The effects of counterfactual comparison on learning and reasoning

Ph.D. Thesis

Benjamin Timberlake
31st cycle, Cognitive and Brain Sciences

Center for Mind/Brain Sciences, CIMeC,
University of Trento

Supervisor:
Prof. Giorgio Coricelli

Co-Supervisor:
Dr. Nadège Bault

UNIVERSITÀ DEGLI STUDI DI TRENTO

CIMeC Center for Mind/Brain Sciences

Abstract

How humans make choices in uncertain and competitive situations is a key determinant of viability and successful living. Improving those choices requires sometimes encountering undesirable outcomes and avoiding them, eventually even anticipating them in novel situations. Learning depends on making choices, encountering errors and updating evaluations of options. Various models extended from the reinforcement learning framework compared to human behavior describe in part how individuals heterogeneously make choices. To peer into the components of these mechanisms, strategic games that emulate real-world situations provide measurable and manageable environments in which to examine slight differences in choice behavior among different people. Such differences may be endogenous to participants (e.g. age or learning disposition) while others derive from external events (e.g. emotional induction or brain stimulation). We contrasted such behavior in three situations involving learning or competition, leveraging differences in age, emotional induction and brain stimulation. We aimed to describe the variations in choice behavior across these differences and investigated, when possible, how prior conditions generated a transfer of learning from one domain to another. The work here builds on recent investigations of neural mechanisms underlying choice behavior during strategic or competitive interaction.

Table of Contents

Structure of this thesis

This thesis comprises four sections, each concerned with modulating processes of learning, reasoning and decision making. The notion of each effort was to examine specific features of the decision process via deviations from baseline, be they induced or selected. In the first, a review chapter, I critically examine the intersections between deficits in moral decision making and regret decision making in psychopathic and brain-lesion patients, finding a number of similarities and mutual deficiencies. The next chapter consists of analysis of a previously conducted experiment that considers age as a natural agent of cognitive change. My analysis examines how the relationship between choice behavior and counterfactual learning appears to change with age. The latter two chapters consist of experimental studies that I conducted in their entireties. The first study consists of influencing counterfactual learning via emotional priming in a large group of subjects and using computational modeling to characterize underlying cognitive mechanisms. The final chapter takes a yet more direct approach to test the causal role of two different brain areas in strategic thinking by employing transcranial brain stimulation in an effort to induce higher levels of strategic thinking in a classic economic strategic game.

General introduction

Humans make innumerable choices during waking life, from simple to complex, leading to patterns of behavior. Those choices can be changed by new information, that is, learning. These can be influenced by other internal and external factors brought to bear on a decision maker. Similar situations with only slight differences can give rise to different decisions and varied patterns of behavior. State of mind, mood, information, social situation, age, imagination — all may be implicated in modulating choice with other conditions held constant. From which hand to open a door with to whether it is safe to cross a road; from which cake to eat to which job to choose — any individual who has made it far enough in life to be studied in a laboratory experiment has made a mountain of advantageous decisions, though without doubt, accompanied by plenty of choices whose outcomes were not immediately beneficial. Yet, as many a motivational speaker will remind his audience, as much learning can come from failure as from success. This idea found generalization in the Rescorla-Wagner model's precept that learning occurs only when events violate expectations.

Yet the remarkable flexibility of learning has allowed adaptation to countless situations. In repeated games, players often exhibit learning behavior in making best choices. Early reinforcement learning models assumed that players responded only to the results of their own choices, repeating choices that lead to success and avoiding those that failed. These models were particularly good at explaining behavior in the context of the bounded rationality of specific decision problems. Later, expanded models such as fictitious play that accounted for foregone choices and in social situations anticipated actions of other players proved more descriptive of observed behavior (Fudenberg and Levine 1995). Yet more recently, adaptive models that account for some mixture of behavior described by more than one model alone have accounted for behavior more precisely. These hybrid models, such as experience-weighted attraction, afford individual-level description of sophistication of learning (Camerer and Ho 1998). More sophisticated players make choices in response to anticipated actions of others. The highest-level players further anticipate how their own actions will influence the strategies of competitors and maximize rewards by sometimes taking short-term losses.

In economic models of choice, agents choose options to maximize long-run reward. Various models applied to behavior, however, can account for long-run reward in different ways. In the short term, in fact, people often give up a higher expected reward in favor of avoiding the risk of loss. In repeated probabilistic tasks in limited-information feedback settings, choices are guided by counterfactual thinking that compares the outcome of a choice to the best outcome that could have been obtained with that choice. In settings in which outcomes of unselected choices are known, however, subsequent choices are guided by the difference between the outcome received and the outcome of a choice not made. This adds to counterfactual thinking element of responsibility, a crucial component of the experience of regret. Anticipated regret is so influential that in choice problems similar to those that have brought regret in the past, people increasingly avoid the choices that present the greatest potential regret – to an even greater extent than they avoid risk (Coricelli, Critchley et al. 2005). A discrete signal in the medial orbitofrontal cortex both accompanies the occurrence of regret during a task as well as its anticipation during the same task, indicating that experiencing regret is adaptive. Could this mechanism that functions so vigorously within one setting carry over and offer its influence in another similar but novel situation? This document proposes to explore that question at several depths.

Proposed circuits of regret learning share some structures and patterns with the systems implicated in moral decision making, in particular the ventromedial prefrontal cortex. We start the consideration of decision-making regret by comparing its implementation and neural correlates to those involved in moral decision making (Chapter 1). Specifically in instances in which normal function has been interrupted in both domains, the processes share some remarkable similarities.

We examined evidence of transfer between a lottery choice task and an instrumental learning task (Chapter 2). Here, we examined choice behavior as an indicator of how learning might vary between younger adults and older adults. We wanted to see if the relationship and potential transfer between tasks varied depending on age. Aging is marked by selective areas of cognitive decline, particularly in the context of decision making (Tymula, Belmaker et al. 2013). Moreover, adults older than 60 have been observed in one study to incorporate counterfactual information in learning to a lesser extent than younger adults (Tobia, Guo et al. 2016). Examining data from two age-segregated cohorts, we investigated how the choice behavior of older adults and younger adults in the lottery task indicated different learning patterns in the second task.

The question of transfer is probed more pointedly in an experiment in which we tried to make people feel very bad right from the beginning, having them lose a stack of money and showing them what they could have won if they had made a different choice (Chapter 3). Our hypothesis was that this induction of regret would elicit counterfactual thinking and make players learn from counterfactual comparison in a different game they played right after this large loss. We employed a limited-space strategic game, and then compared behavior to the fit of several models of reinforcement learning and belief-based learning that incorporate counterfactual thinking and strategic learning to various extents (Sutton and Barto 1998, Camerer 2003, Zhu, Mathewson et al. 2012). Employment of belief-based learning depends largely on understanding the broader structure of a system, to which a person already attuned to regret may be more sensitive if the experience transfers. In the context of the competitive game, we hypothesized that the consideration of the counterfactual demanded by the prior regret outcome would encourage this type of learning.

Transfer can begin with experience in one activity before commencing another. In the two previous overviews, this is realized with the outcome of a decision. That information may then modulate performance in the subsequent task. In a third study, we proposed to skip the step of introducing information with a behavioral situation and, instead, encourage the brain to reach a target state via electrical stimulation (Chapter 4). Different people engage in strategic thinking at various levels of sophistication, and measured brain activity reflects that diversity. Imaging studies have located some neural correlates of mentalizing in frontal cortical areas (Hampton, Bossaerts et al. 2008, Coricelli and Nagel 2009). If those areas are more active during higher levels of strategic thinking, they may well contribute the behavior. We therefore suspect that if these areas are stimulated to higher levels of activity, they could give rise to more strategic thinking.

Chapter 1

Regret, Responsibility and the Brain

Ben Timberlake[1], Giorgio Coricelli[1,2] and Nadège Bault[1]

[1]Center for Mind and Brain Sciences, University of Trento.
[2]Departments of Economics, University of Southern California.

To be published in The Moral Psychology of Regret *(2019), A. Gotlib (Ed.). London, United Kingdom: Rowman & Littlefield International.*

Introduction

Regret describes an emotion that arises from a variety of circumstances. We focus here on a particular type of regret, decision regret, which comes to the study of decision making by way of traditional economics, along with insights from psychology. This is clearly not the only formal description of regret, but it bears resemblance to variations studied in other fields. The benefits of this regret definition are its formalization, its operationalized measurability and its attendant body of literature in neuroimaging. This last is critical for comparison to the neural bases of other phenomena.

Regret refers to a specific set of conditions and responses, which include learning from an imagined alternative outcome that could have been reached through different action by the person feeling the emotion. This arises after an actor or agent has made a choice, sees its outcome, and then realizes that another outcome — the result of a different choice of hers — is more desirable. Decision-based regret or "decision regret" is proportional to the magnitude of the difference between the obtained and missed outcomes. These elements are the definitive components of decision regret: learning, responsibility and counterfactual information. Other emotions may arise from any one or two of these elements, but all three must be present for regret. These situational requirements have long guided the psychological description of regret (Zeelenberg, Beattie et al. 1996, Zeelenberg and Pieters 2007), and they persist in the economic definition of decision regret (Loomes and Sugden 1982). Decision-making studies operationalize this description, using both behavior and a modified utility function to quantify the effects of the emotional experience (Bell 1982, Loomes and Sugden 1982).

Like most decision processes, moral decision-making pits multiple options against one another in an effort to arrive at the most desirable outcome. Moral norms are personal convictions reflecting rules of conduct one ought to adopt in a given situation. They represent socially derived, internalized values attributed to a pattern of behavior thought to be appropriate (Manstead 2000). Moral norms play an important role in decision making because internalized values attributed to a particular course of action are likely to guide behavior. Consequently, behaving in contradiction to one's own moral norms is likely to elicit strong negative emotions. In such a situation, regret is likely to arise, especially if the norm violation results in a negative outcome. Some studies suggest that feelings of regret are anticipated at the prospect of violating one's moral norms (Parker, Manstead et al. 1995). Other studies have shown that anticipated regret and moral norms are confounded in explaining choices, especially those with moral implications (Rivis, Sheeran et al. 2009, Newton, Newton et al. 2013). Despite preliminary evidence from social psychology of a possible overlap between anticipated regret and moral norms, the cognitive mechanisms linking the two concepts have not yet been deeply investigated. Evidence from neuropsychology, however, suggests that the brain mechanisms underlying regret anticipation and the implementation of moral norms might involve similar neural circuits.

By tracing the brain activity associated with moral decision making and decision regret behaviors, it becomes clear that some of the same brain areas are similarly implicated in both processes, suggesting that some connections between the two categories of choices may be identified. Here, we explore this potential connection between moral- and regret-based decisions by reviewing their features and neural bases.

Counterfactual information

Regret arises from comparison to an alternative result: one that has not actually occurred. It requires the imagination of an alternative reality that results from a different choice than the one made. The process of deconstructing the present to imagine a different reality, called counterfactual thinking, is at the core of regret. Counterfactual thoughts are often generated after goal failure (Byrne 2002). The functional role of upward counterfactual thinking, and thus, associated regret, is to learn from mistakes, to generate variant courses of action suspected to prove more successful when similar situations are encountered in the future.

In a simple illustration of the definition and measurement of decision regret, imagine a game of chance: a slot machine. A gambler can pull the lever in exactly one way and take whatever result comes. Win or lose, his actions make no difference (other than the choice to play the game in the first place). Nature, wearing the guise of probability, determines the outcome every time. If he loses, the gambler by definition feels disappointment (and if he wins, satisfaction), but not regret. Now imagine two slot machines next to each other. The gambler must choose one to which to stake his fortunes, yet when he pulls the lever, the wheels spin on both machines, and he can see both outcomes. Now he sees both his actual winnings or losses on the machine he chose, as well as what he would have won or lost had he selected the other machine. If his slot machine loses while the other wins, he can imagine a world in which he made a different, winning, choice. This identification of the counterfactual precipitates regret. A notion, even an imprecise one, that the counterfactual outcome was better may give rise to regret, but the discrepancy between specific values of obtained and foregone allow for clearer interpretation at this point. Simulations of this situation have been used in various experimental settings to measure and compare regret to disappointment (Camille, Coricelli et al. 2004, Nicolle, Bach et al. 2011, Gillan, Morein-Zamir et al. 2014).

Regret is further characterized by a negative-valence error, which differentiates it from relief. The error is the difference between the obtained outcome and the imagined counterfactual outcome. This is an important distinction in regret: that the error must have negative valence, rather than the obtained outcome itself. This underscores the idea that regret is the negative result of comparison between outcomes, which may give rise to changes in behavior. In the slot machine study, even when subjects won with a certain choice but saw that they could have won more had they made a different choice, the net emotional sensation was negative (Camille, Griffiths et al. 2011). People describe their emotions as more negative with a better foregone choice, even when the obtained outcome is the same. This comparison is so clear that the emotion following a good outcome of a choice made (winning $50) compared to a very good outcome of a foregone choice ($200) can be rated even lower than that following a bad obtained outcome (-$50) compared to a very bad outcome avoided (-$200) (Camille et al. 2004). That is, despite winning more money, people said they felt worse—because they compared their winnings with what they could have won had they made a different choice. This ability to imagine an alternative reality after the fact informs decision problems not yet encountered. In fact, after experiencing regret, subjects made choices in subsequent tasks that were consistent with trying to minimize that feeling of regret (Coricelli, Dolan et al. 2007).

Learning value

In a more complex scenario that employs regret in learning, we might assign the two machines different probabilities of paying out. We could task the decision maker with earning the most money and therefore the goal of choosing the right (i.e. more likely) machine to play more often over the course of a number of opportunities. Such a sequential task (as employed in Daw, O'Doherty et al. 2006) allows the exploration of learning and the comparison of various models, which can include those that incorporate regret learning. Lohrenz and colleagues adopt the regret-learning model and rename it "fictive learning" to discard emotional connotations and to maintain only the error signal of an unobtained outcome (Lohrenz, McCabe et al. 2007). Subjects played an investment game, in which the researchers saw that incorporating fictive error (the difference between chosen-obtained and foregone-obtained) over gains better predicted the subject's subsequent bet than simple reward prediction error: the difference between what the subject thought she would win/lose and what she actually won/lost.

In the scenario of sequential choices of two different gambles, the difference between the results of the choice the gambler made and those of the one he did not—precisely the measure we call decision regret—can be described as a signal enlisted to learn to make better choices. That ability depends on computing that difference, then employing it to foresee a possible recurrence before the next choice is made, and finally making a different, presumably better choice (Coricelli, Critchley et al. 2005). Anticipation of regret induces a disposition to change behavioral strategies (Ritov 1996), and characterizes an emotion-motivated learning process in decision making (Zeelenberg, Beattie et al. 1996). In theories of adaptive learning driven by regret-based feedback (Megiddo 1980, Foster and Vohra 1999, Hart and Mas-Colell 2000, Foster and Young 2003, Hart 2005), learning occurs by adjusting the propensity to choose an action according to the difference between the total rewards that could have been obtained with the choice of that action and the realized total rewards. That is, the tendency of choosing machine A depends on how much would have been won by choosing that machine all along compared to how much the gambler has actually won. As gamblers, humans tend to be pretty good at this. Following regret-based learning models, decision makers converge to optimal choices (Coricelli and Rustichini 2010).

Responsibility

People show strong regularities in the nature of the event they “undo” when reflecting on a bad situation. One of these regularities, the agency effect, is particularly at stake in the experience of regret: though people feel regret both for actions taken and inaction – and although nostalgia and autobiographical retrospection often highlight missed opportunities – people in fact more often generate counterfactuals that undo some undertaken action, rather than inaction (Byrne 2002). Thus, people have greater regret for actions they have taken, more so than for those they failed to take—at least in the short term. When no action could have been taken to prevent a bad outcome, and in the absence of agency, people report feeling disappointment rather than regret. Disappointment is also elicited by counterfactual thought, though the critical outcome must be due to circumstances beyond the agent's control, absolving him of responsibility. The key distinction is this: Disappointment arises from recognizing that a better outcome might have come given the same choice; regret, from identifying a better outcome given a different choice (Zeelenberg, van Dijk et al. 1998). Both emotions come from examining outcomes and seeing that a better one could have been obtained, but regret is associated with the responsibility of having caused the sub-optimal outcome by taking a specific action. Because regret comes with the outcome of a forgone choice, it does bring with it greater information, but its effect on subsequent decisions amounts to more than simply the addition of that data. Rather, the increased information allows for the recognition of agency, along with counterfactual comparison.

Zeelenberg and colleagues sought to differentiate regret from both disappointment and a general sense of happiness by repeating and expanding on studies by Connolly, Ordoñez, and Coughlan (1997). They asked college students to consider scenarios in which fictional college students changed their class assignments — either by their own choice or by computer fiat. The results of these changes for the fictional students range from improvement to neutral to downgrade. The subjects rated how the fictional students would feel along scales measuring happiness, regret and disappointment, as well as to what extent students in the stories were responsible for their outcomes. The researchers found that happiness tracked outcome but not responsibility, while disappointment and regret were assessed inversely depending on level of responsibility: that is, the more responsibility subjects perceived, the greater the amount of regret they believed the character would feel in downgrade outcomes.

Children as young as 5 seem to have some grasp of their agency. In a choice task involving two boxes containing different amounts of stickers, children reported greater happiness or unhappiness when they chose which box to open than when the choice was determined by an experimenter or a roll of dice (Weisberg and Beck 2012). Though it was long unclear at what age the notion of personal responsibility in choices emerges, recent research suggests that agency does not influence the emotional response to outcomes in children younger than 6 (Guerini, FitzGibbon et al. in press). Using a modified Wheels of Fortune task (with stickers rather than money as the winnings) on children between ages 3 and 10, Guerini and colleagues found that children were more sensitive to the outcomes of the choice they made than those the computer made for them — but only in trials with complete feedback, and only significantly for children ages 6 and older. That is, both counterfactual outcome and responsibility were required in order for the child to feel the outcome with greater magnitude. In trials with just partial feedback, the children's sensitivity to outcomes was similar when they made the choice and when the computer made the choice — situations that generate disappointment rather than regret. This evidence of differentiation at young ages further supports the necessary role of agency in regret.

Neural Circuits of Regret

The comparison between the outcome of a choice and the foregone outcome of an alternative option triggers specific brain responses. The ventromedial prefrontal cortex (vmPFC) encodes the difference between what has been obtained and the outcome of the non-chosen option (Coricelli, Critchley et al. 2005). The vmPFC is a functional area that includes the anatomical medial orbitofrontal cortex (mOFC), an area that encompasses the most central parts of both hemispheres at the very front of the brain. The vmPFC is believed to hold on to reward value over time, possibly through tonic activity, then to send that signal to other areas involved in choice, like the dorsolateral prefrontal cortex and the medial caudate (Hampton, Bossaerts et al. 2006, Behrens, Hunt et al. 2008). Findings from neuroimaging studies support the understanding that responsibility is a necessary component of experiencing regret. Indeed, during the lottery task, activity of the OFC in response to a gain or a loss was modulated by the outcome of the non-chosen lottery (Coricelli, Dolan et al. 2007). However, when the outcome of the non-chosen lottery remained unknown, the counterfactual process between losses (or wins) and any missed outcome of the chosen lottery was accompanied by a weaker effect in OFC activity. Thus, the OFC appears to encode the counterfactual comparison between obtained and unobtained outcomes, but only when the result comes from a choice, rather than misfortune. vmPFC signals the value of the obtained outcome compared to that of the non-obtained outcome, suggesting that these regret signals are related to the way the brain evaluates choices and their consequences. It exhibits activity that correlates with regret at all stages of the choice process: preference, expectation and reward (Montague, King-Casas et al. 2006).

Correlates of regret have also been measured in parts of the brain considered to have key roles in assessing and communicating the value of choice (Nicolle, Bach et al. 2011). In neuroimaging studies, the anterior cingulate cortex (ACC) and hippocampus have also shown increased activity correlated with regret during choice tasks (Coricelli, Critchley et al. 2005). The hippocampus, a cortical folding below the cerebral cortex, is implicated in consciously accessible declarative memory, which is important for making future decisions based on past events (Coricelli, Dolan et al. 2007), such as trying to avoid previously encountered sub-optimal outcomes. This ability to guide future actions is a key component in anticipating regret based on experience.

The vmPFC increased activity during the reported experience of regret reoccurs in the period just before making subsequent choices—the period leading up to a decision in which regret would be anticipated (Coricelli, Dolan et al. 2007). Because the signal measured in the vmPFC appears in other areas, this reoccurrence suggests that the measurement is not merely of happiness, nor simply an outcome value (Coricelli, Critchley et al. 2005, Van Hoeck, Watson et al. 2015). It suggests that regret is computed by one brain area and then conveyed to others that modulate and implement it in subsequent decisions. Critically, the differentiation of experience and anticipation is clear, though they both involve the vmPFC/mOFC (Coricelli, Critchley et al. 2005). Thanks to that error signal, along with the opportunity to make a different choice, modeling regret anticipation is a reliable predictor of choice probability in certain sequential decision tasks (Coricelli, Critchley et al. 2005, Marchiori and Warglien 2008). Marchiori and Warglien found that incorporating a regret signal into even a simple learning neural network better predicted human behavior than long-employed models like reinforcement learning and a hybrid model that combines reinforcement learning with a player's beliefs about other players. Coricelli and colleagues observed that, as players experienced more regret in complete-feedback trials of a sequential Wheels of Fortune task, they decreasingly chose options more likely to lead to regret. They also saw that the more a given choice had lead to regret before, the less likely the subject was to choose it again (Coricelli, Critchley et al. 2005). Regret, then, is not merely a negative emotion, but a calculated signal that guides agents away from choices that could reproduce that signal. This effort to minimize regret is a key differentiator in its role as a learning mechanism: the emotional experience alone would have little meaning beyond sensation, were it not to guide future behavior.

The examination of choice behavior of patients with lesions in the vmPFC reveals insight into the causal link between regret-related brain activity and behavior. vmPFC patients are typically described as making disastrous life decisions despite apparently intact cognitive abilities. A famous example is the case of ERV, a patient who had a successful career and stable marital life before he developed a meningioma compressing his OFC. He then lost his job and, against his family's advice, invested all his savings in a business partnership with a man of questionable reputation. He went bankrupt, got divorced and then a month later married a prostitute, a union that lasted just six months. Yet he passed all neuropsychological tests of intellectual, memory and verbal skills with normal scores (Damasio, Tranel et al. 1990). Alongside such calamities in their daily lives, experimental evidence shows that people with vmPFC lesions display abnormal emotions elicited by reward and punishment (Bechara, Tranel et al. 1996, Bechara, Tranel et al. 2000). Careful investigation of the underlying computational deficits has revealed a general deficit in integrating values attributed to various actions with the current goals (Camille, Griffiths et al. 2011), function that has been assigned to the vmPFC in brain imaging studies. Patients are able to assign a subjective value to options; however they will not commit to the option with the highest value. Additionally, vmPFC lesions result in an inability to feel regret after a bad choice, and consequently in anticipating future regret during the decision process (Camille, Coricelli et al. 2004). Both reported subjective ratings of the outcome of their choices and the associated skin conductance responses of vmPFC patients were different from that of controls. Behavior of vmPFC patients was not significantly changed by knowing the outcome of the alternative option, an absence of the signature feature of regret. While healthy control subjects changed their choices to avoid regret over the course of the task, vmPFC patients did not.

While the fMRI and lesion studies mentioned above have identified common neural mechanisms for experienced and anticipated regret, more recent findings suggest that people with psychiatric and neurological dysfunction can exhibit one stage of the process but not another (Gillan, Morein-Zamir et al. 2014, Levens, Larsen et al. 2015). Although brain areas associated with the several stages of processing and anticipating regret overlap, they are not coextensive. Damage to the vmPFC may allow the recognition and experience of regret but not its application to future decisions (Levens, Larsen et al. 2015). Various dysfunctions of this regret mechanism offer at least partial explanations of the behavior of people with evidence of neurological disorders. Both obsessive-compulsive disorder patients and people with high indications of psychopathy report feeling regret more keenly but do not avoid it in future choices to the same extent as healthy subjects (Hughes, Dolan et al. 2013, Gillan, Morein-Zamir et al. 2014).

Moral Decision Making

The vmPFC, which represents a crucial portion of a proposed regret circuit, also plays a key role in some emotional components of moral decisions (Moll, Oliveira-Souza et al. 2002, Blair 2007, Koenigs, Young et al. 2007). Brain imaging studies of moral decision making have implicated some of the same areas and networks in the frontal cortex that are associated with emotion and deliberation – often finding these regions to be in competition during difficult choices. A study of moral judgment (without any decision component) implicated the mOFC as part of a neural circuit that showed higher activity when subjects read sentences with a moral component. The same areas, which also included the temporal pole and the superior temporal sulcus, did not show higher activation when subjects read statements with emotional components but no moral element (Moll, Oliveira-Souza et al. 2002). Researchers have developed a range of these problems to probe the spectrum of moral decision making, and this has yielded distinct differences in choice and brain activity. Among the most well-known set of dilemmas is the family that arises from the trolley problem. Subjects read about a hypothetical situation in which they are standing next to a set of railroad tracks, while some distance away, a group of workers is standing on the track. The subjects are told that they see a streetcar coming down the tracks with no chance of stopping before striking and killing the five workers. The subjects are told they are standing next to a lever, which, if they pull it, will switch the car and send the train onto a side track, where there is a lone worker who will be struck and killed. Though this would be a difficult situation in real life, in the hypothetical, it is characterized as easy and impersonal — because the subject's level of involvement from the consequences is distant and most people presented with the question answer quickly and in the same manner (Greene, Nystrom et al. 2004). Most people choose to pull the lever, making a simple utility calculation (Greene, Sommerville et al. 2001). A variant of this dilemma that brings the decision closer to the subject, however, is the footbridge problem. Now, the subject is on a bridge over the railroad tracks. He can still see the workers, and there is still a street car barreling toward them, but instead of a switch, the subject has the opportunity to save the workers by pushing a large person, who is also on the bridge, off the bridge and into the path of the street car, saving the five workers but killing the innocent person. Given simple calculation of number of people saved versus killed, these situations are identical. Yet according to measures of three features of these dilemmas identified by Greene (2007): expectation of bodily harm, agency of actor and specificity of victims, some dilemmas are more “up close and personal.” The “closeness” of the action brings the emotional salience of the problem into conflict with the pure utilitarian calculation. This antagonism seems to be carried out in the brain in both processes and areas that bear resemblance to the experience of regret (Koenigs, Young et al. 2007).

Another family of moral decisions brings an even sharper contrast. It starts with the easily solved infanticide dilemma, which poses the question of whether or not a teenage mother should kill her unwanted newborn baby. The prospect of killing a baby in service of discomfort is easily rejected, and subjects respond quickly and uniformly in the negative. Brain imaging during this decision showed lower levels of activity in the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (dlPFC), suggesting little conflict between the overwhelming emotional aversion to the choice to kill the baby and the low level of utility. Subjects also consider a more difficult analogue of this problem: the crying baby dilemma, in which subjects are asked to imagine a group of people hiding from a group of outlaws. Among the people hiding are a mother and her newborn baby, which begins to cry, which could alert the outlaws to the presence of the hiding people, resulting in the death of all of them, including the baby. Subjects are asked if it is morally permissible for the mother to smother her baby to death, saving the people but killing her own baby. Here, the calculation leads to a simple utilitarian conclusion that more people are saved by killing the baby. Yet this stands in conflict with the stark emotional opposition to killing a baby.

Observations in other brain areas support this framework. Greene and colleagues observed increased activity in ACC and dlPFC during more difficult dilemmas like the crying baby and the footbridge problems, as compared to easier dilemmas. They argue that this indicates that the ACC detects these conflicts and that the dlPFC then deliberates and resolves them. Supporting this proposal, the dlPFC shows even greater activity when the problem results in a utilitarian judgment that violates personal morality. But it is also possible that the dlPFC instigates a period of cognitive control, delaying the decision to allow the ACC enough time to employ a utilitarian cognitive response, thus overriding a more immediate affective response (Greene, Nystrom et al. 2004). If the ACC is a general arbiter of antagonism, then it is no surprise that it would be more active both in cases of difficult moral dilemmas and for discrepancies between predictions and realities, as in experiences of regret. This shared step in decision making connects the two processes and suggests that cognitive resolution of conflicts of any type may be handled with some similarity.

Notably, the several types of moral dilemma—personal and impersonal, distant and close—incorporate degrees of action, though Greene et al. (2004) differentiate between the greater agency of “authoring” and the impersonal deflection of a threat, described as “editing”. Regret similarly requires a personal agency—that responsibility attenuated only if the choice giving rise to the emotion is shared with others (Nicolle, Bach et al. 2011). The role of responsibility links the two considerations and carries the question of decision-making regret to a moral level. The more a person gauges himself responsible for an outcome, the more keenly he feels regret (Frijda, Kuipers et al. 1989). Both ranges of moral decision — those that favor utilitarian decisions and those with a greater emotional component — employ brain areas that compose part of the regret circuit. This observation suggests that the ability to feel accountable for one's choice and the phenomenon of feeling regretful in the case of a bad decision might be the premises for making non-utilitarian decisions in moral dilemmas. It does not prove the existence of a causal link between the two. Nevertheless, gathering evidence approaching a causal link, we report the cases of two different populations of patients — patients with lesions in the vmPFC and psychopaths — which exhibit a co-occurrence of difficulties with all previously mentioned processes.

OFC Lesions Modulate Regret and Morality

Patients with particular types of brain damage can demonstrate how those portions of the brain are implicated in specific processes. Brain lesions are disactivations of sections of the brain due to events like tumors, stroke or head injury. Depending on the type of precipitating event, lesions may occur in similar regions. Their specific location, while not uniform, can be established for each patient through the use of anatomical MRI and other brain scanning techniques. By comparing the behavior of healthy controls to that of patients with lesions in the same region, the role of that brain area in the process can be described. So people with lesions to areas implicated in moral decision making or regret decision making may exhibit behavior significantly different from people whose brains are fully functional in that region. Similarly, people with psychological disorders, which have brain-based causes and therefore cognitive implications, may exhibit similar types of different behavior from healthy controls.

Patients with lesions in the vmPFC, like those who demonstrated difficulty with applying anticipated regret, also exhibit trouble in following social norms. Both types of unusual decision outcome accompany damage to the vmPFC, implicating this area in a key role of both moral and regret choice. Specifically, when presented with the footbridge problem, which demands proximate action, most healthy people cannot overcome the emotional aversion of the proposition. Conversely, vmPFC patients — whose lesions deactivate portions of this brain area — exhibited utilitarian behavior, choosing to sacrifice one life in favor of five, a decision that appears to consider only the final tally of the choice and to ignore the emotional aspects (Koenigs, Young et al. 2007). In a battery of hypothetical situations, these patients were presented with choices of sacrificing one life to save multiple other lives. Among the best-known non-emotionally salient dilemmas is the trolley problem, in which the trolley is diverted by a lever onto a track with one person, avoiding the death of five. In this dilemma, vmPFC lesion patients make the choice to pull the lever about as often as healthy controls do, making a pure calculation about the impersonal action of pulling a lever. Given that these patients had impaired autonomic activity in response to emotionally charged pictures, the authors conclude that the problem in generating "normal" moral judgments come from impaired emotional processing. This was supported by two other studies showing that vmPFC patients do not experience aversive emotional responses to moral violations (Ciaramelli and Pellegrino 2011, Gu, Wang et al. 2015). When a personal element is involved, healthy people choose to intervene much less frequently (Greene, Sommerville et al. 2001). Not so lesion patients, who continue to make the utilitarian choice at about the same rate as they did in the less-emotional impersonal scenario (Koenigs, Young et al. 2007).

Importantly, vmPFC lesions also impair the experience of self-conscious emotions such as shame or embarrassment (Beer, Heerey et al. 2003). Moreover, the social behavior of lesion patients in social-norms reinforcing games has been compared to that of psychopaths (Koenigs, Kruepke et al. 2010). It should also be noted that we do not suggest that the moral dilemmas described elicit regret. Rather, because the outcome of the choice has consequences for other people, the anticipated negative counterfactual emotion involved in these situations would better be described as remorse or guilt: cognitively distinct from regret (Baskin-Sommers, Stuppy-Sullivan et al. 2016). Nonetheless, the results from the vmPFC patient studies mentioned here suggest that taking responsibility for one's own actions, questioning oneself, feeling regret and reinforcing social norms rely on the same neural circuitry.

Psychopathy

Psychopathy is characterized by diminished inhibitory control, impulsive behavior and violence. Notably, the psychiatric condition is also attended by unusual morality judgment, including the conflation of conventional and moral violations (Blair 1995). While healthy people see great differences in a conventional violation such as wearing inappropriate clothes in public and a moral violation such as hitting another person, psychopaths see less difference between the two types of transgression. Psychopaths are also more tolerant of moral transgressions against other people, which may stem from a lack of sufficient aversion to distress in others (Blair 2007). They display a similar deficiency for aversion in cost-benefit choice series.

The impaired decision making by people with psychopathic tendencies has long been attributed to their curtailed experience of emotions involving responsibility (Koenigs, Kruepke et al. 2012), but recent studies suggest that the breakdown in learning via regret happens further downstream, at the point of employing regret values in subsequent choices (Hughes, Dolan et al. 2013, Gillan, Morein-Zamir et al. 2014, Baskin-Sommers, Stuppy-Sullivan et al. 2016). This would suggest that people with psychopathy do indeed feel regret but do not incorporate the signal into future decisions, a model consistent with some findings about the moral decision making of psychopaths. Considering the implication of the vmPFC is such feed-forward mechanisms, the breakdown may well stem from a diminished vmPFC, which in psychopathic individuals, has been shown to be reduced in every dimension: volume, thickness and surface area (Yang, Raine et al. 2005, Baskin-Sommers, Neumann et al. 2016). If other considerations are equal, healthy people make the choice that carries the least expected regret, sometimes even at the cost of profit. Yet the higher people scored on a psychopathy scale, the less likely they were to avoid regret in a repeated wheels of fortune task (Baskin-Sommers, Stuppy-Sullivan et al. 2016). It was not simply that the missed opportunity bothered them less – they reported negative emotions at about the same level as controls, and sometimes even more. In fact, the highest scorers on the psychopathy self-report scale reported negative emotions after a bad outcome comparison, yet they seemed to ignore that information. The bad outcome comparison that serves as a signal to healthy people was not being used by the people with psychopathy. Their behavior indicated that they employed only the simpler signal of expected value. This suggests some link between psychopathy and regret avoidance, though a study that searched explicitly for such a connection in criminal offenders did not find one (Hughes, Dolan et al. 2013).

People with psychopathic indications are thus apparently capable of imagining alternative realities and generating and experiencing the negative emotion associated with the comparison to actual reality, suggesting that psychopathy is characterized not by a deficit of emotion but by weakened general cognitive processes like the ability to maintain previous counterfactual information and to apply it to subsequent decisions. So if these people were experiencing the emotion but apparently not employing it in choice tasks immediately following arousal, it raised the possibility that the information was not being applied to guide future choice in the manner of predictive models.

The understanding of moral processing in psychopaths is not well understood. Though people with psychopathy have long been observed to engage in amoral behavior, the mechanism of that deficiency has only recently been explored. Psychopathy has been ascribed to a depleted ability to empathize with a person being harmed as well as a deficient mechanism to inhibit violence (Blair 1995). In a study by Blair, criminal offenders considered several scenarios of moral and conventional violations set in a school, showing that psychopaths significantly did not differentiate permissibility between the two types of violations, while non-psychopaths did. Blair rejects several models in which psychopaths experience moral emotions but do not employ them in mentalization or fail to take perspectives of others. Rather, he proposes a fault in a separate system, a “violence inhibition mechanism.” Cima and colleagues (2010), by contrast, argue that while people with psychopathic traits may have some emotional deficits, enough emotion is preserved (or in fact may be unnecessary) to make similar moral judgments to healthy controls. The fact that they can identify the rightness or wrongness of moral actions, but then by definition act in contrivance, indicates that they may simply not care about morality, the study suggests. This would again be consistent with psychopaths experiencing regret but not applying it to subsequent choices. Whatever emotional component that is lacking in people with psychopathy may be the element responsible for the application of the moral understanding toward future decisions.

Yet by refining groups of people by placement on the psychopathy scale and with greater precision in the moral dilemmas presented, Koenigs and colleagues find that a counterfactual mechanism may indeed be at fault for some abnormal moral choices by people with psychopathy (2012). Using inmates from a Wisconsin prison, the study considered only those participants who scored in the highest and lowest portions of psychopathy indications, further refining the high scorers in terms of assessed anxiety in consideration of a theory that psychopathy is too broad a term for several possible conditions. Using the same situations as in the Greene study, both high-anxious psychopaths and non-psychopaths endorsed the utilitarian outcome of personal dilemmas with approximately the same lower frequency. But low-anxious psychopaths judged the utilitarian choice acceptable more often than either other group. The finding suggests that some subtypes of people with psychopathic indications resolve the emotion-utility conflict in a similarly unusual manner to that with which psychopathic people eschew regret. Where the breakdown occurs in either population and in either mechanism — or even the certainty that the causes are the same — is still up for debate: psychopaths and lesion patients may experience emotion less, or they may experience emotion and simply not apply it. Either way, it is clear that people with psychopathic tendencies do not change their choice behavior in emotional situations to the same extent that healthy people do, both after experiences that typically generate regret and when confronted with moral dilemmas.

A Social Dimension of Regret and Agency

The consideration of others connects with regret not only in representing levels of responsibility. The regret circuit co-locates with neurological phenomena that involve consideration of others via social versus private situations (Bault, Joffily et al. 2011, Zhu, Mathewson et al. 2012). Studies on levels of strategic thinking have shown higher levels associated with the same areas as counterfactual emotions like regret (Bault, Joffily et al. 2011). In an experimental game called the "beauty contest" or guessing game, the choices a player makes indicate the extent to which he is thinking about other players and how much he thinks they are thinking about him. Increased amounts of this recursive thinking are associated with higher levels of brain activity in the mOFC (Coricelli and Nagel 2009), the location of most of the vmPFC, a key component of the regret circuit. As with so many co-located brain activities, however, it is necessary to note that anatomical proximity does not necessarily indicate a functional relationship. Nevertheless, the notion of thinking about the activity in other brains (in the case of the recursive thinking demanded in the beauty contest) is different from other types of input in a similar way that the calculation and experience of counterfactual-based emotions (as in the case of regret) varies from other input—that is, it is largely internal.

Studies have associated the vmPFC/mOFC with thoughts about others (Frith and Frith 1999, Gallagher and Frith 2003, Hampton, Bossaerts et al. 2006, Suzuki, Jensen et al. 2016).

These areas become active not only when thinking about others—when evaluating violations of social norms, for example—but also when it comes to representing our own mental state, including emotion (Gallagher and Frith 2003). When subjects were directed to think about a friend or someone who was similar to them, the vmPFC showed stronger activations (Mitchell, Macrae et al. 2006). Given the vmPFC/mOFC association with processing information relevant to the self, Mitchell and colleagues suggest that thinking about related others may depend on self-evaluations in the vmPFC. This introduces the possibility of a connection between internal and external considerations: between regret's internally oriented self-evaluation and thoughts about others.

In fact, despite regret's essential interior aspect, it has been shown to be modulated by the actions of others. If an individual experiences regret that comes as the partial result of the actions of others, the brain appears to shift some of the blame for the less-then-optimal outcome to these others—thus reducing at least the anticipation of regret (Nicolle, Bach et al. 2011). As described above, measurable regret is defined by the notion of agency. It is usually addressed in a polar manner, however: with agency, the negative feeling associated with a different outcome is regret; and in its absence, disappointment (Zeelenberg, van Dijk et al. 1998). But within those categorizations, there appears to be room for gradation. Nicolle et al. had participants complete a task in which they made similar gambling choices as in standard regret tasks, but on some trials, the choice was determined not by the participant alone, but by vote (they were told) of a group of which they were a member, ranging from 2 to 8 people in all. In this case, the participant's action alone did not determine the choice and its attendant result. The measured effect saw reduced activity in the amygdala, compared to trials in which the participant was solely responsible for choices. The amygdala, implicated in emotional memory, is associated with activity involving personally relevant information. It is also known to integrate the relationship between stimulus and reward and to send it on to the vmPFC, where the information is used in subsequent choices (Coricelli, Critchley et al. 2005). So increased activity during instances of regret in which the participant is the only decision maker suggests a kind of "self-blame regret", Nicolle and colleagues argue. The diminished sense of responsibility attenuates the negative feeling of regret, and that consequently appears also to dampen the learning effect. A better response in an alternative reality becomes clearer in the amygdala with greater individual responsibility. A related question, unexplored to this point, is how, if at all, shared responsibility for positive outcomes might modulate brain activity compared to that of negative outcomes, or for positive outcomes that result from solo choices.

Conclusion

The goal of any decision process is to arrive at the optimal outcome, given the conditions. But when several important factors come into conflict in a decision, the brain must mediate among them. Separately, the processes for moral decision making and choices involving decision regret have been further explored via brain imaging and lesion studies. These have shown that segments of these processes share some anatomy and even similar dysfunction among people with psychopathy or lesions to the vmPFC. Our understanding of both systems still needs clarity before they can be considered to play any part in each other, but some recent research proposes frameworks that hint at how they might be joined. Blair argues that the learning systems in the vmPFC are the foundations of moral decisions that concern harm to other people (2007). These same systems undergird error signals that include decision regret, showing heightened activity during both the experience and anticipation of regret. The work on moral decisions by Greene and colleagues suggests that the vmPFC might serve in a regulatory role, delaying decisions during high-conflict or difficult dilemmas — especially those involving competition between emotional and utilitarian outcomes.

Moll and de Oliveira-Souza, however, push back on the Greene model, saying this conflict framework is too complex. They hold instead that the lesions attenuate the prosocial influence of the vmPFC, thus allowing utilitarian decisions without the interference of emotion. The inverse logic is that in healthy people, by contrast, the vmPFC encourages greater consideration of other people, in contrivance of purely numeric considerations. Yet this runs against the tonic activity of the vmPFC that maintains value information during a series of choices. Moral and regret decision processes appear to share patterns, but if those are reflections of shared pathways in the brain, studies to this point present contradicting roles for these areas.

Those who see the greatest connections between learning signals and moral decisions include Thomas and colleagues, who argue that the vmPFC's role is similar across reasoning processes — including moral and complex decision making. In their model, the vmPFC integrates emotion into judgments of complicated decisions, acting as adjudicator when considering future consequences (Thomas, Croft et al. 2011). The vmPFC would be responsible for assimilating the emotional effects of regret experience or imagination of harm to another into a decision that would otherwise address only the utilitarian concerns of economic value or number of people protected from harm. Such a broad function could incorporate either of the Greene or Moll/de Oliveira-Souza proposals.

Separating these competing goals and observing how special populations deviate in their decisions from they typical allows us to see that regret and morality are at least occupying some of the same space in the brain. Moral decisions play serious emotional consequences against preserving the lives (or limbs) of others. Similarly, decision regret pits the possible emotional pain of making a sub-optimal choice against maximizing gains. In both cases, the effort to avoid negative emotions comes into competition with the achieving the most utilitarian outcome. Though the implications of moral versus economic decisions are on different scales, the human brain appears to process similarly some portion of them. Crucially, they both require the previous experience or understanding of emotional outcomes and the incorporation of their possible reoccurrence into a new decision. Thus, these complex types of decision require the ability to consider the impact of the choice before it is made — they demand the conception of realities both encountered and imagined. These processes use the past and a conceptual future to put new realities in conflict with each other to judge one the most desirable.

Author note

This work was supported by a European Research Council Consolidator Grant "Transfer Learning within and between brains" (TRANSFER-LEARNING; agreement No. 617629).

Chapter 2

The effect of aging on regret in decision making

Motivating questions

How do the experience and anticipation of regret during choices evolve with age?

Is the propensity to experience regret after a bad decision linked to counterfactual learning?

Introduction

Older adults make different choices than younger adults under certain conditions, even in identical situations, suggesting an internal change in the decision process correlated with aging. These changes are evident in myriad behavioral decision-making studies, most often to worsening effect (Riggle and Johnson 1996, Denburg, Cole et al. 2007, Löckenhoff and Carstensen 2007). Older adults make hastier investment decisions and have more difficulty justifying those choices (Shivapour, Nguyen et al. 2012). Older adults are also more prone to be duped by scams and are more susceptible to deceptive advertising (Yoon, Cole et al. 2009). Aging is marked by selective areas of cognitive decline, including performance in episodic memory and executive function. While some mental processes remain stable throughout adulthood, others change in ways that result in less-advantageous outcomes (Tymula, Belmaker et al. 2013). Performance in both long-term memory and working memory tasks, which depend on processing capacity, appear to decline beginning almost as soon as adulthood is reached (Park, Lautenschlager et al. 2002). Due to these several cognitive constraints, older adults are more likely to rely on heuristic processing to make decisions (Riggle and Johnson 1996).

Such age-related decline in tasks that involve reward and learning are consistent with reduced density of dopamine receptors in brain areas implicated in encoding reward prediction error and in learning. Midbrain dopamine neurons have been robustly demonstrated to encode for reward prediction error (Schultz, Tremblay et al. 1998, Bayer and Glimcher 2005). Older adults exhibit lower dopamine transporter density in the striatum, which correlated with reduced performance in tests assessing episodic memory and executive function (Erixon-Lindroth, Farde et al. 2005, Troiano, Schulzer et al. 2010). Through the combined lenses of imaging and behavioral studies, these changes can be read as markers of cognitive decline with age.

Purely cognitive processes like working memory and reward-based learning have been studied thoroughly, if not exhaustively, but the effect of aging on other aspects of learning and decision making are less well explored – in particular the influence, if any, of emotional affect. Because the neuronal decline is different in areas implicated in emotion, they may attenuate – or exacerbate – the decline seen in strictly cognitive tasks. Cognitively enhanced emotions such as regret and envy employ counterfactual reasoning, an additional vector of learning processes (Coricelli and Rustichini 2010). Importantly, they appear to follow discrete pathways, suggesting that their influence on learning over the course of aging may be not only different from simpler models, but also from each other.

Slipping performance accompanies age in tasks that call on episodic memory (recall of words and figures and for face recognition) as well as those that employ executive functioning (visuospatial working memory and verbal fluency), which correlate with reduced density of striatal dopamine transporter, a key neurotransmitter in discerning reward (Erixon-Lindroth, Farde et al. 2005, Troiano, Schulzer et al. 2010). Age-related decline also attends structural connections between first-order reward-processing areas in the striatum and basal ganglia to higher-order areas like the prefrontal cortex (Samanez-Larkin, Levens et al. 2012). When these connections are depleted, the processes that represent value and reward predictions are attenuated, impeding identification and exploitation of rewarding decisions and more successful strategies (O'Doherty 2004). The combined effect of a reduction in dopamine density and diminished structural connection in the reward system suggest at least a partial explanation for changing choice behavior in older adults.

Similarly impeding these processes, the level of detail of both past events and future scenarios declines in older adults (Addis, Wong et al. 2008). Suddendorf and Corballis argue that mental time travel plays a crucial role in predicting future situations because it allows the recollection of previous events and the anticipation of outcomes when those events are reencountered (2007). In an evaluation of autobiographical memory, older adults demonstrate a diminished capacity for this form of mentalization, both in self-projection into future events and in situating events in the future with regard to the present (Anelli, Ciaramelli et al. 2016). This trend has bidirectional implications for learning: both a reduction in recall of consequences of previous actions and lesser ability to predict outcomes of future choices.

The types of decision contexts that reveal different performance with age can be characterized by the several types of probability situations that decision makers encounter (Mata, Josef et al. 2011). Mata and colleagues outline that *A priori* probabilities feature known probabilities and are marked by relatively easy mathematical calculations. Statistical probabilities, by contrast, demand an empirical gauge informed by experience. A third type involving rare events brings extreme uncertainty and prompts individuals to make estimates. Changes in cognitive control may modulate the assessment of the type of probability situation and therefore how to respond to it. Declines in working memory make strategy selection and application more difficult, which compounds the effects of aging since older adults tend to rely on simpler strategies that require reduced information search and integration (Chen and Sun 2003). In a quintessential example, in a gambling task, while younger people used more cognitive skills like learning and memory, older adults relied on valence of recent outcomes (Wood, Busemeyer et al. 2005). The decline in higher cognitive function is made more clear when age-related performance differences are compared to performance in decision tasks that do not feature a key learning component, in which older adults and younger adults do not significantly differ (Brand and Markowitsch 2010, Hosseini, Rostami et al. 2010, Mata, Josef et al. 2011). Meanwhile, older adults perform about as well as younger adults in memory tasks that demand only storage, such as short-term memory span tasks, compared to significant deficiency in tasks that require both storage and processing, as in working memory tasks (Bopp and Verhaeghen 2005). Such differences emerge in tasks that call on subjects to learn from feedback over time.

Emotional assistance

The story of aging, however, is not one of broad, continuous decline. Some meta-features of decision making improve with age, such as performance assessment, in which older adults seem to have greater understanding of the limits of their knowledge (Hershey and Wilson 1997). Though much work has been done to specifically characterize cognitive decline that accompanies aging, the effort is not yet exhaustive. In particular, much remains to be understood about the interaction between affective and cognitive processes in learning and decision making. Those processes that remain relatively intact and that may attenuate declines in decision making could be teased out. Key components of choice behavior, such as risk preference, have been measured as not significantly different in older subjects in some contexts (Tymula, Belmaker et al. 2013). Likewise, some affective processes are relatively resistant to effects of age (Carstensen, Turan et al. 2011). Decision-making processes that incorporate both cognitive and affective functions, with their greater and lesser susceptibility to age, may therefore decline to varying extents. In fact, older adults have been shown to focus on positive outcomes and events, perhaps at the expense of comparisons that encourage learning (Mather and Carstensen 2002). We hypothesize that decisions modulated by cognitively enhanced emotions, such as regret, may maintain stability with age, compared to more drastic declines in more purely cognitive processes, such as memory, and that how these emotions stabilize other processes depends on affective valence.

Positivity effect

A striking difference between older and younger adults is the positivity effect: the tendency for older adults to feel positive outcomes more strongly, as well as to recall them better (Reed, Chan et al. 2014). Yet the variable evidence of subsequent studies suggests that this phenomenon does not have a singular effect on learning. Various studies have investigated how age-related depletion in the density of dopamine neurons changes reward-based learning. A key consideration of feedback is valence – both as it applies to reward itself as well as to the various types of errors that feedback informs. While the absolute valence of a reward tends not to be perceived differently depending on age, the valence of its comparison to other amounts in fact may be different for older adults than for younger adults. A small negative result may have a negative prediction error compared to what was predicted, but it may have a positive error if it is compared to some other worse outcome that could have been obtained, given a different result of probability or choice.

Risk and emotion

Older adults make less-advantageous choices under uncertainty, yet depending on the type of risk that accompanies the task, tolerance of risk (measured as the variance of probability of possible outcomes) of older adults may be the same as that of younger adults, or lower, or even higher, as in tasks that call for decisions from experience (Mata, Josef et al. 2011). One explanation for this variability is that older adults may have difficulty representing changing option values, which can lead to inconsistent choices (Mata, Josef et al. 2011, Tymula, Belmaker et al. 2013, Samanez-Larkin and Knutson 2015). In the Iowa Gambling Task (IGT), the highest reward comes from learning that of two risky options, the initially less-attractive, lower-risk choice is more favorable. The Balloon-Analog Risk Task

(BART), by contrast, encourages learning that higher rewards come with riskier choices (Mata, Josef et al. 2011). In some IGT studies, most older adults start out as risk-seeking and become more risk-averse over the duration of the task (Denburg, Cole et al. 2007). Participants choose from four decks of cards in which there are both gain and loss cards. Two contain larger single-win cards and larger losses but average net losses, and the other two, smaller individual gains and losses with average net gains. Most participants start out choosing from the high-gain/high-loss decks, and healthy, unimpaired subjects eventually settling on the decks that provide long-term net gains. That deck also represents the lower-variance and therefore lower-risk deck. The IGT was developed specifically to examine the effect of emotion on cognition, including in an early representative case study of a ventromedial prefrontal cortex (vmPFC) lesion patient in which the authors measured sensitivity to reward, insensitivity to punishment or insensitivity to consideration of consequences (Bechara, Damasio et al. 1994). Individuals with vmPFC lesions, who demonstrate an impaired ability to integrate cognition and emotion, continue selecting the riskier deck in pursuit of high gains (Bechara, Damasio et al. 1997, Bechara, Tranel et al. 2000). A meta-analysis shows that risk-preference differences are more context-dependent: older adults took fewer risks than younger adults in tasks that involved learning, but more risks in tasks that did not (Mata, Josef et al. 2011). When higher rewards require increased risk tolerance, such as in the Balloon-Analog Risk Task (BART), older adults are more risk-averse throughout, in contrast to the IGT, in which higher rewards come from embracing less-lucrative and less-risky options. Other evidence suggests that risk preference may appear to shift due to differentiation in gain and loss contexts (Tymula, Belmaker et al. 2013). These different performances highlight the variability in cognitive demand. Although both tasks engage statistical probabilities and demand experience-based responses, those responses differ depending on outcome valence: in the IGT, participants should learn to avoid the initially attractive risky option to reach the optimal choice, while the BART rewards learning to become more risk-seeking. Complicating this comparison is the

BART's greater demands on calculating statistical probabilities that lead to taking higher risks. A deficiency in this type of learning in older adults means they do not make riskier choices demanded for higher reward (Mata, Josef et al. 2011).

Counterfactual learning

In standard models of reward-based learning, agents receive feedback on choices made, prompting them to adjust future actions. Although informative models can be built based only on the information from the choice made, when the outcomes of alternative choices not taken are known, models that incorporate that information better describe decision behavior. This counterfactual information, along with the comparison between unobtained and obtained outcomes, guide learning (Zeelenberg, Beattie et al. 1996). Counterfactual or fictive learning updates information about potential future choices using information from both obtained and unobtained outcomes. An affective accompaniment to this imagined alternative reality gives rise to negative or positive emotions: regret or relief. The negative feeling of regret prompts aversive behavior that guides learning. Individuals avoid potentially rewarding options if it helps them to avoid anticipated regret (Coricelli, Critchley et al. 2005).

The effect of age on counterfactual learning has been little considered. To date, just one published study addresses this question, moreover, incorporating underlying neural activity (Tobia, Guo et al. 2016). Though the study found differences in counterfactual thinking between the age groups, their characteristics and causes were not as clear. The strategic sequential investment task employed provided counterfactual information as a vector to arrive at more- or less-desirable final states following a three-choice round, experienced 10 times over the course of a block. Older adults on average invested less money and earned less. They selected the path leading to the least-rewarding outcome more often than the other three paths combined, while more than half of the younger adults developed a preference for the most lucrative path. The study found that older adults were more responsive to counterfactual gains, both behaviorally and neuronally, but that this did not lead to more rewarding choices. Computational modeling also showed that older adults explored more, while younger adults made more stable choices, which is more rewarded in this particular task.

Experimental questions

In this study, our goals were twofold: First, we examined the differences in experience of regret between older and younger adults, and second, we explored the connection between experience of regret and counterfactual learning in terms of inter-individual differences, independent of age. We aimed to further explore the characteristics and trajectories of decision-making change in older adults compared to younger adults. As shown in previous studies, the cognitive decline accompanying age is not uniform and seems to hinder various processes differently (Riggle and Johnson 1996, Erixon-Lindroth, Farde et al. 2005, Tymula, Belmaker et al. 2013). Even within a given task, performance with different demands declines to varied extents (Bopp and Verhaeghen 2005, Wood, Busemeyer et al. 2005, Brand and Markowitsch 2010, Hosseini, Rostami et al. 2010, Mata, Josef et al. 2011). The positivity effect in particular indicates that learning and effect in older adults depend on valence. Therefore, we expect different outcomes from the experience of regret, since it is a valence-dependent phenomenon. Yet it is not a direct experience, as with wins or losses, but a comparative one, based on the counterfactual. Would this give rise to the same effect as standard negative outcomes, or would it generate some sort of modulated or even inverted effect? Because older adults report lower emotional response from the experience of negative emotions like regret (Reed, Chan et al. 2014), in the first part of this study, we hypothesized that the emotional effect of regret experience and anticipation would be lower compared to that of younger-adult participants.

We next set out to examine the connection between regret and counterfactual learning. We suspected some relationship between experience and avoidance of regret in one task and the employment of counterfactual learning in another task. However, because the change in counterfactual learning with age is not clear. Due to the asymmetric decline of cognitive processes (Mather and Carstensen 2002, Carstensen, Turan et al. 2011) and the as-yet-unestablished effect of age on counterfactual learning (Tobia, Guo et al. 2016), we realized that any correlations might be evident at an individual level, but not at the level of age groups. We hypothesized that for all participants, individuals who show greater regret sensitivity would also have higher counterfactual learning rates.

To examine the varied effects of regret on older adults as compared to younger adults, we measured the performances in non-counterfactual and counterfactual contexts by age-segregated participants. We employed tasks that allowed discrete characterization of these contexts to examine any differences between age groups. An initial task was selected that would allow us to measure the level of influence of regret anticipation and avoidance for each participant. Then participants would complete a second task in which their learning behavior was described according to several computational models. This would allow us to compare regret sensitivity in the first task to counterfactual learning in the second at group and individual levels.

Methods

One group of 22 adults aged 60 and older (15 female, M_age = 70.1 ± 1.3 years, range 63-86) was recruited in Lyon, France. They were screened for a history of neurological and psychiatric disorders, as well as for depression (score higher than 10 on the Geriatric

Depression Scale, French version, Clément, Nassif et al. 1997) and cognitive impairment (score lower than 24 on the Mini Mental Examination test, French translation, Derouesne, Poitreneau et al. 1999). A group of younger adults comprised 24 participants (11 female, M_age = 24.8 ± 1.8 years, range 18-53). Groups were matched for education (older adults M_edu = 13.2 ± 0.8 years; younger adults M_edu = 14.3 ± 0.4 years).

The experimental session consisted of two main portions: a lottery choice task and a learning/post-learning task.

Lottery choice task

In the first task, participants completed a two-player choice task adapted from Bault, Coricelli et al. (2008). Participants were presented with two Wheel of Fortune lottery circles with possible outcomes of -20, -5, +5, +20 (**Fig. 2-1**). The probability of obtaining each outcome was indicated by different color segments on the wheel. Probabilities were 0.2, 0.5, 0.8. Green indicated probability of a positive outcome; red, negative. In each trial, the expected values of the two lotteries had the same valence, and the difference between the two expected values never exceeded 7 euro.

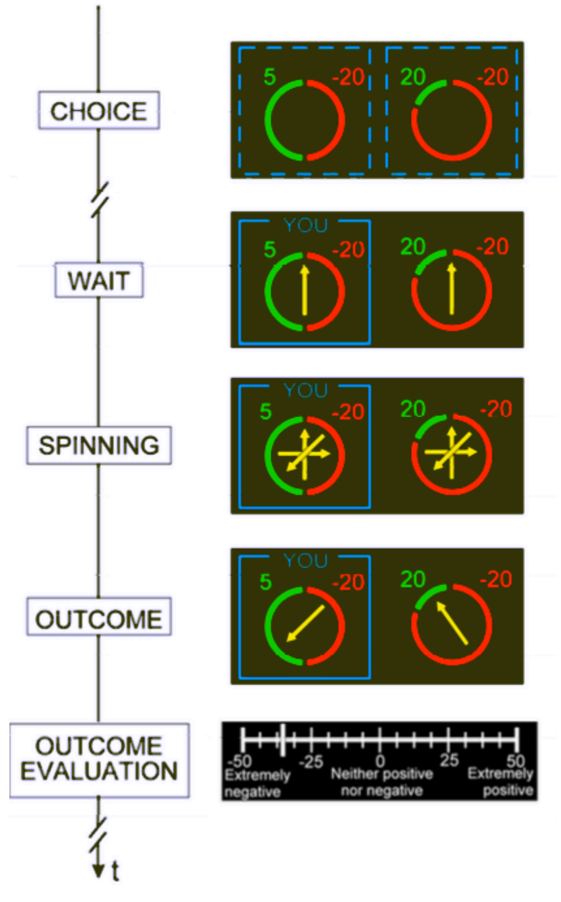


Fig. 2-1. Wheel of fortune lottery task In the complete feedback condition, the participant selects one of the two lottery wheels, then sees both arrows spinning. When the arrows stop, the participant sees both her own, obtained outcome and the outcome of the non-chosen lottery. In the partial-information condition, only the arrow of the selected lottery is shown, and the result of the non-chosen lottery remains unknown.
Adapted from Bault et al. 2019

Participants completed 80 trials with complete feedback. In these, 40 private trials were intermixed with 40 social trials, in which participants saw the outcome of another player's lottery choice. In this study, we consider both types of complete-information trials, as well as 20 trials that provided only partial feedback. In those trials, participants saw only the outcome of the wheel they chose and received no information about the outcome of the unselected lottery. To start each trial, the two wheels were displayed, surrounded by a green dashed square (in private trials). At their own pace, participants chose one by pressing the right or left arrow keys on a keyboard. In complete-feedback trials, arrows inside both wheels would spin at the same time, while in partial-feedback trials, only the arrow in the selected wheel spun. When the arrows stopped, the portion of the wheel indicated the outcome of the trial: green for the positive result, red for the negative. To encourage participants to think of each trial as independent, they were told that 20 trials would be randomly selected to determine payment. After each trial, participants gauged their emotional reaction to the outcome by answering the prompt "How do you feel about the outcome of your choice" by selecting a number between -50 (for "Extremely Negative") through 0 ("Neither Positive nor Negative") to +50 ("Extremely Positive").

Learning task

Participants then performed a two-part probabilistic instrumental learning task adapted from Palminteri et al. (2015). The first section was a learning task that manipulated outcome valence to present either reward or punishment, as well as feedback information (partial or complete), using a 2x2 factorial design (**Fig. 2-2**). Participants completed 192 trials in two blocks of 96 trials each. In each trial, a participant viewed fixed pairs of abstract symbols (Agathodaimon alphabet characters) on a screen and selected one. Each symbol appeared in the same pair, and four fixed pairs were shown 24 times throughout the block. Each pair was tied to one quadrant of the design: reward-partial, reward-complete,

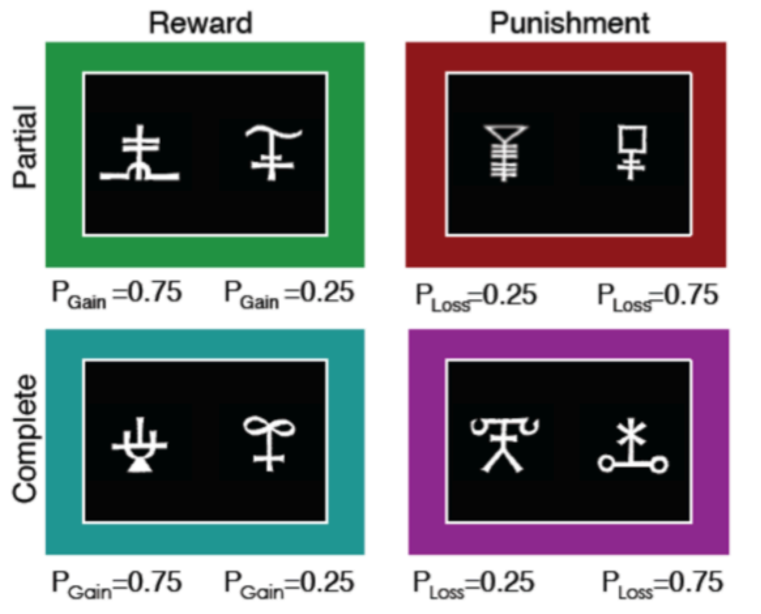


Fig. 2-2 Symbolic learning task Contingencies of the task show probabilities of winning 50 cents in the reward context (left squares) or losing 50 cents in the punishment context (right squares). On each trial, subjects saw one of the four screens, with the color background cueing the specific condition and context pair. In trials featuring the symbol pairs in the partial-feedback condition (top row), the result only of the chosen symbol are shown, while in the complete-feedback condition (bottom row), both results would be shown. Outcomes were independent.
Adapted from Bault et al. (in preparation)

punishment-partial or punishment-complete. They were presented in pseudo-random order. In the reward context, the two outcomes were gaining 50 cents or 0, neither gaining nor losing. In the punishment context, the outcomes were losing 50 cents or 0, neither gaining nor losing. In each pair, one symbol was assigned a 0.75 probability of a positive outcome, and the other was given a 0.25 probability of a positive outcome. Participants were told neither the probability amounts, nor which symbol had a greater probability of a positive outcome. The outcome of each symbol was independent of the other on each trial, so both could yield positive outcomes or both could yield negative outcomes in the same trial.

As in the previous task, participants sometimes saw outcomes of only the symbol they chose, while in other trials, they were shown outcomes of both symbols. Participants were instructed to gain as many points as possible to increase their payment. They were told that only the symbol they chose would count toward their score, even if they saw the outcome of the other symbol. In the second block, eight new symbols replaced those from the first block. Presentation on the screen was counterbalanced within pairs across the block. Values of the symbols were randomly assigned for each participant. Participants completed 4 practice trials, one for each condition, but using symbols not employed in the rest of the experiment.

A trial started with a fixation cross (0.5 s), followed by the symbol pairs. A participant selected the symbol by pressing the corresponding arrow key (self-paced). A red arrow indicated the selection for 1 s. A feedback screen indicating the outcome of either the chosen symbol (partial feedback) or both symbols (complete feedback) appeared for 3 s. In reward trials, appearing below the cue were either a 50-cent coin with the label "+0.5EUR" or a gray square labeled "0EUR". In punishment trials, the outcomes were indicated with either a gray square with the "0EUR" label or a 50-cent coin with an "X" across it and the label "-0.5EUR".

After the learning section, participants took a post-learning test of cue values in which the eight symbols from the second block only were re-presented. They were shown in unfixed pairs, each symbol appearing 4 times with every other symbol. This totaled 112 trials for 28 possible combinations. For each pair, participants were instructed to indicate the symbol they believed had the higher value, based on outcomes from the previous section. Instructions were presented only after the learning section was complete, so as not to prompt attempts to memorize values. Participants were told that symbols would not necessarily appear in the same pairs that had been presented in the previous section. Responses were self-paced, and no feedback was presented. There was no financial incentive in this part, though participants were encouraged to play as if they would be rewarded.

Emotional rating analysis

After each trial in the lottery choice task, participants rated their emotional reaction to the outcome. Trials were categorized as partial-information or complete-information feedback. To analyze ratings, we employed non-parametric tests because we anticipated violation of parametric assumptions. We estimated the significance of differences between behavioral variables and emotional ratings using the Wilcoxon signed rank test (WSRT). We tested differences between groups using the Mann-Whitney U test.

We further analyzed subjective evaluations with mixed-effects linear regressions by age group and by trial information, which allowed us to estimate both random and conditional fixed effects. Only random effects are reported. Parameters were estimated by generalized least squares.

All regressions were run using the statistical software package Stata, Stata Corp., College Station, TX. Other analyses were performed using Matlab, The MathWorks, Natick, MA.

Choice behavior analysis

The choice lottery task yields a range of information to analyze. In order to examine how components of lotteries affected choice, we considered what choice told us about an individual's weighing of expected value, risk, anticipated disappointment and anticipated regret. We further examined the effect these factors had on subsequent choices.

The aspects considered comprised difference in expected value (dEV), anticipated regret (r), and anticipated disappointment (d). These are computed, per Camille et al. (2004) as:

$$dEV = EV_1 - EV_2 = [px_1 + (1-p)y_1] - [qx_2 + (1-q)y_2]$$

$$r = |y_2 - x_1| - |y_1 - x_2|$$

$$d = [|y_2 - x_2|(1-q)] - [|y_1 - x_1|(1-p)]$$

We also considered risk, following Bault, Joffily et al. (2011), and computing it as the difference in standard deviation (dSD):

$$dSD = SD_1 - SD_2 = \sqrt{p(x_1 - EV_1)^2 + (1-p)(y_1 - EV_1)^2}$$
$$- \sqrt{q(x_2 - EV_2)^2 + (1-q)(y_2 - EV_2)^2}$$

Here, x_1, y_1 and x_2, y_2 are the two possible outcomes (x, y) of two lotteries (1, 2). The probability of the first outcome is p or q, while the probability of the second outcome is 1-p or 1-q.

A positive (negative) dEV coefficient indicates that participants were more likely to choose the lottery with the higher (lower) expected value. A positive (negative) regret (r) coefficient indicates that participants anticipated (minimized) regret and chose the lottery with the lower (higher) anticipated regret. In this calculation, participants considered what would happen if they obtained the worst outcome in the chosen wheel, compared to the better outcome in the unchosen wheel. Anticipated disappointment (d), by comparison, involved the consideration of obtaining the lower outcome on the wheel compared to higher outcome on the same wheel. Because this is calculated based on the outcomes of a single wheel, it bears some relation to risk, but as can be seen from the formulae above, it is not the same calculation. The absolute value of the difference between the two outcomes is weighted by the probability of obtaining the lower value, correlating with the notion of avoiding more probable losses. A positive (negative) disappointment (d) coefficient indicates that participants anticipated (minimized) higher and more probable losses disappointment and chose the lottery with the lower (higher) potential disappointment.

We analyzed choice behavior with multi-level mixed logit regressions with participants in groups, which allowed us to estimate both random and conditional fixed effects. Parameters were estimated by maximum likelihood.

To compare the emotional impact of counterfactual outcomes to standard, we took the mean emotional rating of each subject for all complete-feedback trials that had a better (upward-looking) obtained-other outcome than obtained-chosen and subtracted the mean emotional rating for all partial feedback trials with a negative outcome. This regret-disappointment factor is an indication of the relative strength of the counterfactual. We further described the counterfactual by performing a regression of emotional rating on all outcome values in complete trials. The contribution of the obtained-unchosen constitutes a counterfactual coefficient for each subject. We then categorized all subjects into either "weak" or "strong" counterfactual coefficients at the median point (-0.3528, range: [-2.2013, 0.6419]). Lower amounts indicated that greater missed opportunities had a more negative emotional effect, so the lower half composed the strong group.

Learning behavior analysis

From the learning task, we extracted several variables, including earnings and correct choice rate as dependent variables. A "correct" response was determined to be either the more-rewarding choice or the less-punishing choice.

We looked separately at earnings in complete-feedback trials and partial-feedback trials. Because earnings may vary absolutely on an individual basis, we wanted to see how earnings in the two types of trials compared. To obtain an individually internal comparison, we subtracted each participant's mean earnings in partial-feedback trials from the man of earnings in complete-feedback trials to arrive at a regret-disappointment differential score for each participant.

We performed statistical analyses on these variables using Mann-Whitney U-tests, Wilcoxon Signed Rank Tests and groupings by counterfactual-coefficient, age group and between task information feedback type.

Learning computational models

Data was previously analyzed with four reinforcement learning models: Q-learning, counterfactual learning, normalized Q-learning and normalized counterfactual learning (Bault, Palminteri et al. 2018). In the normalized model, learning is considered to occur relative to the average value of context, which allows learning to occur from zero-value outcomes, which are frequently encountered in this task but still provide information. The reinforcement learning models operate on direct experience by updating values only for chosen options based on outcome. Counterfactual models update both the chosen option and the unchosen, when counterfactual information was available.

All four produced a softmax parameter that indicates how selectively an individual discriminated between the two options and a standard learning rate parameter. The two counterfactual models also generated a counterfactual learning rate parameter. The two contextual models also generated a contextual learning rate parameter.

Results

Emotional ratings

As in previous studies, regret in this task is characterized in two ways: by an effect of the outcome of the unchosen lottery on the emotional evaluation, and by a stronger influence on that evaluation of the unchosen lottery outcome in the complete feedback condition than of the unobtained outcome of the chosen lottery in the partial feedback condition. In both instances, it is the imagination of an obtained outcome in an alternative world given a different choice that drives the effect.

Emotional ratings across all trials were not significantly different between older and younger age groups, apart from trials that resulted in regret (complete-information, private trials with a worse outcome than the obtained outcome in the unchosen lottery) (**Fig. 2-3**). Younger adults (YA) reported feeling worse than older adults (OA) (complete-private trials, Mann-Whitney, Z=1.97 , p=0.0484). In situations of large upward comparison – when the obtained outcome is less than the unobtained – OA rated emotions higher, indicating that in some contexts, they experienced regret less than YA (**Fig. 2-4**).

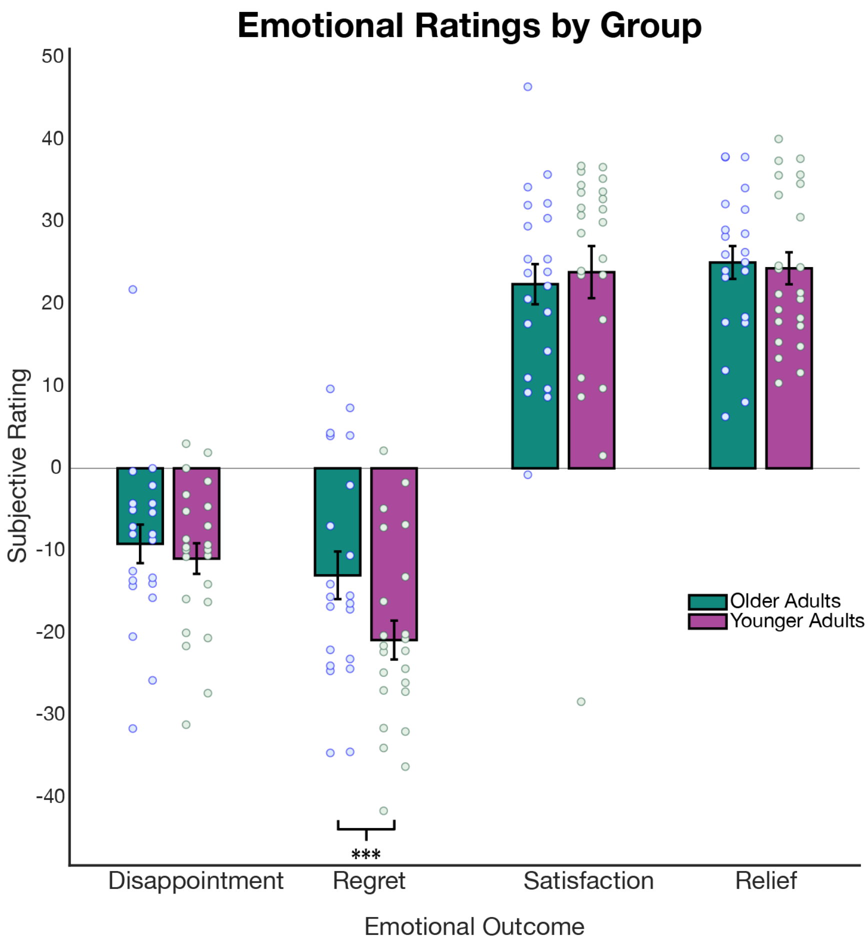
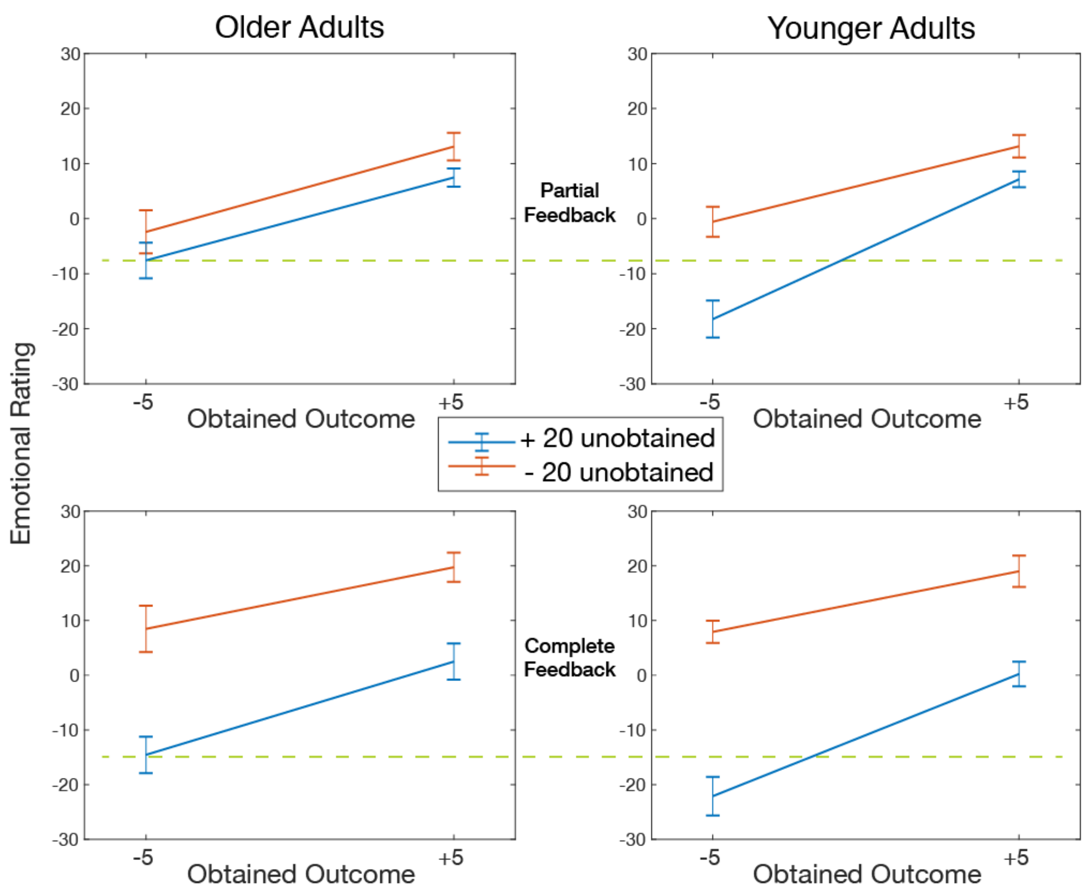


Fig. 2-3. Subjective emotional ratings in lottery task Mean subjective emotional ratings (error bars: s.e.m.) by older adults and younger adults following partial feedback trials ("Disappointment" and "Satisfaction") and complete feedback trials ("Regret" and "Relief").

Fig. 2-4. Emotional ratings comparison by unobtained outcome Mean emotional ratings by older adults and younger adults for two obtained outcomes (–5 and +5) as functions of the unobtained outcome (blue line, +20; red line, +20) in partial and complete feedback conditions. In partial feedback, the unobtained outcome is the unobtained amount in the selected lottery; in complete feedback, it is the obtained amount on the unchosen lottery.



For all participants, emotional reactions were significantly affected by both the amount of the outcome (Subjective rating, Obtained-chosen, partial-feedback trials, Coeff. 1.392, $Z = 41$, $p < 0.001$) and the amount of the unobtained result of the chosen wheel (Subjective rating, Unobtained-chosen, partial-feedback trials, Coeff. -0.102, $Z = -3.17$, $p < 0.01$) in partial-feedback trials. This relationship held true in complete-feedback trials as well, with similar effects of both the amounts of the obtained outcome (Subjective rating, Obtained-chosen, Coeff. 1.345, $Z = 53.78$, $p < 0.001$) and unobtained-chosen (Subjective rating, Unobtained-chosen, Coeff. -0.126, $Z = -5.12$, $p < 0.001$). However, the emotional rating of the obtained outcome was modulated to a greater extent by the obtained outcome of the un-chosen lottery (Subjective rating, Obtained-unchosen, complete-feedback trials, Coeff. -0.305, $Z = 12.68$, $p < 0.001$) than the unobtained outcome of the chosen lottery, demonstrating an amplification effect. That is, the result of the lottery that was not chosen had a greater effect on the emotional rating than the result that was not obtained in the chosen lottery.

Both age groups gave increasingly negative ratings with higher obtained amounts in the un-chosen lottery, meaning the higher the amount they could have obtained with a different choice, the worse they felt. But the effect this had on the rating was significantly different for the two groups (**Table 2-1**), with YA having a greater negative reaction to higher values in the obtained amount on the un-chosen lottery than OA (**Table 2-1**). Here, the negative coefficient indicates the adverse reaction to a higher value. YA had a stronger negative reaction to missed opportunities, i.e. the higher value in the un-chosen lottery. OA, though still reporting a negative emotion, did not experience it as much, indicating that they experienced regret to a lesser extent for the same amount of a missed opportunity. No other value had a significant interaction by group.

In YA, emotional reactions to upward counterfactual comparisons (i.e. relative losses) were significantly stronger in the complete-feedback condition than in the partial-feedback condition (WSRT, Z=-3.77, $p<0.001$). The same comparison was not significantly different among OA (WSRT, Z=-1.76, $p=0.079$).

A key characteristic of regret is the greater affect reported in upward-looking complete-feedback trials compared to upward-looking partial-feedback trials. The individual-level differences between ratings in upward-looking complete-feedback and upward-looking partial-feedback was stronger for YA (all trials, Mann-Whitney, Z=2.11 , $p=0.0345$).

Choice behavior

We tested a model of choice that comprised as choice predictors the difference in expected value of the two lotteries, risk and anticipated regret. Mixed logistic regressions

Table 2-1. Scale: Age group

	Younger			Older		
Subjective ratings	Coeff	Std Error	Z	Coeff	Std Error	Z
obtained-chosen	1.413	0.048	29.49***	1.364	0.050	27.21***
unobtained-chosen	-0.067	0.046	-1.470	-0.118	0.050	-2.34*
obtained-other	-0.457	0.045	-10.14***	-0.290	0.050	-5.8***
unobtained-other	0.068	0.049	1.370	0.042	0.052	0.800
Group	3.285	1.964	1.670			
obtained-chosen X group	-0.049	0.069	-0.700			
unobtained-chosen X group	-0.051	0.068	-0.750			
obtained-other X group	0.167	0.067	2.48*			
unobtained-other X group	-0.026	0.072	-0.360			
constant	2.337	1.358	1.720			
	complete-private feedback trials			complete-private feedback trials		
	Wald Chi2 = 4088.89***			* p < .05; ** p < .01; *** p< .001		

Table 2-1. Scale: Effects of potential outcome amounts on emotional self-evaluation by age group. Mixed-effects linear regressions modeling the effect of lottery components on rating in older adults and younger adults. Both groups had stronger negative emotional ratings with higher amounts obtained in the unchosen wheel, that is, missed opportunities. Younger adults, however, had a significantly stronger negative reaction than older adults.

showed that while both groups sought higher EV when choosing which lottery to play (**Table 2-2**), YA were more likely to do so (**Table 2-2**). Expected Value is the variable with the greatest influence on choice, and higher EV has a greater influence on the choices of YA than on OA.

YA did not significantly account for risk (**Table 2-2**), but OA did, avoiding it across all trial types (**Table 2-2**) to significantly greater extent (**Table 2-2**). That preference is driven largely by complete-feedback trials, because OA consider risk to a significantly higher extent in those compared to partial trials (**Table 2-4**). The same regression for partial-information trials was not significant in either OA or YA. These analyses suggest that OA are more risk averse and less considerate of expected value than YA.

Because the variable *r* represents anticipated regret, a positive coefficient indicates an attempt to avoid regret by choosing the lottery with the smaller difference between worst outcome on chosen and best outcome on unchosen. Both YA and OA minimize regret across

Table 2-2. Choice: EV/risk/regret - Age Group

	Younger			Older		
Choice	Coeff	Std Error	Z	Coeff	Std Error	Z
dev	0.155	0.012	12.86***	0.069	0.012	5.76***
dsd	0.013	0.008	1.550	-0.037	0.008	-4.42***
r	0.015	0.003	4.77***	0.025	0.003	7.57***
Group	-0.011	0.064	-0.170			
dev X group	-0.086	0.017	-5.06***			
dsd X group	-0.050	0.012	-4.26***			
r X group	0.010	0.005	2.21*			
constant	0.051	0.044	1.150	0.040	0.045	0.900
	partial, complete-private, complete-social trials			partial, complete-private, complete-social trials		
	Wald Chi2 = 281.31***			Wald Chi2 = 111.67***		
	* p < .05; ** p < .01; *** p< .001					

Table 2-2. Choice: Effects of expected value, risk and regret on lottery choice by age group. Multi-level mixed logit regression modeling the effect of lottery components on choice in older adults and younger adults. Both older and younger adults made choices at significant levels that favored expected value and minimized anticipated regret. Older adults significantly made choices that minimized risk, while younger adults did not.

Table 2-3. Choice: EV/risk/regret - Younger Adults						
	Partial			**Complete**		
Choice	Coeff	Std Error	Z	Coeff	Std Error	Z
dev	0.121	0.026	4.60***	0.165	0.014	12.08***
dsd	0.018	0.018	0.970	0.011	0.009	1.260
r	0.021	0.007	3.00**	0.013	0.003	3.83***
Feedback	0.064	0.108	0.600			
dev X feedback	0.044	0.030	1.490			
dsd X feedback	-0.006	0.020	-0.300			
r X feedback	-0.008	0.008	-1.000			
constant	-0.000	0.097	0.000	0.064	0.049	1.320
				complete-private, complete-social trials		
	Wald Chi2 = 174.00***			Wald Chi2 = 150.98***		
	* p < .05; ** p < .01; *** p< .001					

Table 2-4. Choice: EV/risk/regret - Older Adults						
	Partial			**Complete**		
Choice	Coeff	Std Error	Z	Coeff	Std Error	Z
dev	0.081	0.027	2.98**	0.067	0.013	4.98***
dsd	0.015	0.019	0.760	-0.050	0.009	-5.31***
r	0.031	0.008	4.07***	0.024	0.004	6.44***
Feedback	0.076	0.112	0.680			
dev X feedback	-0.013	0.030	-0.450			
dsd X feedback	-0.065	0.021	-3.01**			
r X feedback	-0.007	0.008	-0.840			
constant	-0.020	0.100	-0.200	0.056	0.051	1.110
				complete-private, complete-social trials		
	Wald Chi2 = 119.40***			Wald Chi2 = 99.91***		
	* p < .05; ** p < .01; *** p< .001					

Table 2-3. Choice-younger: Effects of expected value, risk and regret on lottery choice by trial type in younger adults. Younger adults made choices at significant levels that favored expected value and minimized anticipated regret in both types of trials. They significantly minimized risk in complete feedback trials only.

Table 2-4. Choice-older: Effects of expected value, risk and regret on lottery choice by trial type in older adults. Older adults made choices at significant levels that favored expected value and minimized anticipated regret in both types of trials. They significantly minimized risk in complete feedback trials only.

both types of trials, but not to any significantly different extent (**Tables 2-3, 2-4**). Surprisingly, a trend shows that higher anticipated regret has a greater influence during choice in partial-feedback trials for both age groups. Although participants would be aware that they would not see the outcome of the unchosen lottery in partial-information trials, they may still anticipate regret as an ordering effect, due to the intermixing of trial types throughout the task. Both groups anticipate regret, with OA doing so to a greater extent across all trials, but not significantly so.

Learning behavior

We compared earnings from the learning task by age group in complete-information trials and found that YA earned significantly more than OA in complete-information feedback trials in the learning task (Mann-Whitney, Z= -2.177, p = 0.0295) (**Fig. 2-5**). Both age groups earned significantly more in complete-information trials than in partial-information trials (YA, WSRT, Z = 3.34, p < .001)(OA, WSRT, Z = 2.24, p = .0249) (**Fig. 2-5**). Then we compared earnings differentials between types of trial. For each subject, we subtracted earnings in partial-feedback trials from earnings in complete-feedback trials. There was no significant difference between OA and YA in this earnings differential in the learning task (Mann-Whitney, Z = -0.82, p = 0.41) (**Fig. 2-6**).

We then considered two types of player measurements from the lottery task, dividing all participants according to a median counterfactual coefficient, as calculated from the lottery task. We compared the two groups, asking if either one earned significantly more than the other in complete-information trials (**Fig. 2-7**) and found that the half of subjects who felt worse about missed opportunities in the lottery task (stronger counterfactual-coefficient) earned significantly more than the weaker group (Mann-Whitney, Z = 1.96, p = 0.0495). We then asked if either group earned more in complete-feedback than in partial-feedback trials

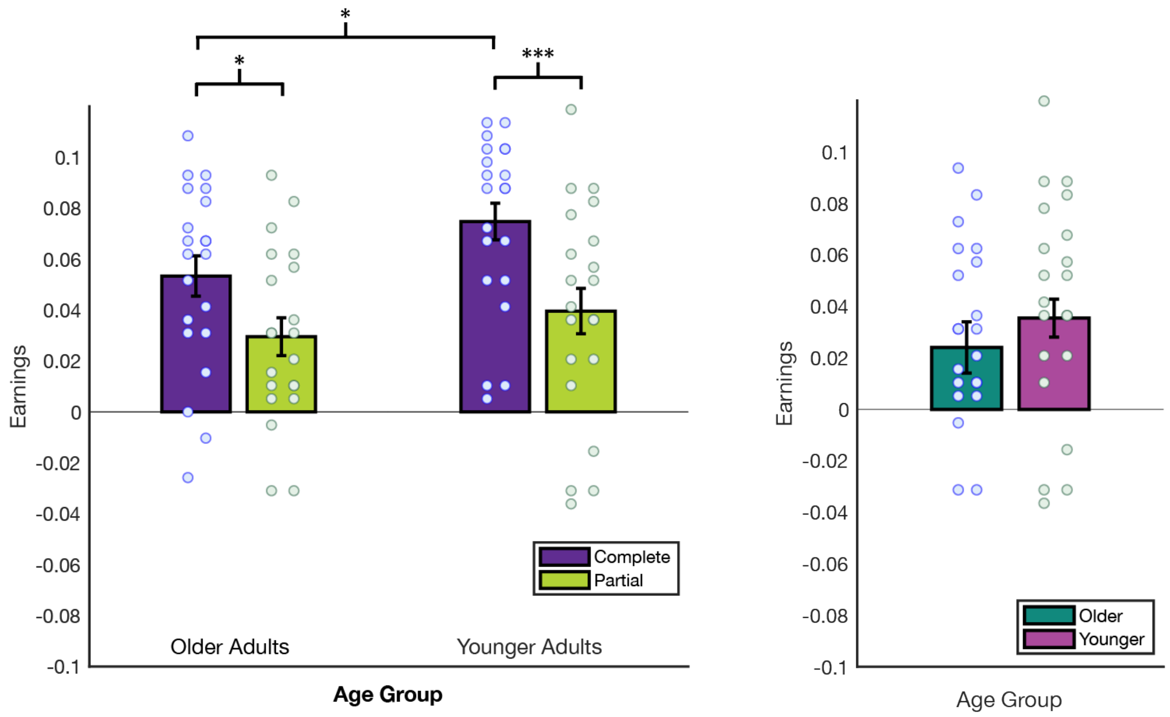
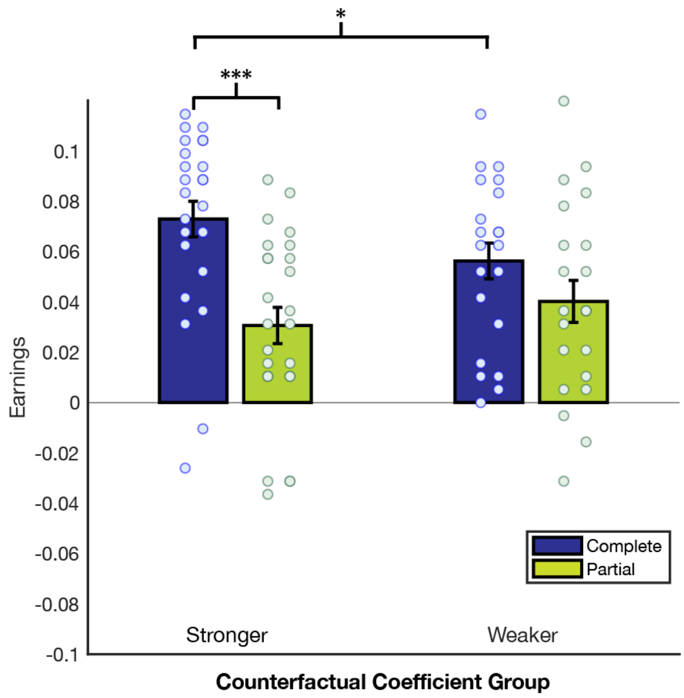


Fig 2-5. Learning task: earnings by feedback and age group Mean earnings (error bars are s.e.m.) of older adults and younger adults in complete feedback trials and partial feedback trials in an instrumental learning task. Circles represent individual mean earnings for the feedback type they overlay.

Fig 2-6. Learning task: earnings differential by age group Mean earning differential (and s.e.m.), calculated as the amount earned in complete feedback trials, less the amount earned in partial feedback trials.

in the learning task (**Fig. 2-7**). The difference between earnings in feedback types was not significant in the weaker counterfactual-coefficient group (WSRT, Z = 1.69, p = .0911. The stronger counterfactual-coefficient group, however, earned significantly more in complete-information trials compared to partial-information trials (WSRT, Z = 3.93, p < .001).

Next we considered the complete-partial earnings differential, as we did age groups. Both counterfactual-coefficient groups on average earned more in complete-feedback trials of the learning task compared to what they earned in partial-feedback trials (**Fig. 2-8**). The stronger counterfactual-coefficient group, however, had a significantly higher differential than that of the weaker counterfactual-coefficient group (WRST, Z = 2.21, p = 0.027). Their earnings differential between complete and partial trials were higher by a larger margin than those who were less unhappy about the counterfactual outcome in the lottery task. That weaker group followed the same trend but did not have a significant difference.



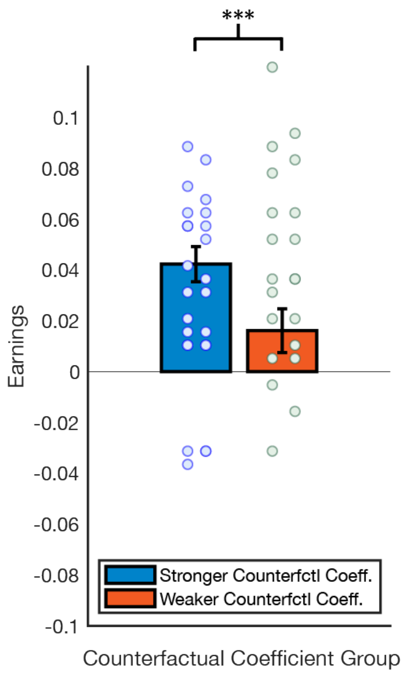


Fig. 2-7. Learning task: earnings by feedback and counterfactual-coefficient group Mean earnings (and s.e.m.) of Stronger and Weaker counterfactual coefficient subjects in complete feedback trials and partial feedback trials in an instrumental learning task. Circles represent individual mean earnings for the feedback type they overlay.

Fig. 2-8. Learning task: earnings differential by counterfactual-coefficient group Mean earning differential (and s.e.m.), calculated as the amount earned in complete feedback trials, less the amount earned in partial feedback trials.

Discussion

Older and younger adults performed a lottery choice task, followed by an instrumental probabilistic learning task. In both tasks, behavior could be guided by measurable factors, such as risk, expected value, anticipated disappointment and anticipated regret. In both tasks different feedback configurations made possible decision comparisons. In partial-information trials, subjects saw the result of their choice only. In complete-information trials, they saw the outcome of the choice they decided against, introducing elements of foregone possibilities and responsibility, therefore enabling counterfactual thinking. In the lottery task, we aimed to gauge individual and group sensitivities to counterfactual outcomes. We asked if comparing these measurements to learning behavior in the subsequent task would reveal differences at an inter-individual level, thus suggesting implications of regret sensitivity on emotion-related counterfactual thinking. We hypothesized that emotional experience and effect in the gambling task would correlate differently with counterfactual learning in the learning task both at an individual level. We categorized subjects by these measurements to gauge how choices in and reactions to the lottery task correlated with performance in the learning task.

Both age and behavior characterization in the lottery choice task yielded significant effects in the learning task. Subjective emotional ratings in the lottery choice task indicate that younger adults felt worse when they saw the better results of a lottery they did not choose. This negative reaction to a missed opportunity is the experience of regret. The affective rating of older adults was also significant, but not as high. Despite this disparity in the experience of regret, we found that both age groups chose to minimize regret more often and at about the same rate. The notable differentiation from the lottery choice task, then, was that while older adults appear to be less disturbed by regret situations, they continue to employ anticipation of regret as an avoidance behavior.

In the learning task, participants faced situations in which they might encounter disappointment and other situations that could bring regret. It is in these complete-information feedback trials that we would expect them to employ anticipated regret as a learning signal. Indeed, both age groups demonstrated greater ability via higher earnings in complete feedback trials over partial feedback trials. Younger adults showed greater mean earnings in complete feedback trials. This may have been an effect of higher earnings across all trials, however, since the differential between mean earnings in complete feedback trials versus partial feedback trials was not significantly different between the two groups.

To gain a better comprehension of how performance in the two tasks interacts, we considered how much the alternative reality of the outcome on the unchosen lottery wheel influenced the emotional rating. We expected this counterfactual coefficient to indicate a strength of experience of regret. Though YA and OA performed differently in the learning task, we did not find a significant difference by age group in the correlation between experience of regret and use of counterfactual learning. We did find, however, that emotional experience correlated with counterfactual learning at an individual level, as we hypothesized. We found that the individuals who felt this more strongly had a similarly significant differentiation between earnings in complete feedback trials in the learning task versus partial feedback trials. The weaker-experience group did not have a significant differentiation, indicating that the lesser experience of regret may lead to a diminished ability to employ counterfactual thinking in learning tasks. The stronger counterfactual group earned more than the weaker group in complete feedback trials, a differentiation confirmed by a significantly higher complete-partial differential in the stronger counterfactual group compared to the weaker group. Individuals who felt worse about regret, according to their ratings during the lottery task, appeared to be more motivated in complete feedback trials during the learning task at a higher rate than those who were less affected by regret. One explanation for this is that the emotional effect that was strong enough to significantly guide choice in the lottery task soon after continued to use counterfactual learning to greater reward by avoiding aversive outcomes.

Aging is accompanied by reduced preference for negative stimuli in both attention and in memory. This well-established positivity effect emerges in middle and late adulthood. Based on the positivity effect, a decreased attention to negativity should lead to reduced experience of regret and therefore reduced anticipation and avoidance of regret. Yet only the first part of that conjecture bears out, raising the question of how important the experience of regret is to its later use in anticipation for avoidance. We suspected that the positivity effect would yield reduced learning from regret situations in older adults. We saw some evidence of that in the lower experience of regret in the lottery choice task among older adults, as well as some indications in the learning task, in which they earned less than younger adults in situations that might employ regret learning. We also considered that the emotional component of counterfactual learning might stabilize declines in performance that have been shown to otherwise accompany aging.

Brassen and colleagues found that, following outcomes that would elicit regret, older adults showed increased activity in the anterior cingulate cortex, an area associated with emotional regulation (2012). The authors propose that this is a cognitive-control mechanism that re-assesses regretful experience as less negative. They also suggest that healthy older adults externalize the causes of regret situations, attributing them to factors they could not control and removing the responsibility that is a key component of regret. This is consistent with the positivity effect, which states that minimizing negative experiences in older adults is not an emotional regulation strategy, but rather goal-oriented cognitive processing (Carstensen and DeLiema 2018). According to the socioemotional selectivity theory, this is consistent with aging, since it is accompanied by changing goals (due to diminishing time horizons) that trigger increasing occurrence of the positivity effect. Brassen and colleagues argue that the positivity effect in general and diminished regret experience in particular are adaptive for emotional well-being in older age. Disengagement from regret constitutes a special case of the positivity effect. In addition to the emotional well-being derived from avoiding negative emotions, disregarding regret experience is protective for older adults because they have reached a time in life when opportunities to undo regretted behavior are diminishing to the point of vanishing (Brassen, Gamer et al. 2012).

Previous studies are mixed in their assessment of how risk preference changes, if at all, with age, and the context of how risk can influence learning and earnings seems to play outsize importance in how it changes. Older adults in this study accounted for risk in the lottery choice task, while younger adults did not, suggesting that minimizing risk becomes more important with age in a task that does not have a tendency of rewarding or punishing risk seeking. Our findings expand on those shown by Tobia and colleagues (Tobia, Guo et al. 2016), who found that older adults were more responsive to counterfactual gains. We found that older adults are conversely less responsive to missed opportunities. We further saw, however, that this does not give rise to any significant difference in choice behavior.

These results suggest differences in the experience and anticipation of regret in decision making and learning. The wheel of fortune lottery is a reliable indicator of preference for risk, anticipation of regret and importance of expected value. To further explore the relation between choice behavior and counterfactual learning, the lottery task could be paired with other learning tasks that can employ regret learning.

Due to the necessity of two types of complete-information feedback trials in the lottery choice task, partial-information feedback trials were limited to just one-quarter of all the complete feedback trials, making regression comparisons to the complete feedback trials less reliable. A future study could increase the number of partial feedback trials.

Chapter 3

Priming regret: inducing counterfactual thinking to influence learning

Experimental questions

Can the experience of complete-information counterfactual, specifically regret, in one task modulate learning in a separate task?

If so, does the transfer encourage more sophisticated learning behavior?

Introduction

In decision making, the brain compares actual outcomes of choices to other possible outcomes, both alternatives from the choice made and those foregone (Loomes and Sugden 1982). This counterfactual information from imagined alternative realities gives rise to a set of emotional situations. Alternative choice outcomes are not always known to us, but when we do see them, and the comparison between those and actual outcomes correspond to the emotions that we label regret or relief (Zeelenberg, van Dijk et al. 1998). These are distinct from the emotions we call disappointment and satisfaction, which are still counterfactual emotions but result in the comparison to alternative outcomes of the same choice due to nature. A signature component of regret is agency: it requires the imagining of an alternative reality that could have been realized via a different choice. The information from these comparisons can guide learning within a task (Zeelenberg, Beattie et al. 1996). But in the framework of Transfer Learning, we ask if this mechanism can cross over from one decision context to another: If learning from counterfactual emotions in one task produces a different state, would that modulate learning in a different task?

In various states, people may be better or worse prepared to learn rapidly (Young and Nusslock 2016). Several learning models with robust reflections of behavioral and brain activity suggest that learning incorporates not only choices made and their actual outcomes, but also paths not taken and imagined rewards or punishments from some alternative reality

(Zeelenberg, Beattie et al. 1996, Camerer and Ho 1998, Zeelenberg, van Dijk et al. 1998, Fudenberg and Levine 1999). The weight given to these alternative realities can vary from session to session, person to person, and even trial to trial. Certain types of thinking may manifest as different behavioral strategies. Counterfactual thinking, including the cognitively enhanced emotions drive learning within the same task (Camille, Coricelli et al. 2004). Traditional models of regret are based on adaptive learning, in which the probability of making a choice varies depending on the difference between actual reward and the rewards that option would have yielded if it had been chosen in the past (Foster and Vohra 1999, Hart and Mas-Colell 2000, Foster and Young 2003).

In a repeated game, a player who recognizes that a different strategy would have brought higher reward if she had made a different choice could change strategies in the next iteration (Coricelli, Dolan et al. 2007). Moreover, simulated behavior in neural networks showed that incorporating regret into choice models yielded improved performance (Marchiori and Warglien 2008). This improvement appears not to be a phenomenon purely of additive information, learning actual and imagined outcomes of a single game, but rather the trigger for a particular learning mechanism. Coricelli and colleagues observed BOLD activity that led them to suggest this integration of cognition and emotion occurs in the orbitofrontal cortex (OFC) following feedback from a decision, in this case a gambling task trial, but before the presentation of a subsequent choice. The supplemental emotional component of regret raises the possibility of sustained affect that may transfer into an unrelated and novel task, and once there, potentially accelerating learning as it has been seen to do within a single repeated game.

After experiencing regret, individuals make choices to avoid the negative feeling, often in violation of normative behavior (Ritov 1996). The emotional motivation of avoiding regret can modulate choices away from purely rational expected utility. In an ongoing, adaptive context, it constitutes a learning behavior (Zeelenberg, Beattie et al. 1996). Punishment as effective learning signal presents a paradox, because once individuals learn to avoid it, the reinforcer is not encountered. Instead, successful avoidance, which is intrinsically neutral, gains a positive sign by way of the counterfactual, comparing the successful, neutral avoidance to another possible outcome that would have been worse (Kim, Shimojo et al. 2006, Palminteri, Khamassi et al. 2015). In adaptive learning models, the regret signal modulates the tendency to make a given choice by comparing the rewards that choice would have brought and actual rewards (Megiddo 1980, Foster and Vohra 1999, Hart and Mas-Colell 2000, Hart 2005). In regret models in repeated games, the probability of switching to another choice varies depending on how much reward that choice would have brought if it had been chosen throughout the game, compared to actual reward (Hart 2005). Players that minimize regret converge on optimal solutions, sometimes more quickly or with fewer loses than those who do not minimize (Coricelli and Rustichini 2010). Because regret carries a connotation of affective influence, some, including Lohrenz and colleagues (2007), differentiate between the emotion regret as described by Bell, Loomes and Sudgen (Bell 1982, Loomes and Sugden 1982) and the signal the later study observes as the result of Fictive Learning. Though both terms are used in the literature, here, we use regret to describe this effect.

Measures of primate dopaminergic neurons examined the temporal difference (TD) model in situations that correspond to disappointment (e.g. a conditioned stimulus not followed by an expected reward) (Dayan 1994, Schultz, Dayan et al. 1997, Schultz, Tremblay et al. 1998, Schultz 2002) and showed that dopamine neurons react not to reward itself but to reward prediction error, the difference between expected reward and received (Schultz et al. 1997). Counterfactual outcomes appear to register in single neurons in the anterior cingulate cortex (ACC) (Hayden, Pearson et al. 2009). In the task, monkeys selected one of eight screen positions in an effort to get a single highest juice reward. Single neurons showed higher activity when the best option was chosen, varying depending on the magnitude of that option in a given trial. Yet even when the monkey failed to choose the best option but saw the magnitude of that missed reward, the ACC neuron (and local population) encoded it at the same magnitude relative to best-option amounts in other trials. More recent work in humans showed that a model combining prediction error and counterfactual learning better predicted striatal dopaminergic activity (Kishida, Saez et al. 2016). This work suggests a more elaborate and complex role for dopamine, encoding not only rewards and losses, but also the results of computations comparing actual outcomes to alternative, imagined realities (Platt and Pearson 2016).

Hsu & Zhu (2012) compared neuronal manifestation of two models of regret-related learning in a competitive game, finding broader evidence for a belief-based Fictitious Play model. One RPE signal corresponded to the difference between the received reward and the highest possible, as with the monkeys, which does not take into account the actions of the opponent. This signal correlated with activity only in the bilateral putamen. The fictitious play-based signal by contrast considered the difference between actual reward and an expected reward that is based on the frequency of the opponent's previous choices. This signal correlated with activity in the bilateral putamen and in an area comprising the mPFC, the OFC and the ACC.

Experience-Weighted Attraction model

Previous research has described learning by comparing players' choice strategies to these various learning models. Reinforcement learning (RL) models best describe players who value strategies that have paid off in the past. Because reinforcement learners pay the most attention to their own choice history, their behavior is often marked by sequentially dependent choices – that is, choices that have led to positive outcomes are most likely to be made again. Though RL has some clear shortcomings because it does not incorporate all available information, it does describe well behavior in a number of mixed-strategy games

(Roth and Erev 1995). Belief-Based learning (BBL) models, by contrast, reflect the choices of players who take into account their opponent's decisions. In BBL, the outcomes of decisions not made are incorporated, taking into consideration beliefs about the actions of another player, allowing subsequent choices that have not been rewarding before or may not even have been made previously.

Players seem to incorporate some combination of these types of play, even varying the amount within a series of choices (Ansari, Montoya et al. 2012). To capture the relative use of each type of learning, Camerer and Ho developed the hybrid Experience-Weighted Attraction (EWA) model, which nests both Reinforcement and Belief-Based learning models. Behavior that reflects either of these models is accounted for in EWA, and EWA's key benefit is its production of parameter ∂ that indicates the relative weight of action values. In RL, the most recently rewarding decisions are most valuable and attractive and therefore continue to be chosen. The BBL model nested in EWA is itself a nested model of several belief models, including fictitious play (Fudenberg and Levine 1998), which incorporates all past actions. But unlike BBL, fictitious play applies no temporal decay. In an update rule for the attraction to strategy *k*, the decay of the strength of past attractions is the weighted parameter Φ, as in

$$A_{ik}^{B}(t) = \frac{\phi_i A_{ik}^{B}(t-1)N_i(t-1) + \pi_i(s_i^k, s_i(t))}{\phi_i N_i(t-1) + 1}$$

where $A_{ik}^{B}(t)$ is the attraction of strategy k to individual *i* after time period *t* and $N_i(t)$ represents updated game experience (Ansari, Montoya et al. 2012). The parameter Φ indicates the level of fictitious play (Φ = 1) versus single-period Cournot belief learning (Φ = 0), and π is the payoff function.

When behavioral data reflects more sequentially dependent play, EWA-modeled behavior indicates RL, whereas when data indicates deletion of dominated strategies, the model should yield BBL. The tendency toward one strategy or the other derives from a weighted combination of the actual payoff and those foregone, as well as an average of all past attractions – rather than a sum of those prior measures (Rapoport and Amaldoss 2000). The EWA dynamically combines the apparent best parts of RL (reinforced chosen strategies) and BBL (consideration of unchosen strategies by all players). It is a flexible model because the extent to which it incorporates these two components varies. And numerous studies have shown that hybrid models like EWA better predict behavioral data than models that employ just one type of learning (Camerer and Ho 1998). This comes out even despite penalties for the higher number of parameters in many versions of EWA. Further, Zhu and colleagues (2012) found that BBL and RL models performed about equally well, but still not as well as EWA.

The constituents of RL and BBL are evident in the EWA update rule for strategy *k* of player *i*:

$$V_i^k(t) = \begin{cases} \dfrac{\phi \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t) \end{cases}$$

(Zhu, Mathewson et al. 2012). The *Φ* parameter discounts previous attractions, and *N(t)* represents the decay of past experience. The *∂* indicates how much weight is given to strategies not taken. Giving them full weight (*∂*=1) would reflect fully belief-based learning, while *∂*=0 indicates no consideration of foregone choices, and therefore reinforcement learning. The *∂* parameter then provides a clear and continuous measure of a player's relative use of RL and BBL. Parameter I makes the switch between the weight of chosen strategy (1) and foregone strategy weight (between 0 and 1). *N(t)* is estimated initially and then is updated each period, according to the decay represented by ρ.

The previous expected reward of a given strategy is depreciated by φ, a conception of the opponent's adaptation speed, and a discount rate for past experience. It is then increased by the reward for that strategy, given the opponent's actual choice in the previous period, and that is divided by all past experience to arrive at the new value of the strategy in question. A small φ means that the player believes her opponent adapts quickly, so previous values are depreciated more quickly. A large ρ, which updates the past-experience discount, indicates a rapid decline of prior beliefs.

The hybrid model reduces to RL when parameters ∂ and ρ are 0 and initial experience N=1. The model is pure BBL when ∂=1 and φ=ρ. So the update to a value of an action is given full weight when it was the one chosen – exactly as it would be in RL. But in BBL, the value is weighted by the beliefs the player has about the future actions of other players (Zhu, Mathewson et al. 2012). So ∂ can be seen to describe the player's tendency toward either RL or BBL.

The Patent Race

The patent race game provides a framework in which to observe iterative thinking in limited strategy space. In this asymmetric configuration, two players compete for a prize in one of two asymmetric roles: one with an endowment of five cards (strong role); the other, four (weak). In each round, the endowment is renewed, and each player must invest from 0 to the full amount of the endowment. The player who invests strictly more wins a prize of 10 cards. Any endowment cards that the player does not invest go into her winnings but do not carry over into the next round’s endowment. In case of a tie, neither player wins the reward but retains that portion of the endowment not invested.

To understand and predict the opponent's choices, players benefit by examining the structure of the game from the beginning, including its asymmetrical aspect. To wit, the strong player might realize that she can win the prize every time by always investing the full endowment. She would lose the entire five-card endowment but win the prize. The weak player might invest some or all of his endowment several times, losing the entire endowment and never winning the prize, before realizing the futility, then reducing his investment to zero so as to retain the entire endowment in his earnings. Seeing this, the strong player might realize that she does not have to invest her entire amount in order to win, leading her to increase earnings by occasionally investing less than the full amount. This then provides openings to the weak player to predict when the strong player will play less than the full amount and to invest more in order to win the prize, even with a smaller endowment.

The strong (weak) player can employ six (five) strategies: one for each possible investment choice. "Strategy" in this case refers simply to the choice of how many cards to invest in each round. Players with more iterative strategic thinking may realize that some strategies almost never make sense in a given role. These so-called iteratively dominated strategies derive from a knowledge of the structure of the game. The strong player may not need to invest her entire amount, but it would never make sense for her to invest 0, thus guaranteeing a payment of 5, since she can guarantee a payment of 10 simply investing the full amount. So 0 is a dominated strategy.

A high-level reasoner will consider whether or not his opponent understands the structure of the game. To that end, he would observe that it never makes sense for the strong player to play 0 cards. If the weak player realizes this, he would see that it would never make sense for him to invest 1, since it would never beat any strategy played by the strong player. If the strong player believes that the weak player understands the structure well enough to reach this level, she may conclude that it never makes sense for her to invest 2 cards. It might result in a win, as well as some retained endowment. But if the weak player is unlikely to play

1, she loses an additional card of earnings by investing two rather than 1. This continues with iteratively eliminated dominated strategies for the weak player comprising 1 and 3, and for the strong player, 0, 2 and 4. In this way, the iteratively eliminated dominated strategies of the strong (weak) player are 0, 2 and 4 (1 and 3).

A player more reliant on reinforcement learning would be slower to adapt to the opponent's behavior, continuing for more rounds to make the choices that brought higher reward more recently. The influence of either some belief-based learning or a tendency to explore new options may eventually induce the player to try a different strategy.

However, if the player observes or assumes that his opponent does not understand the structure of the game – i.e. these dominated strategies – he may well choose the dominated strategies of his role. Rapoport and Amaldoss found that iteratively eliminated dominated strategies were played more often than probability prediction, yet that those higher in the hierarchy of deletion (i.e. those for which the thought process takes longer to get to: 2 and 4 for the strong player, 3 for the weak) were played less frequently than lower-level dominated strategies (2000).

The patent race is particularly well suited to EWA measurements because the asymmetrical endowments generate the different secure strategies of 0 or 5 investment for weak and strong players. These investments act as a sort of "safety net," since they yield the same payoff, regardless of the opponent's choice. When played as a n asymmetric mixed-strategy game, the patent race is shown to employ both RL and BBL (Camerer and Ho 1998, Rapoport and Amaldoss 2000). Yet it is possible that EWA could measure a mixed pool of purely RL players and purely BBL players. To ensure that the two learning methods are both present at various strengths in individuals in one population, the BBL parameter should be distributed along an interval, rather than clustered at either end, which by contrast would indicate two distinct populations (Zhu, Mathewson et al. 2012). Hsu and Zhu characterize the combination of learning types not as a true hybrid, rather as that of two systems in conflict

(2012). Indeed, over the course of the task, the model mix may change in the same individual, employing at first more BBL and, as rhythms of the game and habits of the opponent become clear, relying more on the cognitively less-demanding RL (Ansari, Montoya et al. 2012).

Priming task: Wheel of Fortune

The Wheel of Fortune gambling task adapted by Camille and colleagues (2004) from an earlier task (Mellers, Schwartz et al. 1999) can elicit with little manipulation emotions borne of either responsibility or nature, that is, disappointment or regret. The only change is one of presenting alternative results, so participants primed with either type of emotion undergo largely the same task. The emotion an individual feels depends on both the obtained outcome and one of the unobtained outcomes. Even if the obtained outcome is positive, the prevailing emotion can be negative if an individual sees that the unobtained outcome was better (Camille, Coricelli et al. 2004).

Subjects who obtained an outcome of -50 reported emotional ratings of -20 when the other, unobtained possibility on the wheel was 200 (**Fig. 3-1**). Yet when subjects who obtained -50 saw that the unobtained option on the wheel was -200, they reported net positive emotional ratings. A similar shift was evident in subjects who won 50, in cases in which the unobtained result was 200 the emotion was slightly positive versus highly positive when the unobtained

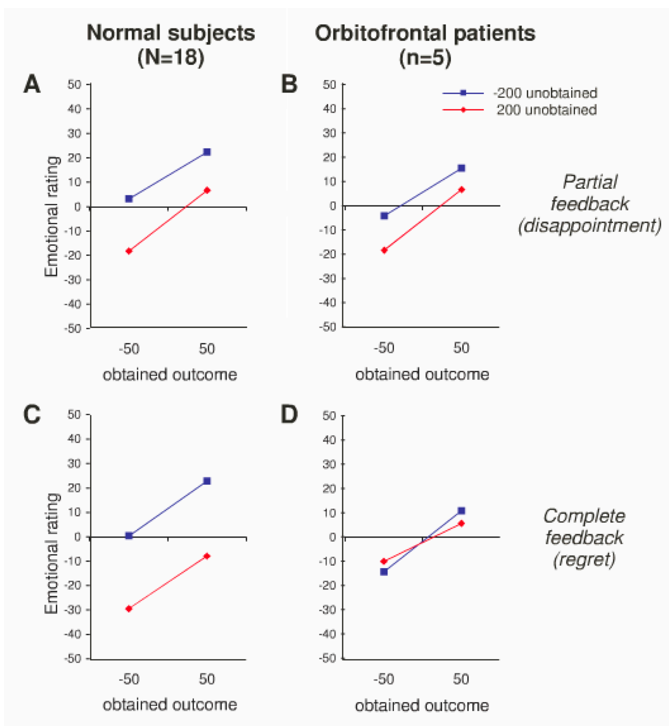


Fig. 3-1 Effect of unobtained outcome on emotional rating of obtained outcome in healthy subjects following (A) partial-information feedback and (B) complete-information feedback. Image adapted from Camille, Coricelli et al. 2004.

result was -200. In complete feedback, the result on the unchosen wheel overwhelmed the effect of the unobtained result on the chosen wheel. The disparity in emotional ratings for the same obtained amount follows the same pattern for the result on the unchosen wheel as for the unobtained amount in the partial feedback condition. In fact the modulating effect of the unobtained outcome on the unchosen wheel is so strong that in complete-information feedback conditions, the relief at obtaining a smaller loss (obtaining -50 on the chosen wheel instead of -200 obtained on the foregone wheel) produces a higher mean emotional rating than 50 obtained on the chosen wheel in light of an unobtained 200 result on the unchosen wheel. Notably, the effect of regret is more pronounced than that of disappointment, even with the magnitudes of obtained and foregone held constant.

Patients with lesions in the orbitofrontal cortex exhibit the same pattern of emotional shift depending on foregone outcome in partial-information feedback conditions. The shift disappears, however, in the complete feedback condition. Both negative emotions at losing and positive emotions with wins are constant, whether the foregone wheel's outcome was better or worse.

Mood priming

Moods, sustained affective states, have been likened to a climate with gradual changes, as contrasted to the more sudden and brief occurrences of emotions, which are comparable to daily weather events (Rottenberg and Gross 2007, Kohn, Falkenberg et al. 2014). Immediate emotional actions are modulated by external events: improved by reward and worsened by loss. They are similarly improved by downward counterfactuals (in which the outcome is better than the alternative) and worsened by upward counterfactuals (seeing that what was foregone was worse than what was obtained) (Markman, Gavanski et al. 1993, Roese 1994, Sanna and Turley 1996, Sanna 1997). The sustained nature of mood, however, allows the effect to run the other direction, forming a potential positive feedback loop between mood, emotion and reward. Indeed, mood and mental simulation are so related that they are both cause and consequence of each other (Sanna 1998).

Mood can change, affecting the valuation of choices (Tamir and Robinson 2007). The outcome of a wheel of fortune (WoF) game changed participants' moods, then influenced their feelings during an immediately subsequent task (Eldar and Niv 2015). Participants chose between pairs of marked slot machines with different but stable probabilities of a reward with the goal of maximizing reward. They then participated in a Wheel of Fortune, an unrelated task with no choice but with a relatively large payout, and those who won reported being in a better mood. Afterward, they played the slot machine task again with different sets of differentiating markings but with (unbeknownst to them) similar probabilities. After the slot machine learning tasks, participants assigned values to all the slot machines they had seen. People who won the wheel of fortune assigned higher values to slot machines they encountered after the WoF game of chance, even though their values were similar to those encountered before the WoF. Participants had no influence over the game of chance, yet the outcome reliably predicted whether they were happier with the slot machines in the later task.

At the neuronal level, positive mood induction is accompanied by cortico-striatal activity during reward anticipation versus loss anticipation, as compared to differences with neutral mood induction (Young and Nusslock 2016). Critically for regret learning, positive mood induction brought greater activity in the vmPFC during anticipation of reward versus anticipation of loss. Notably, those differences are not evident in any of the areas during the outcome phases of win or loss. These measurements suggest that people who are already feeling good assign more importance to positive outcomes, effectively enhancing them. It further suggests that positive feelings may constitute some insulation against the negative feelings required of regret learning. The obverse raises the possibility that the presence of negative feelings could make a person particularly susceptible to learning via anticipation of regret. However, it is possible that instating a negative mood may bring subsequent losses into lower relief, which is consistent with prospect theory (Kahneman and Tversky 1979). That is, if a player is already feeling badly, she will not be "brought down" by subsequent smaller losses. How that high-pass filter might bear on regret anticipation has not been explored.

Hypotheses

As demonstrated in previous studies, the experience of regret and the subsequent anticipation and avoidance of regret is a form of learning (Zeelenberg, Beattie et al. 1996, Foster and Vohra 1999, Hart and Mas-Colell 2000, Camille, Coricelli et al. 2004). In repeated interactions, this adaptive behavior leads to more rewarding outcomes (Coricelli, Dolan et al. 2007, Marchiori and Warglien 2008), as well as more rapid arrival at equilibrium (Coricelli and Rustichini 2010). As described by the EWA model, people may employ either RL or BBL to differing extents while playing an asymmetric repeated strategy game (Camerer and Ho 1998, Rapoport and Amaldoss 2000, Young and Nusslock 2016). The precise ratio of RL to BBL even appears to vary throughout a task (Ansari, Montoya et al. 2012). Yet to this point, the cause of these varieties have not been characterized.

Why do people use RL and BBL to varying extents instead of at the same rates? Why do some people employ BBL more than others? One possibility for the variation between people is their state at initialization, that is, the disposition of the player at the start of the game. Whereas one person might come to a task more naively and consider implications of structure as he goes along, another might be prepared from the outset to consider the entire framework of the task. We hypothesized that players already in the midst of a counterfactual consideration would be more inclined to continue similar consideration. Such a player would more quickly come to understand the structure of the game, and her choices would indicate greater employment of BBL. To test this, we would have all subjects play a strategy game that would measure their learning mix, with one more indicative of behavior arising from counterfactual learning. We would segregate them into groups, priming them with different counterfactual outcomes in an unrelated choice task (or not prime them at all). We randomly assigned various players to one of these five pre-game conditions. Our greatest interest was in the highly salient negative counterfactual regret condition, but we employed two active controls (one each for valence and feedback) and one passive control (no priming). If the player had already experienced a regret situation with large consequences, we hypothesized that it would ready her learning processes in such a condition as to make different choice patterns during a new task. That is, having already engaged in consideration of the better alternatives to her choices, this learning would transfer to a greater preparation to anticipate and avoid regret in an unrelated situation.

Regret is common in the patent race because the game incorporates elements of regret in any round that does not include a perfect win. In most configurations of choice that end in a loss, the player could either have made a different choice to avoid the loss or could have made a different choice to attenuate the loss (i.e. maintain more of the endowment). Likewise, in most strategy configurations that result in a win, the player could have made a different choice to optimize the win, earning more. Any of these non-optimal outcomes results in regret: the recognition that a different choice would have yielded a better outcome. This regret signal, its anticipation and likely avoidance give rise to learning and better understanding the opponent's play. Such increased utilization of the regret signal and more sophisticated understanding should generate a greater incorporation of belief-based learning in the hybrid model of choice behavior.

Methods

We designed an experiment incorporating a well-documented and researched Wheel of Fortune (WoF) lottery task (Mellers, Schwartz et al. 1999, Camille, Coricelli et al. 2004), along with a simple competitive strategy game. Subjects first play the wheel of fortune lottery task and receive either partial feedback or complete feedback (**Fig. 3-2**). The result primes them in different ways: either with counterfactual emotions (regret/relief in complete feedback) or with non-counterfactual emotions (disappointment/satisfaction in the case of partial feedback). They then immediately play the competitive strategy card game called the patent race. The opponent is a learning computer algorithm that responds to the subject's gameplay.

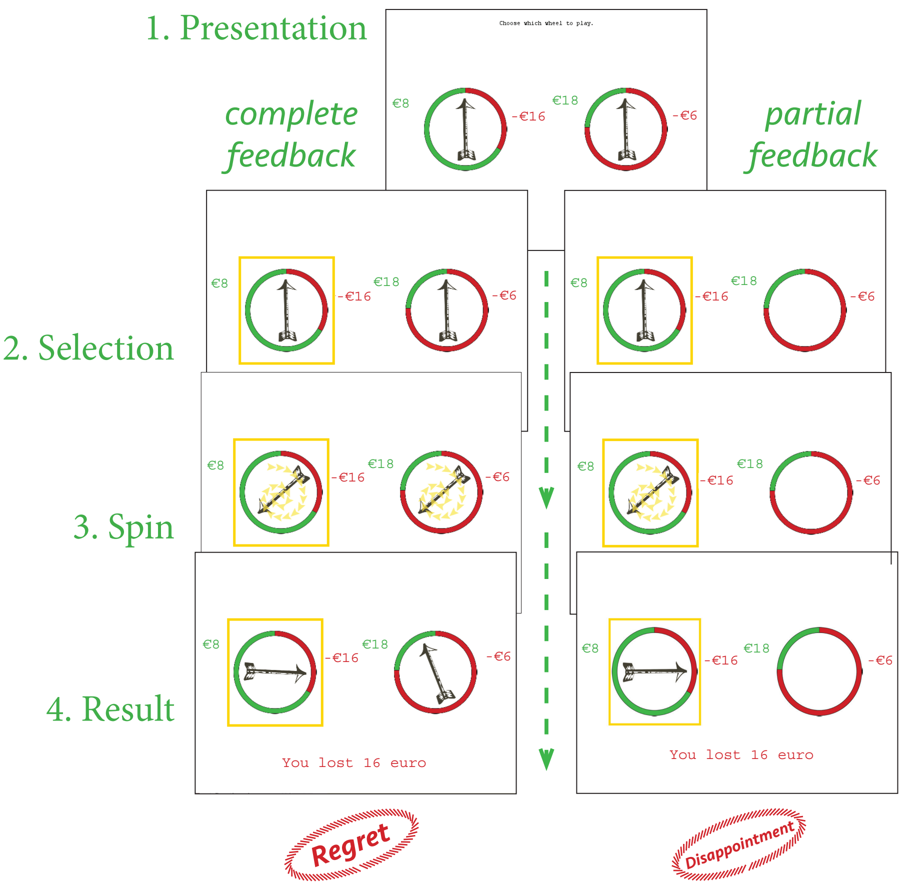


Fig. 3-2 In the Wheel of Fortune (WoF) priming task, participants choose one of two wheels presented to determine their win or loss. Subjects selected for complete feedback see the arrow spin both on the wheel they selected and on the one they did not, even though it has no bearing on the amount they win or lose. In partial feedback, subjects see the arrow spin only on the wheel they chose; they are not informed of the result of the wheel that they did not choose.

The Wheel of Fortune task presents two wheels, each divided into two results sectors (red and green) indicating the probability of that result. Results are matched to value pairs, in this case one positive, one negative and presented in the same color as the corresponding portion of the wheel. Participants completed 10 practice trials with values less than 1 and no currency symbol. They were told these outcomes would not affect their score. When directed, they proceeded to the main lottery task, where they were directed, as before to use arrow keys to select which wheel they wished to play.

All participants saw the same set of wheels: +€8/-€16 at probabilities of .66 and .33, and +€18/-€6 at probabilities .25/.75. Amounts and probabilities were set to have equal expected outcome of 0. The potential regret error differs between the two, making the choice an indicator of either regret tolerance (in complete feedback) or risk tolerance (in partial feedback). In partial feedback sessions, the arrow inside the unselected wheel disappeared, while the arrow inside the selected wheel began to spin, indicating the outcome where it stopped. Outcomes were predetermined by seating position. The number of participants who won or lost corresponded to the presented probabilities. In complete-information feedback sessions, the arrow on the opposing wheel was set to land in the area of opposite valence to that of the obtained outcome.

Participants played the patent race against a computer algorithm that adjusted its play based on the choices of the participant and a belief-based learning “fictitious play” algorithm (see **Box 3-1**). Participants were informed beforehand that their opponent was a learning computer algorithm. Previous studies have had participants play against each other or against randomized round-matched responses from a pool of past human players. We sought consistent opponent play to provide stable progressions of learning over the course of the two blocks. We reasoned that randomized pool play would not provide a sense of consistency and that human versus human interaction would introduce elements of reputation and mentalizing considerations (Zhu, Mathewson et al. 2012). This limited the human player’s

Box 3-1. Computer player value update

Fictive earnings *f(s)* is the amount that would have been earned by playing strategy *s* in the current trial *t*, in which P_t is the human player's investment, given computer endowment *e* and reward *r*.

$$f(s) = e + r - s_t, \text{ if } Pt > st$$

$$f(s) = e - s_t, \text{ if } Pt \leq st$$

Value update $V(s)_t$ is the value of strategy *s* at the end of trial *t,* after being updated by *f(s)*. The learning rate constant *α* [0, 1] determines how much effect new data has on the previous value.

$$V(s)_t = V(s)_t - 1 + \delta \cdot f(s)$$

considerations to recent actions by the computer or its pattern history, which are the levels of differentiation our modeling sought to describe. Though the game features secure strategies for each role, neither has a pure Nash equilibrium because payoff can be increased by changing strategy. The relatively large strategy space provides a breadth of choice and prompts, which affords greater modeling distinction between reinforcement and belief-based behavior.

Regret

Our hypothesis rests on the idea that individuals exposed to regret in the priming task will engage to a greater-extent in belief-based learning during the patent race. One potential vector for this more extensive use of the more sophisticated learning type is an avoidance of regret in the patent race. Regret is a particularly useful signal in the patent race because of the task's asymmetric roles that result in frequent disparities between outcome and regret. We calculate regret error in each round RE_t as the difference between actual earnings and highest-possible earnings:

$$RE_t = V_i^k -_{max} V_t$$

Regret error in round *t* is the reward returned by chosen strategy *k* by player *i* minus the maximum reward returned by any strategy in time *t* if that strategy had been chosen. In rounds in which a player lost but could have played a strategy that would have won, regret is easy to identify: the strong player plays 3 and loses to the weak player playing 4. Here, the strong player has kept 2, while playing 5 would have won him 10, producing a regret error of 8. The weak player, meanwhile, has played the equilibrium strategy and has no regret error. In this scenario, if the strong player had invested just 1, he would still have regret, but it would be lower, having retained 4 cards of the endowment, resulting in a regret error of just 6. In a winning scenario, however, there can still be regret. Suppose the strong player invests all 5, a fairly common investment, and the weak player has realized the frequent inutility of investing anything and so invests 0. The strong player wins, taking 10, but sees that he could have kept even more if he had invested as little as 1 card. The difference and regret error is 4, even though he won the round.

Participants

We recruited 259 healthy volunteers (124 female) via the Cognitive and Experimental Economics Laboratory at the University of Trento, Trento, Italy. Subjects had a mean age of 21.8 ±2.8 with a range of 18 to 38. Subjects were randomly assigned to seating positions, which determined the outcome of the priming experiment.

Procedure

Participants read onscreen instructions (also duplicated in print) on the rules of both the Wheel of Fortune game and the patent race before completing 10 practice rounds of the Wheel of Fortune, which they were told would not affect their payment. They were then told the next Wheel of Fortune task would have the single largest effect on their payment. All experimental subjects were presented with identical probability wheels (probabilities of 66.6/33.3% of €8/-€16 and 25/75%, €18/-€6) and chose either the right or the left.

Subjects undergoing complete feedback priming saw the arrow on the wheel they chose spin, as well as the arrow on the unchosen wheel; those undergoing partial feedback priming saw the arrow on the chosen wheel spin, while the other arrow disappeared. After three revolutions, the arrows stopped spinning and rested on the outcome, then an on-screen alert told the subject whether she had won or lost and the amount. Participants completed an emotional evaluation via an on-screen Likert item, then the Wheel of Fortune outcome was presented again for 10 seconds.

Immediately following, the patent race began (**Fig. 3-3**). Subjects began in either the weak or strong role (132 and 127, respectively). They played 50 rounds of each role, counterbalanced for order. Their opponent was a computer algorithm programmed to use fictive learning to determine its strategy. For the algorithm, each strategy (five strategies in the weak role, playing 0-4, or six in the strong role, 0-5 cards) was assigned an initial value of 5. The function updated the values for each strategy at the end of each round. It calculated the earnings each of its own choices would have brought, had it been played (including the choice actually made), then found the difference between those values and the existing values assigned to each strategy. The function attenuated each difference by multiplying it by a learning rate of 0.5, then added those amounts to the existing values to arrive at a new array of values for the next trial.

Participants were clearly told that they would be playing against a computer. But they began the priming task and subsequent experimental game in groups of 2-4 so that no subject was playing the game alone at any time. Subjects were told the result of the Wheel of Fortune would be added to or deducted from their total, which included a show-up fee of €5. For the patent race, they were told that each card won or retained was worth €0.01.

In each trial, a fixation cross appeared for 4-8 s, followed by the explanation screen, and a graphical illustration of the endowment for the subject and opponent, as well as the possible earnings for the subject. Subjects selected the amount of investment using arrow keys, then confirmed at their own pace. Then, 2-6 s later, the opponent's choice was revealed, along with the subject's earnings. After 50 trials, the subject was informed of earnings for the

Fig. 3-3 Course of play in each of two roles in the patent race. Players invest all or a portion of their endowment (first row blue outlines; selection in filled blue) and keep any uninvested portion (third row, filled blue). The computer's investment is revealed (second row, filled gray), and then the player's prize, if any, is displayed (third and fourth rows, filled green).

round, as well as the opponent's earnings. The subject then switched roles and performed another 50 trials.

Control subjects had no priming and began the card game when they finished reading instructions. We designed a simple visual representation of the card game based on that of Zhu and colleagues (2012), displaying both the player's endowment and the opponent's, along with the potential reward.

Computational Learning models

We analyzed participant choice behavior by performing estimations using four models: Q-learning, a type of reinforcement learning (RL); Counterfactual Q-learning, a type of reinforcement learning that incorporates unchosen outcomes (CRL), Fictitious Play, a belief-based learning model (BBL) (Hampton, Bossaerts et al. 2008); and a hybrid model that nests the RL and BBL models, a simplified version of the experience-weighted attraction model (EWA) (Zhu, Mathewson et al. 2012). The RL and BBL models constitute special cases of EWA, with a parameter ∂ indicating the relative weight of BBL in the examined behavior. EWA updates according to two rules, depending on the player's most recent choice:

$$V_i^k(t) = \begin{cases} \dfrac{\phi \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t) \end{cases}$$

Here, s_i^k represents the strategy (choice) k of player i. $s_i(t)$ is the chosen strategy in period t, so these two equations update differently depending on whether it applies to the chosen action or not. $s_{-i}(t)$ is the strategy played by the opponent in period t. The player's expected reward for playing a given strategy k in period t is $V_i^k(t)$. It is determined by three parameters: θ_i , which depreciates past values at different rates, depending on how fast an adapter the player believes the opponent to be. The key parameter is the δ_i , which determines how much weight an unplayed option has on updated values. If a player believes foregone strategies deliver as much information as those played, then δ_i reaches 1, and the model reduces to the BBL model. At no weight, $\delta_i = 0$, and the model reduces to RL.

For each model and subject by subject, we performed individual maximum likelihood estimation with a grid search over a range of values. We calculated predicted decision probabilities over the full range of each set of parameters and compared them to the subject's actual choices, selecting that set of parameters with the maximum log likelihood. We then performed individual and group level estimations.

Results

We asked whether priming players with exposure to a large gambling loss outcome that typically induces regret would modulate their strategy in a different task played immediately afterward. Subjects were grouped depending on the type of priming they underwent: by positive or negative valence, by complete or partial feedback, and those who received no priming.

We first considered only block 1 of the patent race because it was the closest in time to the priming task. To detect differences between priming groups, we characterized their gameplay, examining what portion of possible earnings they won (**Fig. 3-4 A**) as well as how much regret error they were exposed to, a predictive measure of regret avoidance (**Fig. 3-4 B**). We calculated earnings percentage to normalize between the weak and strong roles. Earnings percentage is calculated as amount won divided by total possible earnings regardless of opponent's strategy: 13 for the weak role and 14 for strong. We analyzed these results across all 50 block 1 trials, as well as in the first 20 trials and the first 5 trials. We

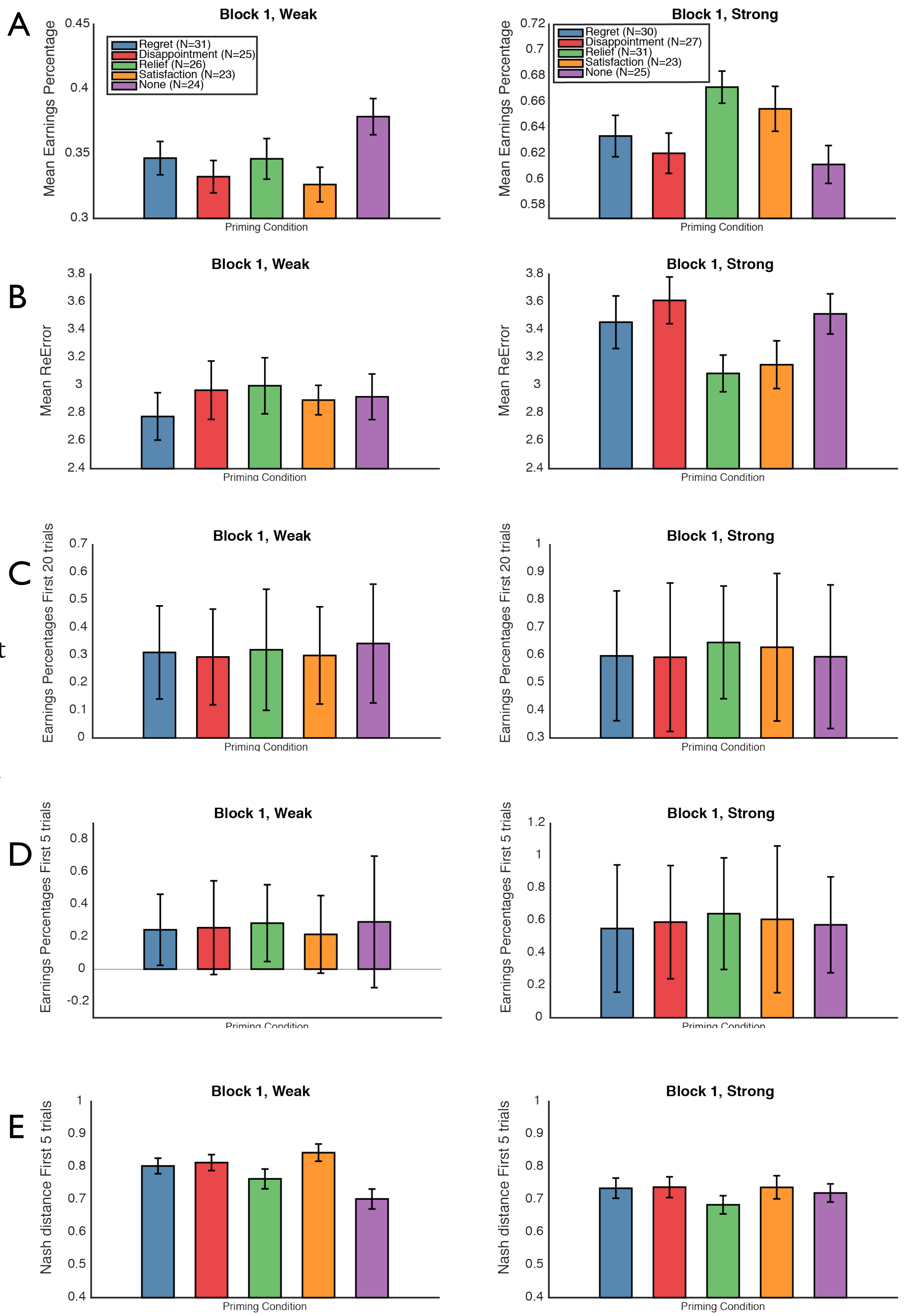


Fig. 3-4 Performance in the patent race game, grouped by priming condition: A) earnings as a percentage of total possible (to allow comparison between strong and weak conditions) across all 50 trials of block 1. B) Mean Regret Error (actual earnings less highest possible earnings) across all block 1 trials. C) Earnings percentage for first 20 trials (during which priming would likely be stronger). D) Earnings percentage for first 5 trials (during which priming would likely be strongest). E) Mean distance from equilibrium prediction in first 5 trials, accounting for dominated strategies. Error bars are s.e.m.

Table 3-1: Predicted strategy probabilities		
	Weak	**Strong**
Investment	*p*	*p*
0	0.6	0
1	0	0.2
2	0.2	0
3	0	0.2
4	0.2	0
5	-	0.6

expected any effect of priming to be strongest in these early trials, so we also calculated for each participant's first five trials the mean distance from Nash equilibrium. Equilibrium probabilities account for the iterated elimination of dominated strategies, as well as the fact that a single strategy is most likely for each role, but not absolute. For each trial, we measured the distance from equilibrium as the probability of not choosing that strategy, or the probability of choosing that strategy subtracted from 1. The predicted probabilities (p) are listed in **Table 3-1**. Distance from each is calculated as 1-p.

In both strong and weak roles, we observed mild trends of regret-primed players earning more than disappointment-primed players, as well as relief-primed players earning more than satisfaction-primed players. Surprisingly, non-primed individuals earned more in block 1 than all priming groups when they started out in the weak role but earned less in block 1 than all primed groups when they started out in the strong role.

Regret error indicated little differentiation between priming types in the weak role. In the strong role, participants primed with the two types of negative outcome had higher average regret error than those primed with positive outcome. Participants who were not primed also had higher average regret error.

There was little differentiation and high variance among all groups and across both roles in earnings percentage during the first 20 trials of the block. The same lack of differentiation characterized the first 5 trials. Analyzing the distance from Nash equilibrium

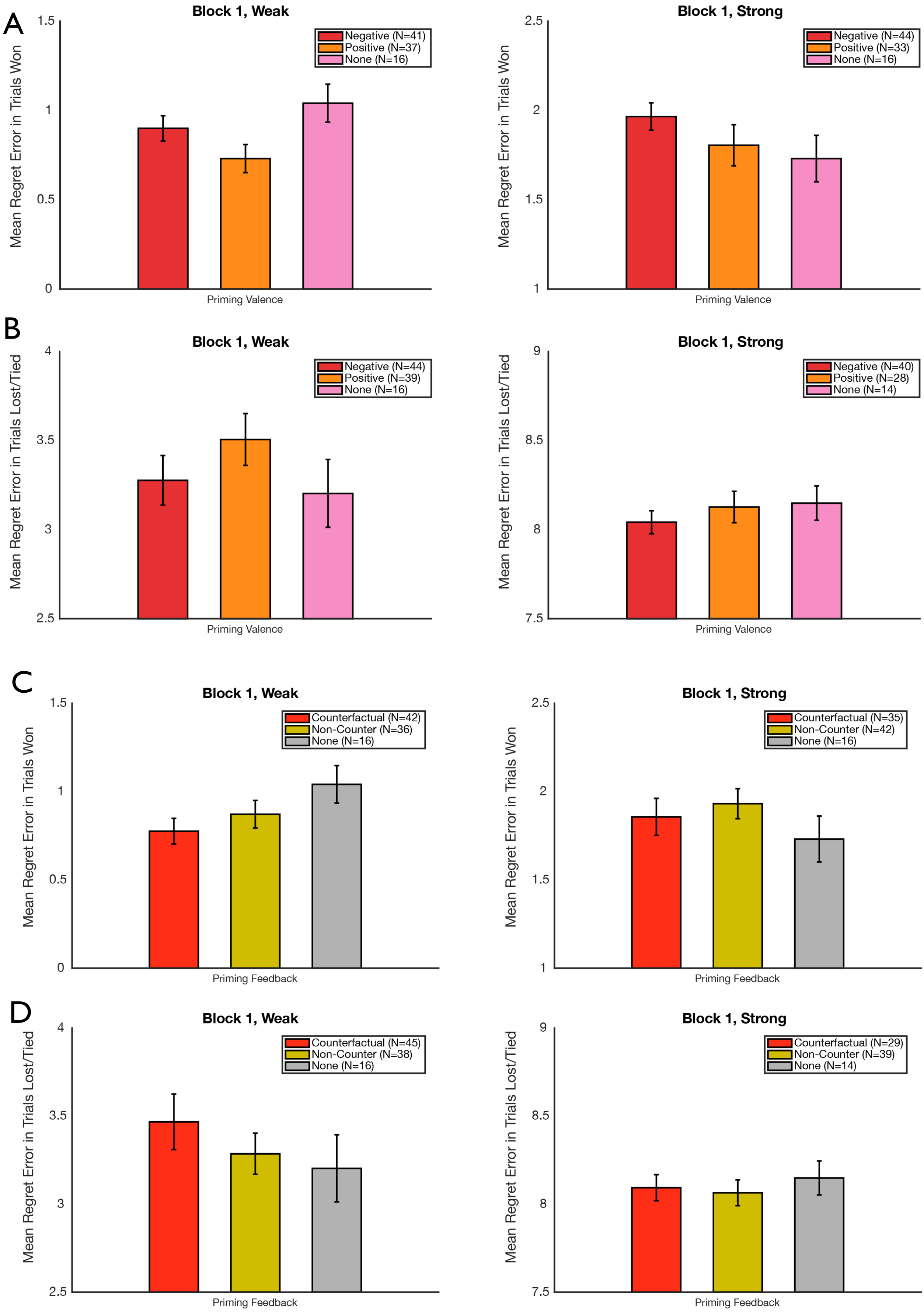


Fig. 3-5 Mean Regret Error (actual earnings less highest possible earnings) in across all block I trials of the patent race game, grouped by priming condition and segregated by trial outcome: A) weak role winning trials B) strong role winning trials D) weak role losing trials E) strong role losing trials. Error bars are s.e.m.

during the first 5 trials, suggested that all four priming categories played expected strategies less frequently than the participants who did not undergo priming.

We then looked at RE segregated by winning trials and losing trials, hypothesizing that across groups, players would generally behave differently in a win versus a loss (**Fig. 3-5**). Because of the tendency to win in the strong role and to lose in the weak role, we considered only subjects with four or more wins or losses for consideration in each category (i.e. if a player won all but two of her rounds in the strong role, the outcomes following her two losses would not be considered in the loss-trials analyses). Of the 192 subjects, 187 were considered in RE-win calculations and 181 in RE-loss calculations. We found no differentiating trend among the five groups.

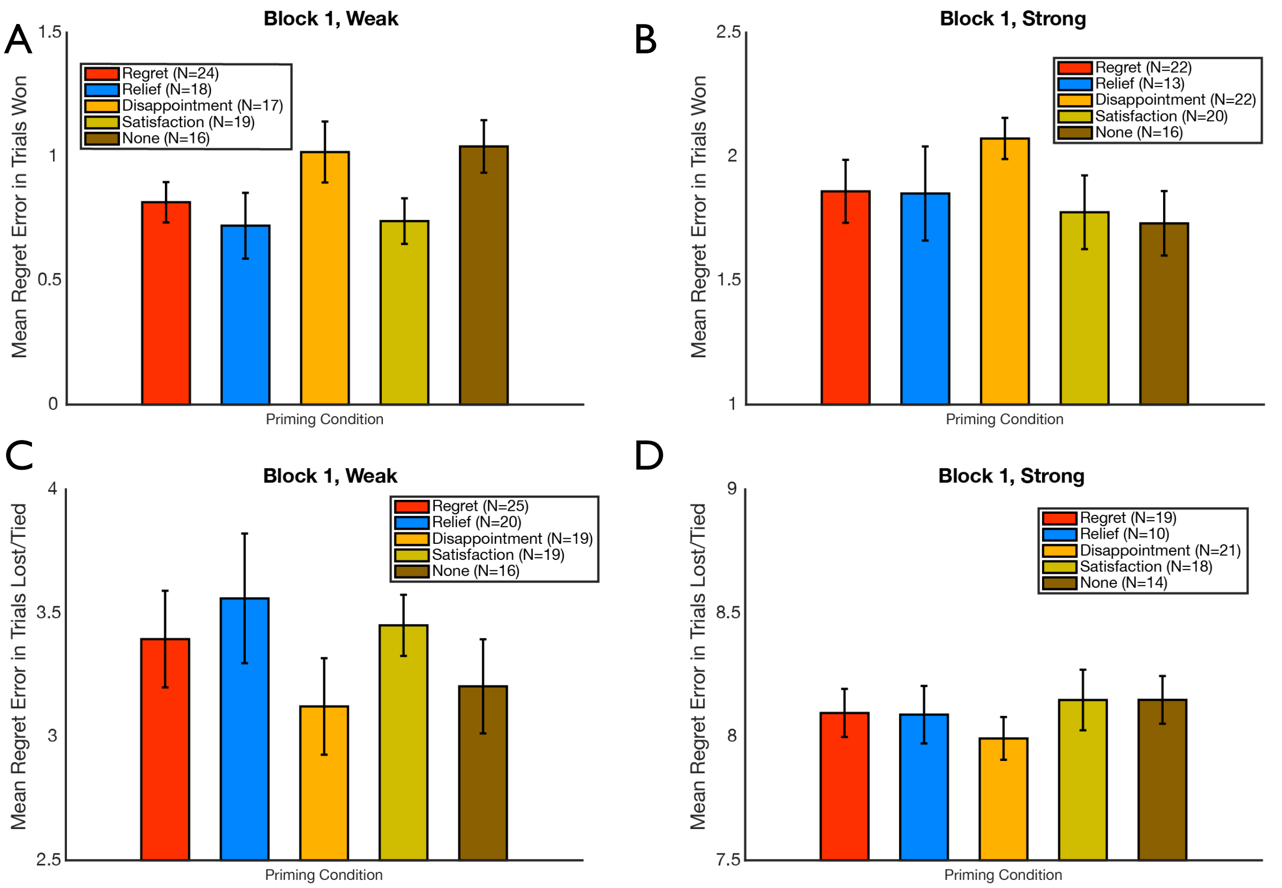


Fig. 3-6 Performance in the patent race game, grouped by priming feedback and valence: A) Mean regret error (RE) in loss trials across all 50 trials of block 1, by priming feedback information level. B) Mean RE in win trials across all 50 trials of block 1, by priming feedback information level. C) Mean RE in loss trials across all 50 trials of block 1, by priming valence. D) Mean RE in win trials across all 50 trials of block 1, by priming valence. Error bars are s.e.m.

Due to similarities in performance across both axes of feedback type (i.e. partial and complete) and feedback valence (i.e. positive or negative), we pooled priming types across those two vectors and analyzed their performance in other measures.

Playing in the weak role, subjects primed with full-information feedback (i.e. regret and relief) exhibited a trend of lower RE in win trials than those who did not undergo priming (**Fig. 3-6 A**). Because the weak role presents few opportunities to win, this may represent successful attempts to eke out wins on opportune occasions. Meanwhile in loss trials, RE trends higher for the full-information feedback priming group. This may suggest similar attempts to win by investing high but losing anyway.

Considering priming by outcome valence, compared to both no-priming and negative-valence subjects, players who had a positive priming outcome trended lower in RE during winning trials and higher in loss trials when playing the weak role. By contrast, negatively primed subjects trended higher in regret error in winning trials and lower in loss trials when playing the strong role. This suggests that they were not efficient in their wins in the strong

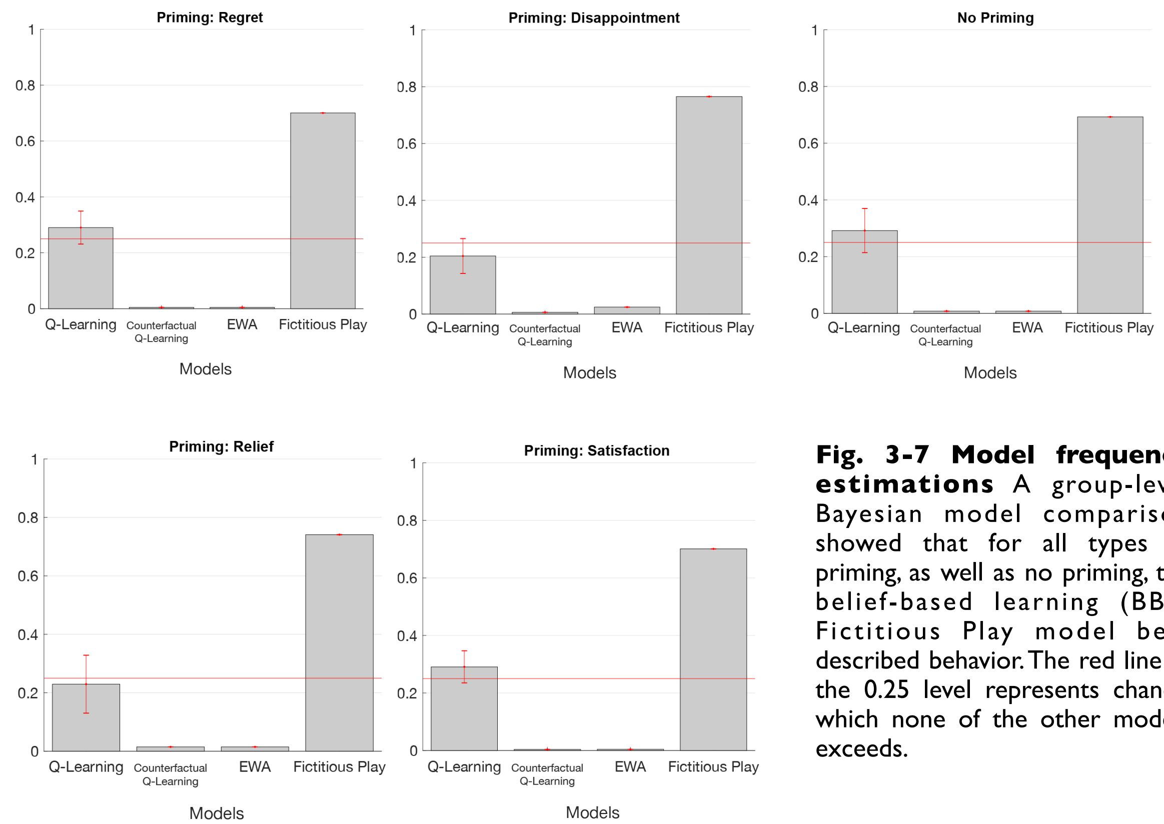


Fig. 3-7 Model frequency estimations A group-level Bayesian model comparison showed that for all types of priming, as well as no priming, the belief-based learning (BBL) Fictitious Play model best described behavior. The red line at the 0.25 level represents chance, which none of the other models exceeds.

role, indicating possible lack of sophisticated understanding of the game or concern about investing too low and losing. Low regret error in losses indicates a slight misreading of the opponent, leading to a loss by a marginal amount.

Modeling

In order to determine the relative contribution of different types of learning, we compared data to four learning models: Reinforcement Learning (RL), Counterfactual Reinforcement Learning (CRL), Belief-Based Learning (BBL) and the hybrid model Experience-Weighted Attraction (EWA). The first two models constitute special cases of EWA, with a parameter ∂ indicating the relative weight of BBL in the examined behavior. We used a group-level Bayesian Model Comparison to compare log likelihoods of model fit to choice behavior. To our surprise, BBL outperformed all other models (**Fig. 3-7**). Employing the learning rate parameter from the BBL model, we compared it among the priming types but found no significant relationships.

Discussion

The manipulation failed to yield any significant difference in measurements of choice behavior. No studies to our knowledge have attempted to change choice behavior in the patent race game, so the field was wide open as to how to attempt the perturbation and how to measure its effects. After a number of considerations how to induce regret, we settled on a large, monetary result in the hopes that its magnitude and salience would be effective. But it is possible that other, more visceral forms of regret, such as autobiographical recollection, or repeated forms, such as several rounds of the lottery, would bring changes to patent race play. It may be that the magnitude of regret was not high enough to yield appreciable changes in the latter task. It is also of course possible that there were effects of the priming but that our measures were not precise enough to record them. Before analyses, however, we considered many approaches for measurement, both in hidden-variable modeling, summary statistics, and calculated variables, along with which portion of the task to approach first. We had little hope, for example, that any effect would persist into the second block of the task but included it for purposes of comparison with the first.

We chose to use an algorithm as the opponent for the purposes of consistency, but its behavior does not match that of human gameplay in that it did not avoid the iteratively eliminated dominated strategies, other than as a result of adaptation to the human choice behavior. This likely gave rise to different play behavior from human subjects than they would have exhibited had they played against other humans. The algorithm could be maintained but its behavior changed by simply reducing the initial values of each dominated strategy in an amount commensurate with its theoretical rate of avoidance. These values are updated each round and represent the relative attractiveness of each strategy and, in part, the likelihood of that strategy being selected by the algorithm. Though it would be important to maintain some balance between consistency and natural play by the opponent, a future study might aim for more human-like choices.

After trying to answer our hypotheses, with the more specific priming types, we pooled priming groups to examine effects more broadly: by valence and by feedback type. Here, as in the more specific groupings, there were no notable trends, nor significant effects. Although these indicators suggested little hope for effects yielded by model estimations, we conducted them and found that even the assumption of the model that would best fit was incorrect. The failure of EWA to better predict subject behavior made our original hypothesis impossible to test. We could not compare the balance of RL and BBL between groups because our analysis told us that BBL alone predicted choice behavior.

In control experiments using an algorithm as the opponent, typically RL has better fit participant behavior, possibly because players view the contest as a simple reward situation rather than a true competition (Zhu, Mathewson et al. 2012). The use of BBL, however, suggests that participants treated the game as a competition in which beliefs about the opponent's actions informed subsequent decisions. This could be an effect of the fictitious play algorithm we employed, which uses a form of belief-based learning to guide its choices. Because of the already large dimensions of the study, we used only one type of play in the opponent algorithm. Employing different algorithms in a future study might indicate if this is a behavior-mirroring effect.

Chapter 4

Electrical brain stimulation effect on level-*k* thinking

Experimental question

How does stimulation of prefrontal cortex areas change the consideration of other players and accounting for their actions in an iterative thinking contest?

Introduction

In competitive situations, choices that accurately account for the actions of others lead to greater success. Because individuals demonstrate ranges of sophistication, this assessment presents several challenges: understanding that there is a range, what its bounds are and where on it competitors lie. The most successful players are not necessarily those who engage at the highest levels of sophistication, but rather those who most accurately assess the level of sophistication of others. Competitive strategic games present a framework to measure these assessments because they call on a player's ability to mentalize, that is, to consider the state of mind of opponents. Correlations of activity in the prefrontal cortex to reasoning levels during such tasks present reasonable candidate targets for manipulation via electrical stimulation.

Level-*k* models assume that individuals heterogeneously employ a cognitive hierarchy of thinking types. These levels correspond to the number of recursions in consideration of the beliefs of others. People at level 0 assume all others are acting randomly, essentially a naive lack of consideration of others. People at level 1 consider the beliefs of the other but no more. Level 2 thinkers believe that other individuals are at level 1 and account for the iterative thinking at that level.

The most rational action in such a situation is to continue as many steps of iterative thinking until reaching the Nash equilibrium of 0, but the fact that people stop well short of this suggests that they are rational within certain limits. This bounded rationality may be due to cognitive limitations and the greater computational demands of each additional step (Camerer, Ho et al. 2004).

John Maynard Keynes encapsulated this notion by describing an old type of contest newspapers used to run. They would print a page full of headshot photos of 100 women, from which readers chose the six prettiest and sent them in. The winner was selected by finding which entry most closely matched the average preferences from all entrants (i.e. the most popular choices). So the task before the reader was not to rely on his own estimation of beauty but to imagine that of all the other entrants, most of whom would be unknown to him and whom he would have to consider in a general way. A player in this game might ignore the method of finding the winner and simply select photos according to her own preference, not considering other players at all. She might rank them according to how she thinks other players would prefer them. And she might rank them according to how she thinks other players will think all other players will prefer them. And so on. In the end, individual assessments of beauty didn't matter at all, rather the ability to gauge how all others would assess beauty (or how all others would assess the assessment of attractiveness).

In a modern laboratory and quantifiable version of this task, participants are instead directed to guess the number that is some fraction of the average of guesses by all participants. Responses to the game can be reasonably described with a cognitive hierarchy model. Given the parameters of a number between 0 and 100 and 2/3 the average of all players, at the lowest level, 0, a player responds randomly, without consideration of the structure of the game or interaction with other players. At one level up, level 1, the player considers all other players to be level-0 players and bases his choice on their random play, guessing their average guess to be something around 50 and then multiplying by 2/3, reaching 33. At level 2, the player figures all the other players are level 1, that they have submitted 33, so she multiplies that figure by 2/3 (effectively, 50 * 2/3 * 2/3) and arrives at 22. Continuing on iterated levels of thinking leads to elimination of dominated strategy until the player eventually arrives at 0, the game's Nash equilibrium.

At level 0, the player is considered naive because he responds without consideration of the structure of the game or other players. Level-1 players employ a model of the game space and respond to the actions they believe other players will take. Notably, there is no opportunity to adapt strategy based on the actions of other players, since those actions are not revealed. Nevertheless, level-2 players iterate a step further by imagining that their own actions are being considered by other players and therefore influencing the choices of the other players. At levels higher than level 2, the players consider to what extent they and their opponents are aware of mutual awareness (i.e. "I know that you know that I know … " etc.).

A key understanding of the level-*k* model is that it does not directly describe strategic sophistication ability, but rather the individual's assessment of others. An individual might be capable of high-level thinking, but if she assesses others as naive, she might only make choices indicative of level-1 thinking. Regardless of the actual outcome, the player's guess indicates her evaluation of the other participants and, therefore, their own k-level of thinking.

The mental calculations required for higher levels of thinking in the beauty contest game demand the multiplication of integers by fractions and fractions by fractions. Poor performance in the beauty contest (BC) might indicate not a low level of reasoning but rather poor mental mathematical abilities. For this reason, the second section of the task comprises calculations of integers multiplied by fractions and integers multiplied by a fraction, then multiplied by the same fraction, direct mimics of the mental calculations required of level 1 and level 2 thinking, respectively.

*Brain areas of level-*k *thinking*

Functional imaging studies have located several areas in the orbitofrontal cortex that covary with components of iterative thinking. The medial prefrontal cortex (mPFC) has long been associated with mentalizing, the consideration of the mental states of others (Frith and Frith 1999, Amodio and Frith 2006, Mitchell, Macrae et al. 2006). Notably, increased activity in the mPFC correlated with computational signals associated with a strategy that incorporates a player's consideration of his own choices on the decisions of others (Hampton, Bossaerts et al. 2008). That study concludes that the mPFC is part of a network that performs computations used in mentalization. Ventral portions of the medial prefrontal cortex (mPFC) in particular has been associated with various self-referential tasks as well as during mentalizing tasks suggest that it may be engaged in assessing the mental states of similar others by referencing the understanding of own personal feelings (Mitchell, Macrae et al. 2006). Similarly ventral areas of the mPFC were particularly active in high-level reasoning players during beauty contest trials against another human, as compared to those against a computer (Coricelli and Nagel 2009). In against-human trials only, higher activity was observed in that area in both high- and low-level reasoners, suggesting the mPFC is a center for strategic thinking about others' behavior. That same study also found higher activity in right and left dorsolateral prefrontal cortex (dlPFC) in high-level reasoning players at greater magnitudes than in low-level reasoning players, implicating the areas in a process of higher-level reasoning about others. They did not observe commensurate activity during mental calculation tasks that made similar computational demands.

To investigate how and at what point these constituent areas have a causal role in the process, they can be electrically stimulated for excitatory on inhibitory neuronal effect, accompanied by measurement of any behavioral changes, in particular the level thinking demonstrated by members of the experimental group.

tDCS

Noninvasive electrical stimulation allows for the modulation of neuronal activity at the regional level. In particular, transcranial direct-current stimulation (tDCS) provides for both excitation and inhibition, depending on the orientation of electrodes. The technique involves a low-level electrical current (often ≤ 2 mA) between two electrodes, from anodal to cathodal, though the brain. The areas under the anodal electrode are generally thought to undergo excitatory stimulation via sub-threshold depolarization of neurons. Crucially, tDCS is not believed to trigger action potentials in neurons, but rather to change the likelihood that an action potential will result in post-synaptic firing (Nitsche, Fricke et al. 2003, Coricelli and Rusconi 2010). The areas underlying the cathodal electrode, near where the current leaves the body are believed to be hyperpolarized, resulting in an inhibitory effect.

Various studies have successfully stimulated the mPFC (Civai, Miniussi et al. 2015, Hämmerer, Bonaiuto et al. 2016, Zheng, Huang et al. 2016) and the dlPFC (Fecteau, Knoch et al. 2007, Fecteau, Pascual-Leone et al. 2007, Boggio, Campanhã et al. 2010, Hecht, Walsh et al. 2010, Minati, Campanhã et al. 2012, Pripfl, Neumann et al. 2013) in decision-making tasks. Though similarly implicated in strategic thinking tasks, to our knowledge, modulation had not been attempted via tDCS. We proposed to influence behavior by altering neuronal activity using anodal tDCS. Because fMRI studies suggest an increase in mPFC and dlPFC activations during higher level-*k* thinking, we aimed to confirm the causal involvement of the areas. We hypothesized that anodal stimulation of the mPFC and the dlPFC would encourage excitatory activity, resulting in higher level-*k* performance during the beauty contest task in more participants, while sham stimulation groups would have fewer members exhibiting higher level-*k* thinking. We expected stimulation and sham groups would have similar outcomes of other measurements of tasks under stimulation, such as of memory and calculation.

To expand upon brain imaging findings of level-k thinking, we applied electrical stimulation to participants playing the same game used in a previous study. Results from that study described the frontal activity during the task (Coricelli and Nagel 2009). The selectively heightened activity of mPFC in trials against other humans and in dlPFC among higher-level-thinking participants suggested that these areas may play some role in generating strategic thinking. We hypothesized that electrical stimulation on the scalp above each of these areas separately would increase neuronal activity, which could in turn give rise to higher levels of iterative thinking. If that result were indeed found, it would suggest a causative role in iterative thinking for the targeted area.

Methods

Participants played the Beauty Contest game against other present participants, during which some underwent stimulation while others had sensors placed on their scalps but experienced sham stimulation. This experiment comprised the iterative thinking task called the Beauty Contest, which was conducted during transcranial direct-current stimulation (tDCS) or sham stimulation. That main task was followed by a calculation task, two digit span memory tasks and finally a series of questionnaires. The Beauty Contest requires participants to guess an average number, but that figure is influenced by both their own selection and the numbers they believe will be chosen by others. Because the target number is modulated by the choices of all other players, the most successful players consider how the other participants will choose. Based on the choices, each participant was assigned a precise level of thinking score and then categorized as high- or low-level thinking.

Participants

We recruited 64 healthy volunteers (32 female) to take part in a two-part study at the Mattarello Research Center of the University of Trento, Italy. Mean age of participants was 23.9

years (4.25, SD). Volunteers gave fully informed consent for the project, which was part of umbrella tDCS project approval from the University Ethical Committee. Each participant was screened to exclude risk of epileptic seizure, psycho-active medication and conditions including psychological or physical illness or history of head injury. For the first experiment, we recruited throughout the university, but after observing difficulty in the mathematical portions of the experiment, for the second experiment, we recruited in areas of the university frequented by students studying science, mathematics and engineering. All sessions were conducted with exactly 8 volunteers.

Experimental Design and Task

Experimenters fitted two electrodes over (experiment 1) the mPFC (Brodmann area 10) and visual cortex (BA17) and (experiment 2) the right and left dlPFC (BA9). After all had been fitted, participants then underwent 30 minutes of trans-cranial direct current stimulation while performing 50 trials of the experimental task. The experimental task consisted of a first session of 26 trials of the Beauty Contest game and then a second session of 24 trials of a mental calculation task. Next, they executed a two-part memory task: Forward Digit Span, measuring short-term memory and consisting of 2 to 14 trials, dependent on performance; and backward digit span, measuring working memory and consisting of 15 trials, regardless of performance. The experiment lasted about 90 minutes, with the 30 minutes of stimulation covering instructions for the beauty contest (so that any effect from stimulation also affected reading and comprehension of the instructions), completion of the beauty contest and completion of the calculation task. The stimulation period ended for 47 participants during the digit span memory tasks. Participants completed tDCS questionnaires, Raven’s Progressive Matrices and a cognitive reflection task to measure inter individual differences unrelated to effects of the stimulation. As it is unknown how long the effects of the stimulation last, we cannot exclude that participants were still under the fading influence of stimulation during these tasks.

The beauty contest game as designed by Coricelli and Nagel (2009) consists of a human condition and a computer condition. In the human condition, participants are directed to select an integer between 0 and 100 (inclusive), with the aim of being closest to a fraction M of the average of the number chosen by all participants: M*mean, in which six values are M<1 (1/8, 1/5, 1/3, 1/2, 2/3, 3/4) and another six values are M>1 (9/8, 6/5, 4/3, 3/2, 5/3, 7/4). We also included no-multiplier control trials in which M=1.

In the computer condition, they are told that the computer will select seven numbers from 0 to 100 at random. Playing only against the computer, and not against other participants, the participant wins if her number is closest to the product of M and the average of all eight numbers (i.e. her number and the seven randomly selected numbers). The same 13 values of M are used in the computer condition for a total of 26 trials. Computer and human condition trials were intermixed.

In the second session, which comprised the calculation trials, participants were instructed to find the product of a two-digit integer N and either M or M*M, in which M was the same set of multipliers used in the beauty contest, other than M=1, which was excluded. Each of the M values, other than M=1, appeared once in the N*M calculation and once in the N*M*M calculation (in which both Ms are the same multiplier). A correct response was judged to be +/-1 around the rounded up and down answer. E.g., if N*M = 22.2, the range 21-24 would be judged correct. Participants received 50 euro cents for each correct response. Participants received no feedback between trials. To avoid behavioral priming, the calculation task followed the beauty contest for all participants. Tasks were presented and responses recorded using MATLAB (The MathWorks, Inc., Natick, MA) using PsychToolBox extensions. Participants were seated at divided computer stations and could not interact with or see each other during the task.

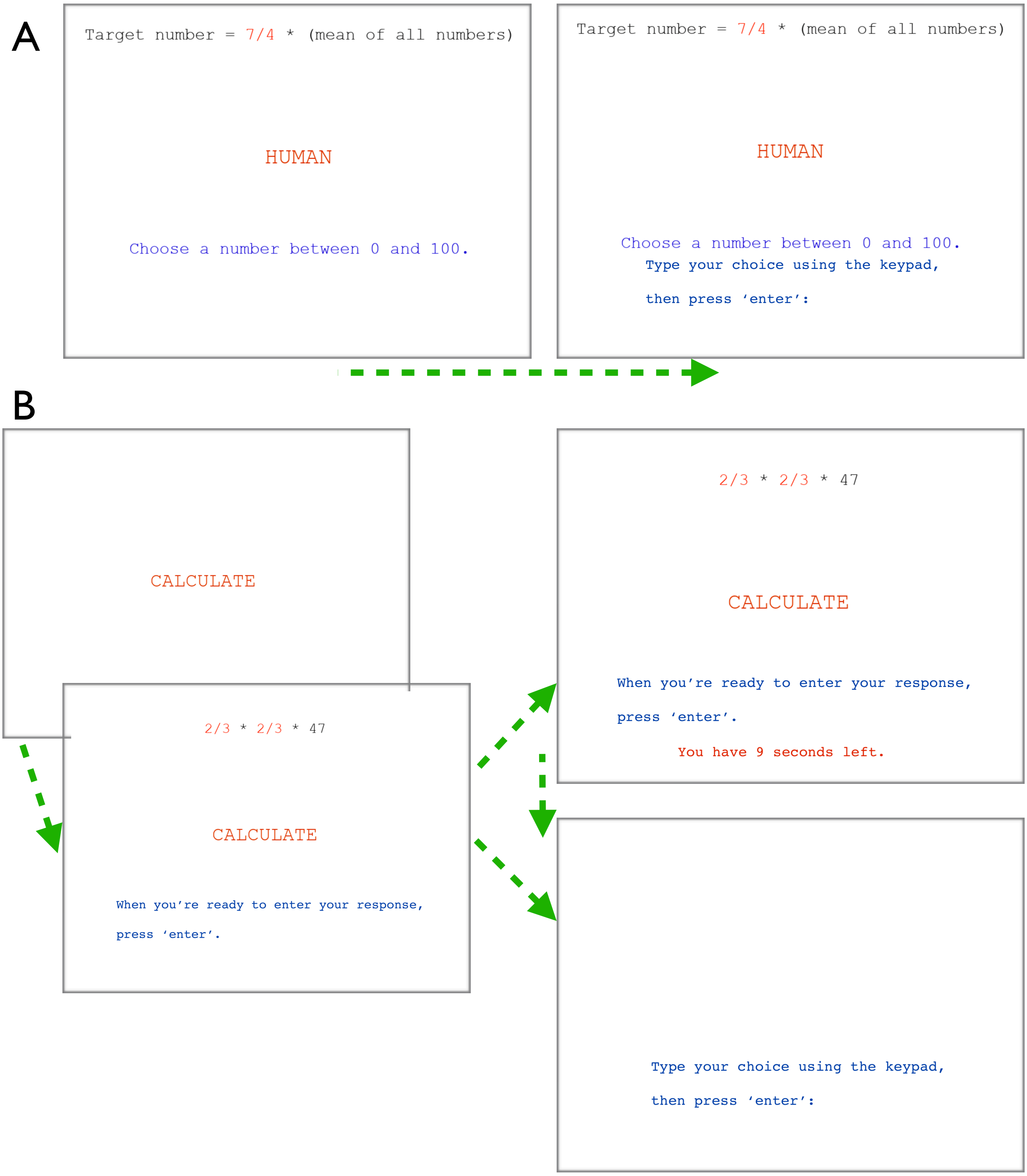


Fig. 4-1 Time course of beauty contest and calculation trials A) The screen progression for the beauty contest shows the condition (Human or Computer) at the beginning of each trial, along with a multiplier, in red, and the instruction. After 1-2 sec, the response prompt appeared. Participants answered at their own pace with no time limit. For each human-condition round, a reward of €5 was paid to a single winner from the session, or in case of a tie, divided evenly. For each computer round, the participant won €1 if she was closest to the target or €0.50 in case of a tie with one of the computer's numbers. B) In the calculation task, participants were instructed to calculate the product of (in M*1 condition) a fraction multiplied by an integer or (in M*2 condition) the product of a fraction, the same fraction and an integer, answering with an integer. The multipliers were the same set as those encountered in the beauty contest. When the prompt appeared, the participant had to press "enter" to continue to the response screen, at which point the prompt disappeared to encourage a response instead of continued consideration. If the participant did not press enter within 21 sec, a warning appeared that there were only 9 seconds remaining. If the participant did not enter a response within 30 seconds, the trial ended and the response was categorized as incorrect. Participants won €0.20 for each correct calculation. A calculation was considered correct if it was within 1 of the correct answer rounded up and rounded down to the nearest integers. All participants performed the calculation task after the beauty contest task.

Time course of experimental tasks

Each trial of the Beauty Contest game (**Fig. 4-1 A**) consisted of an information screen displayed for 1-2 sec, which included the condition of the trial (i.e. human or computer), the multiplier (M) and the instruction. The information screen remained visible, and a response prompt appeared, where the participant could type her response using the computer number pad, followed by the "enter" key. Choice was self-paced with no time limit, and the task continued as soon as the participant responded. Once the response was entered, a fixation cross appeared for 1-3 sec before the next trial began. Participants received no feedback between trials.

In session 2, the calculation task, a similar time course was followed: an information screen with type of calculation (N*M or N*M*M), multiplier M, integer N and instruction (2 sec), followed by the response prompt (**Fig. 4-1 B**). Choice was self-paced, but response time was constrained to 30 sec, with a warning after 21 sec. As soon as the response was entered, a fixation cross appeared for 1-2 sec.

<u>Questionnaires</u>

After the stimulation period, participants were asked four debriefing questions:

1. Please explain your reasoning in your first choice, M = 2/3, in the human condition.
2. Please explain your reasoning in your choice when M = 1/4 in the computer condition.
3. Did you have a general rule for the trials in the human condition?
4. Did you have a general rule for the trials in the computer condition?

<u>Short-term and working memory tasks</u>

Participants then completed a forward digit span task, in which a fixation cross appeared, followed by a series of digits, appearing on screen singly and sequentially for 1 sec each. Once the series was complete, a series of lines in the same amount as the series appeared on screen with the instruction for the participant to enter the sequence in order.

Participants were allowed to make corrections before pressing "enter", which began the next trial. The response was required to have exactly the number of digits as the prompt in order to proceed. The sequence began with a series of three numbers, increasing by one digit on each trial until an incorrect response or after a 9-digit series. At that point, a second sequence began with a series of three digits, increasing each round until an incorrect response or completion of the 9-digit series. We generated pseudo-random series of non-repeating digits, then used the same series and orders for all participants. The score was determined by the longest correctly completed response in either of the two series.

At the completion of the forward digit span task, instructions appeared explaining the backward digit-span task, in which participants were told they would once again see series of numbers, but that in order to correctly respond, they had to enter the series in the reverse order. A fixation cross once again appeared, followed by a series of single digits for 1 sec each, then followed by the series of blank spaces. Participants once again entered their response before typing "enter". In the backward digit span task, all participants completed three series each of lengths increasing from 4 to 8, for a total of 15 series. In the backward task, any number in the correct position was awarded a point. For the backward task, we used a prescribed set of series from Devetag & Warglien (2003).

Payment

Participant performance was financially motivated. Each subject received a €5 show-up fee and €0.20 for each correct calculation. In each 8-subject session, one human trial of the Beauty Contest was chosen at random, and one computer trial was chosen at random. Anonymized results were presented to all subjects on the computer screen. The participants who came closest in each selected trial were awarded an additional €5. In the case of ties, the prize was split evenly among all tying players. Each participant's total appeared on the screen at the end of the session, and they were later paid by bank transfer.

tDCS stimulation

The tDCS procedure applies a weak direct current into and out of the scalp via two electrodes, each sandwiched in saline-soaked sponges and spread with a layer of conductance gel. The constant current is delivered by a BrainSTIM battery-powered stimulator (E.M.S. Medical, Bologna, Italy). If participants reported uncomfortable tickling or itching sensations, experimenters added gel under the electrode to increase contact between electrode and scalp. None complained of pain during the session. During debriefing session, subjects reported mild sensations of tickling, tingling, warmth or pain, mostly at the beginning of the session, but some at the end, and a few in the middle. They reported that the sensations subsided quickly. The direction of the current can have different effects on the target area. Anodal stimulation encourages cortical excitability, while cathodal stimulation inhibits (Nitsche, Doemkes et al. 2007). Participants in experiment 1 were randomly assigned to receive anodal tDCS over the mPFC (N=16, 10 female, mean age=24.4) or sham stimulation (N=16, 7 female, mean age=24.4). Participants in experiment 2 were randomly assigned to receive anodal tDCS over the right dlPFC (N=16, 7 female, mean age=22.4) or sham stimulation (N=16, 8 female, mean age=24.5).

Stimulation and reference points were selected by simulating tDCS stimulation in SimNIBS software (Thielscher, Antunes et al. 2015) with various electrode placements and sizes, along with varied current strengths. Simulations were viewed in GMSH software (Geuzaine and Remacle 2009). Current density, the strength of current divided by the area of the electrode, has an effect on stimulation efficiency. For excitatory purposes, it is desirable to make the anodal electrode smaller to focus current and to make the cathodal electrode larger to diffuse it (Nitsche, Doemkes et al. 2007). We set a current density target of 0.07 mA/cm^2, which meant a decrease in size of the electrode in service of greater precision might require an attendant reduction in current strength, which would reduce both the reach and intensity of the stimulation (**Fig. 4-2**).

In experiment 1, the smaller, anodal electrode (4x4 cm) was placed over the vmPFC at the FPz position, according to the international EEG 10/20 system, and the cathodal electrode (5x7 cm) was placed over the visual cortex at the Oz position. At 16 cm^2 and a current of 1 mA, the current density at the anodal position was 0.0625 mA/cm^2.

In experiment 2, the smaller, anodal electrode (5x5 cm) was placed over the right dlPFC at the F4 position, according to the international EEG 10/20 system, and the cathodal electrode (5x7 cm) was placed over the left dlPFC at the F3 position, both positions calculated using an online location system (Beam, Borckardt et al. 2009). At 25 cm^2 and a current of 2 mA, the current density was 0.08 mA/cm^2.

For stimulated participants, a current was ramped up over 30 sec to 1 mA in experiment 1 and 2 mA in experiment 2, then kept constant for the length of the experimental tasks (no more than 29 min), followed by 30 sec ramping down. In the sham condition, the electrodes were placed as in stimulation condition, but stimulation halted after the 30 sec ramp-up, unbeknownst to the participant. The procedure sometimes produces an itching sensation at the beginning of a session, and sham participants would be exposed to that telltale sign, making it unclear to them if they were being stimulated or not (Gandiga, Hummel et al. 2006).

Experimenters and assistants set scalp locations by measuring fiduciary points, making measurements from those points, then marking on the scalp at electrode locations. Electrodes were held in place by a hairnet and surgical rubber straps. Conductance with the scalp was facilitated by applying a conductive gel to the underside of sponges soaked in a saline physiological solution.

Protocols

We set two stimulation protocols for each experiment: (full) stimulation and sham. For each session, four of each protocol were randomly distributed among participants. Participants were told they might undergo stimulation or sham. The experiment was double blind: neither participants nor experimenters (during the testing and analysis phases) were aware of which protocol was real and which was sham stimulation. Each session's stimulation was initiated and monitored from a central PC (schematic **Fig. 4-2 C, D**).

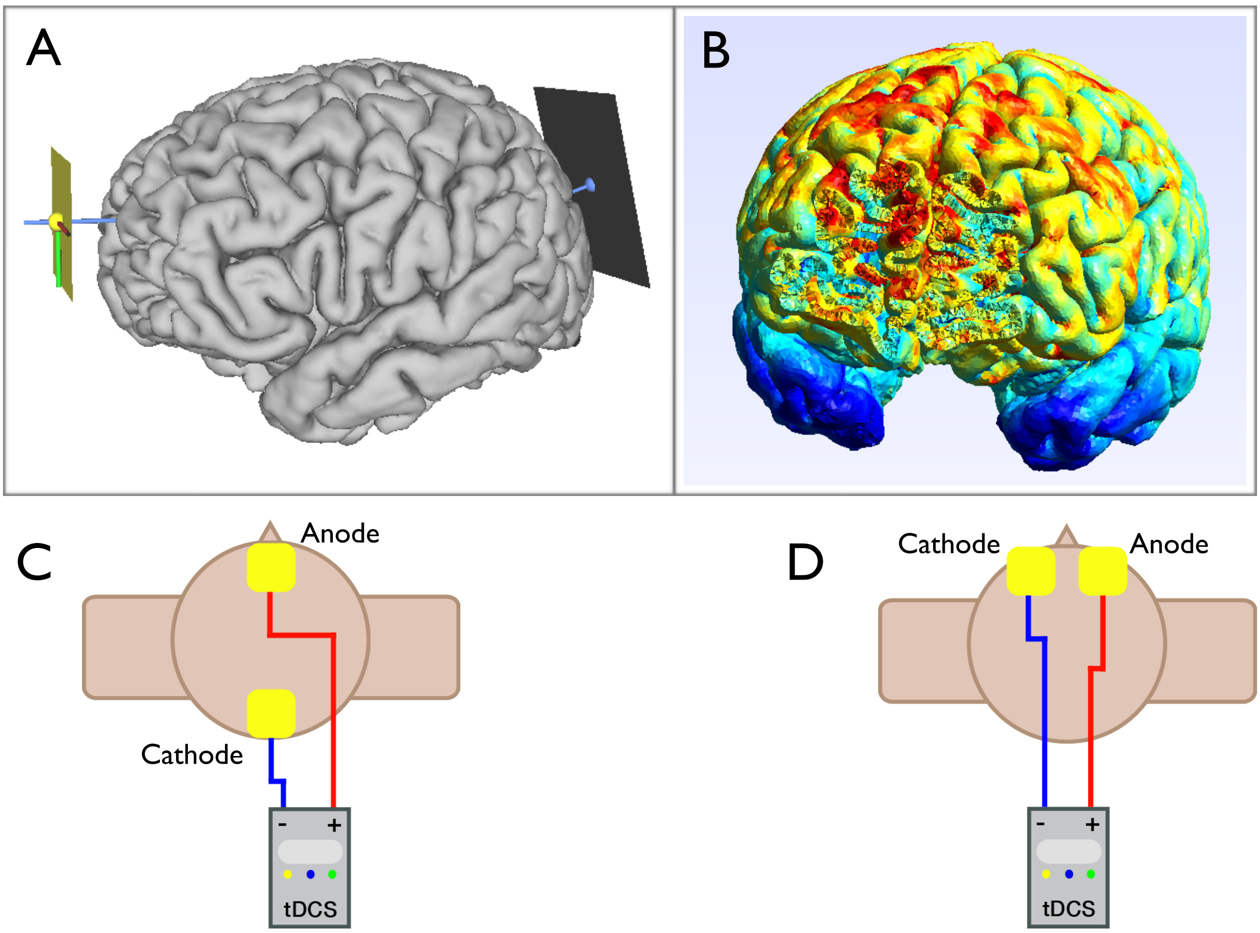


Fig. 4-2 Simulated electrode placement over mPFC and visual cortex and simulated effects of current A) Electrode were placed over a simulated brain according to MNI reference coordinates for mPFC and visual cortex, based on Zheng and colleagues (2016). The sizes of the electrodes and the current were adjusted to achieve a focused simulated stimulation, as viewed in a cross-section brain (B). Schematics of the electrode placement are shown for C) mPFC and visual cortex: 1 mA for 30 min and D) right dlPFC and left dlPFC: 2 mA for 30 min.

Statistical analysis

First, we calculated for each trial the quadratic distance *QDM* between the response and the theoretical level-*k* values based on the Cognitive Hierarchy model using the equation

$$QDM_{ij}k = (x_{ijM} - 50 * M^k)^2,$$

where *x* is the choice of participant *i* in human/computer condition *j* for multiplier *M* (Coricelli and Nagel 2009). The equation is solved for each level of *k* = (0,1,2). The minimum quadratic distance for each level indicated the level-*k* for that trial. When a participant had a majority (seven or more) trials of minimum distance for a level, she was assigned that level-*k*. If the participant did not have seven occurrences of any level, her level was set at level 0, random (**Table 4-1**).

Next, we used subject responses to calculate a precise level-*k* for each trial. The *k* is determined by the number of times the multiplier fraction is applied to the mean of the integer range, which means it can also be found in continuous values using the solving equation

$$k_i = log(r_i/50)/log(M_i)\,,$$

where *r* is the response to multiplier *M*. E.g., for *M*=2/3, a response of 25 indicates a level-*k* of 1.7). A response of 0 presents a calculation problem of precise level-*k*, so in those instances, we used a corrected precise value, in which responses were indexed by adding 1 to the response integer. They were then divided by 50.5 instead of 50, ensuring that responses of 100 would be the same as in uncorrected, since 101/50.5 = 100/50.

Then, for each experiment, we compared level-*k* outcomes between the two stimulation protocols. To accommodate outliers and non-normal distributions, we used the Kruskal-Wallis test to compare each subject's median precise level-*k* for first human trials and then computer trials (illustrated in **Fig. 4-3**). We also considered data from the calculation task, calculating the absolute distance (AD) from the correct answer. We compared AD across

Table 4-1: Modal minimum QD				
Experiment	**1 - vmPFC**		**2 - dlPFC**	
Protocol	*sham*	*full stimulation*	*sham*	*full stimulation*
Versus-human trials	Level 0	Level 0	Level 1	Level 1
Versus-computer trials	Level 2	Level 1	Level 1	Level 1

Table 4-1 Level-*k* instances by stimulation protocol and opponent type For each trial, we solved for the quadratic distance from levels 0, 1 and 2, classifying the choice as the level with the lowest quadratic distance. If a participant had seven or more trials of one level, she was classified as that level. Otherwise, she was classified as level 0 (random). These are the most common level types for each treatment and condition.

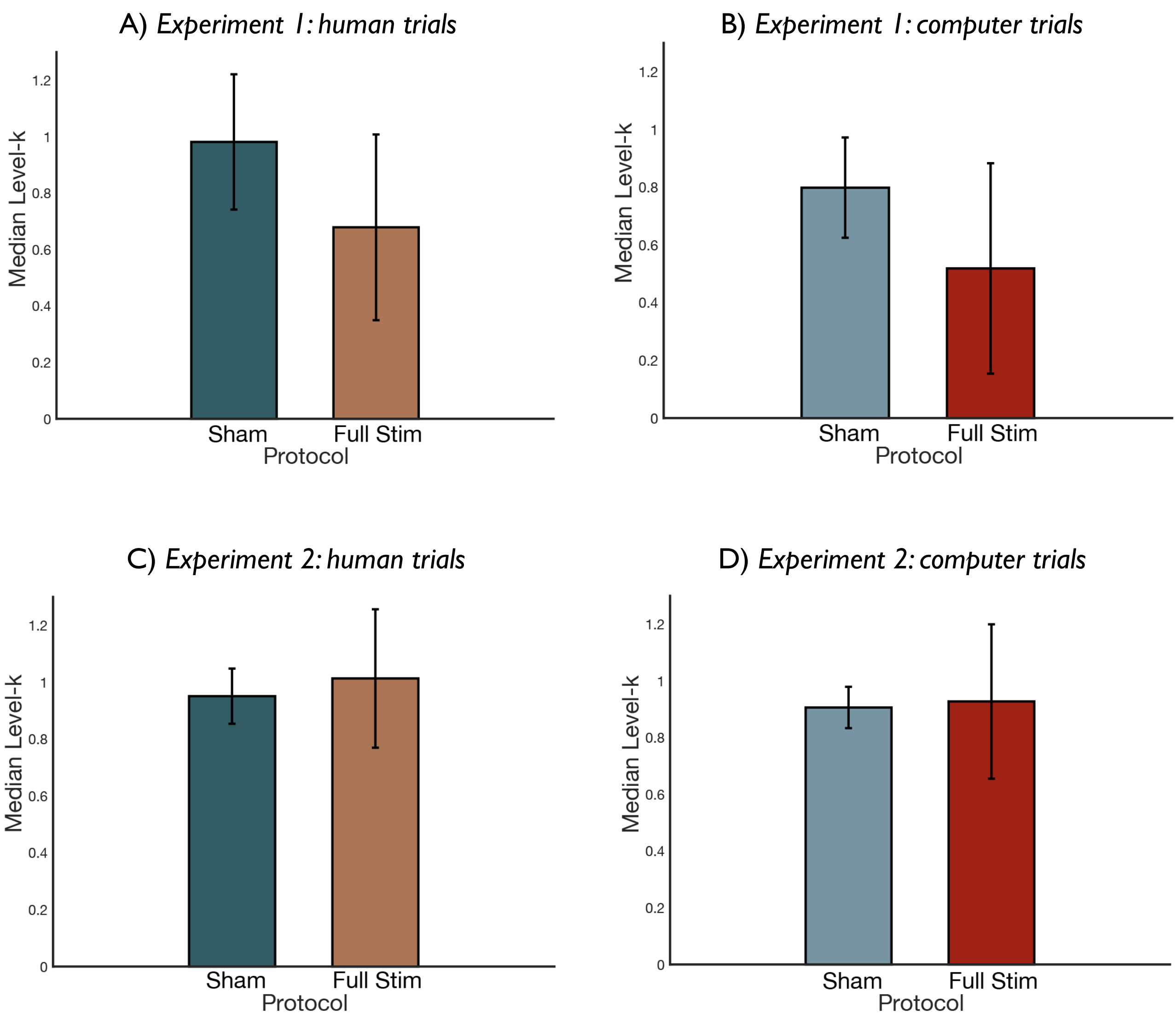


Fig. 4-3 Level-*k* in Beauty Contest by opponent and by stimulation protocol Precise level-k was calculated for each trial and then averaged by participant. These graphs illustrate the median level-*k* for the two experiments. There was no significant difference in the medians. Error bars are s.e.m.

all trials, as well as categorized into trials with a single multiplier (M*1) and trials with a double multiplier (M*2) (**Fig. 4-4**). Higher AD values indicate worse performance. We used mixed ANOVA to test for significance within subjects for multiplier level and between subjects for stimulation protocol, as well as for interaction between the two factors. Because sphericity of data tests failed, we used Greenhouse-Geisser tests for interpretation.

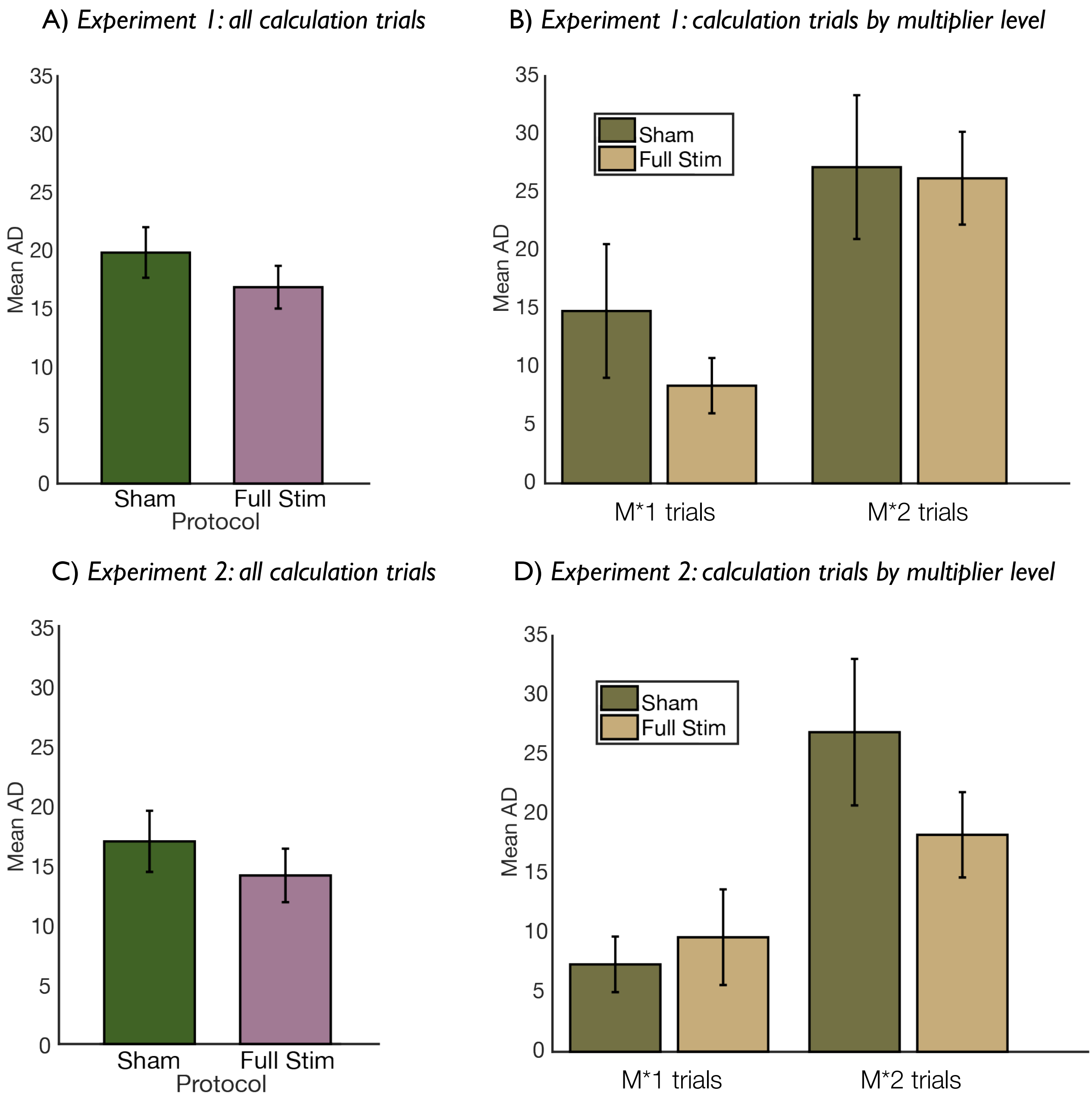


Fig. 4-4 Absolute distances in calculation task by stimulation protocol Mean absolute distance from correct answer in all calculation trials (A, C) by stimulation protocol, and mean distance in M*1 and M*2 trials (B, D), by protocol. Mixed ANOVA showed in both experiments an effect of multiplier level but not stimulation protocol alone, nor the interaction between stimulation protocol and multiplier level. Error bars are s.e.m.

Kruskal-Wallis tests were run with SPSS Statistics (IBM Corp., Armonk, NY). Mixed ANOVA tests were run with Stata, Stata Corp., College Station, TX.

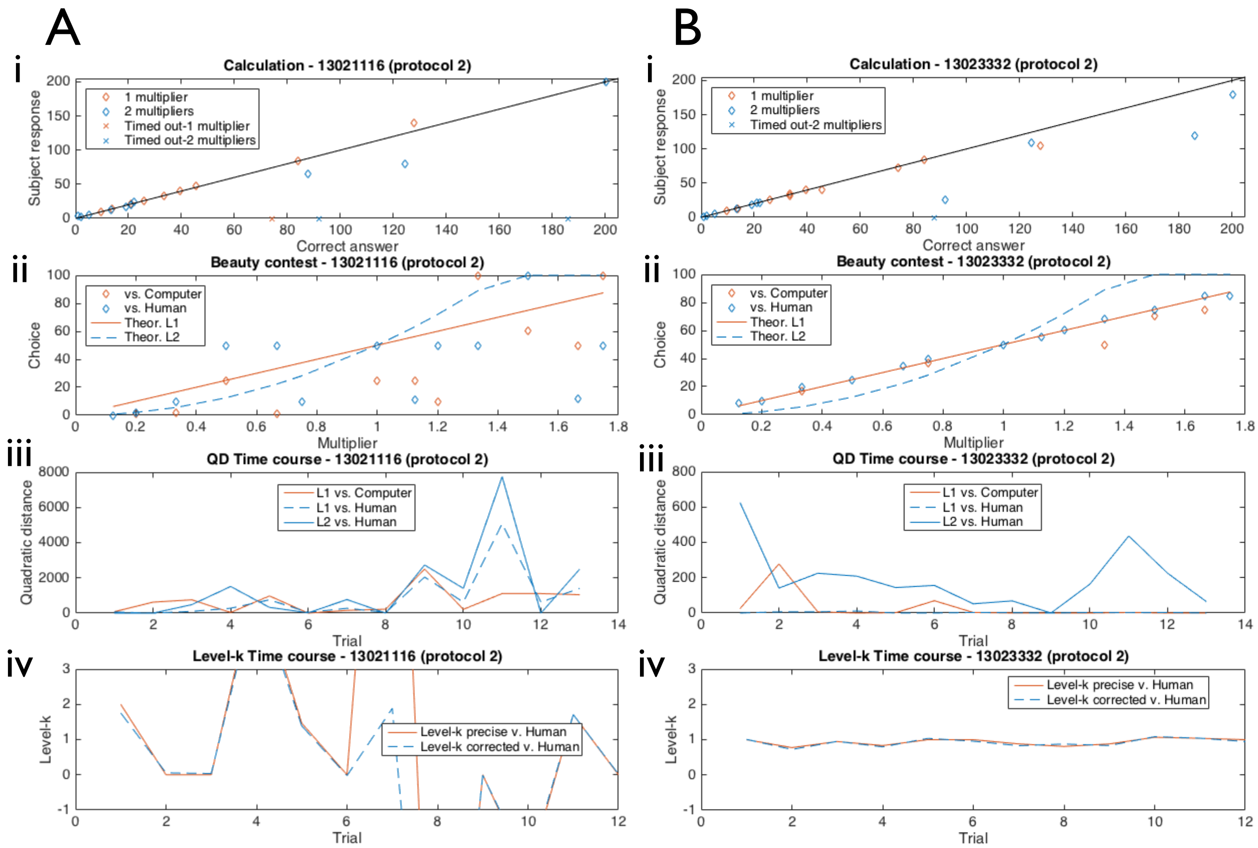
Results

In the first experiment, tests showed a significant effect of the multiplier on calculation performance ($F(1,30) = 32.36$, $p < .001$) but no significance in stimulation protocol, nor in the interaction between multiplier level and protocol. In the second experiment, we found the same pattern: a significant effect of the multiplier ($F(1,30) = 130.28$, $p < .001$) but no significance in stimulation protocol, nor in the interaction.

In experiment 1, a Kruskal-Wallis Test was conducted to examine the differences on level-*k* according to the stimulation protocol undergone. No significant differences were found between the two stimulation protocols in human trials (Chi square = 1.04, p = .309, mean rank sham=18.19, mean rank full=14.81), nor in computer trials (Chi square = 1.74, p = .817, mean rank sham=18.69, mean rank full=14.31). In experiment 2, a Kruskal-Wallis Test was also conducted to examine the differences on level-*k* according to the stimulation protocol undergone. No significant differences were found between the two stimulation protocols in human trials (Chi square = 0.05, p = .821, mean rank sham=16.13, mean rank full=16.88), nor in computer trials (Chi square = 0, p = .955, mean rank sham=16.41, mean rank full=16.59). (**Table 4-2**).

Experiment 1 (vmPFC)				
	chi-square	*p*	mean rank sham	mean rank full
versus humans	1.04	0.309	18.19	14.81
versus computer	1.74	0.187	18.69	14.31
Experiment 2 (dlPFC)				
	chi-square	*p*	mean rank sham	mean rank full
versus humans	0.05	0.821	16.13	16.88
versus computer	0	0.955	16.41	16.59

Table 4-2 Kruskal-Wallis test statistics table



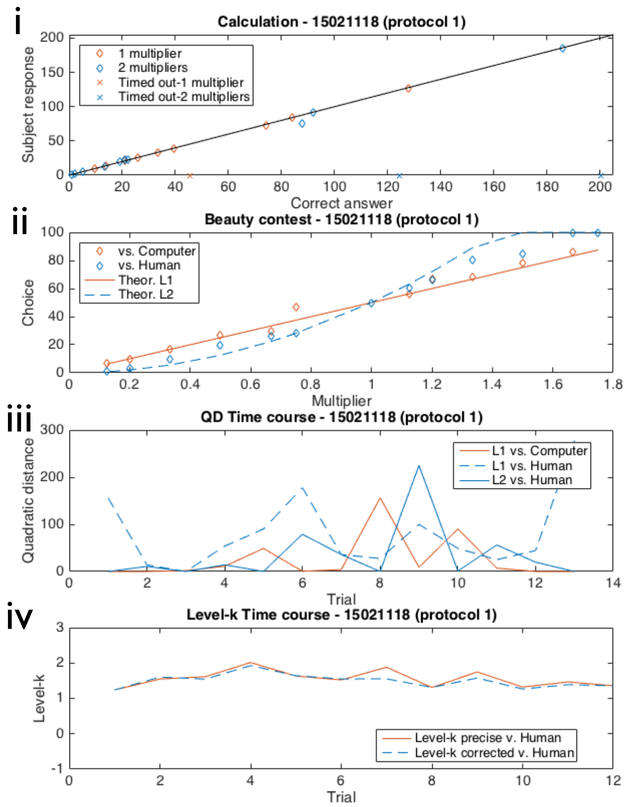


Fig. 4-5 Participant description visualizations Visualizations for typical level-*k* players show 1) performance in the calculation task, with the solid line representing the correct answer and subject responses shown as red diamonds for single-multiplier trials and blue diamonds for double-multiplier trials; 2) beauty contest responses, with the solid red line representing the theoretical level 1 and the dashed blue line, the theoretical level 2, with participant responses shown as red diamonds for computer trials and blue diamonds for human trials; 3) trial-by-trial quadratic distance from level 1 versus computer (red line), level 1 versus humans (dashed blue line) and level 2 versus humans (solid blue line); 4) trial-by-trial level-k precise and corrected versus humans.

A) A level-0 player's results show proficiency in calculations, but irregularity in the beauty contest against both computer and humans. B) A level-1 player has responses against both computer and human along the theoretical L1 line (graph 2) and shows consistent level-1 choices across all trials. C) A level-2 player has versus-human choices closer to the theoretical level 2 dashed blue line and has level-*k* choices above level 1 across all trials.

After finding no significant difference in the experimental task due to stimulation conditions, we explored the data in the hopes of finding guidance for a subsequent experiment. We first plotted calculation answers to ensure that a given participant did not have any mathematical limitations (**Fig. 4-5 [i]**). We then compared participant answers in both human and computer conditions to theoretical level 1 (the target for responses in computer trials) and theoretical level 2 (**Fig. 4-5 [ii]**). Next we plotted the time course of quadratic distances over the course of the task for each opponent type (**Fig. 4-5 [iii]**). And finally, we plotted trial-by-trial level-*k* for human trials only (**Fig. 4-5 [iv]**). We categorized players by their choices over the course of the task (**Fig. 4-6**) in an effort to see if their understanding of the task appeared to change (as illustrated in **Fig. 4-7**). We found no difference in categorization types between the types of stimulation in either experiment. Several players in each experiment, and in both stimulation protocols, appeared to change their levels of thinking over the course of the experiment: some from level-0 to level-1, and some to level-1 to level-2.

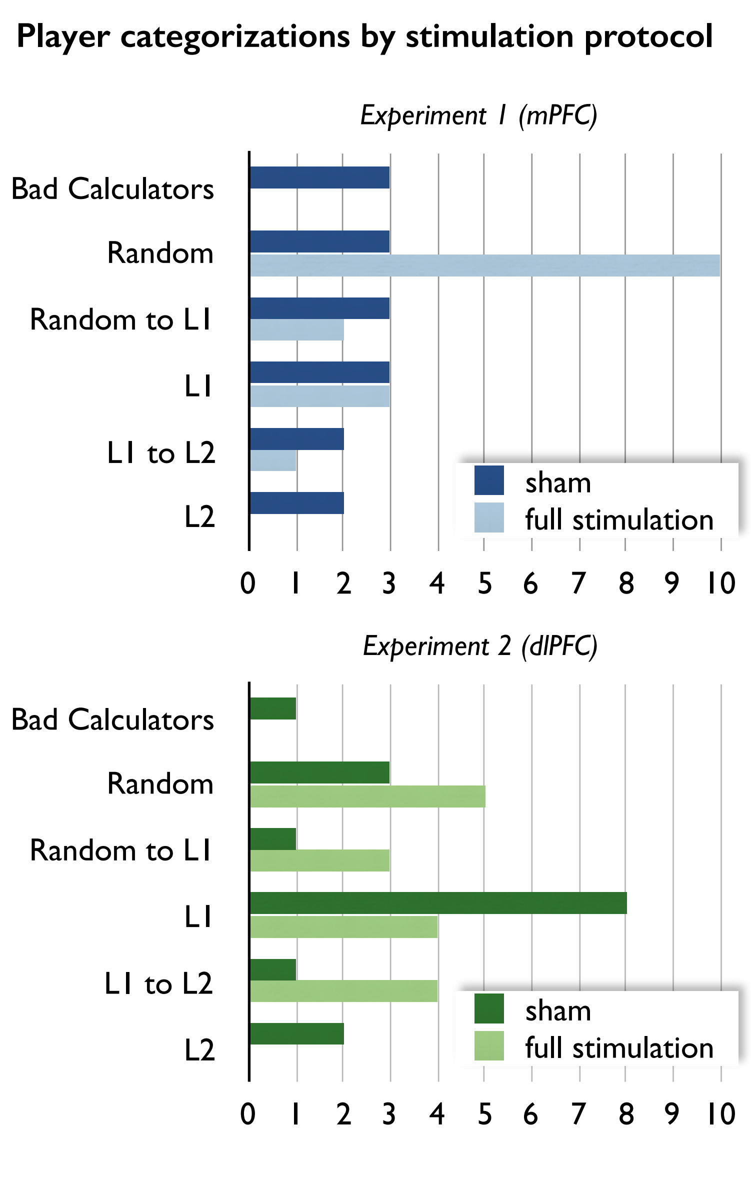


Fig. 4-6 Player categorizations by stimulation protocol

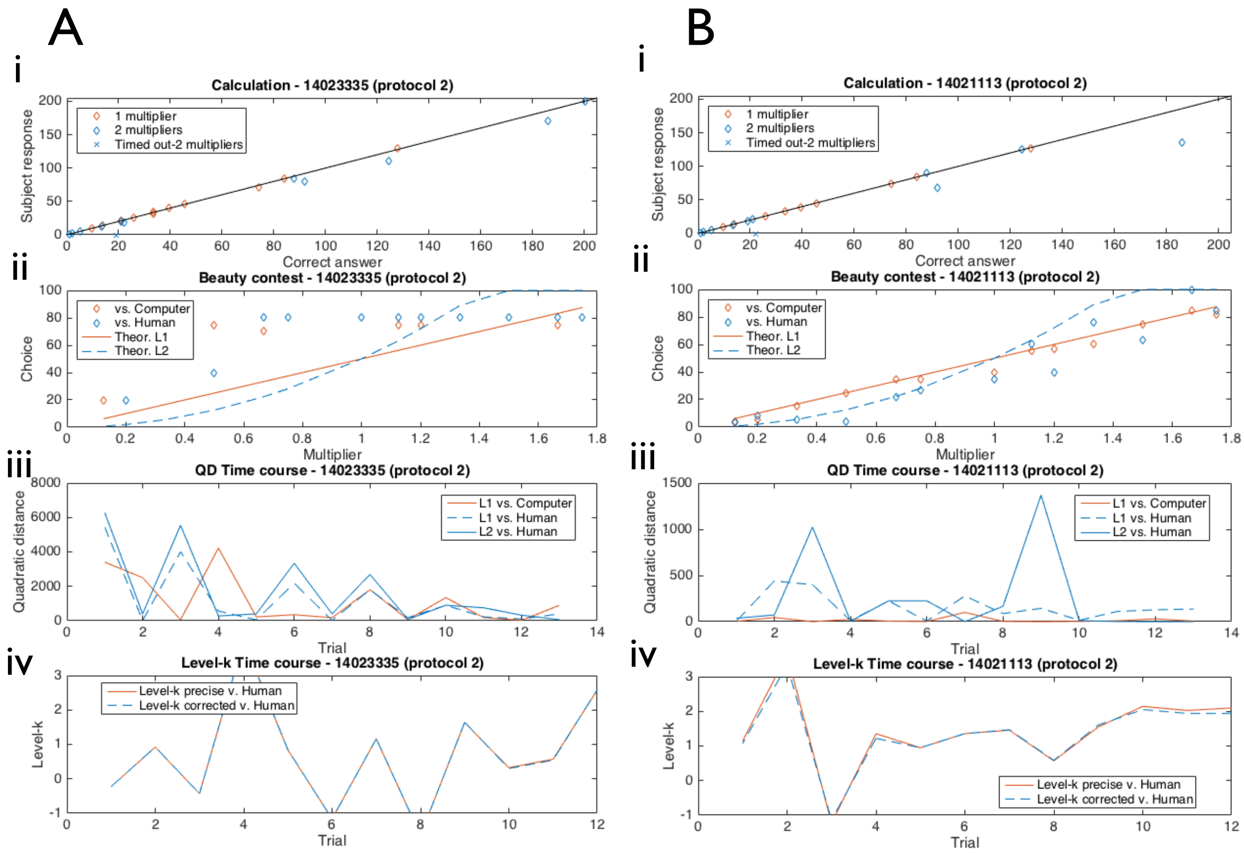


Fig 4-7 Evolving participant description visualizations Some players' choices suggested an understanding of the task that changed throughout the session. A) An example of a player who at first seems to play at level 0, in later trials reduces the quadratic distance to level 1 in both computer and human trials. B) A player who begins making choices around level 1 or level 0 by the late trials of the task makes several level-2 choices in a row.

Discussion

We used tDCS to test 64 volunteers, applying full or sham stimulation over first the mPFC and later the dlPFC while participants played a strategic thinking game against each other. Results showed a mild trend for higher levels of thinking among the sham group in the first experiment (**Fig. 4-3 [A]**), which aimed to stimulate the mPFC with anodal tDCS. Meanwhile, results from the second experiment, in which the dlPFC was targeted, did not indicate even a trend in level-*k* thinking between sham and stimulation (**Fig. 4-3 [C]**).

The lack of significant outcome may indicate a problem with our experimental design. We may have targeted areas with less precision than necessary. That could be improved with a different arrangement of electrodes, as well as by consulting anatomical scans of participants from which to map electrode positioning.

On the other hand, the problem may lie not with design. It may be that even if we successfully stimulated the targeted areas, any change in neuronal activity has insufficient effect on task performance – or that any change is below measurable levels. It is also possible that the areas we targeted for manipulation, while involved in the process of iterative thinking as described in Coricelli and Nagel (2009), do not play a singularly sufficient role in that process. Despite results from previous studies, it is possible that mPFC and dlPFC do not play causal roles in level-*k* thinking.

The tDCS technique itself is not fully proven and may not have an effect on brain activity that can modulate decision making at all. Though tDCS has been used with success in stimulating motor areas, its effectiveness in decision-making studies has been far less demonstrated. The dlPFC has been a promising area in those successful studies, but the mPFC less so.

Problems could lie within the task as well. Though the Beauty Contest game has been successfully deployed by many groups across numerous studies, the presentation of the task proved particularly difficult in this setting. A key component is ensuring the task instructions are clear but without guiding participants to higher levels of thinking. Examples and practice trials have high potential to prompt the higher levels we are trying to measure as arising from stimulation and so were excluded.

The trends we did see in data suggest that stimulation has a surprising effect on level-*k* thinking, that is, attenuation (**Fig. 4-3**). Anodal stimulation, which typically induces cortical excitability, on the mPFC appears to have diminished the number of high-level thinkers. This could suggest greater specificity is needed in identifying and stimulating locations of level-*k* reasoning in the mPFC, possibly at the individual level.

Consistent with previous studies, we positioned electrodes to stimulate the right dlPFC and to diminish current under the reference electrode so as not to inhibit the left dlPFC. If that dispersal was insufficient, it would create an inhibitory effect in the left dlPFC. If the area's role in iterative thinking is bilateral, then an effective stimulation would both excite and inhibit, possibly producing a net-zero change in behavior.

Without feedback, participants should not learn over the course of the task, but with repeated exposure to similar prompts with minimal variations, reasoning may change resulting in players who at first act as level 0 or level 1 players realizing a more sophisticated strategy – what some have claimed is a phenomenon called Epiphany Learning (Chen and Krajbich 2017). Indeed we saw evidence of this phenomenon in nine participants whose behavior indicated a transition from level 0 to level 1 in the course of the task and eight with indications of moving from level 1 to level 2 (examples in **Fig. 4-7**).

In both studies, we found significant effect of the multiplier level in the calculation task, but not of the stimulation protocol alone, nor of its interaction with multiplier level. This suggests that our groups comprised a spectrum of mathematical abilities, with some participants able to calculate well enough on M*1 trials but fewer calculating well on M*2 trials. In a group where most participants were not able to calculate either multiplier level well, there would be no effect of multiplier level on the absolute distance from the correct answer. Furthermore, because this was a control task ensuring that any effect detected in the beauty contest is not due to inhibition or encouragement of calculation abilities, equal performance between stimulation protocols is prerequisite for further inferences from the beauty contest. An effect accompanying stimulation of the mPFC would be a surprise, since Coricelli and Nagel detected no activity in the mPFC associated with the control calculation task (2009).

General Discussion

The moment-to-moment decisions human beings make throughout their waking lives come atop a mountain of prior experience encountered in a variety of domains and situations. Because we know that these occurrences are not independent and come together to form a continuous experience, it is reasonable to suspect that some of these choices and consequences must commingle and inform one another. Yet many implications and mechanisms of transfer are to this point unexplored. We poked at this problem, asking how differences in decision-making conditions might influence changes in later choices and learning. In four chapters, we considered similarities between decision-making regret and moral decision making, differences in learning with age, emotional priming in learning and electrical stimulation of iterative thinking.

Moral decision making

Understanding how humans make low-importance, quantifiable decisions may aid understanding of broader choices with larger impact like moral decisions. The brain processes observed to underly certain types of economic decision making and moral decision making appear to overlap. Perhaps not surprisingly, similar injuries to and deficiencies in the areas implicated in these processes give rise to similar hindrances to those processes. Both people with high pscyhopathy indications and patients with lesions in the ventromedial prefrontal cortex (vmPFC) experience regret but do not apply it as fully to future decisions as healthy subjects. People of both these groups also make more utilitarian moral decisions, rejecting the emotional attenuation seen in the choices of healthy subjects.

The confluence of these deviations from behavior seen in healthy populations suggest possibilities for the vmPFC's particular role. It could be the site for learning that both provides error signals that inform regret and provides foundation for moral decisions regarding others. It may be a location for conflict between utilitarian and emotional considerations; or more basically, a way station that simply delays decisions until the conflict can be worked out elsewhere. It may also be an integrator of emotion into complicated decisions across many domains, putting to use the experiences of the past in consideration of future consequences.

Aging, regret, risk and learning

We examined data from two decision-making tasks designed to compare choice behavior and learning between older and younger adult age groups. In the first task, participants chose between two lotteries with different probabilities of winning or losing uniform amounts. In some trials, only the outcome of the chosen lottery was revealed, while in others, both outcomes were shown. In a subsequent probabilistic learning task, participants selected between pairs of symbols, each of which had hidden probabilities of delivering rewards or punishments. For some symbol pairs, both outcomes were shown, while for other symbol pairs, only the outcome of the chosen symbol was shown. We characterized their choice influences in the lottery task using mixed regressions, and we analyzed learning behavior via computational modeling. Based on the findings of a previous study, we hypothesized that older adults would experience regret to the same extent as younger adults, but that they would anticipate and avoid it in subsequent choices to a lesser extent. We further hypothesized that the two age groups would learn similarly in partial-information feedback contexts but that learning rates would differ in complete-information counterfactual contexts.

We found in fact that in the lottery task, both groups were significantly emotionally affected by complete-feedback negative outcomes, the condition for regret. However, younger adults reported significantly more negative reactions to these outcomes. Yet when it came to anticipating or avoiding regret, both groups incorporated it into their choices, but not differently. The attenuated reaction of older adults to negative outcomes is consistent with a positivity effect that accompanies aging. Older adults pay selective attention away from the negative early in processing of experience. Even later in appraisal, younger adults tend to dwell on negative experiences (Charles and Carstensen 2010, Carstensen and DeLiema 2018). Older adults, while not suffering the negative emotional consequences to the same extent, still applied the experience into future choices at about the same level. This suggests a benefit of age: avoidance of potentially negative outcomes but at lower emotional cost. In the learning task, younger adults had better outcomes in terms of earnings, but it appeared to be due to overall performance, rather than sustained ability in counterfactual learning. Both age groups earned more in complete-feedback trials than in partial-feedback trials, and younger adults earned more in complete-feedback trials than older adults did. The difference between trial types, however, was not significantly different between age groups. This indicates that both groups are more successful when incorporating counterfactual learning and that younger adults simply outperformed older adults generally in the learning task. This fails to support the findings of Tobia and colleagues (2016), who found that older adults were more responsive to counterfactual gains but that this actually hindered subsequent choices. Further analyses should explore the differences in gains between positive and negative counterfactual outcomes between age groups. This experiment would also benefit from extension to neuroimaging to compare to the fMRI results of the Tobia study.

Although our hypotheses did not directly address risk preference, the lottery task incorporated possibilities of both gain and loss, so we considered risk as a regression factor in our analysis. We found that it did not play a significant role in the choices of younger adults but that it did in older adults, who avoided it, and particularly in complete-feedback trials. Their risk aversion and the risk tolerance of younger adults is partially consistent with Tymula and colleagues' general assessments of risk preference variations across adulthood (2013). Our results show lower risk tolerance among older adults after encountering regret situations but risk preference on par with younger adults in partial-feedback contexts. Our results support a similar distinction in risk preference observed between younger adults and older adults with Multiple Sclerosis, in which younger adults were risk neutral in a wheel of fortune lottery, while older adult patients were risk averse (Simioni, Schluep et al. 2012). Another study that addressed both regret and risk showed that healthy older adults did not change risk tolerance after experiencing regret, while younger adults and depressed older adults did (Brassen, Gamer et al. 2012). The regret-eliciting task in this case was a "hot" devil game, like the balloon analogue risk task, in which risk computation is not explicit, and learning to tolerate more risk eventually leads to higher rewards. Our results bolster support for the notion that regret leads to risk-seeking behavior in younger adults but that older adults are resistant to this and may even become more risk-aversive after experiencing regret. A future study with this as hypothesis could more specifically address this possibility.

In tasks where risk is explicitly stated, decision-from-description paradigms, such as the wheel of fortune lotteries, older and younger adults typically perform similarly (Mata, Josef et al. 2011), consistent with our results in partial-feedback trials, but in contrast to increased risk aversion we saw in complete-feedback contexts. Yet older adults reported feeling less badly about regret outcomes. This raises the possibility of a relationship between reduced experience of regret and increased risk aversion, even with stability of regret anticipation. A future study could explore this potential relationship specifically. Other studies should examine risk and regret in variations of paired tasks with attention to the conditions that give rise to regret and to the particular types of risk each task employs.

Regret induction

We had participants play a strategic competitive investment game with asymmetric roles, encouraging variations in strategy to reveal patterns of learning behavior. In previous studies, this game has been used to characterize individual learning, specifically when behavior is compared to a hybrid Experience-Weighted Attraction (EWA) model that nests both reinforcement learning (RL) and belief-based learning. Belief-based learning (BBL) requires an understanding of the structure of the game as well as anticipation of the strategy of the other player. Just prior to playing the game, participants played a wheel of fortune lottery designed to induce regret, relief, disappointment or satisfaction. In repeated games, regret has been shown to influence learning (Camille, Coricelli et al. 2004). Our hypothesis was that those exposed to complete-feedback counterfactual emotions would be primed for thinking about alternative situations rather than only the choice they had made. We suspected that they would engage to greater extents and at greater rates in the more sophisticated BBL than in simpler RL. We would measure these outcomes with a parameter in the hybrid experience-weighted attraction model that indicates the balance of BBL and RL.

A requirement of computational modeling is demonstrating that the model used is the best of those available. Because we planned to use a model that incorporated both RL and BBL, it was necessary to consider those comparatively simpler models on their own. Our model comparison showed, to our surprise, that the BBL model outperformed both RL, EWA and a reinforcement learning model that also incorporates counterfactual outcomes. Our BBL model features a learning rate parameter, but this is distinct from the weighted parameter of the EWA model that indicates relative utilization of RL and BBL. Gauging this weight among different groups was central to our hypothesis. Because we had to reject the use of the EWA model, we could not test our hypothesis. Previous studies that modeled patent race gameplay have consistently found EWA to be the best-fitting model. A possible reason for the unexpected outcome of our model estimations may be the opponent algorithm we programmed. The algorithm was based on fictitious play, a form of belief-based learning. In past studies, players have played against other humans or against pooled responses by humans, in which the opponent's choice was selected from a number of human choices on that trial number (Zhu, Mathewson et al. 2012). Though we believed participants would behave and learn similarly against an algorithm set to play with similar choices to a human, it may have inadvertently prompted mirroring behavior from our participants. It is unlikely that the priming task had such a broad effect, since we detected no pattern across priming types, nor control participants who did not undergo priming.

In future studies, we should consider using a different type of opponent, be it a variety of computerized opponents with varied models and parameters, actual human opponents or the pooled opponent play employed by Zhu and colleagues. We at first rejected using the pooled play because we believed the play would seem disjointed, and we wanted a realistic opponent. We thought that even if participants knew their opponent was an algorithm, they would still use the same learning behaviors to try to win.

Electrical stimulation of sophisticated thinking

In competitive strategic conditions, consideration of others and an ability to gauge their mental state is an advantage. Previous studies show that individuals consider the actions of others to a range of extents: from not at all to assessment of others as also highly considerate. We used a beauty contest calculation competition in which each participant guessed a target number that would be influenced by their own choice, as well as the choices of others. The average of all the choices multiplied by a fraction determined the target number, and so for fractions less than 1, the equilibrium goal was 0. The degree of sophistication of thinking, then, was measurably greater as a participant's choice approached 0. In previous imaging studies, these higher levels of thinking were accompanied by increased activity in the medial prefrontal cortex (mPFC) and in both right and left dorsolateral prefrontal cortex (dlPFC) (Coricelli and Nagel 2009). A well-established limitation of imaging studies is their inability to establish causation, and in the case of fMRI, even sequential order. One aim of stimulation studies is to fill that gap by interrupting or encouraging processes by targeting brain areas during activities whose results are well-studied. If observable behavior is changed, it indicates that the targeted area plays a particular role in the process.

We hypothesized that the "neural signatures" of higher-level strategic thinking in mPFC and dlPFC indicated responsibilities of these areas for more sophisticated reasoning. To test the hypothesis, we targeted those areas using transcranial direct-current stimulation (tDCS). If the group of subjects receiving stimulation had higher levels of thinking or higher levels at higher rates, it would indicate a causative role for the targeted area. Our stimulations, however, did not produce different behavioral results between sham and stimulation. This outcome may be an indication that the targeted areas of dlPFC and mPFC do not play causative roles in higher-level reasoning. The higher activity observed in these areas in fMRI studies during higher-level reasoning simply may not drive the process that gives rise to the behavior.

It would be overreach to say this is conclusive, however, especially given the unproven nature of tDCS as a technique. Though studies targeting motor actions and visual perception have seen successful manipulation via tDCS, the record in strategic decision making is shorter. A mild trend in the stimulation of mPFC was in the opposite direction of our hypothesized result: lower levels of thinking in the group that received full stimulation. This suggests less-sophisticated thinking as a result of excitatory activity in the mPFC. If further study supported this trend, it could, along with the established imaging results, indicate a regulatory or mediating role for mPFC, rather than a generative one. However, some evidence from previous studies indicates that anodal tDCS stimulations for more than 26 minutes at some point cease having excitatory effects and reverse to inhibitory (Thair, Holloway et al. 2017). This is a less likely explanation because we examined trial-by-trial level-*k* activity at the individual level and did not see indications of a reversal from higher- to lower-level activity. If anything, level thinking appeared to increase in some participants over the course of the experiment. Other studies have demonstrated inhibitory effects of cathodal stimulation but no behavioral change due to anodal stimulation, suggesting that targeted brain areas may be active at an optimal level and unable to be further excited (Antal, Nitsche et al. 2001).

Reviews of tDCS studies have found variability depending on age and sex, as well as mental states of alertness, sleep debt, time of day, and even recent caffeine consumption (Krause and Cohen Kadosh 2014). Future studies could control for more of these factors, conducting all experimental sessions at the same time of day, requesting participants abstain from caffeine for a period before the session and asking for information about recent sleep habits.

These multiple attempts to make modulations to learning and decision making processes largely did not produce measurable effects. Though these were often failures to reject null hypotheses, these inquiries together reveal the resilience of a set of learning and thinking processes that can withstand perturbations in the lab.

References

Addis, D. R., A. T. Wong and D. L. Schacter (2008). "Age-related changes in the episodic simulation of future events." Psychological science **19**(1): 33-41.

Amodio, D. M. and C. D. Frith (2006). "Meeting of minds: the medial frontal cortex and social cognition." Nature Reviews Neuroscience **7**(4).

Anelli, F., E. Ciaramelli, S. Arzy and F. Frassinetti (2016). "Age-Related Effects on Future Mental Time Travel." Neural Plasticity: 1-8.

Ansari, A., R. Montoya and O. Netzer (2012). "Dynamic learning in behavioral games: A hidden Markov mixture of experts approach." Quantitative Marketing and Economics **10**(4).

Antal, A., M. A. Nitsche and W. Paulus (2001). "External modulation of visual perception in humans." Neuroreport **12**(16): 3553-3555.

Baskin-Sommers, A., A. M. Stuppy-Sullivan and J. W. Buckholtz (2016). "Psychopathic individuals exhibit but do not avoid regret during counterfactual decision making." Proceedings of the National Academy of Sciences **113**(50): 14438-14443.

Baskin-Sommers, A. R., C. S. Neumann, L. M. Cope and K. A. Kiehl (2016). "Latent-variable modeling of brain gray-matter volume and psychopathy in incarcerated offenders." Journal of abnormal psychology **125**(6): 811.

Bault, N., G. Coricelli and A. Rustichini (2008). "Interdependent Utilities: How Social Ranking Affects Choice Behavior." PLoS ONE **3**(10).

Bault, N., M. Joffily, A. Rustichini and G. Coricelli (2011). "Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process." Proceedings of the National Academy of Sciences **108**(38): 16044-16049.

Bault, N., S. Palminteri, V. Aglieri and G. Coricelli (2018). Reduced value contextualization impairs punishment avoidance learning during aging, Center for Mind/Brain Sciences (Cimec), University of Trento, Trento, Italy Laboratoire de Neurosciences Cognitives, INSERM U960, Ecole Normale Supérieure, Paris, France Institut de Neurosciences de la Timone, UMR 7289, CNRS and Université Aix-Marseille, France Department of Economics, University of Southern California, USA.

Bayer, H. M. and P. W. Glimcher (2005). "Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal." Neuron: 129-141.

Beam, W., J. J. Borckardt, S. T. Reeves and M. S. George (2009). "An efficient and accurate new method for locating the F3 position for prefrontal TMS applications." Brain Stimulation **2**(1): 50-54.

Bechara, A., A. R. Damasio, H. Damasio and S. W. Anderson (1994). "Insensitivity to future consequences following damage to human prefrontal cortex." Cognition **50**(1-3): 7-15.

Bechara, A., H. Damasio, D. Tranel and A. R. Damasio (1997). "Deciding Advantageously Before Knowing the Advantageous Strategy." Science **275**(5304): 1293-1295.

Bechara, A., D. Tranel and H. Damasio (2000). "Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions." Brain **123**(11): 2189-2202.

Bechara, A., D. Tranel, H. Damasio and A. R. Damasio (1996). "Failure to Respond Autonomically to Anticipated Future Outcomes Following Damage to Prefrontal Cortex." Cerebral Cortex **6**(2): 215-225.

Beer, J. S., E. A. Heerey, D. Keltner, D. Scabini and R. T. Knight (2003). "The regulatory function of self-conscious emotion: Insights from patients with orbitofrontal damage." Journal of Personality and Social Psychology **85**(4): 594.

Behrens, T. E. J., L. T. Hunt, M. W. Woolrich and M. F. S. Rushworth (2008). "Associative learning of social value." Nature **456**(7219): 245-249.

Bell, D. E. (1982). "Regret in Decision Making under Uncertainty." Operations Research **30**(5): 961-981.

Blair, R. J. (1995). "A cognitive developmental approach to mortality: investigating the psychopath." Cognition **57**(1): 1-29.

Blair, R. J. R. (2007). "The amygdala and ventromedial prefrontal cortex in morality and psychopathy." Trends in Cognitive Sciences **11**(9): 387-392.

Boggio, P., C. Campanhã, C. A. Valasek, S. Fecteau, A. Pascual-Leone and F. Fregni (2010). "Modulation of decision-making in a gambling task in older adults with transcranial direct current stimulation." European Journal of Neuroscience **31**(3): 593-597.

Bopp, K. L. and P. Verhaeghen (2005). "Aging and verbal memory span: a meta-analysis." The journals of gerontology. Series B, Psychological sciences and social sciences **60**(5): 33.

Bopp, K. L. and P. Verhaeghen (2005). "Aging and Verbal Memory Span: A Meta-Analysis." The Journals of Gerontology: Series B.

Brand, M. and H. J. Markowitsch (2010). "Aging and decision-making: a neurocognitive perspective." Gerontology **56**(3): 319-324.

Brassen, S., M. Gamer, J. Peters, S. Gluth and C. Buchel (2012). "Don't Look Back in Anger! Responsiveness to Missed Chances in Successful and Nonsuccessful Aging." Science **336**(6081): 612-614.

Byrne, R. (2002). "Mental models and counterfactual thoughts about what might have been." Trends in Cognitive Sciences **6**(10): 426-431.

Camerer, C. and T.-H. Ho (1998). "Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity, and Time-Variation." Journal of Mathematical Psychology **42**(2-3): 305-326.

Camerer, C. F. (2003). "Behavioural studies of strategic thinking in games." Behavioural studies of strategic thinking in games.

Camerer, C. F., H. T. Ho and K. J. Chong (2004). "A Cognitive Hierarchy Model of Games." The Quarterly Journal of Economics **119**(3): 861-898.

Camille, N., G. Coricelli, J. Sallet, P. Pradat-Diehl, J.-R. Duhamel and A. Sirigu (2004). "The involvement of the orbitofrontal cortex in the experience of regret." Science (New York, N.Y.) **304**(ii): 1167-1170.

Camille, N., C. A. Griffiths, K. Vo, L. K. Fellows and J. W. Kable (2011). "Ventromedial frontal lobe damage disrupts value maximization in humans." The Journal of neuroscience : the official journal of the Society for Neuroscience **31**(20): 7527-7532.

Carstensen, L. L. and M. DeLiema (2018). "The positivity effect: a negativity bias in youth fades with age." Current Opinion in Behavioral Sciences **19**: 7-12.

Carstensen, L. L., B. Turan, S. Scheibe, N. Ram, H. Ersner-Hershfield, G. R. Samanez-Larkin, K. P. Brooks and J. R. Nesselroade (2011). "Emotional experience improves with age: evidence based on over 10 years of experience sampling." Psychology and aging **26**(1): 21-33.

Charles, S. T. and L. L. Carstensen (2010). "Social and Emotional Aging." Annual Review of Psychology **61**(1): 383-409.

Chen, W. and I. Krajbich (2017). "Computational modeling of epiphany learning." Proceedings of the National Academy of Sciences **114**(18): 4637-4642.

Chen, Y. and Y. Sun (2003). "AGE DIFFERENCES IN FINANCIAL DECISION-MAKING: USING SIMPLE HEURISTICS." Educational Gerontology **29**(7): 627-635.

Ciaramelli, E. and G. d. Pellegrino (2011). "Ventromedial Prefrontal Cortex and the Future of Morality TL - 3." Emotion Review(3): 308-309.

Cima, M., F. Tonnaer and M. D. Hauser (2010). "Psychopaths know right from wrong but don't care." Social Cognitive and Affective Neuroscience **5**(1): 59-67.

Civai, C., C. Miniussi and R. I. Rumiati (2015). "Medial prefrontal cortex reacts to unfairness if this damages the self: a tDCS study." Social Cognitive and Affective Neuroscience **10**(8): 1054-1060.

Clément, J. P., R. F. Nassif, J. M. Léger and F. Marchan (1997). "[Development and contribution to the validation of a brief French version of the Yesavage Geriatric Depression Scale]." L'Encéphale: 91-99.

Connolly, T., L. D. Ordóñez and R. Coughlan (1997). "Regret and Responsibility in the Evaluation of Decision Outcomes." Organizational Behavior and Human Decision Processes **70**(1): 73-85.

Coricelli, G., H. D. Critchley, M. Joffily, J. P. O'Doherty, A. Sirigu and R. J. Dolan (2005). "Regret and its avoidance: a neuroimaging study of choice behavior." Nature Neuroscience **8**(9): 1255-1262.

Coricelli, G., R. J. Dolan and A. Sirigu (2007). "Brain, emotion and decision making: the paradigmatic example of regret." Trends in Cognitive Sciences **11**(6): 258-265.

Coricelli, G. and R. Nagel (2009). "Neural correlates of depth of strategic reasoning in medial prefrontal cortex." Proceedings of the National Academy of Sciences **106**(23): 9163-9168.

Coricelli, G. and E. Rusconi (2010). Probing the decisional brain with rTMS and tDCS. A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide. M. Schulte-Mecklenbeck, A. Kuehberger and R. Ranyard. Hove, UK, Psychology Press: 205-222.

Coricelli, G. and A. Rustichini (2010). "Counterfactual Thinking and Emotions: Regret and Envy Learning." Philosophical Transactions of the Royal Society B: Biological Sciences **365**(1538).

Damasio, A. R., D. Tranel and H. Damasio (1990). "Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli." Behavioural Brain Research **41**(2): 81-94.

Dayan, P. (1994). "Computational modelling." Current Opinion in Neurobiology **4**(2): 212-217.

Denburg, N. L., C. A. Cole, M. Heandez, T. H. Yamada, D. Tranel, A. Bechara and R. B. Wallace (2007). "The Orbitofrontal Cortex, Real-World Decision Making, and Normal Aging." Annals of the New York Academy of Sciences **1121**(1): 480-498.

Derouesne, C., J. Poitreneau, L. Hugonot, M. Kalafat, B. Dubois and B. Laurent (1999). "[Mini-Mental State Examination:a useful method for the evaluation of the cognitive status of patients by the clinician. Consensual French version]." Presse médicale (Paris, France : 1983): 1141-1148.

Devetag, G. and M. Warglien (2003). "Games and phone numbers: Do short-term memory bounds affect strategic behavior?" Journal of Economic Psychology **24**(2): 189-202.

Eldar, E. and Y. Niv (2015). "Interaction between emotional state and learning underlies mood instability." Nature Communications **6**: 6149.

Erixon-Lindroth, N., L. Farde, T.-B. Wahlin, J. Sovago, C. Halldin and L. Bäckman (2005). "The role of the striatal dopamine transporter in cognitive aging." Psychiatry Research: Neuroimaging **138**(1).

Fecteau, S., D. Knoch, F. Fregni, N. Sultani, P. Boggio and A. Pascual-Leone (2007). "Diminishing Risk-Taking Behavior by Modulating Activity in the Prefrontal Cortex: A Direct Current Stimulation Study." The Journal of Neuroscience **27**(46): 12500-12505.

Fecteau, S., A. Pascual-Leone, D. H. Zald, P. Liguori, H. Théoret, P. S. Boggio and F. Fregni (2007). "Activation of Prefrontal Cortex by Transcranial Direct Current Stimulation Reduces Appetite for Risk during Ambiguous Decision Making." The Journal of Neuroscience **27**(23): 6212-6218.

Foster, D. P. and R. Vohra (1999). "Regret in the On-Line Decision Problem." Games and Economic Behavior **29**(1-2): 7-35.

Foster, D. P. and H. P. Young (2003). "Learning, hypothesis testing, and Nash equilibrium." Games and Economic Behavior **45**(1): 73-96.

Frijda, N. H., P. Kuipers and E. t. Schure (1989). "Relations among emotion, appraisal, and emotional action readiness. TL - 57." Journal of Personality and Social Psychology **57**(2): 212-228.

Frith, C. D. and U. Frith (1999). "Interacting minds – a biological basis." Science **286**(5445).

Fudenberg, D. and D. Levine (1999). "Learning in games." European Economic Review **42**(3-5): 631-639.

Fudenberg, D. and D. K. Levine (1995). "Consistency and cautious fictitious play." Journal of Economic Dynamics and Control **19**(5-7): 1065-1089.

Fudenberg, D. and D. K. Levine (1998). "The theory of learning in games." MIT press **2**.

Gallagher, H. L. and C. D. Frith (2003). "Functional imaging of theory of mind." Trends in cognitive sciences **7**(2): 77-83.

Gandiga, P. C., F. C. Hummel and L. G. Cohen (2006). "Transcranial DC stimulation (tDCS): A tool for double-blind sham-controlled clinical studies in brain stimulation." Clinical Neurophysiology **117**(4): 845-850.

Geuzaine, C. and J. F. Remacle (2009). "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities." International Journal for Numerical Methods in Engineering **79**(11): 1309-1331.

Gillan, C. M., S. Morein-Zamir, M. Kaser, N. A. Fineberg, A. Sule, B. J. Sahakian, R. N. Cardinal and T. W. Robbins (2014). "Counterfactual Processing of Economic Action-Outcome Alternatives in Obsessive-Compulsive Disorder: Further Evidence of Impaired Goal-Directed Behavior." Biological Psychiatry **75**(8): 639-646.

Greene, J. D. (2007). "Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains." Trends in Cognitive Sciences **11**(8): 322-323.

Greene, J. D., L. E. Nystrom, A. D. Engell, J. M. Darley and J. D. Cohen (2004). "The Neural Bases of Cognitive Conflict and Control in Moral Judgment." Neuron **44**(2): 389-400.

Greene, J. D., R. B. Sommerville, L. E. Nystrom, J. M. Darley and J. D. Cohen (2001). "An fMRI investigation of emotional engagement in moral judgment." Science (New York, N.Y.) **293**(5537): 2105-2108.

Gu, X., X. Wang, A. Hula, S. Wang, S. Xu, T. M. Lohrenz, R. T. Knight, Z. Gao, P. Dayan and P. R. Montague (2015). "Necessary, Yet Dissociable Contributions of the Insular and Ventromedial Prefrontal Cortices to Norm Adaptation: Computational and Lesion Evidence in Humans TL - 35." The Journal of Neuroscience **35**(2): 467-473.

Guerini, R., L. FitzGibbon and G. Coricelli (in press). "The role of agency in regret and relief in 3- to 10-year-old children."

Hämmerer, D., J. Bonaiuto, M. Klein-Flügge, M. Bikson and S. Bestmann (2016). "Selective alteration of human value decisions with medial frontal tDCS is predicted by changes in attractor dynamics." Scientific Reports **6**(1).

Hampton, A. N., P. Bossaerts and J. P. O'Doherty (2008). "Neural correlates of mentalizing-related computations during strategic interactions in humans." Proceedings of the National Academy of Sciences of the United States of America **105**(18).

Hampton, A. N., P. Bossaerts and J. P. O'Doherty (2006). "The Role of the Ventromedial Prefrontal Cortex in Abstract State-Based Inference during Decision Making in Humans." The Journal of Neuroscience **26**(32): 8360-8367.

Hart, S. (2005). "Adaptive Heuristics." Econometrica **73**(5): 1401-1430.

Hart, S. and A. Mas-Colell (2000). "A Simple Adaptive Procedure Leading to Correlated Equilibrium." Econometrica: 1127-1150.

Hayden, B. Y., J. M. Pearson and M. L. Platt (2009). "Fictive Reward Signals in the Anterior Cingulate Cortex." Science **324**(5929): 948-950.

Hecht, D., V. Walsh and M. Lavidor (2010). "Transcranial Direct Current Stimulation Facilitates Decision Making in a Probabilistic Guessing Task." The Journal of Neuroscience **30**(12): 4241-4245.

Hershey, D. A. and J. A. Wilson (1997). "Age Differences in Performance Awareness on a Complex Financial Decision-Making Task." Experimental Aging Research **23**(3): 257-273.

Hosseini, S. M., M. Rostami, Y. Yomogida, M. Takahashi, T. Tsukiura and R. Kawashima (2010). "Aging and decision making under uncertainty: Behavioral and neural evidence for the preservation of decision making in the absence of learning in old age." NeuroImage: 1514-1520.

Hosseini, S. M., M. Rostami, Y. Yomogida, M. Takahashi, T. Tsukiura and R. Kawashima (2010). "Aging and decision making under uncertainty: behavioral and neural evidence for the preservation of decision making in the absence of learning in old age." NeuroImage **52**(4): 1514-1520.

Hsu, M. and L. Zhu (2012). "Learning in games: neural computations underlying strategic learning." Recherches économiques de Louvain **78**(3): 47-72.

Hughes, M. A., M. C. Dolan and J. C. Stout (2013). "Regret in the context of unobtained rewards in criminal offenders." Cognition and Emotion **28**(5): 913-925.

Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica **47**(2).

Kim, H., S. Shimojo and J. P. O'Doherty (2006). "Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain." PLoS biology **4**(8).

Kishida, K. T., I. Saez, T. Lohrenz, M. R. Witcher, A. W. Laxton, S. B. Tatter, J. P. White, T. L. Ellis, P. E. M. Phillips and R. P. Montague (2016). "Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward." Proceedings of the National Academy of Sciences **113**(1).

Koenigs, M., M. Kruepke and J. P. Newman (2010). "Economic decision-making in psychopathy: A comparison with ventromedial prefrontal lesion patients." Neuropsychologia **48**(7): 2198-2204.

Koenigs, M., M. Kruepke, J. Zeier and J. P. Newman (2012). "Utilitarian moral judgment in psychopathy." Social cognitive and affective neuroscience **7**(6): 708-714.

Koenigs, M., L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser and A. Damasio (2007). "Damage to the prefrontal cortex increases utilitarian moral judgements." Nature **446**(7138): 908-911.

Kohn, N., I. Falkenberg, T. Kellermann, S. B. Eickhoff, R. C. Gur and U. Habel (2014). "Neural correlates of effective and ineffective mood induction." Social Cognitive and Affective Neuroscience **9**(6): 864-872.

Krause, B. and R. Cohen Kadosh (2014). "Not all brains are created equal: the relevance of individual differences in responsiveness to transcranial electrical stimulation." Frontiers in systems neuroscience **8**: 25.

Levens, S. M., J. T. Larsen, J. Bruss, D. Tranel, A. Bechara and B. A. Mellers (2015). "What might have been? The role of the ventromedial prefrontal cortex and lateral orbitofrontal cortex in counterfactual emotions and choice TL - 54." Neuropsychologia **54 VN - readcube.com**.

Levens, S. M., J. T. Larsen, J. Bruss, D. Tranel, A. Bechara and B. A. Mellers (2015). "What might have been? The role of the ventromedial prefrontal cortex and lateral orbitofrontal cortex in counterfactual emotions and choice TL - 54." Neuropsychologia **54**.

Löckenhoff, C. E. and L. L. Carstensen (2007). "Aging, emotion, and health-related decision strategies: Motivational manipulations can reduce age differences." Psychology and Aging **22**(1): 134.

Lohrenz, T., K. McCabe, C. F. Camerer and P. R. Montague (2007). "Neural signature of fictive learning signals in a sequential investment task." Proceedings of the National Academy of Sciences of the United States of America **104**(22): 9493-9498.

Loomes, G. and R. Sugden (1982). "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." The Economic Journal **92**(368): 805.

Manstead, A. S. R. (2000). The role of moral norm in the attitude–behavior relation. Applied social research. Attitudes, behavior, and social context: The role of norms and group membership. D. J. Terry and M. A. Hogg. Mahwah, NJ, Lawrence Erlbaum Associates: 11-30.

Marchiori, D. and M. Warglien (2008). "Predicting Human Interactive Learning by Regret-Driven Neural Networks." Science **319**(5866): 1111-1113.

Marchiori, D. and M. Warglien (2008). "Predicting human interactive learning by regret-driven neural networks." Science **319**(2001): 1111-1113.

Markman, K. D., I. Gavanski, S. J. Sherman and M. N. McMullen (1993). "The Mental Simulation of Better and Worse Possible Worlds." Journal of Experimental Social Psychology **29**(1): 87-109.

Mata, R., A. K. Josef, G. R. Samanez-Larkin and R. Hertwig (2011). "Age differences in risky choice: a meta-analysis." Annals of the New York Academy of Sciences **1235**(1): 18-29.

Mather, M. and L. L. Carstensen (2002). "Aging and Attentional Biases for Emotional Faces." Psychological Science **14**(5): 409-415.

Megiddo, N. (1980). "On repeated games with incomplete information played by non-Bayesian players." International Journal of Game Theory **9**(3): 157-167.

Mellers, B., A. Schwartz and I. Ritov (1999). "Emotion-based choice." Journal of Experimental Psychology: General **128**(3).

Minati, L., C. Campanhã, H. D. Critchley and P. Boggio (2012). "Effects of transcranial direct-current stimulation (tDCS) of the dorsolateral prefrontal cortex (DLPFC) during a mixed-gambling risky decision-making task." Cognitive Neuroscience **3**(2): 80-88.

Mitchell, J. P., C. N. Macrae and M. R. Banaji (2006). "Dissociable medial prefrontal contributions to judgments of similar and dissimilar others." Neuron **50**(4): 655-663.

Moll, J., R. d. Oliveira-Souza, I. E. Bramati and J. Grafman (2002). "Functional Networks in Emotional Moral and Nonmoral Social Judgments TL - 16." NeuroImage **16 VN - readcube.com**(3): 696-703.

Moll, J., R. d. Oliveira-Souza, I. E. Bramati and J. Grafman (2002). "Functional Networks in Emotional Moral and Nonmoral Social Judgments TL - 16." NeuroImage **16**(3): 696-703.

Montague, P. R., B. King-Casas and J. D. Cohen (2006). "Imaging Valuation Models in Human Choice." Annual Review of Neuroscience **29**: 417-448.

Newton, J. D., F. J. Newton, M. T. Ewing, S. Burney and M. Hay (2013). "Conceptual overlap between moral norms and anticipated regret in the prediction of intention: Implications for theory of planned behaviour research." Psychology & Health **28**(5): 495-513.

Nicolle, A., D. R. Bach, C. Frith and R. J. Dolan (2011). "Amygdala involvement in self-blame regret." Social Neuroscience **6**(2): 178-189.

Nitsche, M. A., S. Doemkes, T. Karaköse, A. Antal, D. Liebetanz, N. Lang, F. Tergau and W. Paulus (2007). "Shaping the Effects of Transcranial Direct Current Stimulation of the Human Motor Cortex." Journal of Neurophysiology **97**(4): 3109-3117.

Nitsche, M. A., K. Fricke, U. Henschke, A. Schlitterlau, D. Liebetanz, N. Lang, S. Henning, F. Tergau and W. Paulus (2003). "Pharmacological Modulation of Cortical Excitability Shifts Induced by Transcranial Direct Current Stimulation in Humans." The Journal of Physiology **553**(1): 293-301.

O'Doherty, J. P. (2004). "Reward representations and reward-related learning in the human brain: insights from neuroimaging." Current opinion in neurobiology **14**(6): 769-776.

Palminteri, S., M. Khamassi, M. Joffily and G. Coricelli (2015). "Contextual modulation of value signals in reward and punishment learning." Nature Communications **6**: 8096.

Park, D. C., G. Lautenschlager, T. Hedden, N. S. Davidson, A. D. Smith and P. K. Smith (2002). "Models of visuospatial and verbal memory across the adult life span." Psychology and Aging **17**(2): 299.

Parker, D., A. S. R. Manstead and S. G. Stradling (1995). "Extending the theory of planned behaviour: The role of personal norm." British Journal of Social Psychology **34**(2): 127-138.

Platt, M. L. and J. M. Pearson (2016). "Dopamine: Context and counterfactuals." Proceedings of the National Academy of Sciences **113**(1): 22-23.

Pripfl, J., R. Neumann, U. Köhler and C. Lamm (2013). "Effects of transcranial direct current stimulation on risky decision making are mediated by 'hot' and 'cold' decisions, personality, and hemisphere." European Journal of Neuroscience **38**(12): 3778-3785.

Rapoport, A. and W. Amaldoss (2000). "Mixed strategies and iterative elimination of strongly dominated strategies: An experimental investigation of states of knowledge." Journal of Economic Behavior & Organization **42**(4): 483-521.

Reed, A. E., L. Chan and J. A. Mikels (2014). "Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information." Psychology and Aging **29**(1): 1.

Riggle, E. D. B. and M. M. S. Johnson (1996). "Age difference in political decision making: Strategies for evaluating political candidates." Political Behavior **18**(1): 99-118.

Ritov, I. (1996). "Probability of Regret: Anticipation of Uncertainty Resolution in Choice." Organizational Behavior and Human Decision Processes **66**(2): 228-236.

Rivis, A., P. Sheeran and C. J. Armitage (2009). "Expanding the Affective and Normative Components of the Theory of Planned Behavior: A Meta-Analysis of Anticipated Affect and Moral Norms." Journal of Applied Social Psychology **39**(12): 2985-3019.

Roese, N. J. (1994). "The functional basis of counterfactual thinking." Journal of Personality and Social Psychology **66**(5): 805.

Roth, A. E. and I. Erev (1995). "Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term." Games and economic behavior.

Rottenberg, J. and J. J. Gross (2007). "Emotion and Emotion Regulation: A Map for Psychotherapy Researchers." Clinical Psychology: Science and Practice **14**(4): 323-328.

Samanez-Larkin, G. R. and B. Knutson (2015). "Decision making in the ageing brain: changes in affective and motivational circuits." Nature Reviews Neuroscience **16**(5).

Samanez-Larkin, G. R., S. M. Levens, L. M. Perry, R. F. Dougherty and B. Knutson (2012). "Frontostriatal White Matter Integrity Mediates Adult Age Differences in Probabilistic Reward Learning." The Journal of Neuroscience **32**(15): 5333-5337.

Sanna, L. J. (1997). "Self-Efficacy and Counterfactual Thinking: Up a Creek with and without a Paddle." Personality and Social Psychology Bulletin **23**(6): 654-666.

Sanna, L. J. (1998). "Defensive Pessimism and Optimism: The Bitter-Sweet Influence of Mood on Performance and Prefactual and Counterfactual Thinking." Cognition & Emotion **12**(5): 635-665.

Sanna, L. J. and K. Turley (1996). "Antecedents to Spontaneous Counterfactual Thinking: Effects of Expectancy Violation and Outcome Valence." Personality and Social Psychology Bulletin **22**(9): 906-919.

Schultz, W. (2002). "Getting Formal with Dopamine and Reward." Neuron **36**(2): 241-263.

Schultz, W., P. Dayan and R. P. Montague (1997). "A neural substrate of prediction and reward." Science **275**(5306): 1593-1599.

Schultz, W., L. Tremblay and J. R. Hollerman (1998). "Reward prediction in primate basal ganglia and frontal cortex." Neuropharmacology **37**(4-5): 421-429.

Shivapour, S. K., C. M. Nguyen, C. A. Cole and N. L. Denburg (2012). "Effects of Age, Sex, and Neuropsychological Performance on Financial Decision-Making." Frontiers in Neuroscience **6**: 82.

Simioni, S., M. Schluep, N. Bault, G. Coricelli, J. Kleeberg, R. A. Du Pasquier, M. Gschwind, P. Vuilleumier and J.-M. Annoni (2012). "Multiple Sclerosis Decreases Explicit Counterfactual Processing and Risk Taking in Decision Making." PLOS ONE **7**(12): e50718.

Suddendorf, T. and M. C. Corballis (2007). "The evolution of foresight: What is mental time travel, and is it unique to humans?" The Behavioral and brain sciences **30**(3): 299.

Sutton, R. S. and A. G. Barto (1998). Reinforcement learning: An introduction, MIT press Cambridge.

Suzuki, S., E. L. S. Jensen, P. Bossaerts and J. P. O. Doherty (2016). "Behavioral contagion during learning about another agent ' s risk-preferences acts on the neural representation of decision-risk." **113**(14).

Tamir, M. and M. D. Robinson (2007). "The happy spotlight: positive mood and selective attention to rewarding information." Personality & social psychology bulletin **33**(8): 1124-1136.

Thair, H., A. L. Holloway, R. Newport and A. D. Smith (2017). "Transcranial Direct Current Stimulation (tDCS): A Beginner's Guide for Design and Implementation." Frontiers in Neuroscience **11**: 641.

Thielscher, A., A. Antunes and G. B. Saturnino (2015). "Field Modeling for Transcranial Magnetic Stimulation: A Useful Tool to Understand the Physiological Effects of TMS?**All authors contributed equally to the work." 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) **2015**: 222-225.

Thomas, B. C., K. E. Croft and D. Tranel (2011). "Harming Kin to Save Strangers: Further Evidence for Abnormally Utilitarian Moral Judgments after Ventromedial Prefrontal Damage." Journal of Cognitive Neuroscience **23**(9): 2186-2196.

Tobia, M. J., R. Guo, J. Gläscher, U. Schwarze, S. Brassen, C. Büchel, K. Obermayer and T. Sommer (2016). "Altered behavioral and neural responsiveness to counterfactual gains in the elderly." Cognitive, Affective, & Behavioral Neuroscience **16**(3): 457-472.

Troiano, A. R., M. Schulzer and F. F.-R. L. Synapse (2010). "Dopamine transporter PET in normal aging: dopamine transporter decline and its possible role in preservation of motor function." Synapse.

Tymula, A., L. A. Belmaker, L. Ruderman, P. W. Glimcher and I. Levy (2013). "Like cognitive function, decision making across the life span shows profound age-related changes." Proceedings of the National Academy of Sciences **110**(42): 17143-17148.

Van Hoeck, N., P. D. Watson and A. K. Barbey (2015). "Cognitive neuroscience of human counterfactual reasoning." Frontiers in Human Neuroscience **9**.

Weisberg, D. P. and S. R. Beck (2012). "The development of children's regret and relief." Cognition & Emotion **26**(5): 820-835.

Wood, S., J. Busemeyer, A. Koling, C. R. Cox and H. Davis (2005). "Older Adults as Adaptive Decision Makers: Evidence From the Iowa Gambling Task." Psychology and Aging **20**(2): 220.

Yang, Y., A. Raine, T. Lencz, S. Bihrle, L. LaCasse and P. Colletti (2005). "Volume Reduction in Prefrontal Gray Matter in Unsuccessful Criminal Psychopaths." Biological Psychiatry **57**(10): 1103-1108.

Yoon, C., C. A. Cole and M. P. Lee (2009). "Consumer decision making and aging: Current knowledge and future directions." Journal of Consumer Psychology **19**(1): 2-16.

Young, C. B. and R. Nusslock (2016). "Positive mood enhances reward-related neural activity." Social Cognitive and Affective Neuroscience **11**(6): 934-944.

Zeelenberg, M., J. Beattie, J. v. d. Pligt and N. K. d. Vries (1996). "Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making TL - 65." Organizational Behavior and Human Decision Processes **65 VN - readcube.com**(2): 148-158.

Zeelenberg, M., J. Beattie, J. v. d. Pligt and N. K. d. Vries (1996). "Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making TL - 65." Organizational Behavior and Human Decision Processes **65**(2): 148-158.

Zeelenberg, M. and R. Pieters (2007). "A Theory of Regret Regulation 1.0." Journal of Consumer Psychology **17**(1).

Zeelenberg, M., W. W. van Dijk and A. S. R. Manstead (1998). "Reconsidering the Relation between Regret and Responsibility." Organizational Behavior & Human Decision Processes **74**(3).

Zeelenberg, M., W. W. van Dijk, A. S. R. Manstead and J. van der Pligt (1998). "The Experience of Regret and Disappointment." Cognition and Emotion **12**(2).

Zeelenberg, M., W. W. van Dijk, J. van der Pligt, A. S. R. Manstead, P. van Empelen and D. Reinderman (1998). "Emotional Reactions to the Outcomes of Decisions: The Role of Counterfactual Thought in the Experience of Regret and Disappointment." Organizational Behavior & Human Decision Processes **75**(2).

Zheng, H., D. Huang, S. Chen, S. Wang, W. Guo, J. Luo, H. Ye and Y. Chen (2016). "Modulating the Activity of Ventromedial Prefrontal Cortex by Anodal tDCS Enhances the Trustee's Repayment through Altruism." Frontiers in Psychology **7**: 1437.

Zhu, L., K. E. Mathewson and M. Hsu (2012). "Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning." Proceedings of the National Academy of Sciences **109**(5): 1419-1424.