation of PMd and PRR neurons as a function of bias degree. Neuronal populations in both brain areas on average showed increasing activities with increasing bias towards neurons' PD (*Bias-in*) and decreasing activities with increasing bias towards the OD (*Bias-out*) (Figure 6A). Confirmed by GLMM, activities in response to bias towards the PD_{max} showed a significant increase that depended on bias degree (PMd: estimate ± SE = 8.11 ± 0.64, t-statistic = 12.63, $p < 0.001$; PRR: estimate ± SE = 7.08 ± 0.55, t-statistic = 12.89, $p < 0.001$) while the OD showed a significantly negative slope (PMd: estimate ± SE = -6.14 ± 0.54, t-statistic = -11.42, $p < 0.001$; PRR: estimate ± SE = -5.09 ± 0.50, t-statistic = -10.09, $p < 0.001$; Figure 6B). No significant slope was detected for the orthogonal directions in PMd ($p > 0.1$). PRR showed slight significant slopes (Op: $p < 0.05$; Om: $p < 0.01$) but no significant graded effect when tested post-hoc. No significant difference between both brain areas was found in this respect ($p > 0.05$). Also, the gradation between bias conditions for both brain areas looks very similar. When the data from each monkey were analyzed separately, subtle differences in non-linear dependencies of planning activities during the memory period from motor-goal predictability remarkably mirrored each monkey's pattern of choice biases (Supp. Figure 3). Notably, the two reach goal directions orthogonal to the two potential motor goals mark fixed points of the directional population tuning, i.e., they are not modulated by motor-goal predictability. Figure 6B confirmed that neural responses to orthogonal reach goals were mostly

unaffected by bias degrees.

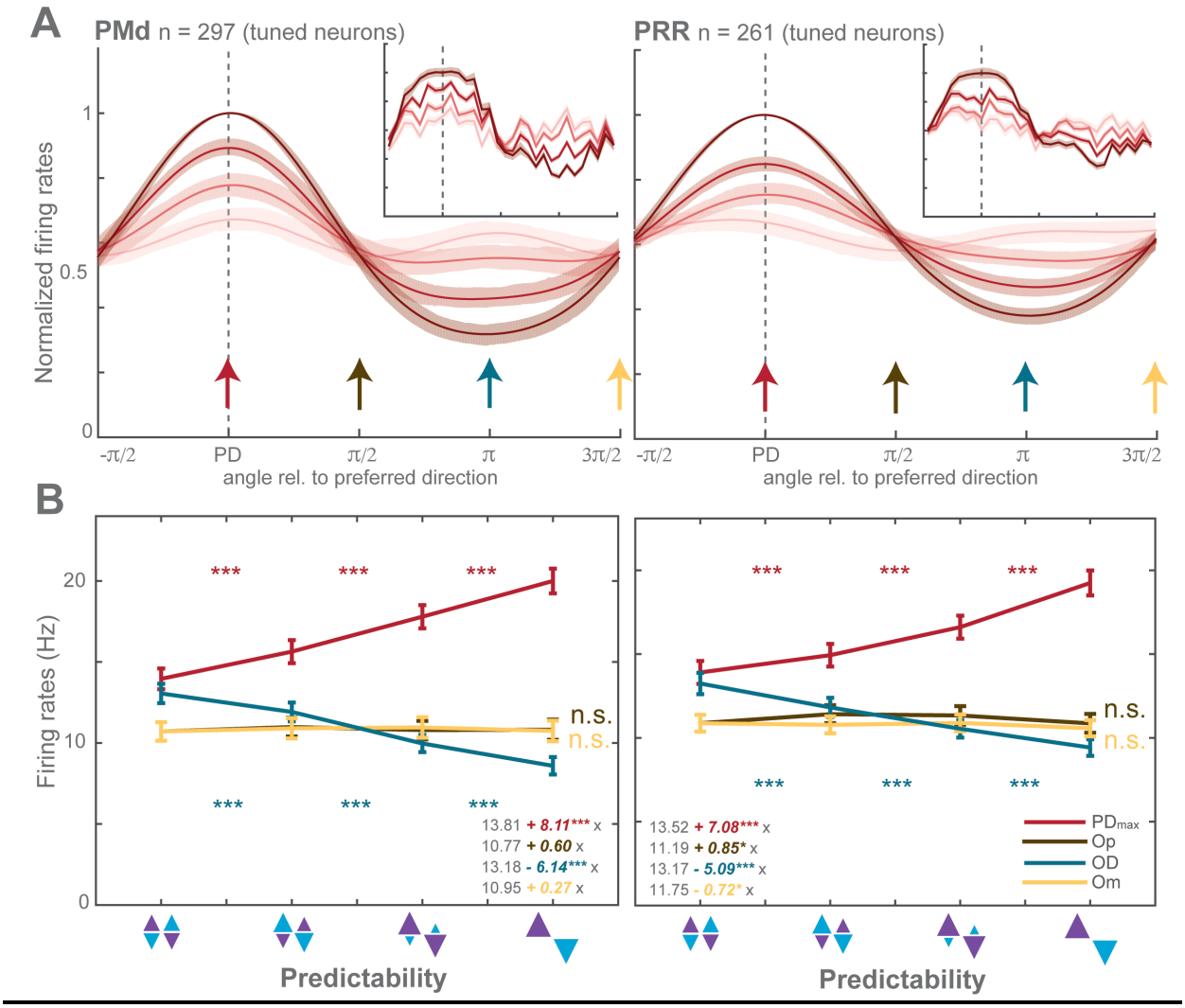


FIGURE 6 | Analyses of graded modulation. (A) Average normalized tuning function during the late memory period (300ms) of all motor-goal tuned PMd (left column) and PRR (right column) neurons. In the insets are shown unsmoothed population tuning (see *Methods*). Shaded areas correspond to standard errors of the mean. Note that vertical arrows are for illustrative purpose; red arrows in (A) refer to PD whereas red traces in (B) refer to PD_{max}. **(B)** Multiple comparison of (non-normalized) average firing rates at maximum direction (PD_{max}; red), opposite direction (OD; blue), and two orthogonal directions (Op (brown) and Om (yellow) for $PD_{max} + \pi/2$ and $PD_{max} - \pi/2$, respectively). Each panel displays fitted equation using GLMM for each trace. Error bars depict standard errors. (*p < α_{corr} at 5%, **p < α_{corr} at 1% , ***p < α_{corr} at 0.1%; *Bonferroni*-corrected t-test)

Individual neurons could contribute in different ways to the population-level results which showed increasing responses when there is a bias towards the PD and decreasing responses for a bias towards the OD, while orthogonal directions were on average unaffected. The bipolar modulation pattern could either be implemented in an equivalent fashion at the level of individual neurons. Alternatively, separate groups of neurons could be responsible for the enhancement when bias direction and PD match, on the one hand, and the reduction when bias direction and OD match, on the other hand. At the single

unit level, we found that the majority of PMd (60%) and PRR (76%) neurons were either up-regulated in *Bias-in* trials or down-regulated in *Bias-out* trials, but not both. Only one fourth of PMd neurons and one tenth of PRR neurons showed graded modulation in both directions.

Figure 7 shows the lack of co-emerging biases at the PD_{max} and OD responses; most neurons coalesced along the PD_{max}+ and OD- axes and not in the upper-right quadrant along the unity line, which would have to be expected in case of correlated up- and down-regulation. The results strongly suggest a separate down-regulation of the explicitly disregarded motor-goal option (as opposed to unaffected neutral positions) and up-regulation of the preferable motor-goal location by barely overlapping groups of neurons.

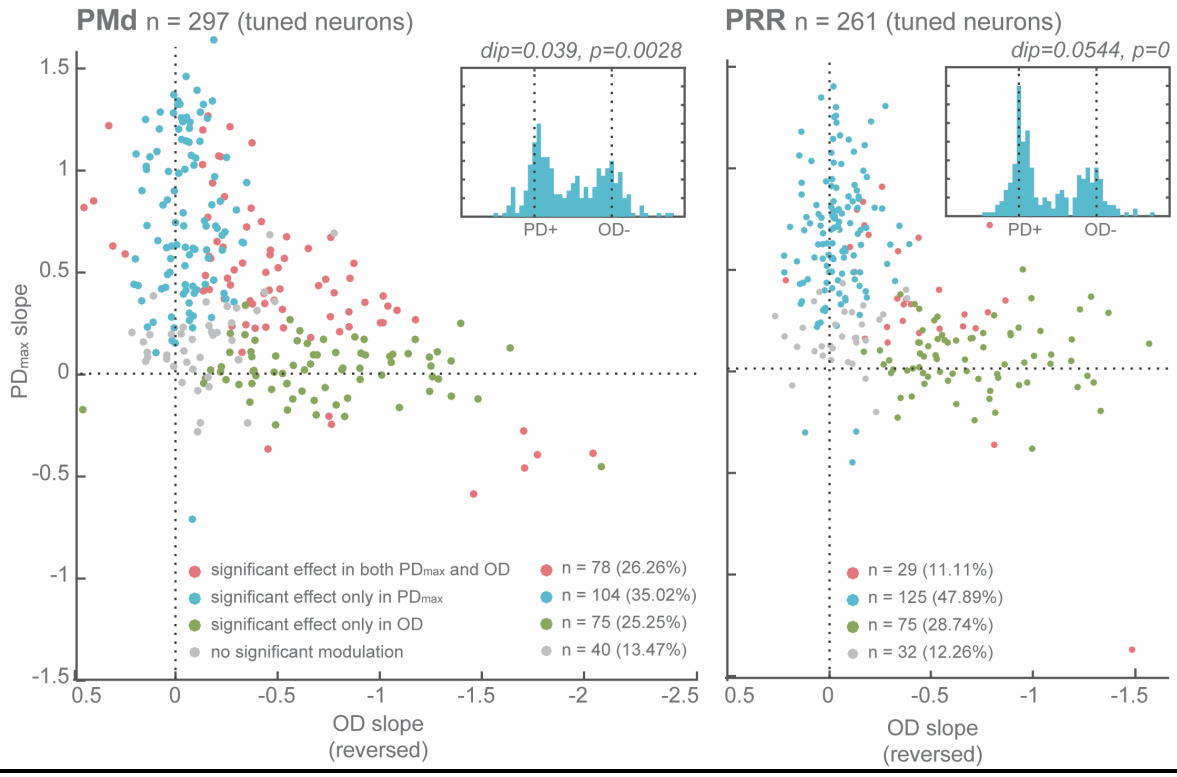


FIGURE 7 | Neural gradedness: slopes of activity modulation with bias conditions in the maximum direction (PD_{max}) and the opposite direction (OD) of PMd (left) and PRR (right) neurons. The upper-right quadrant reflects increasing activities (positive slope) of PDmax responses and decreasing activities (negative slope; inversed x-axis) of OD responses. Red data points: neurons that showed significant modulation of both PDmax and OD responses with bias conditions; blue: significant bias-dependent modulation of PDmax responses; green: significant modulation of OD responses; grey: no significant modulation. The insets show the angular distribution of all neurons in each area, with the Hartigans' dip test results.

Analysis of neural co-activation

In principle, it was possible that the monkeys in each trial, even in zero-bias trials, prematurely generated an unambiguous movement plan towards one of the two potential reach goals soon after having seen the pre-cue and corrected the plan if needed once the final instruction was provided with the go-cue. If, additionally, they alternated between opposing premature plans, then on average neural directional selectivity would deceptively result in graded levels of activities and bimodal tuning during ambiguous movement preparation. Therefore, for demonstrating proper dual encoding of two potential motor goals simultaneously, it needs to be ruled out that the bimodal population tuning in the zero-bias conditions (Figure 6A) is an effect of averaging neural responses over trials with randomly alternating bias of the animal towards only one of the two potential motor goals in each trial. To achieve this, we sought evidence for a trial-by-trial co-encoding of two opposing reach directions during planning by analyzing response correlations in oppositely tuned neurons. We computed trial-by-trial signal correlation of all pairs of motor-goal tuned neurons recorded simultaneously in the same experimental session. In case of alternating premature plans, we expected to observe that neuron pairs with opposite PD_{max} should show negative signal correlation during motor planning throughout all bias conditions. This is because the neurons are oppositely tuned and whenever the animal has an unambiguous plan towards only one direction, only one neuron can be active and the other inactive, i.e. they are anti-correlated. If, however, the monkeys planned both alternative motor goals, then in trials in which the two plans overlapped with the PDs of the two neurons these should be active, and in trial with potential motor goal orthogonal to the PDs they should be both less active. Hence, the signal correlation should become less negative or even positive in the zero-bias condition, since in this case correlation marks a mixture of expected weak signal correlations and unknown (positive or negative) noise correlations induced by the trial-to-trial direction-independent neuronal response fluctuations. Some pairs of opposing neurons in PMd and PRR did show positive correlation in low bias conditions (Figure 8A), but not in biased conditions, consistent with the idea of proper co-encoding of two potential motor goals. Example pairs of neurons are shown in Figure 8B. On average over all neurons, GLMM results confirmed significant decrease in signal correlation between opposing neurons

with decreasing bias degree (PMd: estimate ± SE = -0.22 ± 0.02, t-statistic = -11.36, p < 0.001; PRR: estimate ± SE = -0.14 ± 0.03, t-statistic = -5.12, p < 0.001). PMd showed slightly stronger decrease than PRR (interaction between *Bias* and brain areas (PMd - PRR): estimate ± SE = -0.08 ± 0.03, t-statistic = -2.25, p < 0.05; Figure 8C).

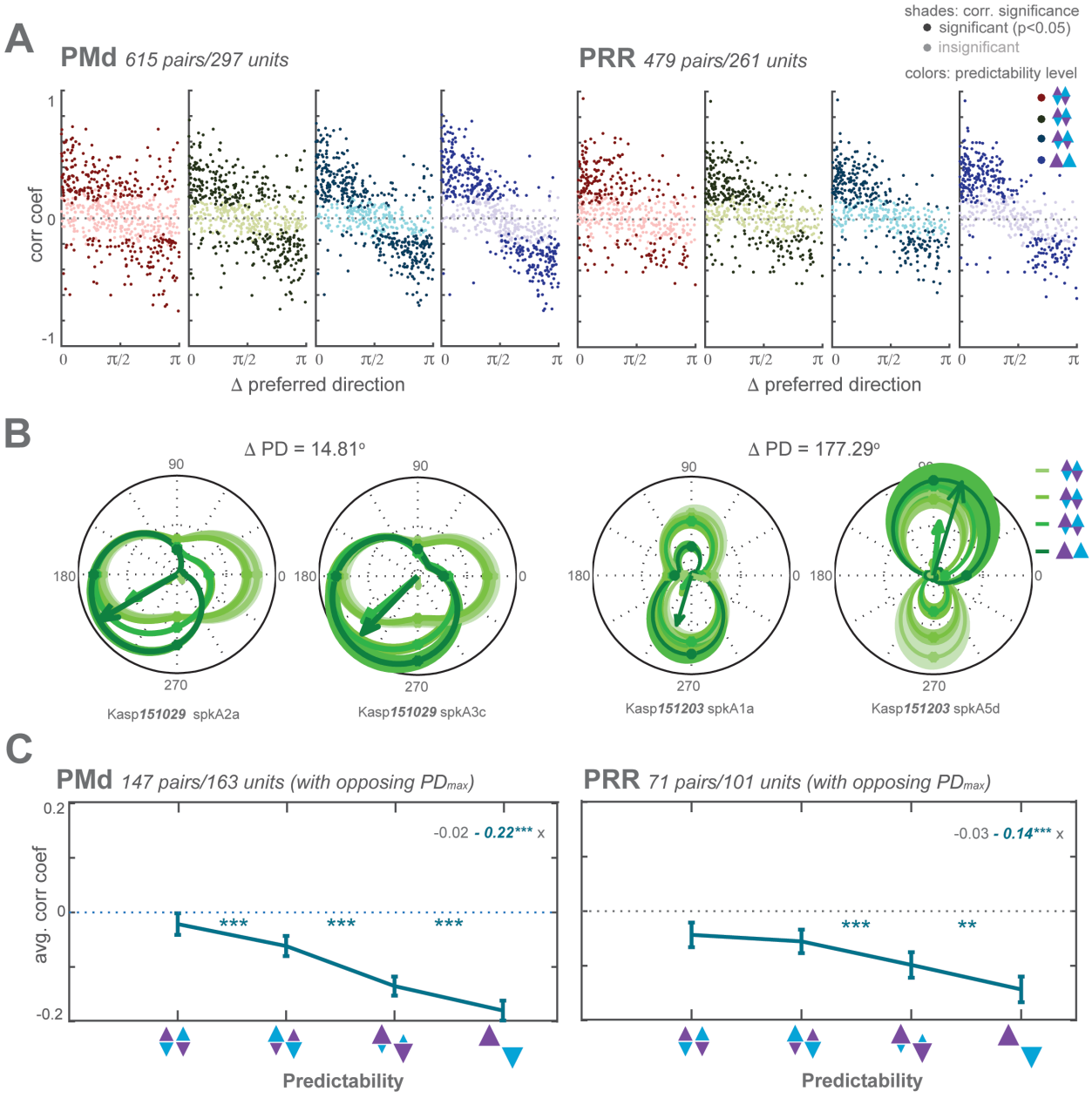


FIGURE 8 | Neural coactivation analysis. (A) Trial-by-trial signal correlation between each pair of PMd (left) and PRR (right) neurons recorded in the same experimental session, plotted against difference in PD between neurons at each bias degrees (low - high bias degrees: red - green - blue - purple data points). Dark color points depict significantly positive or negative correlation. Some neurons with large distance in PD (right half of each plot) showed positive signal correlation in zero-bias condition (dark red dots on the right half above the zero-line). **(B)** Examples of neuron pairs with small (left) and large (right) difference in PDs (low - high bias degrees: light - dark traces). **(C)** Multiple comparison of average signal correlation in all bias degrees. Each panel displays fitted equation using GLMM for each trace. Error bars depict standard errors. (*p < α_{corr} at 5%, **p < α_{corr} at 1% , ***p < α_{corr} at 0.1%; *Bonferroni*-corrected t-test)

Choice predictive activities & ROC

In the conditions with high motor goal predictability, the average activities in OD trials were even lower than the activities at the orthogonal directions. In the context of drift-diffusion models one might wonder how such low level of baseline activity prior to the demanded choice commitment could allow the observed *against* choice behavior, which the monkeys showed more often (Figure 3D) than they selected orthogonal directions (< 0.1%), which were not parts of the valid options.

We investigated which cognitive state led to *against* choices by analyzing the planning activities in only the choice trials. We sorted the trials according to subsequent choice: *Reach-in* (reach direction matched neurons' PD_{max}) vs. *Reach-out* (reach direction away from the PD_{max}). Note that the combinations of *Bias-in* followed by *Reach-in* and *Bias-out* followed by *Reach-out* were *follow* trials whereas the combinations of *Bias-in* followed by *Reach-out* and *Bias-out* followed by *Reach-in* constituted *against* trials. In *follow* trials, which formed the majority of the choice trials, GLMM results confirmed positive biasing effect on neural activities in case of *Bias-in* as observed in previously in the grand average results (*Bias-in* & *Reach-in*: PMd: estimate ± SE = 8.25 ± 0.64, t-statistic = 12.92, $p < 0.001$; PRR: estimate ± SE = 7.11 ± 0.56, t-statistic = 12.76, $p < 0.001$) and the opposite effect in case of *Bias-out* (*Bias-out* & *Reach-out*: PMd: estimate ± SE = -6.40 ± 0.54, t-statistic = -11.84, $p < 0.001$; PRR: estimate ± SE = -4.98 ± 0.49, t-statistic = -10.13, $p < 0.001$). Importantly, prior to *against* choice, the biasing effect declined but remained significant (Interaction between *Bias* and *Congruency* in *against* trials (*Bias-in* – *Bias-out*): PMd: estimate ± SE = 4.05 ± 0.69, t-statistic = 5.89, $p < 0.001$; PRR: estimate ± SE = 3.32 ± 0.81, t-statistic = 4.11, $p < 0.001$) and did not reverse, i.e. the biasing effect in *Bias-in* trials were still positive and *Bias-out* trials still negative [*Bias-in* & *Reach-out*: PMd: estimate ± SE = 4.70 ± 0.57, t-statistic = 8.18, $p < 0.001$; PRR: estimate ± SE = 2.81 ± 0.46, t-statistic = 6.14, $p < 0.001$; *Bias-out* & *Reach-in*: PMd: estimate ± SE = -3.10 ± 0.45, t-statistic = -6.87, $p < 0.001$; PRR: estimate ± SE = -4.03 ± 0.73, t-statistic = -5.52, $p < 0.001$] (Figure 9A).

Furthermore, we tested the discriminability of *against* and *follow* choices by assessing the area under the ROC for each unit separately in Bias-in and Bias-out trials (ROC within condition; Reach-in vs. Reach-out discrimination) based on trial-by-trial activity levels

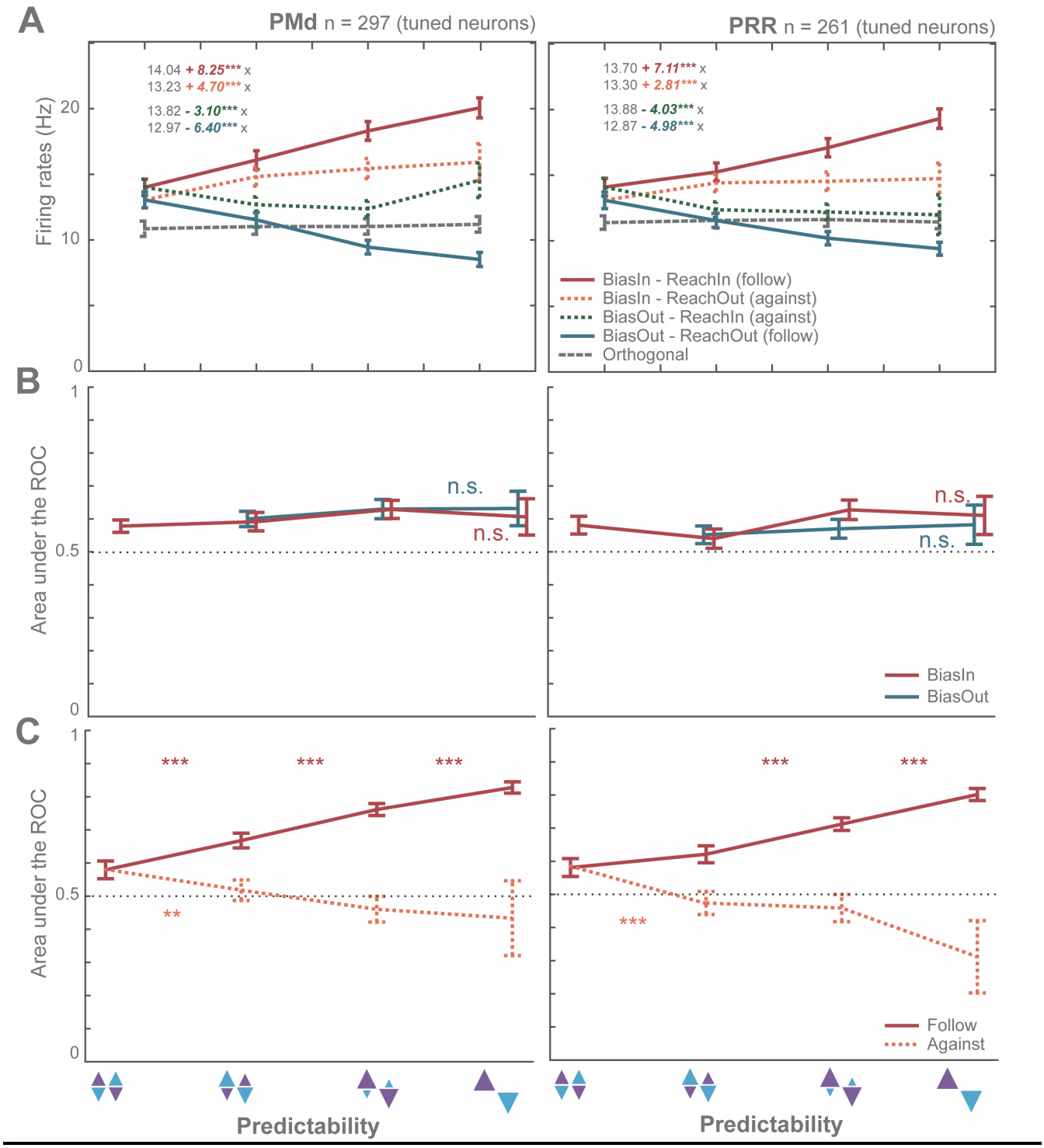


FIGURE 9 | Analyses of choice predictive response. **(A)** Multiple comparison of average firing rates at PD_{max}, split into *Reach-in* and *Reach-out* (*Bias-in* & *Reach-in* (*follow*): solid red; *Bias-in* & *Reach-out* (*against*): dashed orange), OD (*Bias-out* & *Reach-out* (*follow*): solid blue; *Bias-out* & *Reach-in* (*against*): dashed green), and orthogonal directions (dashed grey). Each panel displays fitted equation using GLMM for each trace. **(B)** Average area under the ROC curves discriminating *follow* and *against* trials in all bias levels (*within-condition* ROC). **(C)** Average area under the ROC curves discriminating *Bias-in* and *Bias-out* trials (*between-condition* ROC). Error bars depict standard errors. (*p < α_{corr} at 5%, **p < α_{corr} at 1% , ***p < α_{corr} at 0.1%; *Bonferroni*-corrected t-test).

during late motor planning prior to the go-cue. The average discriminability stagnated around 0.6 across bias levels with insignificant biasing effect in both areas ($p > 0.9$; GLMM; Figure 9B). On the other hand, when we tested ROC separately in *against* and

follow trials (ROC between conditions; *Bias-in* vs. *Bias-out* discrimination), we retrieved the expected ROC discriminability being low between both options in zero-bias conditions and increasing with increasing bias degree (Figure 9C).

Neural choice signal latencies

Receiving the color-neutral rule-cue, the animals were signaled to make a choice immediately. We investigated whether PMd and/or PRR neurons showed movement-related activities that signaled the chosen option prior to the overt movement initiation. Individual neurons which did not reach predefined significance criteria could still have contributed to a choice signal at the population level. To not reject these contributions, we used a multivariate approach including all recorded neurons. The Euclidean distances of the trajectories in the neural state space between opposite reach goals of each bias degree and reach congruency is shown in Figure 10A-B. The neural distance measure looks overall similar to the population activity levels (Figure 5B) but has the additional advantage for the analysis of the movement period that all neural distance traces reached clear movement-related peaks useful to assess peak latencies (which was not the case for the difference between Reach-in and Reach-out activity levels in PMd spike densities; data not shown).

Neural response latencies were measured as the time at which the neural distance curves reached their maximal velocity or the peak (see Methods). The maximal velocity latencies were faster in PMd than in PRR (latencies PMd < PRR; negative differences) in all cases, confirmed by permutation tests (Figure 10C-left). Additionally, the neural peak latencies, the time at which the neural distance started to level out, of PMd also preceded those of PRR (Figure 10C-right).

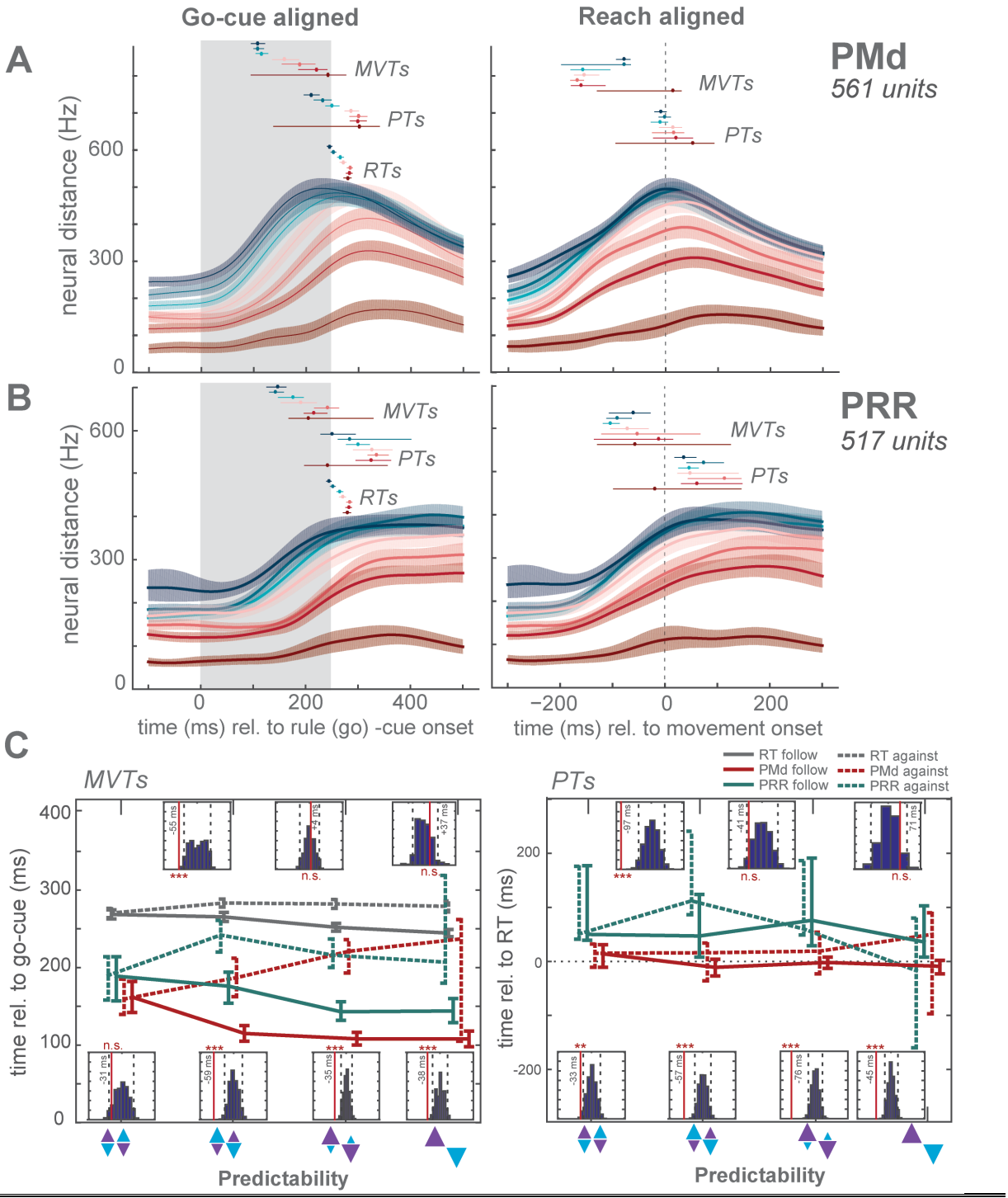


FIGURE 10 | Analyses of choice signal latencies. (A) Neural distance of PMd and **(B)** Neural distance of PRR aligned to go-cue onset (left) and movement onset (right) at different bias degrees (shades) and split into *follow* (blue) – *against* (red) trials. Only choice trials are shown. Reaction times (RTs), peak times (PTs), and maximal velocity times (MVTs) are shown above the curves with corresponding colors. Shaded areas represent go-cue duration when the neural distances are aligned to go-cue and vertical dashed lines represent the movement-onset time when the neural distances are aligned to the movement onset. **(C)** Maximal velocity times (MVTs; left) and peak times (PTs; right) of PMd and PRR neural distances relative to go-cue and reach onset, respectively. RTs are also shown with MVTs. Histograms depict permutation test results for PMd and PRR latency difference at each bias degree, positions according to bias degrees on the x-axis (bottom – *follow* trials, top – *against* trials). Each histogram is centered at zero. The red lines depict the actual latencies between PMd and PRR (PMd - PRR) at each condition. Pairs of dashed lines in each histogram represent the interval covering 95% of the permutation distribution. (*p < α_{corr} at 5%, **p < α_{corr} at 1% , ***p < α_{corr} at 0.1%; permutation test).

Discussion

In response to a parametric manipulation of *prior* motor-goal probability in a rule-based action selection task, monkey behavior showed decreased error rates (ER), faster reaction times (RT), and increased choice probabilities (CP) in a graded manner in favor of the biased reach option. The opposite was true for reaches against the biased option. Behavioral costs and benefits of *prior* predictability occurred despite both motor-goal options providing equal reward at the time of commitment to the choice and they occurred in instructed as well as choice trials. In parallel with behavioral indications of the bias, neural activities in PMd and PRR showed graded directional selectivity, mirroring gradual choice bias between two rule-based reach alternatives at the level of the neural population. At the level of single units, most PMd and PRR neurons showed bias-dependent modulation at either the preferred direction or the opposite direction, but rarely both. In addition, the graded choice signal latencies in PMd always preceded latencies in PRR and depended on motor goal predictability similar to RTs. The results suggest that *prior* motor-goal probability can induce graded levels of motor planning activities and that imbalanced planning can subsequently drive choice bias in reward neutral decisions. Our results are consistent with the idea of shifting the baseline in drift-diffusion models (DDM) in sensorimotor cortices as a neural mechanism of *prior*.

Encoding of potential motor-goals in cortical sensorimotor areas

In line with previous studies (Cisek 2007; Klaes et al. 2011; Sugrue et al. 2005; Platt & Glimcher 1999; Yang & Shadlen 2007; Lindner et al. 2010), we observed parallel activity of two alternative spatially selective neural populations in the fronto-parietal reach areas PRR and PMd when the monkey was provided with two potential reach options. We ensured that the graded neural activities reflect genuine motor planning information due to three central features of our experimental design: First, we spatially dissociated the memory of the visual cue from the motor-goal by introducing the *cw* and *ccw* rotation rules, analogous to memory-guided anti-reaches (Crammond & Kalaska 1994; Gail & Andersen 2006; Gail et al. 2009; Westendorff et al. 2010) and anti-saccades (Munoz & Everling 2004; Zhang & Barash 2004; Gottlieb & Goldberg 1999). We thereby ruled out graded visual memory encoding of target stimuli with variable stimulus value or salience.

Second, we combined a majority of instructed trials with randomly interleaved choice trials to encourage sustained preliminary planning of two potential actions (Klaes et al. 2011; Klaes et al. 2012). We thereby avoided immediate pre-mature choices of the animal, which would likely have occurred if we only had memory-guided choice trials. In addition, by sequentially providing parts of the information, we avoided dealing with lack of planning, which could have occurred if we provided all the information only once before the decision time. Third, in contrast to most other binary choice studies, we sampled the neurons' spatial selectivity profiles not just for trials with each neuron's PD coinciding with the two potential motor goal positions, but also two orthogonal positions (four directions total). This allowed us to quantify the bimodality of the neural population responses as a contrast between potential motor goal locations and neutral locations (Cisek & Kalaska 2005; Klaes et al. 2011). We could thereby show that the observed responses in the zero-bias trials were spatially selective for the two potential motor goals rather just marking a spatially unspecific response enhancement. In combination, the three measures ensured that the observed spatially selective neural responses reflected spatial properties of the preliminarily planned action during the instructed delay between pre-cue and final instruction, i.e. that we observed encoding of potential motor-goals.

Graded predictability induces graded levels of behavioral bias and action plans

With our task design we successfully controlled the monkeys *prior* expectation about the valid rule, and hence the action likely to be taken, in a graded fashion trial by trial. Differently to previous studies on potential reach goal encoding (Cisek & Kalaska 2005; Klaes et al. 2011), the probability of the later *cw*/*ccw*-rule instructions was reliably announced by a pre-cue. We previously found that, in humans, this leads to a trial-by-trial manipulation of subjects' *prior* expectations on which rule to apply and hence the action to be taken (Suriya-Arunroj & Gail 2015). Such manipulation of *prior* rule expectation in a rule-based decision-making task is equivalent to *prior* expectations on the appearance of a stimulus in perceptual decision-making (Mulder et al. 2012; Leite & Ratcliff 2011; Maddox & Bohil 1998; Simen et al. 2009). We confirmed the effectiveness of the *prior* manipulation in our monkey by showing graded *prior*-dependent ERs and RTs in instructed trials and, most importantly, *prior*-dependent CPs plus RTs in choice trials. In our task design, the *prior* information was not essential for human and monkey subjects to

successfully complete the trials; they received a clear instruction at the end of each instructed trials and they received equal reward for both options in all choice trials. The observed behavioral bias suggested that once the *prior* information is available, we unavoidably make use of them and plan our actions accordingly.

From the biased behavior we conclude that the monkeys in the time following the pre-cue and preceding the rule-cue planned both movements with graded bias towards the more probable goal, in the way we intended. The striking similarity in *prior*-dependent RT costs and benefits between instructed and choice trials further supports the notion that the observed behavioral outcome is due to effects that are common to both instructed and choice trials, hence should occur before the final rule instruction.

We strongly believe that our graded *prior* manipulation induced biased action plans and not biased attentional effect. First, attention facilitates the selection of relevant sources of information by sensitivity control (Gottlieb & Balan 2010; Knudsen 2007). As we never showed visual target at the reach goal location, enhanced stimulus sensitivity at the reach goal location had no behavioral relevance. Furthermore, a previous study (Lebedev et al. 2004) showed that, when monkeys were faced with two saccade alternatives: one physically marked and one remembered, the majority of prefrontal neurons represented attended saccade locations, i.e. the cued ones, and not the remembered ones. Following this idea, in absence of physical targets in our task, the PMd and PRR activities observed during the memory period likely did not represent attended reach targets, but intended reach goals. Second, the color contrast we used for our rule-cues were far beyond discrimination threshold and we do not believe that higher performance observed when monkeys were instructed to reach following the initial bias was due to monkeys' better perceptual discrimination of the color of the biased rule-cue. Third, recent progress on studies of distinction between feature-based attention and feature-based expectation made the case that, in the ideal observer framework, stimulus probabilities should not enhance discrimination sensitivity but, instead, adjusts the detection criterion towards the more expected category (Summerfield & Egner 2016; Lauwereyns 2011).

We therefore conclude that, in response to different biasing pre-cues, the monkeys develop and maintain graded levels of potential action planning during the delay period,

allowing us to interpret the observed neural responses as reflectance of biased action planning.

In a related previous study of our lab, one dataset from two monkeys showed strongly biased movement planning activity, matching subsequent choice bias, while another, separately recorded dataset from the same two animals showed balanced planning activity, mirroring subsequent unbiased choice (Klaes et al. 2011). Our current results go beyond the previous findings in four important ways. First, both potential motor goals in the current study had to be inferred from a spatially incongruent pre-cue. The behavioral bias observed here is thereby due to our controlled manipulation of *prior*, not due to uncontrolled differences in task-difficultly between congruent and incongruent stimulus-response mappings such as those found in pro- and anti-reach tasks. Also, there is no risk that the observed neural responses reflected a spatial memory of the pre-cue which may have interfered with spatial neural encoding of the future motor goal in either of the two potential motor-goal locations. Second, by introducing a trial-by-trial bias we could confirm that it is the same population of PMd and PRR neurons that is active during both balanced or partially biased planning of two potential motor goals and planning of unambiguously selected motor goals. In addition, we could reveal single neurons' heterogeneous modulation patterns in response to probabilistic bias manipulation. Third, by explicitly cueing the choice trials with a color-neutral rule-cue, we could compare RTs between choice and instructed trials to demonstrate their identical dependency on *prior*. Previously such a comparison was prevented by the fact that choice trials were indicated to the monkeys by omitting the presentation of the rule-cue, creating a discrepancy between instructed and choice trials and a risk of animals postponing their response in choice trials while waiting if a rule-cue was yet to come. Forth, we can here correlate graded behavioral bias effects with graded changes in neural response. We could thereby more explicitly test the hypothesis that motor planning is not the consequence of choice, i.e. does not reflect a binary outcome of a motor goal selection process, but rather reflects an inherent competitive process leading or contributing to the selection process (affordance competition; Cisek 2007).

Neural response latencies reflect reaction time differences

An open question in sensorimotor research is the mutual role of parietal and frontal lobe sensorimotor structures in motor planning and choice (Hanks et al. 2015). With graded *prior* manipulation, we also observed graded motor goal representation in both brain areas, PRR and PMd. Our analysis of graded directional modulation showed no obvious differences between the two brain areas in the encoding of potential motor goals in the late memory period. PMd and PRR neural latencies captured RT pattern equally well, i.e. reduced latencies in *follow* trials and increased latencies when the more firmly planned movement had to be countermanded.

In contrast, the analysis of neural response latencies suggests major differences in the onset latencies of neural choice signals between PRR and PMd. Inconsistent with previous studies (e.g. Roitman & Shadlen 2002), we found motor-related peaks in premotor (in all except *follow* reaches) and parietal neurons after the movement initiation in monkey H. In contrast to experiments requesting eye movement responses, our monkeys performed more extended and effortful reach movements. This provided us with a finer time resolution to study neural latencies in relation to movement initiation between brain areas. The observed late-coming peak activities in both areas argue against considering peak activities as decision commitment signal in sensorimotor areas.

Conclusion

The predictability of a required action that was provided by *prior* probability of a rule-cue was translated into different levels of preparedness represented by different levels of motor planning activities in sensorimotor areas. Subsequently, these graded levels of preparedness in sensorimotor areas result in graded behavioral bias in monkeys, as previously described in humans (Suriya-Arunroj & Gail 2015). Our observed neural encoding during the instructed delay is consistent with the idea that *prior* probability is reflected in a shifted baseline in the framework of drift diffusion models (DDM: e.g. Ratcliff 1978; Smith & Ratcliff 2004), leading to higher choice probability and shorter reaction times of biased action. When equipotent potential action alternatives are available, both, PMd and PRR showed dual representation of both action goals,

suggesting a stage of processing at which the two alternative response options are not yet integrated into a single decision variable, but at which both options might be engaged in a competitive "race". Parallel representation of available movement options is believed to be the process through which action cost can be integrated into the decision mechanism (Cisek 2012). This is in contrast to situations of good-based decision-making in which action-independent choices might be accomplished outside sensorimotor structures, e.g. prefrontal cortex, before the chosen option is passed to the sensorimotor cortices for planning of a corresponding action (Padoa-Schioppa 2011). Competitive motor planning in sensorimotor cortex hence likely contributes to action selection in ambiguous rule-based choice situations, independently of target value, which is consistent with the affordance competition model of decision-making.

References

Ames, K.C., Ryu, S.I. & Shenoy, K.V., 2014. Neural dynamics of reaching following incorrect or absent motor preparation. Neuron, 81(2), pp.438–451.

Andersen, R.A. & Cui, H., 2009. Intention, Action Planning, and Decision Making in Parietal-Frontal Circuits. Neuron, 63(5), pp.568–583.

Churchland, M.M. & Shenoy, K.V., 2007. Temporal Complexity and Heterogeneity of Single-Neuron Activity in Premotor and Motor Cortex. Journal of Neurophysiology, 97(6), pp.4235–4257.

Cisek, P., 2007. Cortical mechanisms of action selection: the affordance competition hypothesis. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1485), pp.1585–1599.

Cisek, P., 2012. Making decisions through a distributed consensus. Current Opinion in Neurobiology, pp.927–936.

Cisek, P. & Kalaska, J.F., 2005. Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. Neuron, 45(5), pp.801–814.

Coallier, É., Michelet, T. & Kalaska, J.F., 2015. Dorsal premotor cortex: neural correlates of reach target decisions based on a color-location matching rule and conflicting sensory evidence. Journal of Neurophysiology, p.jn.00166.2014.

Crammond, D.J. & Kalaska, J.F., 1994. Modulation of preparatory neuronal activity in dorsal premotor cortex due to stimulus-response compatibility. Journal of Neurophysiology, 71(3), pp.1281–1284.

Dorris, M.C. & Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. Neuron, 44(2), pp.365–378.

Gail, A. & Andersen, R.A., 2006. Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. Journal of Neuroscience, 26(37), pp.9376–9384.

Gail, A., Klaes, C. & Westendorff, S., 2009. Implementation of Spatial Transformation Rules for Goal-Directed Reaching via Gain Modulation in Monkey Parietal and Premotor Cortex. The Journal of Neuroscience, 29(30), pp.9490–9499. Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1095-09.2009.

Gold, J.I. & Shadlen, M.N., 2007. The Neural Basis of Decision Making. Annual Review of Neuroscience, 30(1), pp.535–574.

Gottlieb, J. & Balan, P., 2010. Attention as a decision in information space. Trends in Cognitive Sciences, 14(6), pp.240–248.

Gottlieb, J. & Goldberg, M.E., 1999. Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. Nature Neuroscience, 2(10), pp.906–912.

Hanks, T.D. et al., 2015. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature, 520(7546), pp.220–223.

Hartigan, J.A. & Hartigan, P.M., 1985. The Dip Test of Unimodality. The Annals of Statistics, 13(1), pp.70–84.

Johnson, M.T. et al., 1999. Visuomotor processing as reflected in the directional discharge of premotor and primary motor cortex neurons. Journal of Neurophysiology, 81(2), pp.875–894.

Klaes, C. et al., 2011. Choosing Goals, Not Rules: Deciding among Rule-Based Action Plans. Neuron, 70(3), pp.536–548.

Klaes, C. et al., 2012. Sensorimotor Learning Biases Choice Behavior: A Learning Neural Field Model for Decision Making. PLOS Computational Biology, 8(11), pp.e1002774–19.

Knudsen, E.I., 2007. Fundamental components of attention. Annual Review of Neuroscience, 30, pp.57–78.

Kurata, K., 1991. Corticocortical inputs to the dorsal and ventral aspects of the premotor cortex of macaque monkeys. Neuroscience research, 12(1), pp.263–280.

Lauwereyns, J., 2011. The Anatomy of Bias, MIT Press.

Lebedev, M.A. et al., 2004. Representation of attended versus remembered locations in prefrontal cortex. PLOS Biology, 2(11), p.e365.

Leite, F.P. & Ratcliff, R., 2011. What cognitive processes drive response biases? A diffusion model analysis. Judgment and Decision Making, 6(7), pp.651–687.

Lindner, A. et al., 2010. Human posterior parietal cortex plans where to reach and what to avoid. Journal of Neuroscience, 30(35), pp.11715–11725.

Maddox, W.T. & Bohil, C.J., 1998. Base-rate and payoff effects in multidimensional perceptual categorization. Journal of experimental psychology. Learning, memory, and cognition, 24(6), pp.1459–1482.

Mulder, M.J. et al., 2012. Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. Journal of Neuroscience, 32(7), pp.2335–2343.

Munoz, D.P. & Everling, S., 2004. Look away: the anti-saccade task and the voluntary control of eye movement. Nature Publishing Group, 5(3), pp.218–228.

Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. Annual Review of Neuroscience, 34, pp.333–359.

Pastor-Bernier, A. & Cisek, P., 2011. Neural correlates of biased competition in premotor cortex. Journal of Neuroscience, 31(19), pp.7083–7088.

Platt, M.L. & Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. Nature, 400(6741), pp.233–238.

Posner, M.I., Snyder, C.R. & Davidson, B.J., 1980. Attention and the detection of signals. J Exp Psychol, 109(2), pp.160–174.

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

Ratcliff, R., 1981. A theory of order relations in perceptual matching. Psychological Review, 88(6), pp.552–572.

Ratcliff, R., 1985. Theoretical interpretations of the speed and accuracy of positive and negative responses. Psychological Review, 92(2), pp.212–225.

Ratcliff, R., Van Zandt, T. & McKoon, G., 1999. Connectionist and diffusion models of reaction time. Psychological Review, 106(2), pp.261–300.

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

Simen, P. et al., 2009. Reward rate optimization in two-alternative decision making: Empirical tests of theoretical predictions. Journal of experimental psychology. Human perception and performance, 35(6), pp.1865–1897.

Smith, P.L. & Ratcliff, R., 2004. Psychology and neurobiology of simple decisions. Trends in Neurosciences, 27(3), pp.161–168.

Sugrue, L.P., Corrado, G.S. & Newsome, W.T., 2005. Choosing the greater of two goods: neural currencies for valuation and decision making. Nature Publishing Group, 6(5), pp.363–375.

Sugrue, L.P., Corrado, G.S. & Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. Science, 304(5678), pp.1782–1787.

Suminski, A.J. et al., 2015. Temporal evolution of both premotor and motor cortical tuning properties reflect changes in limb biomechanics. Journal of Neurophysiology, 113(7), pp.2812–2823.

Summerfield, C. & Egner, T., 2016. Feature-Based Attention and Feature-Based Expectation. Trends in Cognitive Sciences, 20(6), pp.401–404.

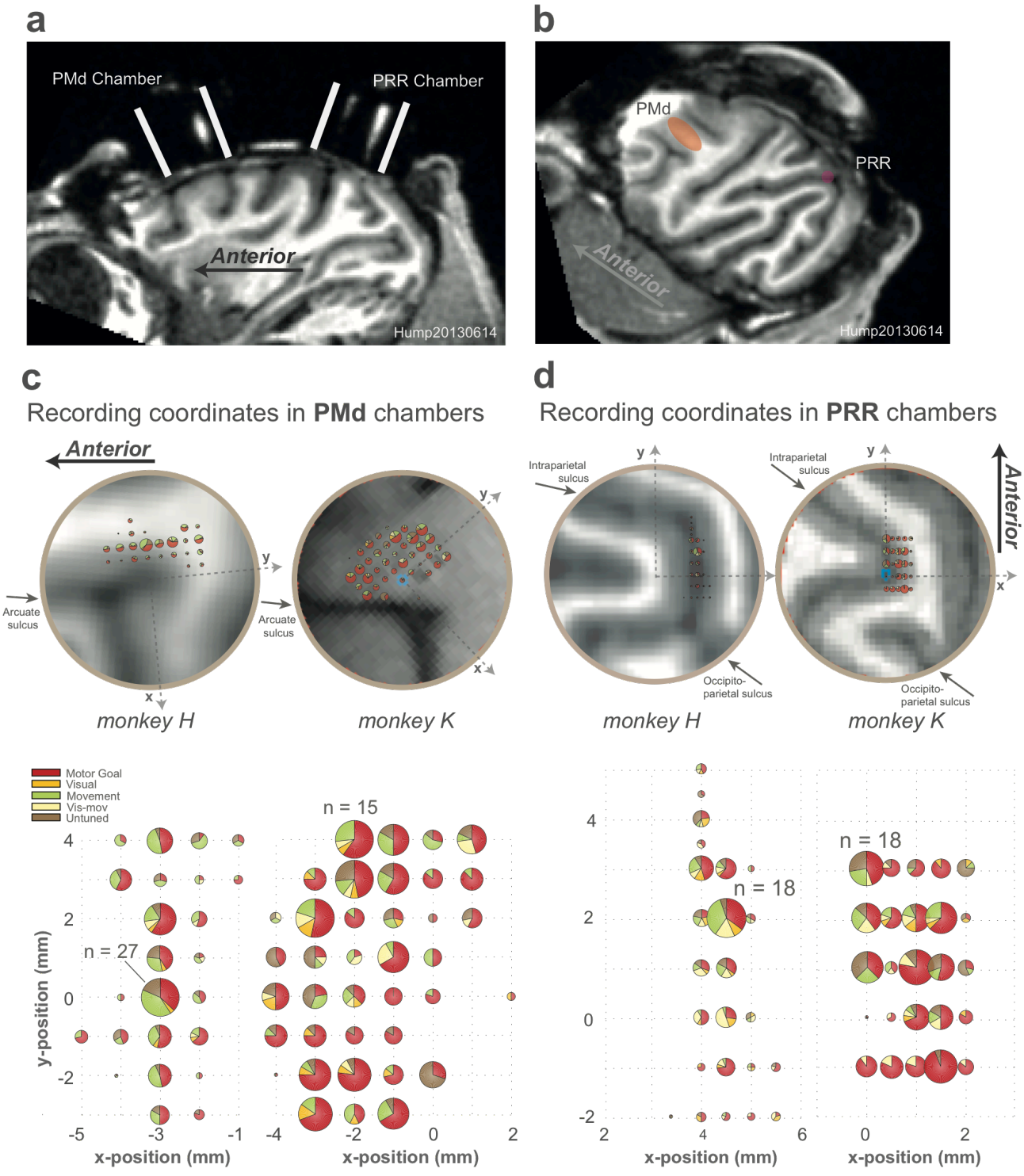
Suriya-Arunroj, L. & Gail, A., 2015. I Plan Therefore I Choose: Free-Choice Bias Due to Prior Action-Probability but Not Action-Value. Frontiers in behavioral neuroscience, 9, p.315.

Westendorff, S., Klaes, C. & Gail, A., 2010. The cortical timeline for deciding on reach motor goals. Journal of Neuroscience, 30(15), pp.5426–5436.

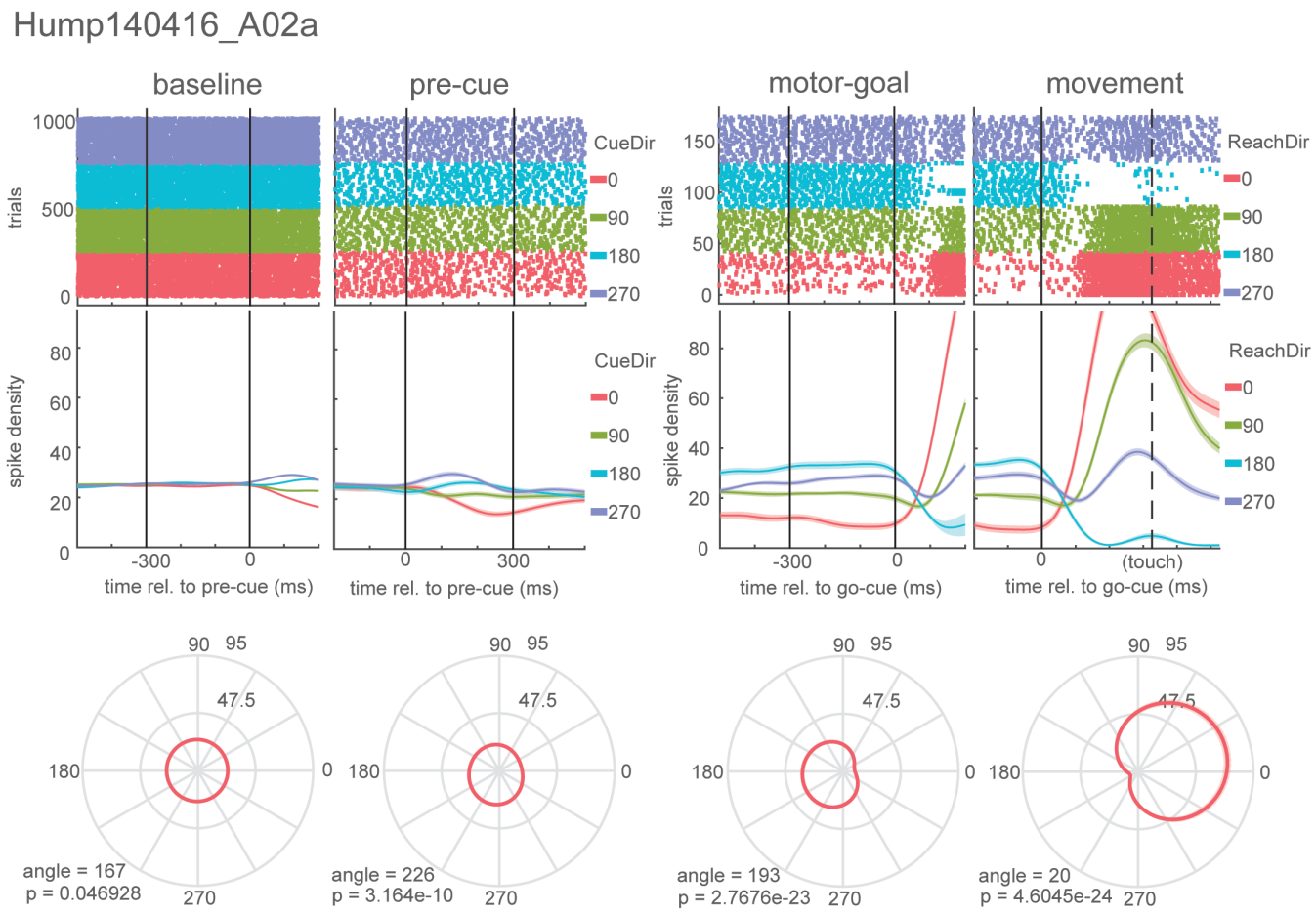
Yang, T. & Shadlen, M.N., 2007. Probabilistic reasoning by neurons. Nature, 447(7148), pp.1075–1080.

Zhang, M. & Barash, S., 2004. Persistent LIP activity in memory antisaccades: working memory for a sensorimotor transformation. Journal of Neurophysiology, 91(3), pp.1424–1441.

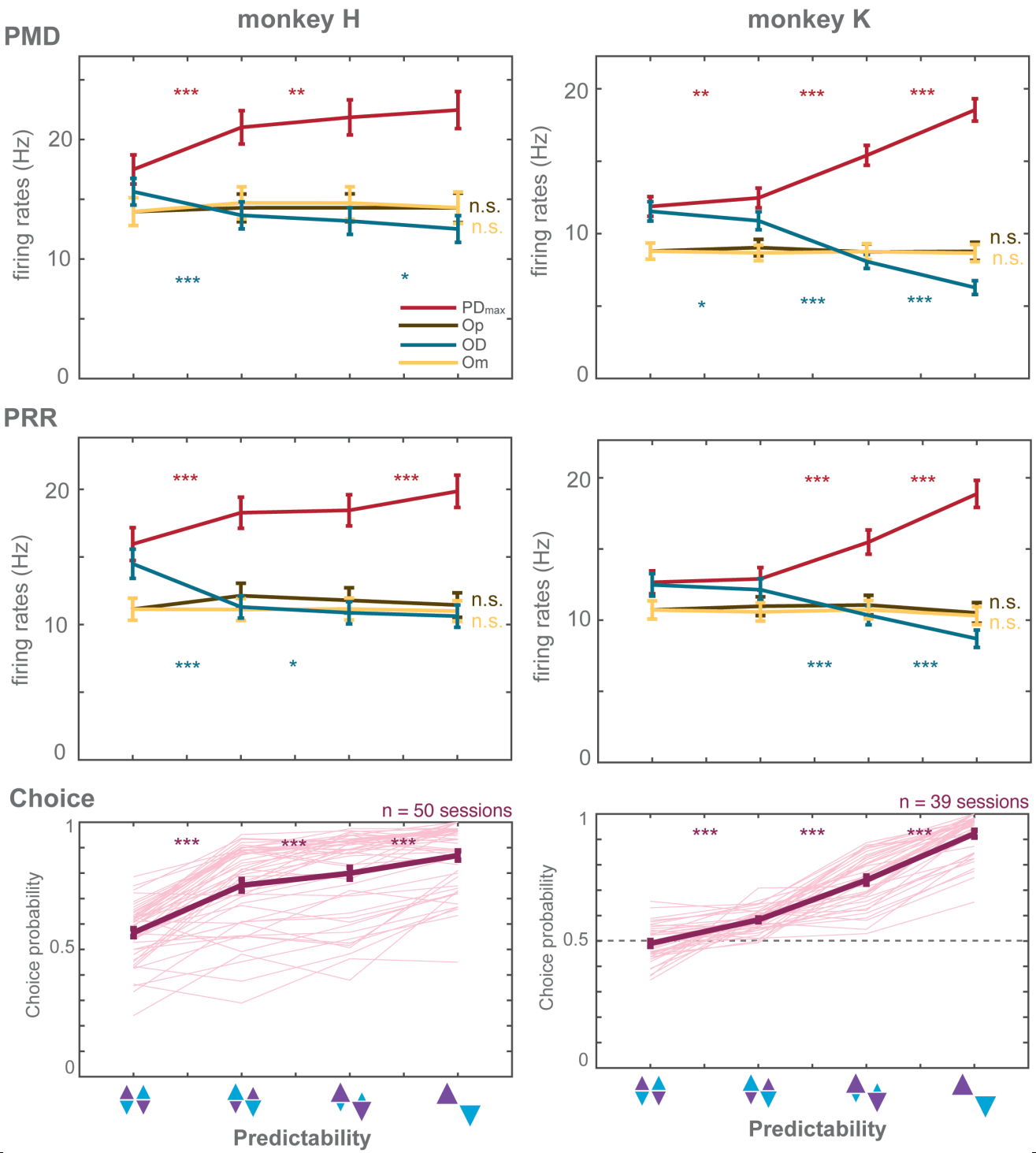
Supplementary figures



SUPP. FIGURE 1 | Chamber positions and recording coordinates. (A) PMd and PRR chambers: rotated view from the monkey's left side to see both chambers on the same plane **(B)** PMd and PRR areas: rotated view to visualize both areas on the same plane. **(C-D)** Recording coordinates in PMd and PRR chambers. The area of each pie chart corresponds to the number of neurons of different types at each coordinate. Note that the sizes of the pie chart are not comparable between areas; the biggest pie chart in PMd represents 23 cells vs. 32 cells in PRR. Motor Goal: motor-goal tuned neurons, Visual: visually tuned neurons, Movement: peri-movement neurons, Vis-mov: neurons showing visual tuning and movement tuning but untuned during memory period (see further details in *Neural data selection and selectivity profiles*).



SUPP. FIGURE 2 | Example tuning profile. (top panel) Raster plots show neural activities at the different cue directions, when aligned to the pre-cue onset (baseline and pre-cue), and reach directions, when aligned to the go-cue onset (motor-goal and movement). **(middle panel)** Spike density functions, in identical arrangement as the rasters are shown. **(bottom panel)** Tuning functions corresponding are shown. Cue-related tuning is computed relative to cue positions (baseline: 300 ms before pre-cue onset; pre-cue: 300 ms after pre-cue onset) whereas movement-related tunings are relative to reach directions (motor-goal: 300 ms before rule-cue onset; movement: rule-cue to movement offset (touch)). The example neuron shows evolving tuning properties across periods with tuning direction changes from 226° and 193° in pre-cue and motor-goal epochs (difference less than 90°) to 20° in movement epoch. p-values are results of Kruskal-Wallis test.



SUPP. FIGURE 3 | Planning activities reflect monkeys' choice bias. (PMd) Multiple comparison of (non-normalized) average firing rates of motor-goal tuned PMd neurons of monkey H (left) and monkey K (right) at maximum direction (PD_{max}; red), opposite direction (OD; blue), and two orthogonal directions (Op (brown) and Om (yellow) for $PD_{max} + \pi/2$ and $PD_{max} - \pi/2$, respectively). **(PRR)** average firing rates of motor-goal tuned PRR neurons, organized in an identical way as PMd. **(Choice)** average choice probabilities (CP) of monkey H (left) and monkey K (right) in all bias levels. (*p < α_{corr} at 5%, **p < α_{corr} at 1% , ***p < α_{corr} at 0.1%; *Bonferroni*-corrected t-test)

General discussion

In this final section, the findings of the human psychophysics and monkey physiology studies are summarized and discussed in a broader context.

The human psychophysical study focused on a comparison between the influence of initial predictability and initial preferability of available options on decision-making when both options became equally valued at the decision moment.

Parts of our results are in line with results of previous studies, in which *prior* probability and potential payoff were used to bias subjects' behaviors, and stronger effects of *prior* bias were found compared to a manipulation of potential payoff (Maddox & Bohil 1998; Simen et al. 2009; Leite & Ratcliff 2011; Mulder et al. 2012). According to the rise-to-threshold model for decision mechanisms, both bias manipulations were suggested to induce a baseline shift with the shift driven by payoff bias being about half of the magnitude of the shift due to *prior* bias (Maddox 2002; Bogacz et al. 2006).

In contrast to existing work, the choice trials in our task design provided a new perspective on subjects' responses that had not been reported before. Graded choice bias and choice reaction times (RT) in the PROB task mirrored the *prior* probability. This means that the preliminary action plan could drive choice bias by shifting the baseline depending on how reliable the plan was. If the findings from other labs hold and the payoff manipulation in the AMNT task induces the same underlying mechanism as the *prior* manipulation in the PROB task but with a smaller effect size, we should have observed at least some degree of RT shortening when subjects chose the biased target in the AMNT task. This was however not the case; the choice RT pattern in AMNT task was bias-independent. We therefore ruled out the previous *halfway shift* explanation. Instead, we argued in favor of a drift rate change to explain the behavioral bias in instructed trials of our payoff manipulation experiment.

Interestingly, we found matching-like behavior resulting from planning but not expected value. This finding suggests that choice bias, accompanied by reduced reaction times, that had previously been interpreted as a 'match' of expected value (Sugrue et al. 2004; Lau &

Glimcher 2008; Platt & Glimcher 1999) could actually be due to preliminary action plans for which previous experiments often did not control.

Finally, it is worth noting that rise-to-threshold models have previously mostly been applied to perceptual decision-making tasks, with temporal integration of evidence. We found that the parameters typically used in these models did not fit well to the data from our experiment, especially the AMNT task. Further extensions of current decision-making models that could better account for the data in the AMNT scenarios would be of great interest and would certainly add another important dimension to our current understanding of decision-making mechanisms.

The monkey physiology experiment also revealed biased behaviors, both in RTs and choice probabilities, in response to predictability. Importantly, the neural activities in PMd and PRR represented reach plans in a graded fashion depending on the bias level that was previously pre-cued and reflecting the degree of later choice bias the monkey displayed. In addition, we were able to confirm that both biased and balanced reach plans as previously shown in separate datasets by Klaes et al. (2011) could be represented by same neural population in both PMd and PRR.

The fact that PMd and PRR neurons encode graded planning activities depending on the level of predictability of each reach goal supports the relevance of sensorimotor areas and motor planning activities to the decision-making process. The good-based model of decision-making, in which the decision process is finalized in the prefrontal cortex and the decision outcome is fed into the sensorimotor areas to plan the movement, would neither predict any representation of motor-goal alternatives in the sensorimotor areas nor any influence of planning on subsequent free-choice behaviors.

We also looked beyond the planning period in our analyses of the relative latencies of movement-related responses in PMd and PRR. The result showed that final motor goal selection appears to be represented earlier in PMd than PRR. This result confirms PMd role in decision commitment.

Overall, it was intriguing to observe behavioral bias in our experimental paradigm, not only in humans but also in monkeys. The subjects could have solved the task perfectly

well without pre-cue sizes taking into account. In instructed trials, they anyway receive the instruction to be followed or, in choice trials, receive reward for either movement with equal probability. The robust bias seen in measured behaviors suggests that predictability is a temping source of information and anticipation is unavoidable; monkeys and humans could not help but plan once the *prior* information was available, even though it was provided as an accessory in the pre-cue.

In everyday life, we often have to decide between options with unclear values, which could be due to minimal contrast of the values or the complexity of values, and in many cases we have no idea whether the decision we have made was right or wrong. We face such 'hard choice' situations, in which we cannot easily rank the options, more often than we might think. Action plans which bias equally valued choices, could be interpreted as our 'default' plans or habits that might be able to bias our everyday hard choice and could contribute to the contemporary idea of 'irrationality' in economic decisions.

Taken together, this dissertation provided both behavioral and neural evidence that action planning can be encouraged by the predictability of future events and that choices can be biased when the movement plan overlaps with one of the movement alternatives. It uncovered novel findings on the influence of *prior* probability and reward expectation on decision, provided novel perspectives on decision-making process, and paved the way for subsequent experiments. This thesis further highlights the relevant role of movement planning in the decision-making process and supports the idea that both processes share common neural mechanisms.

References

Andersen, R.A. & Buneo, C.A., 2002. Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience*, 25, pp.189–220.

Andersen, R.A. & Cui, H., 2009. Intention, Action Planning, and Decision Making in Parietal-Frontal Circuits. *Neuron*, 63(5), pp.568–583.

Arnsten, A.F., 2003. Patricia Goldman-Rakic:A Remembrance. *Neuron*, 40(3), pp.465–470.

Batista, A.P. & Andersen, R.A., 2001. The parietal reach region codes the next planned movement in a sequential reach task. *Journal of Neurophysiology*, 85(2), pp.539–544.

Batista, A.P. et al., 1999. Reach plans in eye-centered coordinates. *Science*, 285(5425), pp.257–260.

Baumann, M.A., Fluet, M.-C. & Scherberger, H., 2009. Context-Specific Grasp Movement Representation in the Macaque Anterior Intraparietal Area. *The Journal of Neuroscience*, 29(20), pp.6436–6448.

Bechara, A. et al., 1997. Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304), pp.1293–1295.

Bogacz, R. et al., 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), pp.700–765.

Bonem, M. & Crossman, E.K., 1988. Elucidating the effects of reinforcement magnitude. *Psychological bulletin*, 104(3), pp.348–362.

Buneo, C.A. et al., 2002. Direct visuomotor transformations for reaching. *Nature*, 416(6881), pp.632–636.

Carello, C.D. & Krauzlis, R.J., 2004. Manipulating intent: evidence for a causal role of the superior colliculus in target selection. *Neuron*, 43(4), pp.575–583.

Cavada, C. & Goldman-Rakic, P.S., 1989. Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. *The Journal of comparative neurology*, 287(4), pp.393–421.

Cavada, C. & Goldman-Rakic, P.S., 2005. Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *The Journal of comparative neurology*, 287(4), pp.422–445. Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2477406&retmode=ref&cmd=prlinks.

Churchland, A.K., Kiani, R. & Shadlen, M.N., 2008. Decision-making with multiple alternatives. *Nature Neuroscience*, 11(6), pp.693–702.

Cisek, P., 2007. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485), pp.1585–1599.

Cisek, P., 2012. Making decisions through a distributed consensus. *Current Opinion in Neurobiology*, pp.927–936.

Cisek, P. & Kalaska, J.F., 2010. Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience*, 33(1), pp.269–298.

Coallier, É., Michelet, T. & Kalaska, J.F., 2015. Dorsal premotor cortex: neural correlates of reach target decisions based on a color-location matching rule and conflicting sensory evidence. *Journal of Neurophysiology*, p.jn.00166.2014.

Colby, C.L. et al., 1988. Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *The Journal of comparative neurology*, 269(3), pp.392–413.

Colby, C.L., Duhamel, J.R. & Goldberg, M.E., 1996. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of Neurophysiology*, 76(5), pp.2841–2852.

Crammond, D.J. & Kalaska, J.F., 1996. Differential relation of discharge in primary motor cortex and premotor cortex to movements versus actively maintained postures during a reaching task. *Experimental brain research*, 108(1), pp.45–61.

Crammond, D.J. & Kalaska, J.F., 1994. Modulation of preparatory neuronal activity in dorsal premotor cortex due to stimulus-response compatibility. *Journal of Neurophysiology*, 71(3), pp.1281–1284.

Dayan, P. & Balleine, B.W., 2002. Reward, motivation, and reinforcement learning. *Neuron*, 36(2), pp.285–298.

Desmurget, M. & Sirigu, A., 2009. A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences*, 13(10), pp.411–419.

di Pellegrino, G. & Wise, S.P., 1991. A neurophysiological comparison of three distinct regions of the primate frontal lobe. *Brain*, 114 (Pt 2), pp.951–978.

Diederich, A. & Busemeyer, J.R., 2006. Modeling the effects of payoff on response bias in a perceptual discrimination task: bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, 68(2), pp.194–207.

Dolan, R. & Sharot, T., 2012. *Neuroscience of Preference and Choice*, Academic Press.

Edwards, W., 1956. Reward probability, amount, and information as determiners of sequential two-alternative decisions. *J Exp Psychol*, 52(3), pp.177–188.

Fagg, A.H. & Arbib, M.A., 1998. Modeling parietal-premotor interactions in primate control of grasping. *Neural networks : the official journal of the International Neural Network Society*, 11(7-8), pp.1277–1303.

Felleman, D.J. & Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, N.Y. : 1991)*, 1(1), pp.1–47.

Fellows, L.K., 2006. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain*, 129(Pt 4), pp.944–952.

Fellows, L.K. & Farah, M.J., 2005a. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral cortex (New York, N.Y. : 1991)*, 15(1), pp.58–63.

Fellows, L.K. & Farah, M.J., 2005b. Dissociable elements of human foresight: a role for the ventromedial frontal lobes in framing the future, but not in discounting future rewards. *Visuomotor Functions of the Posterior Parietal Cortex*, 43(8), pp.1214–1221.

Fellows, L.K. & Farah, M.J., 2007. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral cortex (New York, N.Y. : 1991)*, 17(11), pp.2669–2674.

Franchina, J.J. & Brown, T.S., 1971. Reward magnitude shift effects in rats with hippocampal lesions. *Journal of comparative and physiological psychology*, 76(3), pp.365–370.

Gail, A. & Andersen, R.A., 2006. Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. *Journal of Neuroscience*, 26(37), pp.9376–9384.

Gail, A., Klaes, C. & Westendorff, S., 2009. Implementation of Spatial Transformation Rules for Goal-Directed Reaching via Gain Modulation in Monkey Parietal and Premotor Cortex. *The Journal of Neuroscience*, 29(30), pp.9490–9499. Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1095-09.2009.

Glimcher, P.W. & Fehr, E., 2014. *Neuroeconomics* second edition, Academic Press.

Glimcher, P.W. & Rustichini, A., 2004. Neuroeconomics: the consilience of brain and decision. *Science*, 306(5695), pp.447–452.

Gnadt, J.W. & Andersen, R.A., 1988. Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental brain research*, 70(1), pp.216–220.

Gold, J.I. & Shadlen, M.N., 2007. The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30(1), pp.535–574.

Goldman-Rakic, P.S., 1996. The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 351(1346), pp.1445–1453.

Goodale, M.A. & Milner, A.D., 1992a. Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), pp.20–25.

Goodale, M.A. & Milner, A.D., 1992b. Separate visual pathways for perception and action. *Trends in Neurosciences*.

Hassani, O.K., Cromwell, H.C. & Schultz, W., 2001. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *Journal of Neurophysiology*, 85(6), pp.2477–2489.

Herrnstein, R.J., 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, pp.267–272.

Hollerman, J.R., Tremblay, L. & Schultz, W., 1998. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of Neurophysiology*, 80(2), pp.947–963.

Johnson, P.B. et al., 1996. Cortical networks for visual reaching: physiological and anatomical organization of frontal and parietal lobe arm regions. *Cerebral cortex (New York, N.Y. : 1991)*, 6(2), pp.102–119.

Kawagoe, R., Takikawa, Y. & Hikosaka, O., 1998. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience*, 1(5), pp.411–416.

Klaes, C. et al., 2011. Choosing Goals, Not Rules: Deciding among Rule-Based Action Plans. *Neuron*, 70(3), pp.536–548.

Knutson, B. et al., 2001. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, 21(16), p.RC159.

Kurata, K., 1991. Corticocortical inputs to the dorsal and ventral aspects of the premotor cortex of macaque monkeys. *Neuroscience research*, 12(1), pp.263–280.

Lau, B. & Glimcher, P.W., 2005. Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys. *Journal of the Experimental Analysis of Behavior*, 84(3), pp.555–579.

Lau, B. & Glimcher, P.W., 2008. Value Representations in the Primate Striatum during Matching Behavior. *Neuron*, 58(3), pp.451–463.

Leite, F.P. & Ratcliff, R., 2011. What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making*, 6(7), pp.651–687.

Levy, D.J. & Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, 22(6), pp.1027–1038.

Luppino, G. et al., 2005. Occipital (V6) and parietal (V6A) areas in the anterior wall of the parieto-occipital sulcus of the macaque: a cytoarchitectonic study. *European Journal of Neuroscience*, 21(11), pp.3056–3076.

Maddox, W.T., 2002. Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior*, 78(3), pp.567–595.

Maddox, W.T. & Bohil, C.J., 1998. Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of experimental psychology. Learning, memory, and cognition*, 24(6), pp.1459–1482.

Marconi, B. et al., 2001. Eye-hand coordination during reaching. I. Anatomical relationships between parietal and frontal cortex. *Cerebral cortex (New York, N.Y. : 1991)*, 11(6), pp.513–527.

Maunsell, J.H. & Newsome, W.T., 1987. Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10, pp.363–401.

Mazzoni, P. et al., 1996. Motor intention activity in the macaque's lateral intraparietal area. I. Dissociation of motor plan from sensory memory. *Journal of Neurophysiology*, 76(3), pp.1439–1456.

McPeek, R.M. & Keller, E.L., 2004. Deficits in saccade target selection after inactivation of superior colliculus. *Nature Neuroscience*, 7(7), pp.757–763.

Miller, E.K. & Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, pp.167–202.

Mir, P. et al., 2011. Motivation and movement: the effect of monetary incentive on performance speed. *Experimental brain research*, 209(4), pp.551–559.

Mountcastle, V.B. et al., 1975. Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *Journal of Neurophysiology*, 38(4), pp.871–908.

Mulder, M.J. et al., 2012. Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. *Journal of Neuroscience*, 32(7), pp.2335–2343.

Neumann, Von, J. & Morgenstern, O., 1944. *Theory of games and economic behavior*, Princeton University Press. Available at: https://archive.org/stream/theoryofgamesand030098mbp/theoryofgamesand030098mbp_djvu.txt.

Norman, D.A. & Shallice, T., 1980. Attention to Action: Willed and Automatic Control of Behavior. *Center for Human Information Processing Technical Report no. 99*.

O'Doherty, J.P., 2014. The problem with value. *Neuroscience and Biobehavioral Reviews*, 43, pp.259–268.

O'Reilly, R.C., 2010. The What and How of prefrontal cortical organization. *Trends in Neurosciences*, 33(8), pp.355–361.

Olds, J. & Milner, P., 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology*, 47(6), pp.419–427.

Oliveira, F.T.P. et al., 2010. Transcranial magnetic stimulation of posterior parietal cortex affects decisions of hand choice. *Proceedings of the National Academy of Sciences*, 107(41), pp.17751–17756.

Orban, G.A., 2007. Higher order visual processing in macaque extrastriate cortex. *Physiological Reviews*, 88(1), pp.59–89.

Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annual Review of Neuroscience*, 34, pp.333–359.

Padoa-Schioppa, C. & Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441(7090), pp.223–226.

Pandya, D.N. & Kuypers, H.G., 1969. Cortico-cortical connections in the rhesus monkey. *Brain Research*, 13(1), pp.13–36.

Pastor-Bernier, A. & Cisek, P., 2011. Neural correlates of biased competition in premotor cortex. *Journal of Neuroscience*, 31(19), pp.7083–7088.

Pastor-Bernier, A., Tremblay, E. & Cisek, P., 2012. Dorsal premotor cortex is involved in switching motor plans. *Frontiers in neuroengineering*, 5, p.5.

Pearson, J.M., Watson, K.K. & Platt, M.L., 2014. Decision Making: The Neuroethological Turn. *Neuron*, 82(5), pp.950–965.

Petrides, M., 2002. The mid-ventrolateral prefrontal cortex and active mnemonic retrieval. *Neurobiology of Learning and Memory*, 78(3), pp.528–538.

Petrides, M. & Pandya, D.N., 1999. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience*, 11(3), pp.1011–1036.

Pierce, W.D. & Epling, W.F., 1983. Choice, matching, and human behavior: A review of the literature. *The Behavior analyst / MABA*, 6(1), pp.57–76.

Pisella, L. et al., 2000. An "automatic pilot" for the hand in human posterior parietal cortex: toward reinterpreting optic ataxia. *Nature Neuroscience*, 3(7), pp.729–736.

Platt, M.L. & Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741), pp.233–238.

Puig, M.V. et al., 2014. Dopamine modulation of learning and memory in the prefrontal cortex: insights from studies in primates, rodents, and birds. *Frontiers in neural circuits*, 8, pp.1–15.

Quian Quiroga, R. et al., 2006. Movement intention is better predicted than attention in the posterior parietal cortex. *Journal of Neuroscience*, 26(13), pp.3615–3620.

Rangel, A. & Hare, T., 2010. Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, 20(2), pp.262–270.

Rangel, A., Camerer, C. & Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. *Nature Publishing Group*, 9(7), pp.545–556.

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

Ratcliff, R., 1981. A theory of order relations in perceptual matching. *Psychological Review*, 88(6), pp.552–572.

Ratcliff, R., 1985. Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review*, 92(2), pp.212–225.

Resulaj, A. et al., 2009. Changes of mind in decision-making. *Nature*, pp.1–6.

Rizzolatti, G. & Luppino, G., 2001. The cortical motor system. *Neuron*, 31(6), pp.889–901.

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

Sakagami, M., Pan, X. & Uttl, B., 2006. Behavioral inhibition and prefrontal cortex in decision-making. *Neural networks : the official journal of the International Neural Network Society*, 19(8), pp.1255–1265.

Samuelson, P.A., 1938. A Note on the Pure Theory of Consumer's Behaviour. *Economica*, 5, pp.61–71.

Schieber, M.H., 2000. Inactivation of the ventral premotor cortex biases the laterality of motoric choices. *Experimental brain research*, 130(4), pp.497–507.

Shadlen, M.N. & Kiani, R., 2013. Decision Making as a Window on Cognition. *Neuron*, 80(3), pp.791–806.

Simen, P. et al., 2009. Reward rate optimization in two-alternative decision making: Empirical tests of theoretical predictions. *Journal of experimental psychology. Human perception and performance*, 35(6), pp.1865–1897.

Smith, P.L. & Ratcliff, R., 2004. Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3), pp.161–168.

Snyder, L.H., Batista, A.P. & Andersen, R.A., 1997. Coding of intention in the posterior parietal cortex. *Nature*, 386(6621), pp.167–170.

Snyder, L.H., Batista, A.P. & Andersen, R.A., 2000a. Intention-related activity in the posterior

parietal cortex: a review. *Vision Research*, 40(10-12), pp.1433–1441.

Snyder, L.H., Batista, A.P. & Andersen, R.A., 2000b. Saccade-related activity in the parietal reach region. *Journal of Neurophysiology*, 83(2), pp.1099–1102.

Strait, C.E., Blanchard, T.C. & Hayden, B.Y., 2014. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron*, 82(6), pp.1357–1366.

Sugrue, L.P., Corrado, G.S. & Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science*, 304(5678), pp.1782–1787.

Summerfield, C. & Tsetsos, K., 2012. Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms. *Frontiers in neuroscience*, 6.

Tanné, J. et al., 1995. Direct visual pathways for reaching movements in the macaque monkey. *Neuroreport*, 7(1), pp.267–272.

Thura, D. & Cisek, P., 2014. Deliberation and Commitmentin the Premotor and Primary Motor Cortex during Dynamic Decision Making. *Neuron*, 81(6), pp.1401–1416.

Tobler, P.N., Fiorillo, C.D. & Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715), pp.1642–1645.

Toni, I. et al., 2001. Learning arbitrary visuomotor associations: temporal dynamic of brain activity. *Cortical Control of Higher Motor Cognition*, 14(5), pp.1048–1057.

Ungerleider, L.G. & Mishkin, M., 1982. Two cortical visual systems. In D. J. lngle, M. A. Goodale, & R. J. W. Mansfield, eds. *Analysis of visual behavior*. Cambridge: The MIT Press, pp. 549–586.

Westendorff, S., Klaes, C. & Gail, A., 2010. The cortical timeline for deciding on reach motor goals. *Journal of Neuroscience*, 30(15), pp.5426–5436.

Wise, S.P. & Murray, E.A., 2000. Arbitrary associations between antecedents and actions. *Trends in Neurosciences*, 23(6), pp.271–276.

Wise, S.P. et al., 1997. Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annual Review of Neuroscience*, 20, pp.25–42.

Yang, T. & Shadlen, M.N., 2007. Probabilistic reasoning by neurons. *Nature*, 447(7148), pp.1075–1080.

Abbreviations

ACC	Anterior Cingulate Cortex
AIP	Anterior Intraparietal area
AMNT	(reward) Amount (task)
BMRS	Bias-Minimizing Reward Schedule
Ccw	Counterclockwise
CI	Confidence Interval
COBRA	COmpetition-Between-Reward-and-Acuracy-maximization
CP	Choice Probability
Cw	Clockwise
DDM	Drift-Diffusion Model
dlPFC	Dorsolateral Prefrontal Cortex
DMI	Directional Modulation Index
dmPFC	Dorsomedial Prefrontal Cortex
DTV	Directional Tuning Vector
DV	Decision Variables
ER	Error Rate
FEF	Frontal Eye Field
GLMM	Generalized Linear Mixed Models
IPL	Inferior Parietal Lobule
IT	Inferotemporal cortex
LIP	Lateral Intraparietal area
MDP	Medial Dorsoparietal area
MIP	Medial Intraparietal area
Mot	Motor-goal tuned neurons
Mov	Peri-movement neurons
MRI	Magnetic Resonance Imaging

ms	Millisecond
MT	Movement Time
M1	Primary motor area
OD	Opposite direction
OFC	Orbitofrontal Cortex
PD	Preferred direction
PFC	Prefrontal Cortex
PMC	Premotor Cortex
PMd	Dorsal Premotor Cortex
PMv	Ventral Premotor Cortex
PO	Parieto-Occipital area
PPC	Posterior Parietal Cortex
PROB	Probabilistic (task)
PRR	Parietal Reach Region
RT	Reaction Time
SE	Standard Error
SNc	Substantia Nigra pars compacta
SPL	Superior Parietal Lobule
STD	Standard Deviation
VA	Visual Angle
Vis	Visually tuned neurons
vlPFC	Ventrolateral Prefrontal Cortex
vmPFC	Ventromedial Prefrontal Cortex
VTA	Ventral Tegmental Area
V1	Primary visual area

Curriculum Vitae

Lalitta Suriya-Arunroj

Born 10th March 1985, Udon Thani, Thailand

Education

2010 -2015 **PhD in Neuroscience**, *Sensorimotor Group, German Primate Center*

Göttingen Graduate School for Neurosciences and Molecular Biosciences (GGNB)

Project: Neural basis of rule-based decisions with graded choice biases

Supervisor: *Prof. Dr. Alexander Gail*

2008 - 2010 **MASTER en Neuroscience**, *The Joint Master in Neuroscience (mention bien)*

Equivalent to a Master's degree in Neuroscience *(with honors, rank 2/5)*

University of Strasbourg (France), Basel (Switzerland), and Freiburg (Germany)

Courses: Fundamental Neuroscience, Molecular & Cellular Neurobiology, Developmental Neurobiology, Neuroanatomy and Physiology, Neurobiology of Learning, Neurochemistry, Neuropharmacology, Neurobiology of the Rhythms, Computational Neuroscience, Quantification and Modelisation in Neurobiology

2005 - 2008 **LICENCE en Sciences du vivant**, *biologie et informatique (mention bien)*

Equivalent to a Bachelor's degree in Life Science majoring in Biological and Computational Sciences *(with honors, rank 1/22)*

Université Louis Pasteur, Strasbourg, France

Courses: Biodiversity, Cellular Biology, Developmental Biology, Environmental Biology, Genetics, Biochemistry, Organic Chemistry, Thermodynamics, Fluid Mechanics, Programming, Algorithm, Applied Mathematics.

2003 - 2005 **Intensive courses in French**

Centre de linguistique appliqué, Université de Franche-Comté, Besançon, France

1997 - 2003 **Secondary school**

Udonpittayanukoon School, Udon Thani, Thailand

Laboratory experience

2009 - 2010 **Trainee in the Computational Neuroscience Group, University of Bern**

Working on a model for reinforcement learning in spiking neuron populations

Supervisor: *Prof. Dr. Walter Senn*

2008 - 2009 **Trainee (2 days a week) in INSERM Unit 575, University of Strasbourg**

Working on neurobiology of the cortical development

Supervisor: *Dr. Dominique Bagnard, Assoc Prof*

August 2006 One-month internship at the Centre of Excellence for Vectors and Vector-Borne Deseases, Mahidol University, Bangkok, Thailand
Supervisor: *Dr. Pattamapron Kittayapong, Assoc Prof*

July 2006 One-month internship a research group working on ESTs in cashew nuts, Department of Biology, Khonkaen University, Khonkaen, Thailand
Supervisor: *Dr. Napaporn Tantisuwichwong, Asst Prof*

Courses and workshops

2014 Advances in Multi-Neuronal Monitoring of Brain Activity, *Washington DC, USA*

2014 Graduate Training in Primate Neuroscience (GTPN), *Tübingen, Germany*

2012 workshop 'Scientific integrity & the responsible conduct of research', *German Primate Center, Göttingen, Germany*

2011 EUPRIM-Net course on General Primate Biology, *Göttingen, Germany*

2011 Perception, Action & Control, *Nijmegen, The Netherlands*

2011 Motor Systems (Lecture), *Göttingen, Germany*

2010 Vision to Action course, *Göttingen, Germany*

Conferences and contributions

2015 Neurizons 2015, *Göttingen, Germany*

2015 Thai Student Interdisciplinary Symposium (TSIS) 2015, *Heidelberg, Germany*

2015 the 11th Göttingen Meeting of the German Neuroscience Society, *Göttingen*

2015 the 8th PrimateNeurobiology Meeting, *Göttingen, Germany*

2014 Society for Neuroscience Annual Meeting 2014, *Washington DC, USA*

2014 Bernstein Conference 2014, *Göttingen, Germany*

2014 the 7th PrimateNeurobiology Meeting, *Tübingen, Germany*

2013 Bernstein Conference 2013, *Tübingen, Germany*

2013 Association of Thai Professionals in Europe Conference (ATPER) 2013, *Stockholm, Sweden*

2013 Thai Student Academic Conference (TSAC-TSIS) 2013, *Göttingen, Germany*

2013 the 10th Göttingen Meeting of the German Neuroscience Society, *Göttingen*

2013 the 6th PrimateNeurobiology Meeting, *Göttingen, Germany*

2012 Thai Student Academic Conference (TSAC) 2012, *Volendam, The Netherlands*

2012 the 5th PrimateNeurobiology Meeting, *Tübingen, Germany*

2011 the 9th Göttingen Meeting of the German Neuroscience Society, *Göttingen*

Computer skills

Operating systems		Linux, Windows, MacOS
Programming Languages		C, Tcl/Tk, Python, Perl
Software	- **Desktop applications:**	Microsoft office, Open office, iWork
	- **Image processing:**	Adobe Photoshop, Adobe Lightroom
	- **Maths and statistics:**	Mathematica, Matlab, R
	- **Graphic and publishing editor:**	CorelDraw, Adobe illustrator, Adobe inDesign

Languages

Thai	Native language
French	Competent in speaking, reading and writing – French language certificate: DALF (equivalent to C2 level) obtained in March 2005
English	Competent in speaking, reading and writing – IELTS: 7.5 obtained in August 2009
German	Intermediate level (B1 level)

Extracurricular activities

2011-2013	President of Thai Student Group in Göttingen
2012	Originated "ThaiGö talks" – monthly multidisciplinary seminars among Thai Students in Göttingen
2013	Organizing chair of Thai Student Academic Conference 2013 in Göttingen
2013-2015	Advisory board of Thai Student Association in Germany
2015	Editorial board for article submission of Thai Student Interdisciplinary Symposium (TSIS) 2015 in Heidelberg

Awards: Thaipattana scholarship from the Thai government in 2003 for completing Bachelor's and Master's Degree in France, then extended in 2010 for PhD in Germany

Interests and activities: Primates, the brain, educational systems, photography & image processing, travelling, music