frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 01 June 2012
doi: 10.3389/fnhum.2012.00148

Games people play—toward an enactive view of cooperation in social neuroscience

Denis A. Engemann[1,2]*†, Danilo Bzdok[3,4†], Simon B. Eickhoff[3,4,5], Kai Vogeley[1,2] and Leonhard Schilbach[1,6]

[1] Brain Imaging Lab, Clinic for Psychiatry and Psychotherapy, University Hospital of Cologne, Cologne, Germany
[2] Jülich Research Centre, Institute of Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3), Jülich, Germany
[3] Jülich Research Centre, Brain Network Modelling Group, Institute of Neuroscience and Medicine (INM-1), Jülich, Germany
[4] Department of Psychiatry Psychotherapy and Psychosomatics, University Hospital of Aachen, Aachen, Germany
[5] Institute for Clinical Neuroscience and Medical Psychology, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany
[6] Max-Planck-Institute for Neurological Research, Cologne, Germany

Edited by:
Chris Frith, Wellcome Trust Centre for Neuroimaging at University College London, UK

Reviewed by:
Natalie Sebanz, Central European University, Hungary
Brooks King-Casas, Virginia Tech Carilion Research Institute, USA

***Correspondence:**
Denis A. Engemann, Jülich Research Centre, Institute of Cognitive Neuroscience, Institute of Neuroscience and Medicine-3 (INM-3), Leo-Brandt Str. 5, 52425 Jülich, Germany.
e-mail: d.engemann@fz-juelich.de
† *These authors contributed equally to this work.*

The field of social neuroscience has made considerable progress in unraveling the neural correlates of human cooperation by making use of brain imaging methods. Within this field, neuroeconomic research has drawn on paradigms from experimental economics, such as the Prisoner's Dilemma (PD) and the Trust Game. These paradigms capture the topic of conflict in cooperation, while focusing strongly on outcome-related decision processes. Cooperation, however, does not equate with that perspective, but relies on additional psychological processes and events, including shared intentions and mutually coordinated joint action. These additional facets of cooperation have been successfully addressed by research in developmental psychology, cognitive science, and social philosophy. Corresponding neuroimaging data, however, is still sparse. Therefore, in this paper, we present a juxtaposition of these mutually related but mostly independent trends in cooperation research. We propose that the neuroscientific study of cooperation could benefit from paradigms and concepts employed in developmental psychology and social philosophy. Bringing both to a neuroimaging environment might allow studying the neural correlates of cooperation by using formal models of decision-making as well as capturing the neural responses that underlie joint action scenarios, thus, promising to advance our understanding of the nature of human cooperation.

Keywords: cooperation, stag hunt, game theory, joint action, joint attention, neuroeconomics, shared intentionality, we-mode

Human cultural knowledge and social institutions are unique features that cannot be found in other species. Without continuous cooperative efforts among humans, there were no such things as cars, computers, or algebra. Neither would human beings get married, earn money, vote for presidents or bring about Beethoven's ninth symphony. To cooperate, according to Webster's Third International Dictionary (Gove and Merriam-Webster Inc., 2002), means (1) to work with another or others toward a common end, (2) to act together, and (3) to associate with another or others for mutual (often economic) benefit. The Collins Cobuild Dictionary (1995) additionally highlights that to cooperate entails a willingness to help one's collaborators. Following this common sense definitions, for the current purposes, we will rely on a working definition of cooperation that includes the following three aspects: (1) acting together to pursue a common goal, (2) striving for mutual benefits, and (3) being willing and able to maintain cooperative activities and remedy problems if necessary.

The aim and object of this paper consists in the attempt to discuss human cooperation from the perspectives of different scientific disciplines. We will first briefly review exemplary empirical evidence and concepts on cooperation in anthropology, economics, behavioral psychology, developmental and comparative psychology, as well as philosophy. We then go on to evaluate how much those disciplines have contributed so far to the burgeoning field of neuroeconomics. Finally, we will suggest and comment on possible future avenues of an intensified multi-disciplinary approach to cooperation research and possible refinements in methodology that could help research in this area.

ANTHROPOLOGICAL ORIGINS OF HUMAN COOPERATION

Advanced forms of cooperation are rare in non-primate species (Dunbar, 1993) and probably emerged in non-human primates several million years ago (Cosmides and Tooby, 2005). Increasingly sophisticated social-cognitive capabilities seem indispensable preconditions for the development of more advanced cooperative skills in non-human primates, such as alliance formation and conjoint hunting (Boesch and Boesch, 1989; Boesch, 1994; Boesch and Tomasello, 1998). In particular, primatologists believe that frequent and targeted grooming is an efficient means to facilitate coalition formation (Barrett and Henzi, 2005). Reconciliation is another example of the advanced social skills of non-human primates necessary for maintaining

cooperative long-term relationships with genetically unrelated conspecifics. The uniquely human ability to contemplate others' thoughts, desires, and intentions (i.e., theory of mind) is likely to have paved the way for the development of the sophisticated social skills of humans, such as language and pedagogy (Tomasello, 1999; Csibra and Gergely, 2009; Frith and Frith, 2010; Csibra and Gergely, 2011). The ensuing ubiquitous and uniquely complex cooperative activities of humans entailed a cumulative cultural evolution and allowed for the emergence of large-scale phenomena, such as nations or the internet.

From a neurobiological perspective, accumulating evidence suggests that the comparatively large human brain did not evolve driven mainly by the need to explore and exploit the inert physical environment, but much rather the dynamic social environment (Humphrey, 1976). In support of this "social brain hypothesis," the relative size of the neocortex in primates statistically correlates with the complexity of their social systems, that is, the social group or grooming clique size, the frequency of coalitions, and strategic deception (Dunbar, 1993; Dunbar and Shultz, 2007). This suggests important selection pressures for neural circuits that decode and assess social information efficiently and reliably. For instance, neural mechanisms for recognizing and punishing free riders (i.e., individuals misusing others' cooperative tendencies) are crucial for expelling harmful individuals from the group and hence ensuring cooperation. Mutual social exchange, on the other hand, is evidently beneficial for survival as economic problems may be jointly solved, i.e., hunting prey, defending oneself against predators, or breeding collectively (Burkart et al., 2009; Hrdy, 2009). Therefore, a propensity to cooperate might be speculated to improve, evolutionarily speaking, survival in a group context (Sober and Wilson, 1999; Bacharach et al., 2006), and could be assumed to, neurobiologically speaking, be driven by activity in reward-related neurocircuits.

NEUROECONOMICS AND SOCIAL NEUROSCIENCE

Neuroeconomics has emerged as a multi-disciplinary field in which psychologists, biologists, economists, and neuroscientists join their efforts to investigate the neural basis of decision-making processes that come into play during social interaction. In this context, "social" neuroeconomics have employed paradigms that are often borrowed from behavioral game theory, which provides formal accounts of strategic interaction. As a consequence, cooperation has been mainly construed in a way that emphasizes decision-processes involved in social interactions associated with explicit payoffs. Among the multitude of games game theory employs for describing such interactions, especially social dilemma games such as the Prisoner's Dilemma (PD), social exchange games, such as the Trust Game, and fairness games, such as the Ultimatum Game constitute key tasks in recent neuroimaging studies of cooperation (for a more detailed background see **Figure 1**).

These economic games and related psychological constructs such as trust, social preferences, have been used in conjunction with different research methods: behavioral experiments, neurological lesion studies (e.g., in the ventromedial prefrontal cortex), transient TMS lesion (Knoch et al., 2006), pharmacological manipulation, e.g., Oxytocin, (Kosfeld et al., 2005; Baumgartner et al., 2008), and functional neuroimaging (King-Casas et al., 2005). Moreover, healthy subjects were compared with subjects suffering from psychiatric conditions, e.g., autism-spectrum disorder (Sally and Hill, 2006), borderline-personality disorders (King-Casas et al., 2008), and conduct disorder (Rilling et al., 2007).

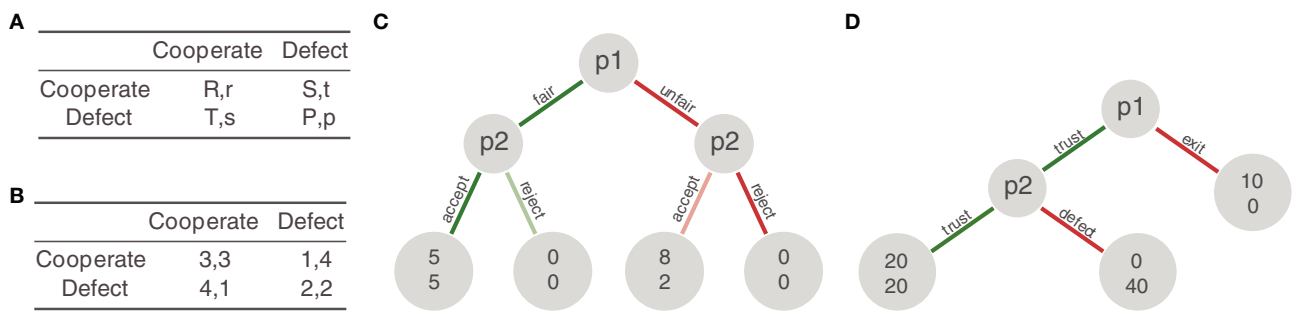

FIGURE 1 | Overview of most commonly employed economic games in social neuroscience. Panel **(A)** depicts the Prisoner's Dilemma in its general form according to which the following relation holds: $T > R > P > S$. The Prisoner's Dilemma involves a decision problem in which individuals can either maximize their own payoff, while potentially harming the co-player, or maximize both players' joint payoff at the risk of significant monetary loss given the possibility that the other does not cooperate. Consequently, it is also called a mixed-motive game and qualifies excellently for investigating the conditions for self-versus socially oriented decisions in cooperation. Panel **(B)** gives an example of the Prisoner's Dilemma's pay-off contingencies. Panel **(C)** schematically depicts the Ultimatum Game (Güth et al., 1982). Here, players have to balance own and the other's payoffs. One player receives money that he may divide up between himself and his co-player. If the co-player accepts the allocation, each of them leaves the game with a monetary reward that corresponds to that allocation. This game classifies as a fairness game because equal allocations indicate the other's preferences and leaning toward punishing unfair behavior. This is even more obvious for the related Dictator Game in which the player may literally dictate the allocation of the money, while the co-player can only accept the allocation. Fair offers and the related replies are colored green, unfair offers and related replies red. Panel **(D)** illustrates the Trust Game. This game constitutes an investment or gift exchange game that can be formally framed as a repeated PD (Binmore, 1987). Initially, players start with a certain endowment. One player then decides whether to keep his endowment or transmit it to the other player. The second player then decides to either keep this gift or to send it back to the first player. Importantly, each time the gift is transmitted its amount doubles, and hence, both players are best off sending the gift back and forth. Cooperative decisions are colored green, uncooperative decisions red.

Most neuroimaging studies in the field have concentrated on disentangling the functional profiles of brain regions involved in economic games according to preferences, reward, and decision behavior. Regarding the relevant inferential cognitive processes, activity change in the dorsomedial prefrontal cortex (dMPFC) has been linked to the evaluation of longer-term pay-off schedules. Additionally, concomitantly increased brain activity in the dMPFC, precuneus, and temporo-parietal junction is believed to reflect the integration of others' mental states during cooperation (van den Bos et al., 2009). Regarding the relevant intuitive and affective processes, activity in the ventral striatum, especially in the nucleus accumbens, and the dorsal striatum (Rilling et al., 2002; de Quervain et al., 2004) are acknowledged to be related to the intrinsically hedonic value of mutual cooperation, while the orbitofrontal cortex has been linked with the desire for revenge toward unfair partners (Singer et al., 2006; for an overview see Fehr and Camerer, 2007).

Similarly, the amgydala was observed to be involved in trust (Bzdok et al., 2011a,b), reaction to unfair offers, and fear of betrayal (Baumgartner et al., 2009). Moreover, the (dorsal) anterior cingulate cortex and anterior insula in tandem were related to anticipatory emotions associated with risk evaluation (Chang et al., 2011) and encountering unfair versus fair offers (Sanfey et al., 2003; Tabibnia et al., 2008). Ensuing prepotent emotional states and behavioral tendencies such as self-regarding preferences (Fehr and Camerer, 2007) might be over-ridden by top-down modulation from the dorsolateral and the ventrolateral prefrontal cortex (Sanfey et al., 2003; Tabibnia et al., 2008; Suzuki et al., 2011), such as when abiding by social norms (Knoch et al., 2010) or taking others' welfare into account.

It is notable that the experimental games employed in neuroeconomics are not only useful to set up effective protocols for the study of cooperation. Rather, the game-theoretic framework also provides analytic solutions that mark choices in these games in which the payoffs cannot be further improved given that the other players' strategies are fixed and provided that the players are perfectly informed. Interestingly, these equilibrium solutions sometimes predict human behavior almost perfectly, for instance in competitive markets, while in other situations, in particular, during "face to face" interactions, they fail to do so. This is especially the case for most experiments relying on the PD, the Ultimatum Game or the Trust Game. While traditional economic thought would predict self-interest to dominate decisions in those games, people show robust cooperation in laboratory and field settings suggesting that social preferences are equally important in cooperative decision-making (Fehr and Schmidt, 1999; Colman, 2003; Bacharach et al., 2006; Fehr and Camerer, 2007; Tuomela, 2007; Camerer, 2008). For example, in a meta-analysis covering 35 years of experimental work based on the PD, Sally (1995) concluded that even in so called one-shot games (with presumable absence of long-term commitments to the other player), most people exhibit a remarkable tendency to cooperate.

Modern game theoretical research, however, took this frequently observed mismatch between predicted and actual behavior as an opportunity for refining and extending the classical game theoretic framework by acknowledging findings from behavioral experiments. For instance, preference assumptions underlying the decision models have been modified to fill the gap between equilibrium predictions and behavioral data. More specifically, Fehr and colleagues (1999, 2007) proposed a utility function in the context of human cooperation termed "inequity aversion" that penalizes inequities in the player's and their co-players' payoffs. More colloquially, this model assumes that players are only fully rewarded if the outcomes are fair to both players. This not only helps to explain why individuals cooperate in one-shot games but also why they engage in punishing others for unfair offers. In another derivative of this methodology, clinical populations were classified by their cooperative biases, which can be quantified in terms of a given tendency (not) to deviate from the classical model that is preferring selfish behavior, in other words, "deviation from deviation." Recent findings show that Borderline patients, exhibit difficulties in maintaining a cooperative strategy as their partners lower their investments (King-Casas et al., 2008), whereas patients with conduct-disorders exhibit a stronger tendency to defect in social dilemma games (Rilling et al., 2007). In this sense, Kishida et al. (2010) refer to "game probes" as they propose to exploit this discrepancy as a quantitative, dimensional measure in psychiatric diagnostics.

Both cases are pertinent examples of how the systematic mismatch between classical equilibrium predictions and observed human behavior promoted qualitative and quantitative models about the cognitive mechanisms underlying human cooperation. However, the focus on economic approaches often results in blurring or neglecting other facets of cooperation. This is especially true for aspects of social interaction, such as the actual challenge of successfully executing a jointly intended cooperative action. The following sections aim at summarizing and integrating research traditions sensitive to these socio-cognitive dimensions.

BASICS IN JOINT ACTION: COORDINATING BEHAVIOR AND SHARING TASKS

Besides striving for mutual benefits, another important facet of cooperation is acting together in form of joint action. This facet has received considerable attention in behavioral and philosophical research but much less in neuroimaging and economics (Schilbach et al., 2012). One central proposition motivating joint action research is that cooperation cannot be reduced to single cooperative choices but also relates to concrete multi-agent activities in which actions are interdependent and in which a continuous flow of coordination and mutual adjustment is thus relevant. This is the case for many instrumental activities carried out in small- to medium-sized groups, such as hunting, cutting trees, or fighting together against a common enemy. Here, individuals work together to materialize a common external goal. But this also holds true for rather cultural activities such as dancing, singing, or playing board games, where the activity itself constitutes the goal. As a consequence, joint action research aims at studying how individuals bring about such tasks or playfully act in concert. This does not imply that those activities cannot be analyzed by game theory or that decisions play no role in joint action. This focus rather motivates a research orientation that emphasizes details different from those put forward by neuroeconomics,

namely, implict synchronization tendencies, anticipatory mechanisms, motor resonance, common action representations, and shared intentions. We will suggest in the course of this review that both perspectives are not necessarily incompatible with each other.

Despite marked differences regarding the proposed constituents of joint action, current research can be summarized by the following coarse taxonomy of joint action (Bratman, 1992; Searle, 1995; Tuomela, 2000, 2007; Tomasello et al., 2005; Knoblich and Sebanz, 2008; Pacherie, 2011). First, joint action implies at least two agents intentionally acting together, that is, consciously pursuing compatible goals. In cases of mutual task dependence, this implicates the agents' ability to coordinate with each other, mostly controlled automatically and without conscious awareness. This mainly includes perceptual sensitivity and behavioral responsiveness to the other's actions and mental states. Second, joint action, in many cases, implicates rather explicitly shared mindsets and motivations including specific beliefs, desires, goals, and intentions. Of note, "shared" emphasizes that those mindsets and motives can be actively expressed by communicative gestures and verbal behavior. This distinction, in particular, responds to the fact that the reasons for which individuals act together often significantly differ and may include full-blown cooperative, but also selfish or simply socially compatible but actually private motives. We will address this point in greater detail toward the end of this review in the section termed "modes of cooperation and we-thinking."

In this section, a number of current findings on joint action predominantly related to "implicit" processes in joint action will be reviewed, including automatic behavioral coordination and action co-representations. The next section will tap into the more explicitly processed facets of joint action by reviewing findings and concepts from developmental and comparative psychology. The subsequent three sections will review current neuroimaging research pursuing rather integrative approaches with regard to the behavioral and cognitive facets of joint action and cooperation. A final section will examine recent theoretical contributions from social philosophy and economics with regard to their potential of bringing together together the different variants and aspects.

AUTOMATIC COORDINATION OF BEHAVIOR

Research on behavioral coordination in joint action demonstrates that individuals show a strong propensity to synchronize their behavior in the presence of others. This can even be observed in experimental settings that are highly unlikely to elicit explicit reflection of mutual actions. For example, Richardson and colleagues (2007) had subjects sit in front of each other on rocking-chairs. Those chairs were shaped in a way to bias for different rocking frequencies. Nevertheless, synchronization of the subjects' rocking frequencies was observed in this scenario. Further studies prompting individuals to coordinate their behavior explored cognitive mechanisms for mutual adjustments. In a behavioral study by Knoblich and Jordan (2003) subjects shared control over a tracking-device that had to be kept aligned with a horizontally moving object. Crucially, this rather difficult task could not be achieved individually. The results suggested that individuals indeed solve coordination problems by anticipating the others' moves without relying on explicit communication. A joint tapping paradigm extends these findings (Konvalinka et al., 2010): subjects had to synchronize to external beats or to their partner. The results suggested that anticipation and adjustment performed by each individual are necessary but not sufficient for high synchronization performance. Instead, anticipation and adjustment need to be bidirectional, i.e., mutual, for accomplishing joint action. First insights into the neural basis of coordination in joint action were provided by Newman-Norlund and colleagues (2008). In their fMRI experiment, single and joint action versions of a virtual lifting task were compared, in which subjects had to adjust their actions in order to prevent a ball from slipping off a bar. The results include increased activity in the inferior frontal gyrus (IFG; pars opercularis) and the posterior superior temporal sulcus (pSTS), which are both believed to be part of the putative human mirror neuron system (Keysers and Perrett, 2004; Oberman et al., 2007; Schulte-Rüther et al., 2007; Schilbach et al., 2010; Wang et al., 2011). This has been taken to suggest that when human actors achieve mutual coordination they rely on a motor representation of their partner's ongoing action.

Taken together, these findings shed light on advanced coordination skills that go beyond mere temporal estimation and prediction as suggested by the individual-joint comparisons. However, some caution is warranted regarding the temptation to take these examples as instantiations of mental-state coordination or full-blown cooperation, as these studies do not permit any principled conclusions about the nature of the entities coordinated nor the agents' attitudes toward their acting together. The following two sections will address these nuances in greater detail.

COMMON ACTION AND TASK REPRESENTATIONS

Despite providing first insights into the cognitive and neural processes that underlie joint action, these tasks put special emphasis on behavioral coordination. Accordingly, insights into the coordination of mental states in joint action can be drawn from another set of studies that specifically investigated how people share representations of their partner's actions, that is, how they form action co-representations. Studies on action co-representation characteristically exploit the so-called Simon-Effect (Simon, 1969). This effect is elicited in subjects that are asked to respond spatially to non-spatial features of stimuli while ignoring the location of the stimulus presentation. Characteristically, in tasks that elicit the Simon-Effect actual stimulus location affects reaction times. For instance, when subjects respond with left button presses to green and with right button presses to red stimuli, they tend to be significantly slower when the green stimulus appears on the right side (incongruent response) as compared to the left side (congruent response). In a seminal behavioral study by Sebanz et al. (2003) a joint action version of the Simon-Task, also called "interactive" or "Social Simon-Task" was established. In their setting, task rules were distributed among two subjects sitting alongside each other in order to reduce the tasks to individual go/nogo, which were performed independently but in parallel. That is, the task was to respond to certain stimuli, e.g., green ones, but not to others, e.g., red ones, which are presented at different locations (left versus right). Importantly, a Simon-Effect (increased reaction times during incompatible trials) was observed in this

joint task as compared to the same task in the absence of a second actor, indicating mutual co-representation of actions among the subjects.

In a number of follow-up experiments (Sebanz et al., 2005) it was demonstrated that individuals co-represent not only their partners' actions but also the rules guiding their actions. These cognitive entities were teased apart by implementing different tasks for both players, which sometimes necessitated the same and sometimes necessitated different responses to the same stimuli. Concretely, one subject had to respond to the direction of a stimulus, whereas the other subject had to respond to its color. The Social Simon-Effect, that can be regarded as a quantitative behavioral marker for joint action, was markedly stronger when both tasks required different responses by the subjects. This indicates a high sensitivity to the partner's task, even if it was irrelevant to the subject's task. Interestingly, the number of neuroimaging studies that make use of Social Simon-Tasks is still limited. In a replication of Sebanz's and colleagues' study (2003) with concomitant EEG recordings during a joint action go/nogo task, negative ERPs in the parietal lobe were observed for the joint nogo as compared to the single nogo trials (Sebanz et al., 2006). As suggested by the authors, these findings might reflect intensified inhibitory processes in response to the challenge of disentangling one's own and the partner's representations during reciprocally dependent action. In the same vein, Sebanz et al. (2007), using fMRI versions of the joint go/nogo tasks (Sebanz et al., 2005), linked increased activation of the dorsal anterior cingulate cortex and the ventral medial frontal gyrus to intensified self-processing during joint action as compared to individual action.

Taken together, these studies suggest that individuals are highly sensitive to their partners' actions and mental states, even when mutual coordination is not relevant for achieving a given task. More recent studies following this methodological track suggest that the set of mental entities that can be tracked using the interactive Simon Task is to be extended to the personal relationship between the actors (Hommel et al., 2009) and cooperative intentions (Ruys and Aarts, 2010; Iani et al., 2011). On the other hand, the Social Simon-Effect does not involve conscious or explicit processing of the mental states governing ones partner's behavior. Recent findings suggest this effect to be grounded on low-level saliency mechanisms rather than higher-level representational processing (Vlainic et al., 2010; Dolk et al., 2011). Capturing the explicit dimension of cognitive processes subserving joint action, at least to some extent, requires to permit subjects to interact and express themselves in a less constrained fashion. The subsequent section discusses findings from developmental and comparative psychology highlighting the role of shared intentions and reward in joint action and cooperation.

JOINT ACTION FROM A COMPARATIVE AND DEVELOPMENTAL PERSPECTIVE: SHARING INTENTIONS AND COMMITMENTS

Complementary to the rather implicit processes of adjustment during joint action investigated above, more elaborate forms of joint action have been investigated by recent studies (Warneken et al., 2006; Gräfenhain et al., 2009; Callaghan et al., 2011; Hamann et al., 2011) from developmental and comparative psychology. In a study by Warneken et al. (2006) experimenters tried to engage human-raised adult chimpanzees as well as 18-month-old and 24-month-old toddlers in shared instrumental activities and social games. In the former acting together lead to material rewards, whereas the latter aim at maintaining and enjoying the shared activity *per se*. For example, in a typical instrumental task, food or toys were hidden inside a long tube with two handles and could only be released by the subjects' and experimenters' combined efforts. In contrast, a typical social game was constituted by a trampoline which the experimenter and subject could utilize in concert to make a ball jump up and down. Chimpanzees and children displayed substantial coordination skills in all instrumental tasks. Contrarily, it was hardly possible to make chimpanzees engage in social games, which appeared to be intrinsically rewarding to 18- and 24-month-old human toddlers. Finally, as activities got spontaneously interrupted by the experimenter, children but not chimpanzees, tried to reengage the experimenter rather than trying to complete the task individually or engage in other activities.

This observed effect is noteworthy, as it cannot be attributed to a lack of cognitive or motor capacities in chimpanzees. Rather, great apes are likely able to understand others' goals and even, to some extent, others' knowledge (Hare et al., 2001). Moreover, the example demonstrates that apes are capable coordinators when food rewards are expected. Comparing instrumental activities with social games, thus suggests that children, but not chimpanzees, exhibit an intrinsic motivation to collaborate. The findings eluded to above also highlight the importance of distinguishing games from instrumental activities and demonstrate the limits of current neuroeconomic approaches relying on instrumental payoff in investigating the putative reward mechanisms underlying cooperation.

Interestingly, when looking at specific social motives that might explain the observed differences, these seem to go beyond altruistic helping (helping irrespective of external rewards), which can be reliably elicited in chimpanzees (Warneken et al., 2007) and, therefore, is not specific to humans. However, another series of studies by Gräfenhain et al. (2009) suggests that it might be the mutual commitment to the joint activity that entails the difference between chimpanzee and human cooperation. This proposal is further motivated by philosophical analyses by Gilbert (1990), Bratman (1992), and Tuomela (2000, 2007) on shared intentions. The basic idea is that people form joint commitments as they act together. Hence, they strongly expect each other to fulfill their respective roles and try to help each other when problems arise during cooperation. Consequently, if one of the agents interrupts his participation in the shared task, her partner should not only be surprised, resulting in updated beliefs about the world, but should also exhibit disappointment, resentment, and other indicators for normative charge (Rakoczy et al., 2008). The latter may then encourage attempts to reengage the cooperator, given the two individuals sharing social commitments to the joint activity.

This line of thought led to the experiments by Gräfenhain et al. (2009) which were based on social games similar to the ones from Warneken et al. (2006), which, however, could be carried out either conjointly or individually. In one characteristic game, child and experimenter sat alongside each other in front

of a box. The box was equipped with one handle for each of the two players that sets a rabbit free. That is, the other's participation was not necessary to play the game, and moreover, the experimental setting allowed taking the other's role. This experimental detail constitutes a decisive variation of the games used by Warneken and colleagues that ultimately allowed the researchers to modulate explicitly the children's joint commitments while playing social games. Assuming that joint commitments arise out of, within and by social interaction, they were varied by either positively engaging the child before and during the play (contingent acting, mutual gaze, smiling at the child) versus neutrally acting in a rather parallel, unrelated fashion (no contingent actions, smile not directed at child). In subsequent interruption periods, around 2-year-old children more often attempted to reengage the experimenter and less often continued the activity on their own when playing in the commitment-facilitating condition compared to the neutral condition. Furthermore, in a subsequent experiment Gräfenhain and colleagues (2009) nicely demonstrated that young children anticipate the experimenter's expectations toward the child depending on their commitments. As a second experimenter tried to engage them into a more interesting game, the children displayed leave-taking behavior (e.g., signs of inner conflict and at later developmental levels verbal justifications) in the commitment-facilitating but not in the neutral condition.

What makes these findings particularly intriguing is, that the partner, i.e., the experimenter, was not relevant for the task in an instrumental sense and thus potential "social tool explanations" of the child's responses to interruption can be ruled out (see Hamann et al., 2011 for additional behavioral and Callaghan et al., 2011 for cross-cultural evidence). Besides again highlighting the importance of social games as a means to access the intrinsic nature of human motivations underlying cooperation, these findings also emphasize the importance of social gaze and contingent social interaction for establishing joint commitments.

In conclusion, developmental and cross-species research strongly suggests a unique cognitive and motivational infrastructure in humans which relies on sharing intentions and forming joint commitments in order to support cooperative joint actions. Consequently, specific reward-related and emotional neural circuits might be expected to be involved. It is to the discussion of this topic that we now turn.

JOINT ACTION, COMPETITION VERSUS COOPERATION, AND REWARD

Shared intentions and reward are yet two other critical psychological entities when contemplating human cooperation. In this and the following section we will briefly introduce and discuss findings related to those two concepts in probably more elaborate forms of joint action. Neuroscientific research recently started to investigate the role of shared intentions in joint action, despite the methodological difficulties that almost naturally arise from investigating complex notions of cooperation using neuroscientific methods (Decety et al., 2004; de Bruijn et al., 2009; Newman-Norlund et al., 2009; Ruys and Aarts, 2010; Koban et al., 2010; Iani et al., 2011; Radke et al., 2011). One viable approximation to testing shared intentions in cooperation includes biasing the participant's interpretation of a shared activity either toward a cooperative or toward a competitive setting. This manipulation can be achieved when modifying the structure of the task, e.g., coupling versus uncoupling rewards (de Bruijn et al., 2009; Radke et al., 2011), designating the one winner who first completes the task (Becchio et al., 2008), or more directly by instructing the partner to behave cooperatively versus uncooperatively (Decety et al., 2004).

For example, in one of Decety's et al. (2004) fMRI experiments, individuals played a simple board game together with a confederate, who either tried to help the subjects to complete the game or tried to block the subjects' moves. Further, both conditions were compared to individual game performance. The contrast comparing joint and individual action revealed increased activation in the superior frontal gyrus, the superior parietal lobe and the anterior insula. This pattern of neural activity might be attributed to higher executive demands of joint activities for which coordination matters as compared to individual activities. Additionally, this interpretation is consistent with findings from research on shared action and task representations. Yet, networks specifically more active during cooperative versus competitive conditions have also been identified. During cooperation, the medial orbitofrontal cortex, the posterior cingulate, as well as bilateral anterior insula increased in activity. Consistent with above mentioned ideas, cooperation might thus be intrinsically rewarding and might automatically raise expectations about ones partner. During competition, however, a network including the inferior parietal cortex, the medial prefrontal cortex and the superior frontal gyrus exhibited increased activity. This possibly reflects that not only cooperation but also competition encourages mental-state ascription reflecting the strategic aspects of competition, being a potential link to the game theoretic paradigms discussed above.

Moreover, the aspect of processing unexpected events during cooperation, e.g., including errors or interruptions, was, at least in approximation, singled out by recent neuroimaging studies that investigated processing one's own and one's partners errors during cooperative versus competitive joint action. In two studies by de Bruijn and colleagues (de Bruijn et al., 2009; Radke et al., 2011) subjects played the so-called cannon-shooting game on a computer, either alone, in cooperation or in competition with their partner. This game requires precisely aligning a cannon to hit a given target. Importantly, it was played in such a way that participants were aware of their own and the other's error. Cooperation versus competition was established by coupling versus decoupling participants' overall outcome based on their respective performance. The functional analyses revealed increased neural activity in the MPFC during cooperation as well as in competition when focusing on the observation of errors that only affected the other compared to errors that only affected oneself. This result is interesting in the context of Decety's et al. (2004) study, where the same region was associated with competitive processing only. Thus, during joint action the MPFC might subserve cognition associated with scanning potential threats to one's own plans and predicting the other's behavior, irrespectively of whether he or she is a competitor or rather an incompetent collaborator.

Despite providing valuable insights into intentional processing during joint action, it becomes increasingly apparent that the neuroimaging studies face considerable problems in teasing apart intrinsic versus extrinsic motivation in cooperation, as activities are externally rewarded in most cases. Therefore, the neural substrates of commitment-disclosing behavior during joint action introduced by Warneken, Tomasello and colleagues as important indicators of cooperation (Tomasello et al., 2005; Warneken et al., 2006; Gräfenhain et al., 2009) still remain unknown (see **Figure 2** for an illustration of the current neuroscientific coverage of psychological constructs in the context of cooperation).

JOINT ACTION AND REWARD IN THE LIGHT OF JOINT ATTENTION

A promising contribution to the importance of social gaze and intrinsic reward in joint action has been made by Schilbach and colleagues (2009), who instantiated episodes of visual joint attention during fMRI scanning. The phenomenon of joint attention has been intensely discussed by Tomasello and colleagues (Tomasello, 1999; Tomasello et al., 2005; Tomasello and Carpenter, 2007) as one important example for shared intentionality, a bundle of abilities and motivations subserving the coordination and sharing of mental states. These authors proposed that the evolution of human cooperation and culture critically depended on the emergence of joint attention. In a large number of studies, they were able to show that children around their 9th month of age are able to make others follow their own gaze. Moreover, most children learn to exactly discern what others want them to attend to when looking at various objects or at compound objects, which emphasizes the intentional nature of joint attention (see Carpenter et al., 1998 for an exhaustive treatment of the topic). Finally, recent research indicates that early joint attention is even predictive of speech and theory of mind performance (Aschersleben et al., 2008). Despite its importance, the neuroscientific understanding of joint gaze and attention is still limited, which can be attributed to the methodological challenge of investigating eye-movements and naturalistic social interaction in an fMRI environment. In order to establish visual joint attention, in the experiment by Schilbach et al. (2009) the methodology was based on an eyetracking algorithm that allows detecting fixations and to adjust stimulus presentation accordingly (Wilms et al., 2010). This algorithm was then used to control the gaze interaction between the participant and a fictive confederate represented by a virtual face, in a way that allowed for capturing gaze following and, ultimately, joint attention. The neuroimaging analyses suggest an increased BOLD response in the ventral striatum during joint attention initiated by the participant, as compared to joint attention initiated by the other suggestive of an inherently rewarding experience associated with establishing a shared experience with another person. This was corroborated by making use of correlation analyses with a postscan pleasantness rating, which indicated that participants actually preferred looking at objects together with the virtual other, rather than alone.

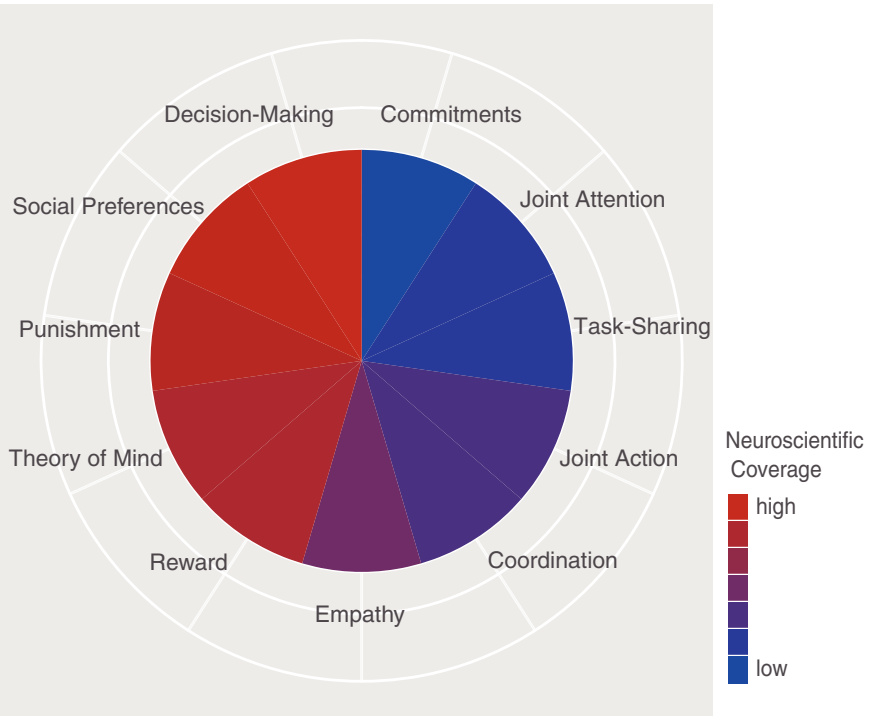

FIGURE 2 | Current neuroscientific coverage of psychological concepts related to cooperation. Purely illustrative depiction that demonstrates to what extent different psychological facets underlying cooperative phenomena have up to now been subject to neuroimaging-based investigations. The measures are based on the authors' subjective impression rather on any objective measure given general inconsistencies in nomenclature and diverging experimental settings throughout existing research on cooperation.

The results and the scope of this study are particularly interesting for at least two aspects relevant to the topic of this paper. First, the task involves externally unrewarded activity and thus the results are in favor of intrinsic rewards driving joint action. Second, joint attention is likely to involve shared intentions and can be understood as a very basic form of joint action (Fiebich and Gallagher, 2012; Pfeiffer et al., 2011). Consequently, and in line with Schilbach et al. (2012) it can be argued that gaze-based interaction paradigms constitute a potential key method for integrating comparative and developmental research with social neureconomics. However, further experiments are needed that clearly establish hypotheses about the intrinsic reward of joint attention in ruling out mere contingency or efficacy-experiences as potential confounds.

To sum up, joint action research elucidates several aspects and preconditions of cooperation in the sense of acting together; some of them even tap into mutual helping and support. It demonstrates individuals' sensitivity to coordinate, to establish common representations of their joint activity in cognitive and motivational terms. Neuroimaging studies complement these findings and corroborate psychological assumptions regarding the role of executive functions, reward processing, action mirroring and mentalizing in joint action. At the same time, it has to be stated that neuroimaging studies have not yet been able to fully capture the details revealed by developmental studies based on interaction-based methodologies. In order to fill this gap, future neuroimaging paradigms should employ externally unrewarded social games and include naturalistic interactions that allow for mutual interventions (reengagement attempts, criticism, teaching, reassuring) as exceptions (interruptions, problems, errors) arise during joint action. The latter point is of considerable significance in unveiling the agent's motives and experiences that govern their mutual cooperation, in other words telling to which extent agents cooperate when acting together. Importantly, this is not just cosmetic in nature, as comparative studies reveal distinctions in ape and human cooperation only at this level.

The studies discussed up to this point capture many important facets of cooperation: striving for mutual benefit, acting together and supporting each other. Nevertheless, it becomes clear that the individual pieces of evidence are difficult to integrate with each other, as the different trends of cooperation research significantly differ in their underlying key concepts (decision making, joint action, shared intentionality) and employed methods (game theory, neuroimaging, action-based approaches). In the following sections, we will discuss conceptual advances that might help to remedy this shortcoming.

NEUROECONOMICS REVISITED: STAG HUNT COOPERATION

As was seen above, game theoretic approaches to cooperation suggest social preferences to guide human decision making in strategic social interaction including material payoffs. Concretely, the findings suggest mechanisms that invoke "social" utility functions, thus coupling personal with social welfare and encouraging cooperative choices by punishment of non-cooperative behavior. In fact, research on social preferences taps into the motivational contingencies in human cooperation. Conceptually, social preferences fulfill a similar role as the joint commitments unveiled by comparative and developmental studies. Both approaches argue for the intrinsic nature of the mechanisms captured by their paradigms. Yet, both perspectives suffer from their own limitations. Also, findings from comparative and developmental research are difficult to compare to findings from neuroeconomics.

Developmental and comparative approaches convincingly rely on naturalistic joint action scenarios in which social interactions during instrumental and non-instrumental activities are compared. Yet, research on social preferences relies on one-shot games in which the social interaction is reduced to making choices known to have consequences for the other's payoff. While the former seem to deemphasize formal analysis, the latter hardly seems to capture ecologically valid scenarios (Schonberg et al., 2011). Studying decision-making using material payoffs thus seems more comparable to instrumental activities but not to the games employed by developmental studies. Moreover, in the neuroeconomic paradigms used to test social preferences selfish and social preferences usually form a potential conflict, at least in the light of the classical game theoretic framework. For example in the PD, uncooperative choices constitute the only Nash-equilibrium but are not Pareto efficient, indicating rational options that, if chosen by all players, would not yield the best possible outcome for all (Myerson, 1997). Therefore, one may also call these games mixed-motive games.

This, however, does not hold for most of the joint action paradigms which may be more appropriately analyzed as common interest games in which one Pareto superior Nash equilibrium exists (Bacharach et al., 2006) implicating congruence between selfish and social preferences, hence, facilitating the formation of shard intentions. One such game is the stag hunt game that can be traced back to a parable by Jean-Jacques Rousseau, in which two hunters can either conjointly hunt a stag together or hunt a hare separately. Stags possess considerable strength and, therefore, hunting them requires successfully combining efforts, but also is most rewarding. Conversely, hunting a hare does not necessitate assistance of the other but yields only limited reward, whereas hunting a stag alone, constitutes the least efficient option (see **Figures 3A,B** for a schematic depiction of the strategic interaction captured by the stag hunt game). From a cooperative stance, thus, the essence of this game is to coordinate each other's action toward hunting a stag together and only to choose hunting hares if it is indicated that hunting a stag is unlikely to succeed.

Another way of looking at this game is to regard the stag choice as payoff dominant, as compared to the risk dominant choice for hunting hares: players can gather sufficient knowledge about each other, e.g., by communication or mentalistic reasoning, to be confident in opting for the stag. Contrastingly, a lack of such knowledge would make hunting hares more advisable. Therefore, proficient stag hunters should be equipped with distinct abilities to assess the other's mental states or even establish shared intentions which should be easier in this strategic interaction due to the absence of principled conflict between social and selfish motives. Interestingly, according to Tomasello and colleagues (Tomasello, 2009; Hamann et al., 2011; Rekers et al., 2011) intensified collaboration, is an essential feature of human cooperation that requires exactly theses capacities. Therefore, typical stag

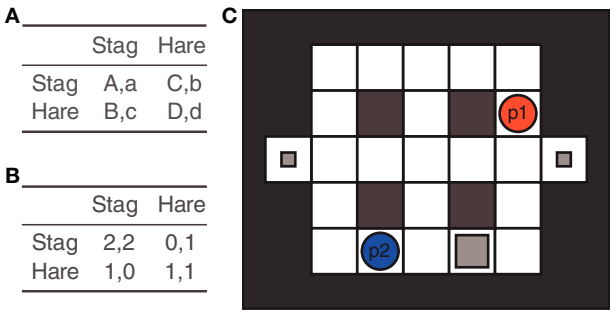

FIGURE 3 | The Stag Hunt game and its current neuroscientific implementation. Panel **(A)** depicts the stag hunt game in its general form according to which the following relation holds: $A > B \geq D > C$. Panel **(B)** gives an example of the stag hunt's payoff matrix. Panel **(C)** shows the Stag Hunt Game as implemented by Yoshida et al. (2008, 2010a,b). Analogous to the original description by—Jean-Jacques Rousseau, the stag yields a higher reward but can only be caught by both players'combined efforts, that is, by conjointly blocking the stag (big gray square) from two sides to prevent its escape. In this implementation of the game, the hares are realized as stationary objects (small gray squares), referred to as "rabbits." This ensures that both players can easily and independently catch them. Rabbits only yield limited rewards. Additionally, the round is over when a hunter caught a rabbit. Similar to a board game players take turns moving their token between fields.

hunt game scenarios, i.e., physical collaboration, presumably were of major importance during the evolution of human cooperation (Skyrms, 2004; Bacharach et al., 2006) and might have paved the way for the emergence of human's unique intention-reading and -sharing aptitude (Tomasello, 2009).

Consequently, paradigms incorporating versions of the stag hunt game (potentially with different pay-off schedules) might help to combine the interactive features of Tomasello's and colleagues' joint action paradigms with the analytic power provided by game theory.

Neural evidence regarding stag hunt interactions, however, is currently rather sparse. In a recent fMRI study comparing high-payoff choices in a stag hunt game and a differently framed but payoff identical lottery game, Ekins et al. (2012) report significantly increased activation in brain regions associated with mentalizing (pSTS, anterior, and posterior cingulate cortex). These findings are in support the "social knowledge" account depicted above and implicate social-intentional processing to support cooperative decisions in stag hunt interactions. However, the evidence is somewhat limited by the lack of whole-brain analyses. Moreover, the employed experimental setup largely resembled typical neuroeconomic paradigms in de-emphasizing the fine-grained details and interactive aspects common to joint action tasks.

A novel approach combining the advantages of neuroeconomic and joint action methods has recently been developed by Yoshida and colleagues in a series of studies (Yoshida et al., 2008, 2010a,b). Here, the stag hunt parable was implemented almost literally as a hunting task, in which players move their respective hunter figure on a labyrinth-like grid to hunt stags or rabbits (see **Figure 3C**). Interestingly this implementation of the game, entails continuous joint action (moving one figure toward a target) as well as strategic choice (moving toward the rabbit versus moving toward the stag together with the other). At the same time, the authors drew on the game theoretic basis of the paradigm, which lends itself to quantitative modeling. In their computational model of theory of mind, Yoshida et al. (2008) aim at explaining the agent's behavior on a trial by trial basis. The model assumed that agents employ higher order belief inference in predicting whether their partner will cooperate, given their own behavior. Moreover, the model issues a cooperation parameter which estimates the probability at which an agent chooses to hunt a stag. This model was then compared to a fixed-strategy model which assumes a constant cooperation rate (Yoshida et al., 2008). Applied to experimental data from the stag hunt paradigm, in which subjects played the hunting task together with a computer agent, the theory of mind model was significantly more predictive than the fixed-strategy model. Interestingly, this only holds for healthy controls but not for subjects suffering from ASD, whose behavior is better characterized by the fixed-strategy model (Yoshida et al., 2010a). These results nicely illustrate the synergic potential of combining joint action paradigms and game theoretic modeling.

This becomes even more evident, when this strategy is combined with neuroscientific methods. In their event-related fMRI study Yoshida et al. (2010b) adopted a model-based analysis approach, which allows for directly regressing the BOLD signal against the parameters provided by the computational model. This approach not only allows inference about the brain areas involved in a given experimental task, but moreover allows to address hypotheses about the computational operations implemented in those areas. The analyses revealed increased activity in the rostral mPFC during movements of the computer-agent as well as the activity in the bilateral ventral striatum. This entices to speculate that mentalizing and reward processes, respectively, might be involved in performing the stag hunt game. Put into practice, the model-based approach allowed for understanding the activity in the rostral mPFC as a function of uncertainty of belief inferences and the activity in the ventral striatum as a function of the outcome. Additionally, increased activity found in the dorsolateral prefrontal cortex was shown to follow a model parameter that described how many levels of recursion where involved when thinking about another agent's.

The modeling strategy employed in this study convincingly allows addressing specific hypotheses regarding the nature of the cognitive processes underlying cooperation and social interaction. Complementary to Yoshida's ToM model, Braun et al. (2009) have proposed a methodology capable of testing joint continuous motor activity for Nash equilibrium solutions. Further success in closing the loop between decision making, motor activity, and joint action is to be expected from future studies combining and exploiting related methodologies in a neuroimaging environment.

Interestingly, the neural correlates underlying the payoff and agent-movement events closely matched those reported for the self-initiated joint attention episodes from Schilbach et al. (2009). Although only speculation, this might indicate a common neural basis subserving stag hunt cooperation and basic social interactions, such as joint attention. This would support Tomasello's

(Tomasello et al., 2005; Tomasello, 2009) assumption of an internal link between stag hunt cooperation as well as the special cognitive and motivational capacities, which he summarized as "shared intentionality."

Moreover, the stag hunt paradigm might also allow studying responses to exceptions of the event-flow in the way demonstrated by Warneken and colleagues (Warneken et al., 2006, 2007; Hamann et al., 2011); and should thus allow detailing the neural basis of joint commitments. Concretely, one could investigate the neural responses to violations of expectation, that is, when subjects recognize the other's choice to hunt a hare individually versus to assume commitment to hunt a stag. Manipulation of the others' reputation, sympathy or interpersonal responsiveness might help creating joint commitments and social expectations. Although more difficult to control, one might alternatively vary to what degree the players can communicate (e.g., verbal versus nonverbal-gaze-based) during the game to study the effects of communication on stag hunt solutions. Verbal responses to exceptions might then serve as additional indicators of the underlying joint commitments. Thus, if implemented in an fMRI or a neurophysiological paradigm that elicits, e.g., verbal protests, this might provide a rather concrete link between the different construals of cooperation discussed so far.

Figure 4 illustrates the neuroimaging findings from the studies discussed in the course of this work.

MODES OF COOPERATION AND WE-THINKING

In this final section we will discuss recent theoretical developments and demonstrate how they can provide a framework that permits an integrated view of cooperation at both the behavioral and the neural level.

The analysis of cooperation by Tuomela (2000, 2007) and the theory of team-reasoning by Bacharach (Bacharach, 1999; Bacharach et al., 2006) formulate precise assumptions about the cognitive architecture and reasoning categories underlying different cooperative behaviors. Further, both theoretical approaches are designed in a comprehensive fashion that allows for a wide range of applications. In particular, Tuomela (2007) proposed that cooperation comprises all activities in which agents share and jointly pursue goals, whether specified as concrete ends involving high levels of behavioral coordination or as group norms and ideals that do not exactly specify how to bring about the implicit common goal. Another important aspect of Tuomela's theory is the distinction between "i-mode" and "we-mode" cooperation, which refers to the mindset involved in joint actions. According to this idea, agents might construe their shared activity as either involving commitment and giving rise to strong mutual expectations—a stance toward the joint action he terms we-mode. Or agents might cooperate as a matter of fact while not sharing psychological attitudes, goals and commitments at a deeper level beyond the concrete situation—a stance toward the

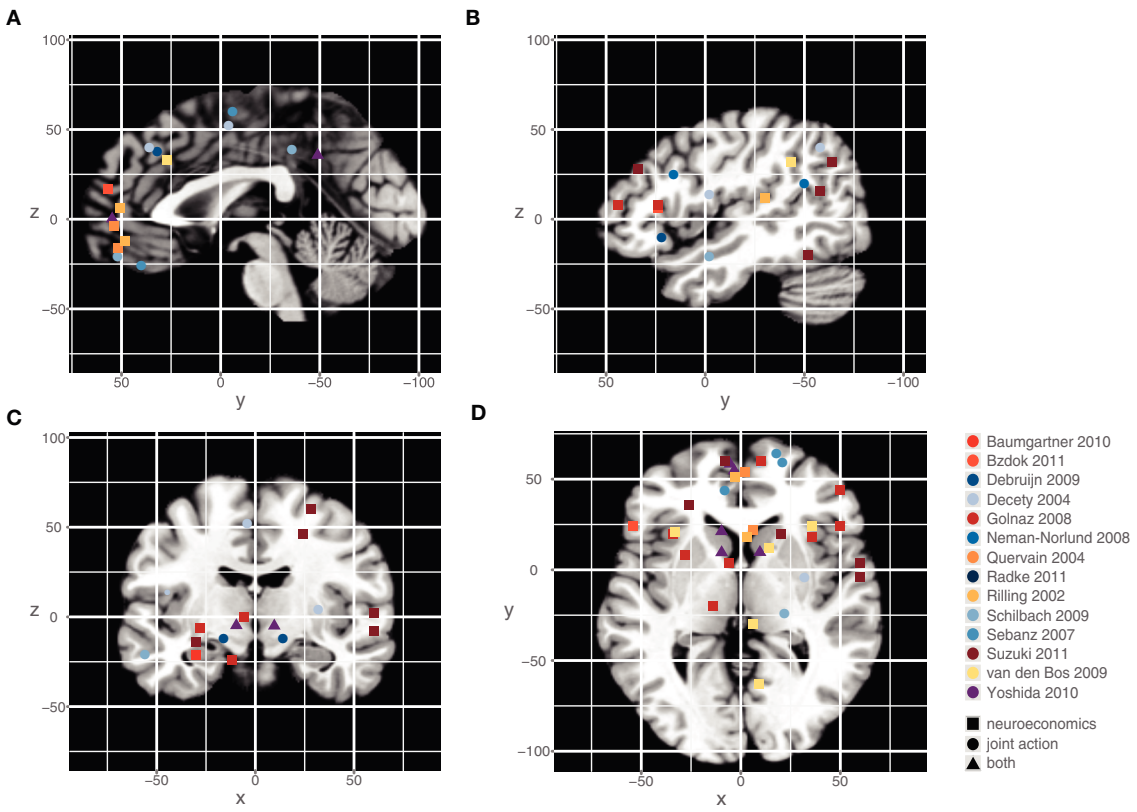

FIGURE 4 | Neural activation patterns observed in neuroeconomic and joint action imaging studies. The peak locations of the metabolic change during relevant tasks were rendered on sagittal **(A,B)**, coronal **(C)** and axial slices **(D)** of a T1 single-subject template. The color of the symbol indicates the publication from which the coordinates were drawn. The shape of the symbol describes the study category. Coordinates in MNI space.

shared activity named i-mode. In the i-mode, agents might, for example, stop cooperating as soon as costs increase.

Concretely, Tuomela argues that both aspects constitute independent dimensions yielding a taxonomy that allows describing the structure of cooperation on a task dimension (dependence versus independence) and a motivational dimension (individual [i-mode] versus social [we-mode] commitments). Applied to our purposes, Tuomela's theoretical framework allows separating the different facets of cooperation discussed so far. For example, the rabbit-game from Gräfenhain's and colleagues (2009, see above for discussion) experiments constitute an independent task while at the same time involving a pronounced we-mode in the joint commitment condition[1]. In contrast, in Newman-Norlund's and colleagues (2008) virtual-lifting task, which also was discussed above, the subjects' tasks are dependent while it remains unclear whether the subjects operate in a we-mode or in the i-mode.

Consequently, "we-mode" and "i-mode" can be conceived as mindsets or schemes between which individuals can undulate, reflecting the ubiquitous fact that humans tend to cooperation selectively. Based on a comprehensive review of the relevant literature, Tuomela (2007) further contends that the we-mode is a uniquely human phenomenon that emerges early in ontogeny (as indicated by developmental research reviewed above). However, while models have recently been proposed to incorporate these ideas into neuroscientific research (Becchio and Bertone, 2004; Adenzato et al., 2005; Becchio et al., 2006), this has so far hardly been put into practice. We propose that the we-mode theory might be a useful concept when it comes to systemizing cooperation research. Moreover, the specific assumptions of the we-mode theory are interesting by themselves and might help to design experiments that systematically tease apart different psychological dimensions constitutive for the wide range of cooperative phenomena. Whether such a thing as the we-mode really exists is an empirical question. However, gathering systematic evidence for or against this concept might help to better understand whether different facets of cooperation discussed in this paper share a common psychological and neural ground. Conversely, they might also constitute distinct phenomena that are only associated by conventional use of language.

The innovative potential of this direction of thought is further illustrated by Bacharach's (Bacharach, 1999; Bacharach et al., 2006) theory of team-reasoning, which is highly related to Tuomela's we-mode theory. Accordingly, thinking can either operate individually, as analyzed by classical economic thought, or socially, as described by his alternative game theoretic model of team-reasoning. When thinking as a team, individuals may overcome social dilemmas by modifying their frame of reasoning to employing what he calls profile selection: instead of worrying about the other's potential lack of cooperation, when thinking as a team, individuals conceive of themselves as parts of a group and assume their role in selecting the option that has the highest outcome profile from the team's perspective. Moreover, Bacharach argues, that a skill to team-reason developed during human evolution and constitutes the key capacity underlying cooperative solutions of mixed-motive games such as the PD and common interest games as the stag hunt, again, suggesting a common biological basis for different kinds of cooperation. Moreover, Bacharach's theory of we-thinking is ontologically more parsimonious as team-reasoning is basically a consequence of framing which neither requires sophisticated mentalizing nor complex normative entities (Bacharach et al., 2006; Pacherie, 2011). Such a theoretical framework thus more easily allows for including developmental and certain clinical populations lacking full-blown mentalizing capacities into the family of cooperators and joint actors.

Taken together, the benefit of a theoretical perspective in neuroscience acknowledging the notion of we-thinking, as highlighted by Tuomela and Bacharach, would clearly provide a rich framework for cooperation research, whether focusing on decision making or joint action. Paradigms capitalizing on this framework may thus constitute a promising direction to assess behavioral, goal-related, and motivational aspects of cooperation.

CONCLUSIONS

Human cooperation is a highly complex phenomenon. Hence, it can and should be viewed from various angles and dissected by diverse scientific disciplines. Anthropological research emphasizes that social selection pressures have shaped human evolution and have led to the emergence of cooperative social systems that appear to be without parallel in the animal kingdom. Experimental psychology has proposed the concept of joint action as one paramount aspect of human cooperation, which refers to the automatic synchronization of behavior during coordinated action execution. Comparative investigations stress that sharing mental states during cooperation is more prevalent in human children than in great apes and altruistic punishment is probably characteristic of human but not non-human primates. Cross-cultural studies likewise suggest that children's capacity and propensity for interpersonal cooperation is an inter-ethnically stable human trait. Finally, the advances in imaging neuroscience have begun to allow mapping the neural correlates and brain networks that subserve decision-making during cooperation tasks.

However, neuroeconomic research on cooperation has so far been mostly based on a small number of paradigms that emphasize material payoffs and decision-making, hereby often disregarding other aspects of naturalistic cooperation. We contend, however, that the stag hunt game, describing a highly under-researched strategic interaction, lends itself to the integration of game theory with findings from joint action research representing the interactive and embedded nature of cooperation. Further progress in this area of research, we hold, will be made by employing stag hunt paradigms to link decision-making with other socio-cognitive momentums, such as joint

[1]This at least holds for the child, whereas, the experimenter's motivation underlying his or her play might be due to his individual commitments to playing a larger paper-writing and publishing game, as kindly pointed out by one of the reviewers. It has to be added that this larger game can of course be played in a we-mode, as in the best case the authors constitute a team. Even competition in a defined game implicates acknowledging the constitutive rules defining that game (Tomasello, 2009; Searle, 1995) and hence implicates a minimum of cooperation.

attention, gaze communication, intrinsic motivation, and social commitments. Needless to say, this move does not imply ignoring the computational and mathematical advances in neuroeconomic cooperation research. On the contrary, we believe that a model-based approach which potentially allows for regional mapping of computational mechanisms, will play an important role in the development of this field. In this way, (social-) philosophical and economic proposals for taxonomically categorizing cooperative phonomena might be efficiently tested for neurobiological pertinence.

Taken together, employing more paradigms based on games and mutual benefits in neuroeconomics might help to link up with psychological research on cooperation, to go beyond mere decision-making aspects during cooperation and to promote computational modeling in the context of ecologically valid cooperation-scenarios. The field of neuroeconomics should integrate, and thus, directly profit from the rich scientific legacy of surrounding theoretical and experimental disciplines in order to most comprehensively capture human cooperation, and, finally, establish a truly social field of neuroeconomics.

REFERENCES

Adenzato, M., Becchio, C., Bertone, C., and Tuomela, R. (2005). "Neural correlates underlying action-intention and aim-intention," in *Poster Presented at the Cognitive Science 2005 Conference,* (Stresa, Italy).

Aschersleben, G., Hofer, T., and Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Dev. Sci.* 11, 862–668.

Bacharach, M. (1999). Interactive team reasoning: a contribution to the theory of co-operation. *Res. Econ.* 53, 117–147.

Bacharach, M., Gold, N., and Sugden, R. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory.* Princeton: Princeton University Press.

Barrett, L., and Henzi, P. (2005). The social nature of primate cognition. *Proc. Biol. Sci.* 272, 1865–1875.

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., and Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron* 64, 756–770.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., and Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* 58, 639–650.

Becchio, C., Adenzato, M., and Bara, B. G. (2006). How the brain understands intention: different neural circuits identify the componential features of motor and prior intentions. *Conscious. Cogn.* 15, 64–74.

Becchio, C., and Bertone, C. (2004). Wittgenstein running: neural mechanisms of collective intentionality and we-mode. *Conscious. Cogn.* 13, 123–133.

Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition* 106, 894–912.

Binmore, K. (1987). Modeling rational players, I. *Econ. Philos.* 3, 179–214.

Boesch, C. (1994). Cooperative hunting in wild chimpanzees. *Anim. Behav.* 48, 653–667.

Boesch, C., and Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Tai National Park. *Am. J. Phys. Anthropol.* 78, 547–573.

Boesch, C., and Tomasello, M. (1998). Chimpanzee and human cultures. *Curr. Anthropol.* 39, 591–614.

Bratman, M. E. (1992). Shared cooperative activity. *Philos. Rev.* 101, 327–341.

Braun, D. A., Ortega, P. A., and Wolpert, D. M. (2009). Nash equilibria in multi-agent motor interactions. *PLoS Comput. Biol.* 5:e1000468. doi: 10.1371/journal.pcbi.1000468

Burkart, J. M., Hrdy, S. B., and van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evol. Anthropol. Issues News Rev.* 18, 175–186.

Bzdok, D., Langner, R., Caspers, S., Kurth, F., Habel, U., Zilles, K., Laird, A., and Eickhoff, S. B. (2011a). ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Struct. Funct.* 215, 209–223.

Bzdok, D., Langner, R., Hoffstaedter, F., Turetsky, B. I., Zilles, K., and Eickhoff, S. B. (2011b). The modular neuroarchitecture of social judgments on faces. *Cereb. Cortex* 22, 951–961.

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T., Tomasello, M., and Collins, W. A. (2011). Early social cognition in three cultural contexts. *Monogr. Soc. Res. Child Dev.* 76, 1–42.

Camerer, C. F. (2008). "Behavioral game theory and the neural basis of strategic choice," in *Neuroeconomics. Decision-Making and the Brain,* eds P. W. Glimcher, E. Fehr, A. Rangel, C. Camerer, and R. A. Poldrak (London: Academic Press), 193–206.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., and More, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, 1–174.

Chang, L. J., Smith, A., Dufwenberg, M., and Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron* 70, 560–572.

Collins Cobuild. (1995). *Collins Cobuild English Dictionary.* London: HarperCollins.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behav. Brain Sci.* 26, 139–198.

Cosmides, L., and Tooby, J. (2005). "Neurocognitive adaptations designed for social exchange," in *The Handbook of Evolutionary Psychology*, ed D. M. Buss (Hoboken, NJ: Wiley), 584–627.

Csibra, G., and Gergely, G. (2009). Natural pedagogy. *Trends Cogn. Sci.* 13, 148–153.

Csibra, G., and Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 1149–1157.

de Bruijn, E. R. A., de Lange, F. P., von Cramon, D. Y., and Ullsperger, M. (2009). When errors are rewarding. *J. Neurosci.* 29, 12183.

de Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science* 305, 1254.

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., and Meltzoff, A. N. (2004). The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* 23, 744–751.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., and Liepelt, W. (2011). How "social" is the social simon effect? *Front. Psychol.* 2:84. doi: 10.3389/fpsyg.2011.00084

Dunbar, R. (1993). Coevolution of neocortex size, group size and language in humans. *Behav. Brain Sci.* 16, 681–735.

Dunbar, R. I. M., and Shultz, S. (2007). Evolution in the social brain. *Science* 317, 1344–1347.

Ekins, W. G., Caceda, R., Capra, C. M., and Berns, G. S. (2012). You can't gamble on others: dissociable systems for strategic uncertainty and risk in the brain. *Experimental Economics, Accounting, and Society: A Conference in Memory of John Dickhaut.*

Fehr, E., and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.

Fehr, E., and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114, 817.

Fiebich, A., and Gallagher, S. (2012). Joint attention in joint action. *Philos. Psychol.* (in press).

Frith, U., and Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 165–175.

Gilbert, M. (1990). Walking together: a paradigmatic social phenomenon. *Midwest Stud. Philos.* 15, 1–14.

Gove, P. B., and Merriam-Webster Inc. (2002). *Webster's Third New International Dictionary of the English Language, Unabridged: A Merriam Webster*. Springfield, MA: Merriam-Webster.

Gräfenhain, M., Behne, T., Carpenter, M., and Tomasello, M. (2009). Young children's understanding of joint commitments. *Dev. Psychol.* 45, 1430–1443.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388.

Hommel, B., Colzato, L. S., and van den Wildenberg, W. P. M. (2009). How social are task representations? *Psychol. Sci.* 20, 794–798.

Hamann, K., Warneken, F., Greenberg, J. R., and Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature* 476, 328–331.

Hare, B., Call, J., and Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Anim. Behav.* 61, 139–151.

Hrdy, S. B. (2009). *Mothers and Others: The Evolutionary Origins of Mutual Understanding.* Cambridge, MA: Belknap Press of Harvard University Press.

Humphrey, N. K. (1976). "The social function of intellect," in *Growing Points in Ethology,* eds P. P. G. Bateson and R. A. Hilde (Cambridge, MA: Cambridge University Press), 303–317.

Iani, C., Anelli, F., Nicoletti, R., Arcuri, L., and Rubichi, S. (2011). The role of group membership on the modulation of joint action. *Exp. Brain Res.* 211, 1–7.

Keysers, C., and Perrett, D. I. (2004). Demystifying social cognition: a hebbian perspective. *Trends Cogn. Sci.* 8, 501–507.

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., and Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science* 321, 806–810.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Kishida, K. T., King-Casas, B., and Montague, P. R. (2010). Neuroeconomic approaches to mental disorders. *Neuron* 67, 543–554.

Knoblich, G., and Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1006–1016.

Knoblich, G., and Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 2021–2031.

Knoch, D., Gianotti, L. R. R., Baumgartner, T., and Fehr, E. (2010). A neural marker of costly punishment behavior. *Psychol. Sci.* 21, 337–342.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., and Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.

Koban, L., Pourtois, G., Vocat, R., and Vuilleumier, P. (2010). When your errors make me lose or win: event-related potentials to observed errors of cooperators and competitors. *Soc. Neurosci.* 5, 360–374.

Konvalinka, I., Vuust, P., Roepstorff, A., and Frith, C. D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Q. J. Exp. Psychol.* 63, 2220–2230.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. *Nature* 435, 673–676.

Myerson, R. B. (1997). *Game Theory. Analysis of Conflict.* Cambridge, MA: Harvard University Press.

Newman-Norlund, R. D., Bosga, J., Meulenbroek, R. G. J., and Bekkering, H. (2008). Anatomical substrates of cooperative joint-action in a continuous motor task: virtual lifting and balancing. *Neuroimage* 41, 169–177.

Newman-Norlund, R. D., Ganesh, S., van Schie, H. T., de Bruijn, E. R. A., and Bekkering, H. (2009). Self-identification and empathy modulate error-related brain activity during the observation of penalty shots between friend and foe. *Soc. Cogn. Affect. Neurosci.* 4, 10–22.

Oberman, L. M., Pineda, J. A., and Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Soc. Cogn. Affect. Neurosci.* 2, 62–66.

Pacherie, E. (2011). Framing joint action. *Rev. Philos. Psychol.* 2, 173–192.

Pfeiffer, U. J., Timmermans, B., Bente, G., Vogeley, K., and Schilbach, L. (2011). A non-verbal turing test: differentiating mind from machine in gaze-based social interaction. *PLoS ONE* 6:e27591. doi: 10.1371/journal.pone.0027591

Radke, S., de Lange, F. P., Ullsperger, M., and de Bruijn, E. R. A. (2011). Mistakes that affect others: an fMRI study on processing of own errors in a social context. *Exp. Brain Res.* 211, 1–9.

Rakoczy, H., Warneken, F., and Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Dev. Psychol.* 44, 875–881.

Rekers, Y., Haun, D., and Tomasello, M. (2011). Children, but not chimpanzees, prefer to collaborate. *Curr. Biol.* 21, 1756–1758.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., and Schmidt, R. C. (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Hum. Mov. Sci.* 26, 867–891.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Rilling, J. K., Glenn, A. L., Jairam, M. R., Pagnoni, G., Goldsmith, D. R., Elfenbein, H. A., and Lilienfeld, S. O. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* 61, 1260–1271.

Ruys, K. I., and Aarts, H. (2010). When competition merges people's behavior: interdependency activates shared action representations. *J. Exp. Soc. Psychol.* 46, 1130–1133.

Sally, D. (1995). Conversation and cooperation in social dilemmas. *Ration. Soc.* 7, 58–92.

Sally, D., and Hill, E. (2006). The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27, 73–97.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Schilbach, L., Eickhoff, S. B., Cieslik, E., Shah, N. J., Fink, G. R., and Vogeley, K. (2010). Eyes on me: an fMRI study of the effects of social gaze on action control. *Soc. Cogn. Affect. Neurosci.* 6, 393–403.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R., and Vogeley, K. (2009). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 3, 37–50.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (2012). Toward a second-person neuroscience. *Behav. Brain Sci.* (in press).

Schonberg, T., Fox, C. R., and Poldrack, R. A. (2011). Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends Cogn. Sci.* 15, 11–19.

Schulte-Rüther, M., Markowitsch, H. J., Fink, G. R., and Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. *J. Cogn. Neurosci.* 19, 1354–1372.

Searle, J. (1995). *The Construction of Social Reality.* New York, NY: The Free Press.

Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition* 88, B11–B21.

Sebanz, N., Knoblich, G., and Prinz, W. (2005). How two share a task: corepresenting stimulus-response mappings. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 1234–1246.

Sebanz, N., Knoblich, G., Prinz, W., and Wascher, E. (2006). Twin peaks: an ERP study of action planning and control in co-acting individuals. *J. Cogn. Neurosci.* 18, 859–870.

Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., and Frith, C. D. (2007). Is it really my turn? an event-related fMRI study of task sharing. *Soc. Neurosci.* 2, 81–95.

Simon, J. R. (1969). Reactions toward the source of stimulation. *J. Exp. Psychol.* 81, 174–176.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure.* Cambridge, MA: Cambridge University Press.

Sober, E., and Wilson, D. S. (1999). *Unto others: The Evolution and Psychology of Unselfish Behavior.* Cambridge, MA: Harvard University Press.

Suzuki, S., Niki, K., Fujisaki, S., and Akiyama, E. (2011). Neural basis of conditional cooperation. *Soc. Cogn. Affect. Neurosci.* 6, 338–347.

Tabibnia, G., Satpute, A. B., and Lieberman, M. D. (2008). The sunny side of fairness - preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychol. Sci.* 19, 339–347.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press

Tomasello, M. (2009). *Why We Cooperate.* Cambridge, MA: The MIT Press.

Tomasello, M., and Carpenter, M. (2007). Shared intentionality. *Dev. Sci.* 10, 121–125.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691.

Tuomela, R. (2000). *Cooperation: A Philosophical Study.* Springer: Netherlands.

Tuomela, R. (2007). *The Philosophy of Sociality.* Oxford, UK: Oxford University Press.

van den Bos, W., van dijk, E., Westenberg, M., Rombouts, S. A., and Crone, E. A. (2009). What motivates repayment? neural correlates of reciprocity in the trust game. *Soc. Cogn. Affect. Neurosci.* 4, 294.

Vlainic, E., Liepelt, R., Colzato, L. S., Prinz, W., and Hommel, B. (2010). The virtual co-actor: the social simon effect does not rely on online feedback from the other. *Front. Psychol.* 1:208. doi: 10.3389/fpsyg.2010.00208

Wang, Y., Ramsey, R., and de C Hamilton, A. F. (2011). The control of mimicry by eye contact is mediated by medial prefrontal cortex. *J. Neurosci.* 31, 12001–12010.

Warneken, F., Chen, F., and Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Dev.* 77, 640–663.

Warneken, F., Hare, B., Melis, A. P., Hanus, D., and Tomasello, M. (2007). Spontaneous altruism by chimpanzees and young children. *PLoS Biol.* 5:e184. doi: 10.1371/journal.pbio.0050184

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., and Vogeley, K. (2010). Tools of the trade it' s in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* 5, 98–107.

Yoshida, W., Dolan, R. J., and Friston, K. J. (2008). Game theory of mind. *PLoS Comput. Biol.* 4:e1000254. doi: 10.1371/journal.pcbi.1000254

Yoshida, W., Dziobek, I., and Kliemann, D. (2010a). Cooperation and heterogeneity of the autistic mind. *J. Neurosci.* 30, 8815–8818.

Yoshida, W., Seymour, B., Friston, K. J., and Dolan, R. J. (2010b). Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* 30, 10744–10751.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 January 2012; accepted: 10 May 2012; published online: 01 June 2012.

*Citation: Engemann DA, Bzdok D, Eickhoff SB, Vogeley K and Schilbach L (2012) Games people play—toward an enactive view of cooperation in social neuroscience. Front. Hum. Neurosci. **6**:148. doi: 10.3389/fnhum.2012.00148*