

The Prefrontal Cortex and Goal-Directed Social Behavior

Aron K. Barbey and Jordan Grafman

Abstract

This chapter develops an integrative cognitive neuroscience framework for understanding the social functions of the lateral prefrontal cortex (PFC), reviewing recent theoretical insights from evolutionary psychology and emerging neuroscience evidence to support the importance of this region for orchestrating social behavior on the basis of evolutionarily adaptive social norms. The chapter begins by reviewing the evolutionary foundations of normative social behavior, surveying contemporary research and theory from evolutionary psychology to suggest that widely shared norms of social exchange are the product of evolutionarily adaptive cognitive mechanisms. It then reviews the biology, evolution, and ontogeny of the human PFC, and introduces a cognitive neuroscience framework for goal-directed social behavior on the basis of evolutionarily adaptive social norms represented by the lateral PFC. It examines a broad range of evidence from the social and decision neuroscience literatures demonstrating that evolutionarily adaptive social norms of obligatory, prohibited, and permissible behavior are mediated by the lateral PFC.

Keywords: prefrontal cortex, social functions, cognitive neuroscience framework, goal-directed social behavior, evolution

Introduction

A primary assumption in cognitive neuroscience is that the brain has evolved to solve adaptive problems encountered by our human ancestors. Throughout evolutionary history, a foremost adaptive challenge for our species was living and interacting with other people. To survive and reproduce, our human ancestors had to select mates, form alliances, and compete for limited resources. They also needed to learn social norms and standards of conduct, as violations of these rules might have been severely punished, resulting in banishment from society. Accordingly, just as the brain has evolved mechanisms for perception, memory, language, and thought, it is likely that there are also evolutionarily

adaptive mechanisms that enable humans to coexist with others. Indeed, it is through cooperation that evolution constructs new levels of organization, from genomes and cells to the formation of multicellular organisms, social insects, and complex human societies (Nowak, 2006).

The neuroscientific study of social cognition reflects the interdisciplinary nature of modern science, with investigators from diverse academic disciplines (including anthropology, evolutionary psychology, social psychology, political science, behavioral economics, and decision neuroscience) exploring the unique social nature of human experience through a multifaceted lens (for a social neuroscience review, see Cacioppo et al., 2002). This interdisciplinary

enterprise has made considerable progress in understanding the involvement of the prefrontal cortex (PFC) in social cognition (Amodio & Frith, 2006; Barbey, Krueger, & Grafman, 2009; Krueger, Barbey, & Grafman, 2009; Wood & Grafman, 2003). Accumulating evidence suggests that representations within the lateral PFC enable people to orchestrate their thoughts and actions in concert with their intentions to support goal-directed social behavior (Fiddick, Spampinato, & Grafman, 2005; Berthoz, Armony, Blair, & Dolan, 2002; Rilling et al., 2008; Buckholz et al., 2008; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Weissman, Perkins, & Woldorff, 2008; Damasio, Tranel, & Damasio, 1990; Bechara, Damasio, Damasio, & Anderson, 1994; Rolls, Hornak, Wade, & McGrath, 1994; Bechara, Damasio, & Damasio, 2000; LoPresti et al., 2008; Ruby & Decety, 2004). Despite the pivotal role of this region in guiding social interactions, fundamental questions remain concerning the functional organization and forms of social knowledge represented within the lateral PFC. We develop an integrative cognitive neuroscience framework for understanding the social functions of the lateral PFC, reviewing recent theoretical insights from evolutionary psychology and emerging neuroscience evidence to support the importance of this region for orchestrating social behavior on the basis of evolutionarily adaptive social norms.

We begin by reviewing the evolutionary foundations of normative social behavior, surveying contemporary research and theory from evolutionary psychology to suggest that widely shared norms of social exchange are the product of evolutionarily adaptive cognitive mechanisms. We then review the biology, evolution, and ontogeny of the human PFC, and develop a cognitive neuroscience framework for goal-directed social behavior on the basis of evolutionarily adaptive social norms represented by the lateral PFC. We examine a broad range of evidence from the social and decision neuroscience literatures demonstrating that evolutionarily adaptive social norms of obligatory, prohibited, and permissible behavior are mediated by the lateral PFC. Accumulating evidence suggests that behaviorally guiding principles for social inference are functionally organized along the dorso-ventral axis of the lateral PFC, whereby obligatory or prohibited action sequences recruit the ventrolateral PFC (vIPFC) (Fiddick et al., 2005; Berthoz et al., 2002; Rilling et al., 2008; Monti, Osherson, Martinez, & Parsons, 2007; Kroger, Nystrom, Cohen, & Johnson-Laird, 2008; Heckers, Zalesak, Weiss,

Ditman, & Titone, 2004; Goel, Buchel, Frith, & Dolan, 2000; Goel & Dolan, 2004; Noveck, Goel, & Smith, 2004) and permissible forms of behavior engage the dorsolateral PFC (dlPFC) (Buckholz et al., 2008; Greene et al., 2004; Weissman et al., 2008; Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; Thomson, 1976; Volz, Schubotz, & von Cramon, 2004; Kroger et al., 2008). Adaptive behavior guided by both categories of inference recruits the anterolateral PFC (alPFC), which represents the highest level of a rostro-caudal hierarchy characterized by multiple forms of social exchange (Rolls et al., 1994; Bechara et al., 2000; LoPresti et al., 2008; Ruby & Decety, 2004; Badre, 2008; Botvinick, 2008; Koechlin & Summerfield, 2007; Christoff & Keramatian, 2007; Christoff et al., 2001; Christoff, Ream, Geddes, & Gabrieli, 2003; Smith, Keramatian, & Christoff, 2007). We illustrate how this framework supports the integration and synthesis of a diverse body of neuroscience evidence, and draw conclusions about the role of the lateral PFC in social cognition more broadly, contributing to social knowledge networks by representing widely shared norms of social behavior and providing the foundations for moral, ethical, and political systems of value and belief.

Evolutionary Foundations of Social Exchange

Social exchange is an essential aspect of life in all human cultures, promoting the survival of individuals who cooperate for mutual benefit—one providing a benefit to the other conditional on the recipient's providing a benefit in return. From our earliest ancestors to present day, social exchange has facilitated access to sustenance, protection, and mates, and enabled people to live healthier and longer lives (Cohen, 2004; Silk, Alberts, & Altmann, 2003). Social exchange interactions are therefore an important and recurrent human activity occurring over a sufficiently long time period for natural selection to have produced specialized cognitive and neural adaptations (Isaac, 1978; Brosnan & de Waal, 2003). Evolutionary psychologists have proposed that social exchange embodies cognitive mechanisms designed to promote the survival of our species, representing normative social behavior that develops in all healthy humans and is mediated by evolutionarily adaptive neural systems (Maynard Smith, 1982; Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone, Cosmides, Tooby,

Kroll, & Knight, 2002; Sugiyama, Tooby, & Cosmides, 2002; Trivers, 1971; Axelrod & Hamilton, 1981; Platt & Griggs, 1993; Gigerenzer & Hug, 1992).

An empirical case for this proposal has been established on the basis of behavioral and neuroscience research elucidating the role of evolutionary design features in shaping cognitive and neural mechanisms for social exchange (Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002). Game-theoretic models predict that for social exchange to persist within a species, members of the species must detect cheaters (i.e., individuals who do not reciprocate) and direct future benefits to reciprocators rather than cheaters (Trivers, 1971; Axelrod & Hamilton, 1981). Accumulating evidence supports this proposal, demonstrating that the mind embodies functionally specialized cognitive mechanisms for detecting cheaters (Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002) that operate according to behavior-guiding principles in the form of a conditional rule: If *X* provides a requested benefit to *Y*, then *Y* will provide a rationed benefit to *X*. A conditional rule expressing this kind of agreement to cooperate is referred to as a *social contract* and represents a normative standard for social behavior (e.g., the normative belief that mutual cooperation is obligatory and cheating prohibited).

A primary method for investigating conditional reasoning about social contracts is Wason's four-card selection task (Wason, 1966; 1983; Wason & Johnson-Laird, 1972). In the classic version of this task, participants are shown a set of four cards, placed on a table, each of which has a number on one side and a colored patch on the other. The visible faces of the cards show a 3, 8, red, and brown. Participants are then asked which card should be turned over to test the truth of the proposition that "If a card shows an even number on one face, then its opposite face shows a primary color (i.e., red, green, or blue)." Thus, participants in the Wason selection task are asked to identify possible violations of a conditional rule of the form "If *P* then *Q*." Conditional rules describing some state of the world using abstract or descriptive content typically elicit a correct response (*P* and *not-Q*) from only 5 to 30 percent of subjects tested. This finding has been observed even when the rules tested are familiar, or

when participants are taught logic or given incentives (Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Cheng & Holyoak, 1985). In contrast, 65 to 80 percent of participants generate correct responses when the conditional rule expresses a social contract and a violation represents cheating (Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002). This pattern of performance has been widely observed in industrialized nations (Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002) and has been found even among isolated, non-literate hunter-horticulturalists (Sugiyama et al., 2002). Cognitive experiments have demonstrated that this improved level of performance is sensitively regulated by the series of variables expected if this were a system optimally designed to reason about obligatory and prohibited forms of social behavior, rather than to support a broader class of inferences (Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002; Gigerenzer & Hug, 1992; Cheng & Holyoak, 1985).

Social contracts therefore represent behavior-guiding principles for evolutionarily adaptive forms of social exchange and are critical for drawing inferences about necessary courses of action concerning socially obligatory or prohibited behavior. From an evolutionary perspective, normative standards for necessary forms of social exchange can be distinguished from a broader class of inferences concerning possible or permissible courses of action. Whereas social norms for necessary behavior are central for the organization of society, representing strictly enforced rules for cooperation, the division of labor, and the distribution of resources, social norms for permissible behavior are critical for achieving adaptive goals within society, representing non-punishable courses of action that enable individuals to explore opportunities for reward, formulate plans for achieving social goals, and gain access to available resources (Maynard Smith, 1982; Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002; Trivers, 1971; Axelrod & Hamilton, 1981; Platt & Griggs, 1993; Gigerenzer & Hug, 1992; Krueger et al., in press; Krueger, Moll, Zahn-

Heinecke, & Grattans for reasoni prohibited) versus action have their architecture of the distinct cognitive wiring about necessity. While cognitive forms of inference behavior (Maynard 1989; Tooby & Cosmides 1989; 1992; 2005 2002; Sugiyama et al. & Hamilton, 1981 & Hug, 1992), not by these systems, adaptive neural between these two

An Evolutionary Architecture for Social Behavior

An emerging body of directed social behavior in the PFC, which is particularly sensitive to past experience environment also behavior-guiding have evolved mechanisms and consequences past experiences, principles that are available in similar actions are likely to

We propose that social inference adaptive social norms respond to a dynamic strength and pattern brain region. In addition, the PFC be complex behavior guiding principles event complexes 2009).

An SEC is a structured in sequence knowledge, more norms, event features. Aspects of social but are encoded in are encoded and a

Heinecke, & Grafman, 2007). Evolutionary adaptations for reasoning about necessary (obligatory or prohibited) versus possible (permissible) courses of action have therefore fundamentally shaped the architecture of the mind, producing functionally distinct cognitive and neural mechanisms for reasoning about necessary and possible states of affairs. While cognitive and neural mechanisms for these forms of inference emerged from goal-directed social behavior (Maynard Smith, 1982; Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002; Trivers, 1971; Axelrod & Hamilton, 1981; Platt & Griggs, 1993; Gigerenzer & Hug, 1992), non-social inferences are also shaped by these systems, relying upon an evolutionarily adaptive neural architecture that distinguishes between these two fundamental classes of inference.

An Evolutionarily Adaptive Neural Architecture for Goal-directed Social Behavior

An emerging body of evidence suggests that goal-directed social behavior centrally depends on the PFC, which is particularly important for grouping specific experiences of our interactions with the environment along common themes, that is, as behavior-guiding principles. To this end, our brains have evolved mechanisms for detecting and storing complex relationships between situations, actions, and consequences. By gleaning this knowledge from past experiences, we can develop behavior-guiding principles that allow us to infer which goals are available in similar situations in the future and what actions are likely to bring us closer to them.

We propose that behavior-guiding principles for social inference take the form of evolutionarily adaptive social rules that, when activated, correspond to a dynamic brain state signified by the strength and pattern of neural activity in a local brain region. In this sense, over the course of evolution, the PFC became capable of supporting more complex behaviors. We have labeled these behavior-guiding principles for social inference *structured event complexes* (SECs, Barbey et al., in press; 2009).

An SEC is a goal-oriented set of events that is structured in sequence and represents thematic knowledge, morals, abstractions, concepts, social norms, event features, event boundaries, and grammars. Aspects of SECs are represented independently but are encoded and retrieved as an episode. SECs are encoded and activated on the basis of simulation

mechanisms (Barbey & Patterson, in press; Barbey & Barsalou, 2009; Barsalou, 2008; Barsalou, Niedenthal, Barbey, & Ruppert, 2003; Barsalou, Simmons, Barbey, & Wilson, 2003). It is widely known that experience in the physical and social world activates feature detectors in relevant features maps of the brain (for a review of feature maps in vision, see Zeki, 1993). When a pattern becomes active in a feature map during perception or action, conjunctive neurons in an association area capture the pattern for later cognitive use. Increasing evidence suggests that behavior-guiding principles for social inference are mediated by higher-order association areas localized within the lateral PFC (Figure 23.1).

Decades of neuroscience research have demonstrated that the lateral PFC is comprised of neurons that are exquisitely sensitive to behaviorally informative associations (for a review, see Miller, 2000). This work has focused on the lateral PFC because it represents a site of convergence of the information needed to synthesize multimodal information from a wide range of brain systems. The lateral PFC consists of three major subregions that each emphasize processing of particular information based on their interconnections with specific cortical regions (see Figure 23.2; for a review, see Miller, 2000). Ventrolateral areas are more heavily interconnected with cortical regions for processing information about visual form and stimulus identity (inferior temporal cortex), supporting the categorization of environmental stimuli in the service of goal-directed behavior. Dorsal portions of the lateral

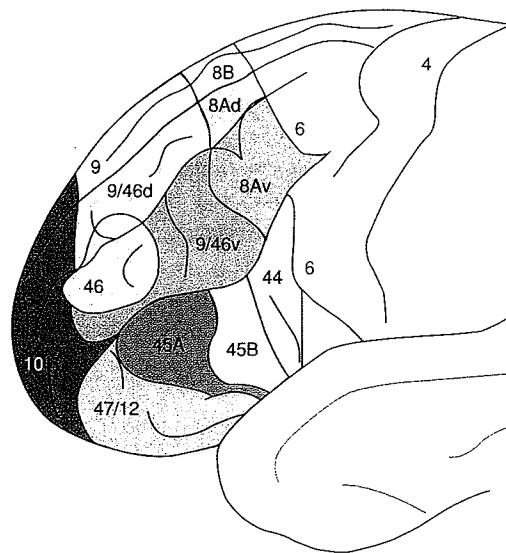
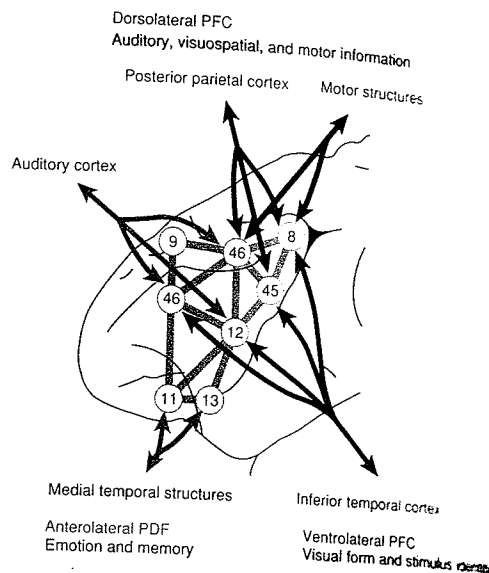


Fig. 23.1 Brodmann map of the lateral prefrontal cortex. Reproduced with permission from Ramnani & Owen (2004).

Fig. 23.2 Integrative anatomy of the macaque monkey prefrontal cortex. Numbers refer to subregions within the lateral prefrontal cortex defined by Brodmann. Adapted with permission from Miller (2000).



PFC are heavily interconnected with cortical areas for processing auditory, visuospatial, and motor information, enabling the regulation and control of responses to environmental stimuli. Finally, the anterolateral PFC is indirectly connected (via the ventromedial PFC) with limbic structures that process internal information, such as emotion, memory, and reward (Goldman-Rakic, 1987; Pandya & Barnes, 1987; Fuster, 1989; Barbas & Pandya, 1991). Together, lateral PFC subregions mediate essential elements of the external and internal environment, enabling goal-directed behavior.

Once feature maps within modality-specific regions are captured by a set of conjunctive neurons in the lateral PFC, the set can later activate the pattern in the absence of bottom-up stimulation, producing a simulation of the event sequence (Barbey & Barsalou, 2009; Barsalou, 2008; Barsalou, Niedenthal, et al., 2003; Barsalou, Simmons, et al., 2003). For example, on entering a familiar situation and recognizing it, a simulation that represents the situation becomes active. Typically, not all of the situation is perceived initially. A relevant person, setting, or event may be perceived, which then suggests that a particular situation is about to play out. The simulation can be viewed as a complex configuration of multimodal components that represent the situation (including agents, objects, actions, mental states, and background settings). Because part of this pattern matched the current situation initially, the larger pattern became active in memory. The remaining parts of the pattern—not yet observed in the situation—constitute inferences, namely predictions about what will occur next.

To the extent that the simulation is entrenched in memory, pattern completion is likely to occur automatically. As a situation is experienced repeatedly, its simulated components and the associations linking them increase in potency. Thus when one component is perceived initially, these strong associations complete the pattern automatically. Social norms of behavior represent deeply entrenched simulations, whose learned associations are the product of evolutionarily adaptive cognitive and neural mechanisms. For example, evolutionarily adaptive norms for social exchange concerning obligatory actions (i.e., reciprocal altruism) and prohibited behavior (i.e., cheating) derive from extensive experience spanning our evolutionary history and have therefore fundamentally shaped the cognitive and neural mechanisms that mediate social exchange. The observed role of simulation mechanisms for social inference in non-human primates supports this account (Gil-da-Costa et al., 2004), suggesting that modality-specific simulations represent continuity of social information processing across the species (Barsalou, 2005). According to this framework, social interactions initially match modality-specific representations in one or more simulations that have become entrenched in memory. Once one of these wins the activation process, it provides inferences via pattern completion (Anderson, 1995). Simulations representing necessary (obligatory or prohibited) courses of action motivate expectations concerning specific actions the perceiver and recipient “must” take, whereas simulations for possible (permissible) forms of behavior represent a broader range of outcomes, motivating expectations about

courses of action. The unified and possible simulation—representing the orchestrated recruitment of social inference development.

Research in many of the PFC emerged from followed by dorsal (Figure 23.3). From an evolutionary lateral PFC for the formation of the vlPFC for necessary action, providing of conduct of society. Social enabled the reasonable outcome native forms

Anterolateral PFC
The late anterolateral PFC that this ventrolateral prefrontal

Fig. 23.3 Ontogenetic order of their development. Modified with permission

courses of action the perceiver and recipient “may” take. The unfolding of inferences about necessary and possible states of affairs—realized as a simulation—represents behavior-guiding principles for the orchestration of social thought and action. The recruitment of specific lateral PFC subregions for social inference is determined by the evolution, development, hierarchical structure, and anatomical connectivity of the PFC.

Research investigating the evolution and ontogeny of the PFC suggests that the lateral PFC initially emerged from ventrolateral prefrontal regions, followed by dorsolateral, and then anterolateral cortices (Figure 23.3; Fuster, 1997; Flechsig, 1901; 1920). From an evolutionary perspective, the emergence of lateral PFC subregions reflects their relative priority for the formation of organized social groups, with the vlPFC signaling the onset of social norms for necessary (obligatory or prohibited) courses of action, providing the foundations for standards of conduct that are central for the organization of society. Social norms for permissible behavior later enabled the representation of a broader range of possible outcomes, supporting the assessment of alternative forms of goal-directed behavior within the

dIPFC. Finally, the evolution of the alPFC enabled processing of higher-order relations and reasoning about complex forms of social behavior involving necessary and possible courses of action. Consistent with its evolutionary development, the ontogeny of the lateral PFC reflects the importance of first representing social norms for necessary behavior (i.e., fundamental rules the child must obey), followed by an understanding of permissible courses of action (e.g., guided by judgments of equity and fairness), and finally higher-order inferences involving both forms of representation (Santrock, 2005).

An emerging body of evidence further demonstrates that the anterior-to-posterior axis of the lateral PFC is organized hierarchically, whereby progressively anterior subregions are associated with higher-order processing requirements for planning and the selection of action (for recent reviews, see Badre, 2008; Botvinick, 2008; Koechlin & Summerfield, 2007; Ramnani & Owen, 2004). Thus, processes within the lateral PFC respect the hierarchical organization of this region, with progressively anterior regions representing simulations that support higher-order inferences incorporating both necessary and possible states of affairs.

Dorsolateral PFC

The myelination of dorsolateral PFC subregions suggests that this area emerges after early ventrolateral PFC subregions.

Anterolateral PFC

The late myelination of the anterolateral PFC suggests that this area emerges after ventrolateral and dorsolateral prefrontal regions.

Ventrolateral PFC

The early myelination of ventrolateral PFC subregions suggests that this region is one of the first prefrontal areas to emerge during development.

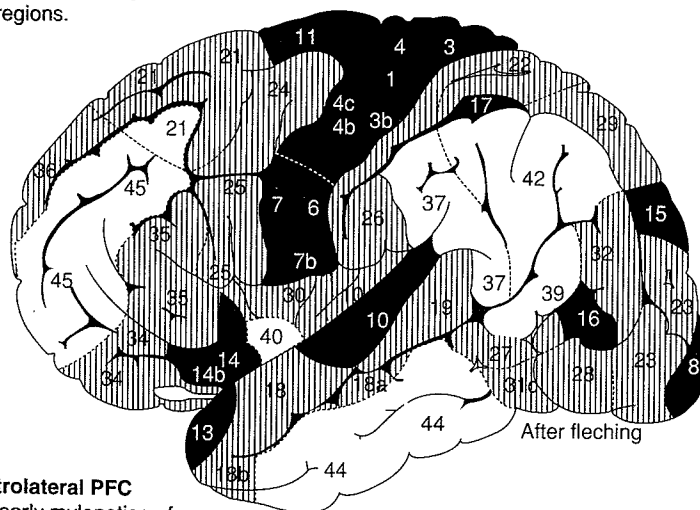


Fig. 23.3 Ontogenetic map of the prefrontal cortex according to Flechsig (1901; 1920). The numeration of the areas indicates the order of their myelination.
Modified with permission from Flechsig (1920).

The connectivity of lateral PFC subregions represents evolutionarily adaptive neural systems for goal-directed social behavior. From an evolutionary perspective, behavior requested by members of high social status represents necessary courses of action that a lower-ranking individual must follow. This provides one explanation for why neural systems for identifying the social status of individuals (based on representations of visual form and stimulus identity) are anatomically connected with ventrolateral prefrontal regions for drawing inferences about necessary courses of action. In contrast, social norms for possible (permissible) behavior are central for achieving adaptive goals within society (Maynard Smith, 1982; Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002; Trivers, 1971; Axelrod & Hamilton, 1981; Platt & Griggs, 1993; Gigerenzer & Hug, 1992), providing one explanation for why dorsolateral prefrontal regions for drawing this type of inference are anatomically connected with brain regions for the regulation and control of behavior. Finally, adaptive behavior guided by both categories of inference draws upon higher-order representations that incorporate multiple forms of social exchange and therefore recruits regions of the antelateral PFC that enable complex representations (e.g., incorporating emotion and memory).

Lateral Prefrontal Contributions to Goal-directed Social Behavior

We review emerging evidence from the social and decision neuroscience literatures demonstrating (1) the involvement of the vLPFC when reasoning about necessary (obligatory or prohibited) courses of action, (2) the recruitment of the dlPFC for drawing inferences about possible (permissible) states of affairs, and (3) activation in the aLPFC for higher-order inferences that incorporate both categories of knowledge (Figure 23.4). The simulation architecture underlying these forms of inference further predicts the recruitment of broadly distributed neural systems, incorporating medial prefrontal and posterior knowledge networks representing modality-specific components of experience.

Ventrolateral Prefrontal Cortex

An increasing number of social neuroscience studies have shown that social norms for necessary (obligatory or prohibited) courses of action are represented by the vLPFC (areas 44, 45, and 47; Figure 23.4b). Fiddick et al. (2005) observed activity

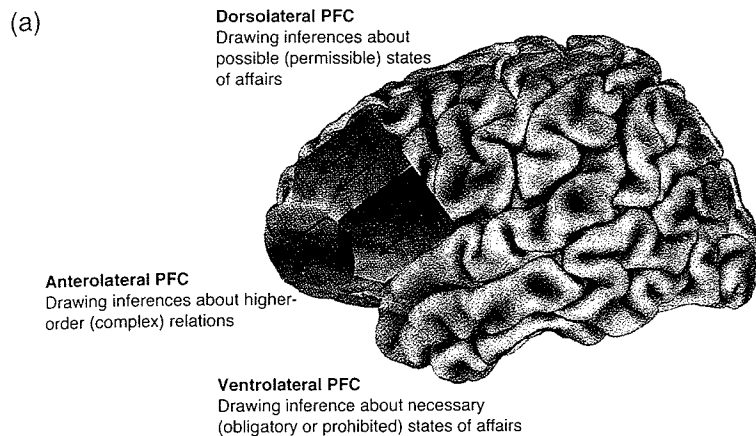
within the bilateral vLPFC (area 47) for social exchange reasoning, employing stimuli consisting primarily of social norms for obligatory and prohibited courses of action. Converging evidence is provided by Berthoz et al. (2002), who demonstrated recruitment of the left vLPFC (area 47) when participants detected violations of social norms stories representing obligatory and prohibited courses of action (e.g., the decision to "spit out food made by the host"). Similarly, Rilling et al. (2008) reported activation within the left vLPFC (area 47) when participants detected the violation of obligatory and prohibited norms of social exchange in a Prisoner's dilemma game (i.e., the failure to cooperate).

The decision neuroscience literature further supports this framework, demonstrating the involvement of the vLPFC when drawing conclusions that necessarily follow from the truth of the premises, that is, for *deductive inference*. Although wide consensus in the literature has not yet been reached, an increasing number of studies report consistent findings when common sources of variability are controlled (regarding the linguistic content, linguistic complexity, and deductive complexity of reasoning problems). A recent series of experiments by Monti et al. (2007) controlled for these sources of variability and provided evidence that the left vLPFC (area 47) mediates representations of the logical structure of a deductive argument (e.g., If *P* or *Q*, then *Not-R/P* Therefore, *Not-R*), supporting the representation of behavior-guiding principles for necessary forms of behavior within this region. Furthermore, a recent study by Kroger et al. (2008) controlled for the complexity and type of calculations that were performed and also observed activation within the left vLPFC (areas 44 and 45) for deductive reasoning (see also Heckers et al., 2004). Converging evidence is provided by Goel and colleagues (Goel et al., 2000; Goel & Dolan, 2004), who have consistently observed activation within the left vLPFC (areas 44 and 45) for deductive conclusions drawn from categorical syllogisms (e.g., All humans are mortal/Some animals are human/Therefore, some animals are mortal). Finally, Noveck et al. (2004) demonstrated recruitment of the left vLPFC (area 47) for drawing deductive conclusions from conditional statements (e.g., If *P* then *Q/P* Therefore, *Q*), consistent with the role of this region for representing behavior-guiding principles in the form of a conditional.

Dorsolateral Prefrontal Cortex

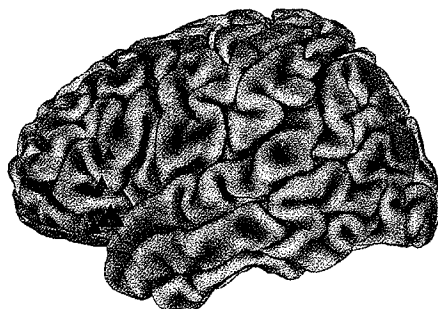
Accumulating evidence demonstrates that the dlPFC (areas 46 and 9) represents behavior-guiding

Fig. 23.4 An evolutionary organization of the lateral prefrontal cortex.



(b) Drawing inferences about necessary states of affairs

- Legend
- ▲ Fiddick et al. (2005)
 - ▲ Berthoz et al. (2002)
 - ▲ Rilling et al. (2008)
 - ▲ Monti et al. (2007)
 - ▲ Kroger et al. (2008)
 - ▲ Heckers et al. (2004)
 - ▲ Goel et al. (2000)
 - ▲ Goel & Dolan (2004)
 - ▲ Noveck et al. (2004)



(c) Drawing inferences about possible states of affairs



- Legend
- Buckholz et al. (2008)
 - Greene et al. (2004)
 - Weissman et al. (2008)
 - Sanfey et al. (2003)
 - Knoch et al. (2006)
 - Prehn et al. (2008)
 - Huettel et al. (2005)
 - Osherson et al. (1998)
 - Kroger et al. (2008)

(d) Drawing inferences about higher-order relations

- Legend
- Ruby & Decety (2004)
 - Kroger et al. (2008)
 - Christoff & Kermatian (2007)
 - Christoff et al. (2001)
 - Christoff et al. (2003)
 - Smith et al. (2007)

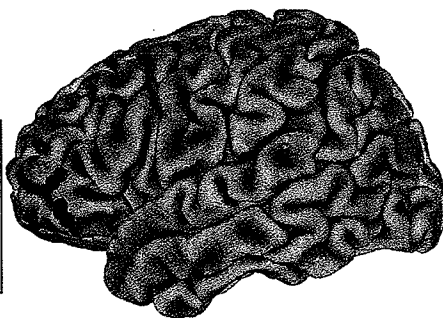


Fig. 23.4 An evolutionarily adaptive neural architecture for goal-directed social behavior. (a) summarizes the functional organization of the lateral PFC, and (b), (c) and (d) illustrate supportive evidence.

principles for evaluating the permissibility or fairness of observed behavior (Figure 23.4c). An early study by Sanfey et al. (2003) reported activity within the right dlPFC (area 46) when participants evaluated the fairness of an offer in an ultimatum game. Knoch et al. (2006) further demonstrated that deactivating this region with repetitive transcranial magnetic stimulation reduced participants' ability to reject unfair offers in the ultimatum game, suggesting that the dlPFC is central for guiding behavior based on evaluations of fairness and permissibility. Converging evidence is provided by Buckholz et al. (2008), who observed activity within the right dlPFC (area 46) when participants assigned responsibility for crimes and made judgments about appropriate (e.g., equitable or fair) forms of punishment in a legal decision-making task. The work of Greene et al. (2004) further suggests that this region is involved in normative evaluations involving conflicting moral goals. These authors employed moral scenarios similar to the famous trolley problem (Thomson, 1976) and assessed trials in which participants acted in the interest of greater aggregate welfare at the expense of personal moral standards. This contrast revealed reliable activation within the right dlPFC (area 46), suggesting that this region is critical for evaluating the permissibility or fairness of behaviors that conflict with personal moral standards (for additional evidence, see Weissman et al., 2008; Prehn et al., 2008).

Further evidence to support this framework derives from the decision neuroscience literature, which demonstrates the involvement of the dlPFC when drawing conclusions about possible or permissible states of affairs. In contrast to deductive inference, conclusions about possible courses of action reflect uncertainty concerning the actions that "should" be taken and/or the consequences that "might" follow, and are referred to as *inductive inferences*. Volz et al. (2004) found that activation within the right dlPFC (area 9) increased parametrically with the degree of uncertainty held by the participant (see also, Huettel, Song, & McCarthy, 2005). Furthermore, Osherson et al. (1998) observed preferential recruitment of the right dlPFC (area 46) when performance on an inductive reasoning task was directly compared to a matched deductive inference task, supporting the role of this region for reasoning about possible (rather than necessary) states of affairs.

Anterolateral Prefrontal Cortex

A large body of social neuroscience evidence demonstrates that the alPFC (areas 10 and 11)—and the orbitofrontal cortex (OFC) more broadly—is

central for social cognition (Figure 23.4d). Studies of patients with lesions confined to the OFC have reported impairments in a wide range of social functions, including the regulation and control of social responses, the perception and integration of social cues, and perspective taking (Rolls et al., 1994; Bechara et al., 2000; LoPresti et al., 2008; Ruby & Decety, 2004). Recent evidence from Stone et al. (2002) further demonstrates that patients with orbitofrontal damage produced selective impairments in reasoning about social contracts, supporting the proposed role of the PFC in social exchange. Bechara et al. (2000) observed profound deficits in the ability of orbitofrontal patients to represent and integrate social and emotional knowledge in the service of decision making. Converging evidence is provided by LoPresti et al. (2008), who demonstrated that the left alPFC (area 11) mediates the integration of multiple social cues (i.e., emotional expression and personal identity), further suggesting that this region supports the integration of multiple classes of social knowledge. Further fMRI evidence was provided by Moll et al. (2006) who reported bilateral recruitment of the OFC (area 11) during a social decision-making task when participants had to evaluate the social contributions of a charitable organization and chose not to make a donation.

Additional support derives from the decision neuroscience literature, which demonstrates that progressively anterior subregions of the lateral PFC (areas 10 and 11) are associated with higher-order processing requirements for thought and action (Barbey, Koenigs, & Grafman, in press; Badre, 2008; Botvinick, 2008; Koechlin & Summerfield, 2007). Ramnani and Owen (2004) reviewed contemporary research and theory investigating the cognitive functions of the alPFC, concluding that this region is central for integrating the outcomes of multiple cognitive operations, consistent with the predicted role of the alPFC for representing higher-order inferences that incorporate both necessary and possible states of affairs (for representative findings, see Kroger et al., 2008; Christoff & Kermatian, 2007; Christoff et al., 2001; 2003; Smith et al., 2007).

Conclusion

We have reviewed converging lines of evidence to support an evolutionarily adaptive neural architecture for goal-directed social behavior within the lateral PFC, drawing upon recent theoretical developments in evolutionary psychology and emerging neuroscience evidence investigating the biology

evolution, ontogeny, and decision making. We have argued that the lateral PFC recruits necessary (obligatory) and possible (permissible) states of affairs, and that these states are utilized (Figure 23.4c).

The review of the lateral PFC interactions and emerging principles for how behavior (obligatory and permissible) behavior theories that control co-ordinated behavior further investigated with inference. Do that control stored in post as an integrative specific representations (2003): Future logical, developmental behavior-guided (spherical) versus of action (Figure 23.4d) principles for thought whether norms should be controlled system evolutionary changes have shaped (Maynard Smith, 1992; 2005; 1992; Sugiyama et al., 1992; Hamilton, 1992). Hug, 1992). gate the role of human belief our understanding evolutionarily adaptive

evolution, ontogeny, and cognitive functions of this region. We have surveyed a broad range of social and decision neuroscience evidence demonstrating that the lateral PFC mediates behavior-guiding principles for specific classes of inference, with the vlPFC recruited when drawing inferences about necessary (obligatory or prohibited) courses of action, engagement of the dlPFC when reasoning about possible (permissible) behavior, and the dlPFC recruited when both categories of inference are utilized (Figure 23.4a).

The reviewed findings elucidate the involvement of the lateral PFC in normative dimensions of social interactions and raise questions for future and emerging programs of neuroscience research. One challenge that awaits future research is to address how behavior-guiding principles for necessary (obligatory and prohibited) and possible (permissible) behavior are represented within dual-process theories that distinguish between automatic versus controlled cognitive processes (Lieberman, 2007; Barbey & Sloman, 2007). Future research should further investigate the cognitive operations that are performed within the lateral PFC to support human inference. Does this region (i) contain mechanisms that control the recruitment of representations stored in posterior cortices (Miller, 2000), (ii) serve as an integrative hub for synthesizing modality-specific representations (Pessoa, 2008), or (iii) store unique forms of knowledge (Wood & Grafman, 2003)? Future research should also address the biological, developmental, and evolutionary principles that account for the observed lateralization of behavior-guiding principles for necessary (left hemispheric) versus possible (right hemispheric) courses of action (Figure 23.4). The proposed evolutionary origins and biological basis of behavior-guiding principles for thought and action motivate the question of whether normative standards for human rationality should be constructed from formal mathematical and logical systems, or instead assessed in terms of the evolutionary conditions and ecological contexts that have shaped the development of the human mind (Maynard Smith, 1982; Cosmides, 1985; 1989; Tooby & Cosmides, 1996; Cosmides & Tooby, 1989; 1992; 2005; Fiddick et al., 2000; Stone et al., 2002; Sugiyama et al., 2002; Trivers, 1971; Axelrod & Hamilton, 1981; Platt & Griggs, 1993; Gigerenzer & Hug, 1992). Finally, future research should investigate the role of the lateral PFC in the formation of human belief systems, which structure and organize our understanding of the social world. From evolutionarily adaptive social norms represented within

the lateral PFC, belief systems for moral (Moll et al., 2005; Kapogiannis et al., 2009), ethical, and political (Zamboni et al., in press) thought are constructed. By investigating the origins of this knowledge—assessing the formation of normative principles for goal-directed behavior, and their expression in moral, ethical, and political thought—the burgeoning field of social cognitive neuroscience will continue to advance our understanding of the remarkable cognitive and neural architecture from which uniquely human systems of value and belief emerge.

Acknowledgments

The authors are supported by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke. Preparation of this manuscript was based on and adapted from research investigating the role of the prefrontal cortex in goal-directed social behavior (Barbey et al., in press; Barbey et al., 2009).

References

- Amodio, D. M. & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–277.
- Anderson, J. A. (1995). *An introduction to neural networks*. Cambridge, MA: MIT Press.
- Axelrod, R. & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211, 1390–1396.
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12, 193–200.
- Barbas, H. & Pandya, D. (1991). Patterns of connections of the prefrontal cortex in the rhesus monkey associated with cortical architecture. In H. S. Levin, H. M. Eisenberg & A. L. Benton (Eds.), *Frontal lobe function and dysfunction* (pp. 35–58). New York: Oxford University Press.
- Barbey, A. K., Koenigs, M., & Grafman, J. (in press). Orbitofrontal contributions to human working memory. *Cerebral Cortex*.
- Barbey, A. K. & Patterson, R. (in press). Neural architecture of explanatory inference in the prefrontal cortex. *Frontiers in Cognitive Science*.
- Barbey, A. K. & Barsalou, L. W. (2009). Reasoning and problem solving: Models. In L. Squire, T. Albright, F. Bloom, F. Gage & N. Spitzer (Eds.), *Encyclopedia of neuroscience* (pp. 35–43). Oxford: Academic Press.
- Barbey, A. K. & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, 30, 241–297.
- Barbey, A. K., Krueger, F., & Grafman, J. (in press). An evolutionarily adaptive neural architecture for goal-directed social behavior. *Trends in Neurosciences*.
- Barbey, A. K., Krueger, F., & Grafman, J. (2009). Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning. *Transactions of the Royal Society: Biological Sciences*.
- Barsalou, L. W. (2005). Continuity of the conceptual system across species. *Trends in Cognitive Sciences*, 9, 309–311.

- 358

- Krueger, J. K., Nyström, L. E., Cohen, J. D., & Johnson-Laird, P. N. (2008). Distinct neural substrates for deductive and mathematical processing. *Brain Research*, 1243, 86–103.
- Krueger, F., Spampinato, M., Barbey, A. K., Huey, T., Morland, T., & Grafman, J. (in press). The role of the frontopolar cortex and inferior parietal cortex in mediating action complexity and duration: A parametric fMRI study. *Neuroreport*.
- Krueger, F., Barbey, A. K., & Grafman, J. (2009). The medial prefrontal cortex mediates social event knowledge. *Trends in Cognitive Sciences*, 13, 103–109.
- Krueger, F., Moll, J., Zahn, R., Heinecke, A., & Grafman, J. (2007). Event frequency modulates the processing of daily life activities in human medial prefrontal cortex. *Cerebral Cortex*, 17, 2346–2353.
- Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, 58, 259–289.
- LoPresti, M. L., Schon, K., Tricarico, M. D., Swisher, J. D., Celone, K. A., & Stern, C. E. (2008). Working memory for social cues recruits orbitofrontal cortex and amygdala: A functional magnetic resonance imaging study of delayed matching to sample for emotional expressions. *Journal of Neuroscience*, 28, 3718–3728.
- Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, England: Cambridge University Press.
- Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, 1, 59–65.
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, 6, 799–809.
- Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., & Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences*, 103, 15623–15628.
- Monti, M. M., Osherson, D. N., Martinez, M. J., & Parsons, L. M. (2007). Functional neuroanatomy of deductive inference: A language-independent distributed network. *NeuroImage*, 37, 1005–1016.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314, 1560–1563.
- Noveck, I. A., Goel, V., & Smith, K. W. (2004). The neural basis of conditional reasoning with arbitrary content. *Cortex*, 40, 613–622.
- Osherson, D. N., Perani, D., Cappa, S., Schnur, T., Grassi, F., & Fazio, F. (1998). Distinct brain loci in deductive versus probabilistic reasoning. *Neuropsychologia*, 36, 369–376.
- Pandya, D. N., & Barnes, C. L. (1987). Architecture and connections of the frontal lobe. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 41–72). New York: The IRBN Press.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9, 148–158.
- Platt, R., & Griggs, R. (1993). Darwinian algorithms and the Wason selection task: A factorial analysis of social contract selection task problems. *Cognition*, 48, 163–192.
- Prehn, K., Wartenburger, I., Meriau, K., Scheibe, C., Goodenough, O. R., Villringer, A., et al. (2008). Individual differences in moral judgment competence influence neural correlates of socio-normative judgments. *Social Cognitive and Affective Neuroscience*, 3, 33–46.
- Rammani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, 5, 184–194.
- Rilling, J. K., Goldsmith, D. R., Glenn, A. L., Jairam, M. R., Elfenbein, H. A., Dagenais, J. E., et al. (2008). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia*, 46, 1256–1266.
- Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry*, 57, 1518–1524.
- Ruby, P., & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Neuroscience*, 16, 988–999.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nyström, L. E., & Cohen, J. D. (2003). The neural basis of decision making in the ultimatum game. *Science*, 300, 1755–1758.
- Santrock, J. W. (2005). *Children* (8th Ed.). New York: McGraw-Hill.
- Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social bonds of female baboons enhance infant survival. *Science*, 302, 1231–1234.
- Smith, R., Keramian, K., & Christoff, K. (2007). Localizing the rostromedial prefrontal cortex at the individual level. *NeuroImage*, 36, 1387–1396.
- Stone, V., Cosmides, L., Tooby, J., Kroll, N., & Knight, R. (2002). Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. *Proceedings of the National Academy of Sciences*, 99(17), 11531–11536.
- Sugiyama, L., Tooby, J., & Cosmides, L. (2002). Cross-cultural evidence of cognitive adaptations for social exchange among the Shiwiar of Ecuadorian Amazonia. *Proceedings of the National Academy of Sciences*, 99(17), 11537–11542.
- Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *Monist*, 59, 204–217.
- Tooby, J., & Cosmides, L. (1996). Friendship and the banker's paradox: Other pathways to the evolution of adaptations for altruism. In W. G. Runciman, J. Maynard Smith, & R. I. M. Dunbar (Eds.), *Evolution of social behaviour patterns in primates and man. Proceedings of the British Academy*, 88, 119–143.
- Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35–57.
- Volz, K. G., Schubotz, R. I., & von Cramon, Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *NeuroImage*, 21, 848–857.
- Wason, P. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, England: Penguin.
- Wason, P. (1983). Realism and rationality in the selection task. In J. St. B. T. Evans (Ed.), *Thinking and reasoning: Psychological approaches* (pp. 44–75). London: Routledge.
- Wason, P., & Johnson-Laird, P. (1972). *The psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Weissman, D. H., Perkins, A. S., & Woldorff, M. G. (2008). Cognitive control in social situations: A role for the dorsolateral prefrontal cortex. *NeuroImage*, 40, 955–962.
- Wood, J., & Grafman, J. (2003). Human prefrontal cortex: Processing and representational perspective. *Nature Reviews Neuroscience*, 4, 139–147.
- Zamboni, G., Gozzi, M., Krueger, F., Duhamel, J., Sirigu, A., & Grafman, J. (in press). Individualism, conservatism, and radicalism as criteria for processing political beliefs: A parametric fMRI study. *Social Neuroscience*.
- Zeki, S. (1993). *A vision of the brain*. Cambridge, MA: Blackwell Scientific Publications.