

Team Reasoning: Controversies and Open Research Questions

Natalie Gold, King's College London

For *Routledge Handbook on Collective Intentionality*, eds. Kirk Ludwig & Marija Jankovic

1. Team reasoning

The theory of team reasoning was developed separately by Robert Sugden (1993, 2000, 2003) and Michael Bacharach (1999, 2006). Its development was motivated by games that are puzzling for orthodox game theory because they have an arguably rational solution, which a substantial number of people play in real life, whose play game theory cannot explain or predict.

One of these games is Hi-Lo, a version of which is shown in Figure 1. There are two pure-strategy Nash equilibria, *(high, high)* and *(low, low)*, and *(high, high)* is strictly better than *(low, low)* for both players. It seems clear that the two players should each play *high*. However standard game theory cannot recommend that. A Nash equilibrium involves a player maximizing her payoff given what the other player is doing. Thus it can only recommend that, if a player expects the other player to play *high*, then it is rational for her to play *high*. However, if she expects the other player to play *low*, then it is rational for her to play *low*. What it is rational for Player 1 to do is conditional on what Player 2 does, and standard game theory gives her no reason to expect Player 2 to play *high* rather than *low*. (For fuller statements of this argument, see Hodgson (1967), Sugden (1993), or Bacharach (2006).)

		Player 2	
		<i>high</i>	<i>low</i>
Player 1	<i>high</i>	2, 2	0, 0
	<i>low</i>	0, 0	1, 1

Figure 1: Hi-Lo

The other game is the Prisoner’s Dilemma, a version of which is shown in Figure 2. The only Nash equilibrium in this game is *(defect, defect)*; indeed *defect* is a dominant strategy, a player is better off playing it whatever the other player does, so standard game theory recommends to each player that she play it regardless of her expectations about the other player. However, *(defect, defect)* leaves both players worse off than if they had played *(cooperate, cooperate)*.

		Player 2	
		<i>cooperate</i>	<i>defect</i>
Player 1	<i>cooperate</i>	3, 3	0, 4
	<i>defect</i>	4, 0	1, 1

Figure 1: Hi-Lo

In standard game theoretic reasoning, an individual player asks “what do *I* want to achieve and what should *I* do to achieve it (given my beliefs about what other players will do)?” The answer is a complete strategy, roughly speaking a contingency plan, which she then carries out. Team reasoning extends the syntax of game theory, to allow players to ask “what do *we* want to achieve and what should *I* do to play my part in achieving it?”. This allows each player to choose the best profile of actions for the team and reach the conclusion that she should choose her component of that profile. Team reasoning de-conditionalizes players’ choices, allowing profile selection.

A player team reasons when she *group identifies*, conceiving of herself as part of a team, where the group is a unit of agency, acting as a single entity in pursuit of some single objective. Then, provided that there is common knowledge that each member group identifies and common knowledge of the best team profile, she will choose her part of the best profile for the group. (For discussion of what happens when we relax the common knowledge conditions and when the team of actors is smaller than the whole group, and how different theorists deal with it, see Gold &

Sugden, 2007b; Gold, 2012.) In the Hi-Lo game it is clear that (*high, high*) would be the best profile of actions for the team so, in the simplest case, if there is common knowledge that each member group identifies and common knowledge that each member aims to choose the best team profile, then each can reason that she should play *high*, as a part of playing (*high, high*). In a similar manner, if the best team profile in the prisoner's dilemma is (*cooperate, cooperate*) and the common knowledge conditions are fulfilled, then team reasoning can predict that players will choose *cooperate* (Gold & Sugden, 2007b).

This is the common core of the theory of team reasoning. It leaves open questions about what exactly it is to group identify, whether it can be rational to group identify, and what the team should take as its goal. These are answered in different ways by Sugden, Bacharach, and others who have written about team reasoning. Other open research questions include how to test the theory, the relation of team reasoning to collective intentions, what is the role of reasoning, and the application of team reasoning to other levels of agency. Each of these is the subject of a section below.

2. Group identification

An individual is said to *group identify* if she conceives of herself as part of a team, where the group is a unit of agency, acting as a single entity in pursuit of some single objective. (Strictly speaking, we should separate “entification”, where an individual conceives of a group as a unity of agency, or team, from “identification”, where she identifies herself as a part of that group agent. But we will simplify and just speak of identification here.) A team reasoner must first come to frame a decision as a problem for “us” and then do team reasoning. But theorists disagree about how group identification happens and what exactly it entails. In order to better understand these differences, I will locate them in a more general framework of decision-making.

We can distinguish three steps involved in framing and decision-making: noticing, noticing as choice-relevant, and deciding to act on a reason (Gold, 2012). Standard decision theory only

recognises individual agents, but the setting up of those agents' decision-problems still involves the three steps. There is an infinite number of ways of describing a situation. As finite individuals, we cannot see them all at once. The way that a situation is described or "framed" conveys information about the situation. For a fact to become the basis of an agent's reason for action (her "motivating reason"), first she must notice it. The agent can only reason from premises about the world that are accessible to her, so frames set the parameters of reasoning. But noticing alone is not enough. For the fact to influence an agent's decision, she must also recognise that it is relevant to her choice and is a potential reason for action. Finally, the agent must act on the reason, it must have a "motivational grip".

We can apply this framework to team reasoning. Standard decision theory only recognises individual agents. Once we allow that there can be different levels of agency, it matters whether the individual frames herself and her co-players only as individuals, using an "*I*-frame", or as a team, using a "*we*-frame". First the agent must see the possibility of cooperation and notice the potential for team reasoning, then she must group identify and see team reasoning as choice-relevant; third she must decide to act on team reasoning. So, once the agent notices the possibility of acting as a team, there are two more steps to becoming a team reasoner. We might think of the second of these steps, noticing as choice-relevant, as group identification.

There is disagreement about whether these steps are filled by psychological processes or choices. In Bacharach's theory of team reasoning, group identification is simply a psychological process. Salient features of the situation may promote group identification via *affiliation*, "a psychological process in which a person who does think about a certain group, defined by some shared property, comes to think about it as 'us'" (Bacharach, 1997: 2). Amongst those salient features are the payoffs of the game. A feature's *effectiveness* is its tendency to stimulate group identity. Both Hi-Lo and the Prisoner's Dilemma have a feature that Bacharach called *strong independence*, roughly speaking a Nash equilibrium that is worse than some other outcome in the game from every player's individual point of view. Bacharach speculated that seeing the possibility

of cooperating to achieve (*high, high*) instead of (*low, low*), or (*cooperate, cooperate*) instead of (*defect, defect*), would promote group identification. However, in the Prisoner's Dilemma, the player may also see the *double-crossing feature*, whereby an individual can benefit if she unilaterally deviates from the team reasoning solution, which inhibits group identification.

For Bacharach, group identification then primes team reasoning: another psychological process. In most of *Beyond Individual Choice*, Bacharach assumes that once an individual group identifies, then she will always team reason (although he recognises in a footnote that this is a simplification, Bacharach, 2006: 152). Individuals either frame the decision as a problem for "me", in which case they use individual reasoning, or they frame it as a problem for "us" and use team reasoning. (Hence, Bacharach's theory has been referred to as producing team reasoning as a result of framing, see Gold and Sudden, 2007a, 2007b.) In other work, he allows that both the *I*-frame and the *we*-frame come to mind at the same time (Bacharach, 1997). Never-the-less, reasoning about whether or not to team reason does not enter the picture.

For Sugden, in contrast, both the step from noticing to group identification and the step from group identification to team reasoning are opportunities for choices. *Assurance* is necessary for group identification: "to construe oneself as a member of a team, one must have some confidence that the other members of that team construe themselves as members too" (Sugden 2000: 194). Such assurance could be created by public acts of commitment or induced by repeated experience of regularities of behaviour in a population. But, even once an individual group identifies, she would still have to decide whether team reasoning was a mode of reasoning that she wanted to endorse. Assurance enters the picture again, as endorsing Sugden's "mutually assured team reasoning" is to make a unilateral commitment to a certain form of practical reasoning, but that reasoning does not generate any implications for action unless one has assurance that others have made the same commitment.

This difference between Sugden and Bacharach, between their views of how group identification occurs, also manifests in different ways of conceptualising the relationship between a

team reasoner and the group goal. For Bacharach, a team reasoner who identifies with the group takes on the goals of the group and will aim to do her part in what is best for the group, even when not all of the group are team reasoners. We might say in this situation that the “team” is smaller than the “group” whose interests are promoted by the team. In contrast, for Sugden, a team reasoner who identifies with a group stands ready to do her part in joint actions in pursuit of the group’s objective; but she does not necessarily take this objective as *hers* in the stronger sense of wanting to pursue it even if other members of the group do not reciprocate.

3. The possibility of rational group identification

The theory of team reasoning allows that groups, as well as individuals, can be agents; and it shows how it can be rational to play *high* in Hi-Lo and *cooperate* in the Prisoner’s Dilemma, given that a player group identifies. According to Hurley (1989:145) “an adequate theory of rational choice should address the question of what the unit of agency among those possible should be”. But neither Bacharach nor Sugden thinks that we can make an instrumentally rational choice to team reason.

For Bacharach, group identification is a psychological process that leads to team reasoning. Further, the evaluation of outcomes is always done by an agent within the perspective of one of the two frames, the *I*-frame or the *we*-frame (Bacharach, 1997). The standards of evaluation and goals are co-determined with the unit of agency; there is no agent-neutral perspective from which we can assess the rationality of team reasoning. All Bacharach can say is: given that an individual group identifies, then it is rational for her to team reason.

Even though, for Sugden, team reasoning is a matter of choice, he also does not think that it is a matter of instrumentally rational choice. Sugden does not acknowledge an agent-neutral concept of “rationality”, or even of the “validity” of reasoning. For Sugden, both individual reasoning and team reasoning are simply modes of reasoning that a person might endorse. As with Bacharach, all

goals are the goals of agents, so there is no instrumentally rational way to make assessments prior to determining the unit of agency.

Hurley differs, suggesting that a person should choose the unit of agency that best realises her goals, where these are chosen “according to one’s substantive goals and ethical views” (1989: 147): “As an individual I can recognise that a collective unit of which I am merely a part can bring about outcomes that I prefer to any that I could bring about by acting as an individual unit” (Hurley, 2005a: 203). Thus she makes team reasoning a matter of instrumental rationality.

There are two ways that Hurley’s position could be construed and they are both problematic. Hurley might mean that we should privilege individual level “substantive goals”; or she might mean that morality specifies agent-neutral goals, which ought to be pursued regardless of the unit of agency, but, since thinking usually occurs at the individual level, these will be determined by the individual’s “ethical views”. However, once we acknowledge that there are multiple levels of agency, it is not clear why one should privilege the individual level over the group level for the purpose of making evaluations (Bardsley, 2001).

Elizabeth Anderson and David Gauthier have also expressed the opinion that group identification (Anderson) or cooperation (Gauthier) is rational, but both admit to lacking an argument for this position.

Anderson maintains that morality requires us to transcend our various identities and harmonise their demands, by identifying with a community that comprehends them all, the Kantian Kingdom of Ends. However, she admits to having no argument that this is rationally required (Anderson, 2001: 37).

Gauthier also argues that something like team reasoning could be instrumentally rational. Gauthier has long held that it can be instrumentally rational to cooperate in the Prisoner's Dilemma game (Gauthier, 1986). In a recent re-working of his theory, Gauthier (2013) claims that *Pareto-optimization* is a necessary condition for rationality in multi-player games, where a Pareto-optimizing theory “provides only a single set of directives to all the interacting agents, with the

directive to each premised on the acceptance by the others of the directives to them” (Gauthier 2013: 607). The outcome selected must be both efficient and fair in how it distributes the expected gains of cooperation. Although he does not explicitly use the term “team reasoning”, it is clear that Gauthier's theory is similar to ideas of team reasoning for mutual gain. His goal is to show that “social morality is part of rational choice, or at least, integral to rational cooperation” (Gauthier, 2013: 624). However, he does not provide a connection to practical rationality—his moral justification for team reasoning is that it would pass a contractarian test whereby it is “eligible for inclusion in an actual society that constitutes a cooperative venture for mutual fulfilment” (Gauthier, 2013: 618)—and he concedes that he has not yet been successful in bridging social morality and rational choice.

If team reasoning is rational, it looks like it is not instrumentally rational. Anderson (2001) expresses the hope that we can discover principles of rational self-identification. If we do, for the reasons given above, in the discussion of Hurley, then they will have to be thicker principles of rationality than instrumental ones.

4. The team’s objective

In the Hi-Lo and Prisoner’s Dilemma games above, it is fairly obvious which outcome a team should aim at. In other situations, things may not be so clear. In particular, there may be a question of how the team should treat its members’ different interests and priorities. There are two main approaches to specifying the objectives of the team: to assume that a team, like an individual, has a utility function and to specify how individual team members’ interests should be traded off against each other; or to treat team reasoning as a form of mutual cooperation or bargaining and to describe what team outcomes the individual members would agree to.

Bacharach (2006) sees the team goal as being captured in a team payoff function and asks how the team function should relate to the payoff functions of the individual members (Ch. 2, Section 4.3). He argues that, at a minimum, the team function will be Paretian with respect to the

payoffs of the players, i.e., if every individual agent gets at least as much utility in outcome x than outcome y , and at least one agent does strictly better, then the group function will rank outcome x above outcome y . This will be enough to identify a preferred team outcome in coordination games, where players' interests are aligned, but in mixed-motive games, where there is some conflict of interest—maybe because there are multiple possible cooperative solutions and different individuals do better in each one—we will need to supplement it with an account of how the team function adjudicates between the interests of different members. Bacharach (2006: 88) puts forwards as a “testable hypothesis” that “in circumstances in which nothing is perceived by individual members about other individual members beyond the facts recorded in a bare game representation, principles of fairness such as those of Nash's axiomatic bargaining theory will be embedded in U [the team function].”

One function that is Paretian is the utilitarian function. Many papers on team reasoning talk about maximizing the average of the players' payoffs, which is equivalent to the utilitarian function (Bacharach 1999, 2006; Colman et al., 2008, 2014; Smerilli, 2012). Allowing this sort of trade-off between individual members' interests enables team reasoning to explain cases where one member sacrifices themselves for the rest of the team, for instance the martyrdom of one member for the good of the group. Maybe unsurprisingly, discussions of team reasoning that favour something like the utilitarian function are either psychological theories, such as Bacharach (2006), or moral ones, such as Regan (1980), who assumes that we should aim for a utilitarian outcome and asks what level of reasoning utilitarians should use.

An alternative way of conceptualising team reasoning is as cooperation for mutual advantage (Sugden, 2011). In this interpretation of the theory, an individual would not team reason if it would leave her individually worse off. Sugden (2015) proposes that we measure mutual advantage relative to the payoff that each player can guarantee herself in a game. A strategy profile where each player receives a payoff that is greater than this threshold and where each player's participation is needed to attain these higher payoffs is said to be mutually beneficial. (There are

similarities to Gauthier, 2013. See Karpus and Gold, 2016 for more discussion.) Karpus and Radzvilas (2015) formalise these ideas of mutual advantage and mutual benefit, together with a Pareto efficiency criterion. Misyak and co-authors argue in a similar vein that we-thinking should be construed as *virtual bargaining*, where people reason about what they would agree on, if they were able to bargain. If it is obvious]’ which outcome they would agree on, then they can choose it directly without doing any actual bargaining (Misyak & Chater, 2014; Misyak, Melkonyan, Zeitoun, & Chater, 2014; Chater, Misyak, Melkonyan & Zeitoun, 2016).

The utilitarian function allows the sacrifice of some members for the good of others: as John Rawls (1999:164) put it, utilitarianism neglects the separateness of persons. In contrast, theories of team reasoning as mutual advantage will not allow this. They privilege the individual level of agency, in that the acceptable team outcomes are defined in terms of what would be acceptable to the individual members.

5. Collective intentions and reasoning

The intentions that we have when we act together are collective intentions. If we decide to go to New York together, that you will buy the tickets and I will book the hotel, then there is a sense in which “we intend to go to New York”, as well as a sense in which “I” intend my part in our joint action. Most analyses of collective intentions focus on the properties of collective intentions as mental states, so anything that is distinctive to cooperative activity has to be represented as a distinctive feature of the corresponding intentions. However, Gold and Sugden (2007a) argue that collective intentions are distinguished from individual intentions by reference to the unit of agency in the reasoning that led to the formation of the intention; that collective intentions are the sort of intentions that are formed when people team reason.¹

¹ The idea that collective intentions result from team reasoning may be related to Tuomela’s (2006) idea of reasoning in the *we-mode*, where agents must function as group members and intend for a group reason (see also Hakli, Miller, & Tuomela, 2010).

Pacherie (2011) argues that leading accounts of collective intentions, such as those of Bratman (2009a, 2009b) and Gilbert (1997, 2009), require too much sophistication on the part of agents and therefore cannot explain the intentions behind the cooperative actions of children, who do not have a sophisticated theory of the mental states of others. She argues that Bacharach's theory of team reasoning gives an account of collective intentions that is less cognitively demanding and therefore to be preferred.

The team reasoning account of collective intentions is consistent with evidence from cognitive linguistics, which shows that *we* is composed of notions of *I* and grouphood (Gold & Harbour, 2012). These cognitive requirements would be satisfied by someone who has the individual concept *I* and the notion of grouped implicit in team reasoning.

The idea that collective intentions are the result of team reasoning has been criticised on the grounds that there are spontaneous collective intentions (Tuomela, 2009; Schweikard & Schmid, 2013). There are a variety of things that could be meant by this criticism. One of them is that there are spontaneous collective intentions, which are not formed as the result of conscious reasoning. This raises the question of what is meant by "reasoning" in team reasoning.

Philosophers often take reasoning to mean conscious deliberation (e.g. Pettit, 2007). However, this is not normally how decision theorists think of reasoning. Amongst economists, patterns of consistent behaviour provide evidence of rationality and of the agent's mental states which caused the behaviour; or, for more philosophical theories of rational choice, which relate it to practical reasoning, given people's beliefs and desires one can identify what behaviour would be rational. In both cases the "reasoning" that is imputed to agents is a rational reconstruction based on their mental states and the standard of rationality is an "external" one applied to behaviour, not an "internal" one applied to decision processes.

This is compatible with the way that "reasoning" is used in cognitive science and AI. Cognitive science has abandoned classical logic, as logical inference mechanisms are too slow to model the "automatic" information processing that is antecedent to decision. While manipulation of

propositions may be a suitable model at Marr's (1982) *computational* level, where the goal of the system and the logic behind its output are specified, it is a less good model at the *algorithmic* level, representing and implementing the computational theory, and at the level of *implementation*, or how the algorithm is realised in the brain. Instead we have the idea that people can implicitly reason according to some rule without explicitly being able to articulate the content of their reasoning (Polanyi, 1962),

In the literature on team reasoning, Hurley (2005b) has explicitly taken this line. She investigates local procedures and heuristics from which collective units of agency can emerge, and she argues that the processes in an agent that actually generate his or her rational behaviour need not be isomorphic with the theoretical account of why the behaviour counts as rational. Bacharach also denies conscious access to reasoning processes, saying, "We should not expect people to be able to identify the reasoning principles that govern their conclusions even when these principles are sound" (Bacharach 2006:45). Team reasoning models the computational level: the goal of the system and the logic behind the output. Therefore the claim that team reasoning leads to collective intentions should be understood as the idea that the team goal and logic of team reasoning underpin behaviour resulting from collective intentions.

6. Testing team reasoning

Tests of team reasoning have been tests of behaviour (as would be expected from the discussion in Section 5), testing hypotheses about how team reasoners will behave in various games, usually pitted against a leading alternative theory. In order to predict behaviour, some assumptions must be made about team goals and group identification, as discussed in Sections 2 and 4. It must also be possible to make predictions from the financial payoffs that are observable in the lab.

We would expect that team reasoners would be more likely to cooperate in mixed-motive games, like the Prisoner's Dilemma, than the individual reasoners of rational choice theory.

However, it is difficult to isolate the effect of team reasoning in such games. The predictable effects of group identification include increased inter-personal affect amongst group members and an increased expectation of cooperation by other group members (both of which might also “transform” the financial payoffs, so that the effective payoffs that motivate a player’s behaviour are different from the observed financial outcomes). Each of these, altruism and expectations, is the basis of alternative theories of cooperation. It’s not clear how we can measure which proposed correlates of group identification—team reasoning, in-group altruism, or expectation that others will cooperate—is causal.²

Luckily, co-ordination may be more fundamental than cooperation. Theories of cooperation in the Prisoner’s Dilemma that invoke payoff transformations or expectations and social norms turn *(cooperate, cooperate)* into an equilibrium but do not change the status of *(defect, defect)*, which remains an equilibrium. (For more detailed explanation of this, see Gold and Sugden, 2007b or Gold, 2012.) In other words, coordination is still necessary; often the game becomes a Hi-Lo. Co-ordination games are also attractive for other reasons. There is no problem with identifying the observed laboratory payoffs with the utility payoffs that motivate people because the players’ payoffs are equal in all outcomes, so it is safe to assume that all players have the same preference ordering over the outcomes. This also means that the team’s objective is not controversial in co-ordination games.

However, we’re remarkably good at co-ordination, especially Hi-Lo!³ If the hypothesis is that team reasoning causes co-ordination on *(high, high)*, then the experimenter needs to get a low baseline level of coordination on that outcome in order to see an effect of team reasoning. Hence, in order to test team reasoning, people have tried to make coordination harder. They have done this in three different ways.

² Colman et al. (2008, 2014) use prisoner’s dilemma type games to pit team reasoning against individual reasoning. As well as the confounds identified above, Sugden (2008) has pointed out that subjects’ behaviour in the experiments would also be predicted by another theory of coordination, cognitive hierarchy theory. See Karpus and Gold (2016) for more discussion.

³ See tasks NA6 and NA8 in Bardsley et al. (2010), where 54 out of 56 subjects (96 per cent) chose *Hi*.

Bacharach and Guerra (reported in Ch.4 of Bacharach, 2006) made coordination harder by making one of the low paying options salient. Bacharach took the main competitor theory to be the idea of Harsanyi and Selten (1988) that people play *high* because the high payoff makes it salient, call this *payoff salience*. In that case, it should be possible to induce people to choose *low* by labelling *low* in a manner that makes it more salient. Call this *label salience*. For instance, one could give the two strategies flower names, calling the low paying strategy “rose” and informing people that in previous experiments, “rose” was the most common choice. If that produced a significant amount of *low* choice, then it would be possible to test whether group identity manipulations make *high* choice more likely. Using a design like this, Bacharach and Guerra ran a small pilot that provides modest support for team reasoning.

A second way to make coordination harder is to increase the number of options, as done by Bardsley et al. (2010) and Bardsley and Ule (2014). A third possibility is to make the payoffs asymmetric, as done in two experiments by Faillo, Smerilli, Sugden (2013, 2015). Both sets of experimenters pitted team reasoning against *cognitive hierarchy theory*, where players are individualistic reasoners and where coordination is supposed to be the result of beliefs about the other players. The theory assumes that there is some very basic strategy, such as choosing at random or choosing according to label salience or choosing according to payoff salience—which one depends on the version—that is done by level-zero players. Then level-one players assume all other players are level-zero and play a best response against that strategy; level-two players play a best response to level-one players; and in theory so on ad infinitum, although in practise no-one has found anything higher than level-two reasoners. These four experiments have mixed results, which are hard to explain within a unified theory. For more detailed examination and discussion see Karpus and Gold (2016). But it is hard to conclude anything on the basis of so few experiments. More research is needed.

6. Other levels of agency: the person over time

In decision theory, problems of intertemporal choice are often analyzed as if, at each time t at which the person has to make a decision, that decision is made by a distinct transient agent or *timeslice*, the person at time t . Each timeslice is treated as an independent rational decision-maker, so that “the individual over time is an infinity of individuals” (Strotz, 1955: 179). This does not imply any metaphysical commitments, in particular it is not an endorsement of perdurantism, the view that things really do consist of temporal parts. Rather, it is a natural way of modelling people because the self at a particular time is the locus of choices, experiences, and perceptions.

This generates a neat model of giving in to temptations and procrastination (O’Donoghue and Rabin, 1999): A timeslice makes a choice that benefits itself (going out, eating dessert) but imposes costs on future selves (failing the exam, not getting into that dress); even if the timeslice cares about future selves, that may not be enough to get her to make a sacrifice for them. This way of thinking about the self is not alien to philosophers. It is the view put forward in Parfit (1984), who argues that it is only rational for the current self to take into account her future outcomes to the extent that she is “connected” to the future self.

The decision theoretic picture does not include intentions (and, in this, it differs from Parfit, 1984). Agency is vested in timeslices, who each act on their own preferences, and it is considered naive to think that an earlier timeslice could motivate her later self to do something just by forming a plan. And even agents who are sophisticated enough to recognise their own self-control problems can end up procrastinating (O’Donoghue and Rabin, 1999).

These self-control problems have a similar structure to the puzzles that motivated the development of interpersonal team reasoning. So we can give a similar solution: intra-personal team reasoning, or the idea that the timeslices can see themselves as constituting a “team over time” and act on team reasoning (Gold, 2013). Being a part of a team over time can give an earlier self a reason to take into account the outcomes of future selves even if she does not feel connected to her future self (Gold, 2015). The plan made by the earlier self is the intention and, when a timeslice

identifies with the team over time, she has a reason to carry out her part of the plan, or to act on her prior intention.

7. Conclusion

I have related the motivation for the theory of team reasoning and explained a basic version. I have detailed some controversies within the theory of team reasoning about group identification, the rationality of team reasoning, and the team's objective. I have sketched some open research questions and the attempts that people have made in answering them, covering collective intentions and reasoning, testing team reasoning, and the application of other levels of agency. Once we admit into the theory the possibility that individuals have more than two levels of agency, or are members of many different groups, then the question of which team someone identifies with and the interplay of different memberships becomes important. However, there is little or nothing on this question, which is an obvious direction for future research.

Related Topics

Collective Action and Agency, Non-reductive views of Shared Intention, Reductive views of Shared Intention, Interpersonal Obligations in Joint Action, Shared values, interests, and desires, Joint Commitment, Collective Rationality and Cooperation, Social Groups.

References

- Anderson, E. (2001) "Unstrapping the straitjacket of 'preference': a comment on Amartya Sen's contributions to philosophy and economics," *Economics and Philosophy* 17(01), 21-38.
- Bacharach, M. (1997) "'We' Equilibria: A Variable Frame Theory of Cooperation," Oxford: Institute of Economics and Statistics, University of Oxford, 30.

- Bacharach, M. (1999) "Interactive Team Reasoning: A Contribution to the Theory of Co-operation," *Research in Economics*, 53, 117-147.
- Bacharach, M. (2006) *Beyond Individual Choice: Teams and Frames in Game Theory*, Princeton: Princeton University Press.
- Bardsley, N. (2001) "Collective reasoning," *Critical Review of International Social and Political Philosophy*, 4 (4), 171-192.
- Bardsley, N., Mehta, J., Starmer, C. and Sugden, R. (2010) "Explaining Focal Points: Cognitive Hierarchy Theory versus Team Reasoning," *The Economic Journal*, 120, 40-79.
- Bardsley, N., & Ule, A. (2014) "Focal Points Revisited: Team Reasoning, the Principle of Insufficient Reason and Cognitive Hierarchy Theory," *Munich Personal RePEc Archive Paper No. 58256*.
- Bratman, M. (2009a) "Shared agency," in C. Mantzavinos (ed.) *Philosophy of the social sciences: philosophical theory and scientific practice*, Cambridge, UK; New York : Cambridge University Press, 41–59.
- Bratman, M. E. (2009b) "Modest sociality and the distinctiveness of intention," *Philosophical Studies*, 144(1), 149-165.
- Colman, A. M., Pulford, B. D. and Rose, J. (2008) "Collective Rationality in Interactive Decisions: Evidence for Team Reasoning," *Acta Psychologica*, 128, 387-397.
- Colman, A. M., Pulford, B. D. and Lawrence, C. L. (2014) "Explaining Strategic Cooperation: Cognitive Hierarchy Theory, Strong Stackelberg Reasoning, and Team Reasoning', *Decision*, 1, 35-58.
- Faillo, M., Smerilli, A. and Sugden, R. (2015) "Identifying the roles of level-k and team reasoning in coordination games," unpublished ms.

- Faillo, M., Smerilli, A. and Sugden, R. (2013) "The Roles of Level-k and Team Reasoning in Solving Coordination Games," Cognitive and Experimental Economics Laboratory Working Paper (No. 6-13), Department of Economics, University of Trento, Italy.
- Gauthier, D. (1986) *Morals by Agreement*, Oxford: Oxford University Press.
- Gauthier, D. (2013) "Twenty-Five On," *Ethics*, 123, 601-624.
- Gilbert, M. (2009) "Shared intention and personal intentions," *Philosophical studies*, 144(1), 167-187.
- Gilbert, M. (1997) "What is it for us to intend? In G. Holmström Hintikka and R. Tuomela (eds.). *Contemporary Action Theory. Vol. 2: Social Action*. Dordrecht: Kluwer, 65-85.
- Guarding against temptation: rational intentions and intra-personal team reasoning
- Gold, N. (2013) "Team Reasoning, Framing, and Self-control: An Aristotelian Account. In N. Levy (ed.) *Addiction and Self-Control: Perspectives from Philosophy, Psychology, and Neuroscience*. Oxford: Oxford University Press
- Gold, N. (2012) "Team Reasoning, Framing and Cooperation," in S. Okasha and K. Binmore (eds), *Evolution and Rationality: Decisions, Co-operation and Strategic Behaviour*, Cambridge: Cambridge University Press, chap. 9, 185-212.
- Gold, N. (2015) "Guarding against temptation: rational intentions and intra-personal team reasoning," Unpublished manuscript.
- Gold, N. and Sugden, R. (2007a) "Collective Intentions and Team Agency', *Journal of Philosophy*, 104, 109-137.
- Gold, N. and Harbour, D. (2012) "Cognitive Primitives of Collective Intentions: Linguistic Evidence of Our Mental Ontology," *Mind & Language*, 27(2), 109-134.
- Gold, N. and Sugden, R. (2007b) "Theories of Team Agency," in F. Peter and H. B. Schmid (eds), *Rationality and Commitment*, Oxford: Oxford University Press, 280-312.

- Hakli, R., Miller, K., and Tuomela, R. (2010) "Two kinds of we-reasoning," *Economics and philosophy*, 26(03), 291-320.
- Harsanyi, J., and Selten, R. (1988) *A general theory of equilibrium selection in games*, Cambridge Ma.: MIT Press.
- Hodgson, D. (1967) *Consequences of Utilitarianism*, Oxford: Clarendon Press.
- Hurley, S. L. (1989) *Natural reasons: Personality and polity*, Oxford: Oxford University Press.
- Hurley, S. (2005a) "Rational Agency, Cooperation and Mind-Reading," in N. Gold (ed.), *Teamwork: Multi-Disciplinary Perspectives*, Palgrave MacMillan.
- Hurley, S. (2005b) "Social Heuristics that Make Us Smarter," *Philosophical Psychology*, 18, 585-612.
- Karpus, J. and Gold, N. (2016) "Team Reasoning: Theory and Experiment," in J. Kiverstein (ed.) *Handbook of the Social Mind*. Routledge.
- Karpus, J. and Radzvilas, M. (2015) "Team Reasoning and a Rank-Based Function of Team's Interests," manuscript under review.
- Marr, D. (1982) *Vision: A computational approach*, W.H. Freeman and Company.
- Chater, N., Misyak, J. B, Melkonyan, T. & Zeitoun, H. (2016) "Virtual Bargaining: Building the foundations for a theory of social interaction," in J. Kiverstein (ed.) *Handbook of the Social Mind*. Routledge.
- Misyak, J. B., & Chater, N. (2014) "Virtual bargaining: a theory of social decision-making," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1655), 20130487.
- Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014) "Unwritten rules: virtual bargaining underpins social interaction, culture, and society," *Trends in cognitive sciences*, 18(10), 512-519.
- O'Donoghue, T., & Rabin, M. (1999) "Doing it now or later. *American Economic Review*, 103-124.
- Pacherie, E. (2011) "Framing joint action. *Review of Philosophy and Psychology*, 2(2), 173-192.

- Parfit, D. (1984) *Reasons and persons*, Oxford: Oxford University Press.
- Pettit, P. (2007) "Rationality, reasoning and group agency." *Dialectica*, 61(4), 495-519.
- Polanyi, M. (1962) "Tacit knowing: Its bearing on some problems of philosophy," *Reviews of Modern Physics*, 34(4), 601-615.
- Rawls, J. (1999) *A Theory of Justice*, rev. ed., Cambridge Ma.: Harvard University Press.
- Regan, D. (1980) *Utilitarianism and Co-operation*, Oxford: Clarendon Press.
- Schweikard, D. P. and Schmid, H. B., "Collective Intentionality," in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition), URL = <http://plato.stanford.edu/archives/sum2013/entries/collective-intentionality/>.
- Smerilli, A. (2012) "We-Thinking and Vacillation Between Frames: Filling a Gap in Bacharach's Theory," *Theory and Decision*, 73, 539-560.
- Strotz, R. H. (1955) "Myopia and inconsistency in dynamic utility maximization," *The Review of Economic Studies*, 165-180.
- Sugden, R. (1993) "Thinking as a Team: Towards an Explanation of Nonselfish Behavior," *Social Philosophy and Policy*, 10, 69-89.
- Sugden, R. (2000) "Team Preferences," *Economics and Philosophy*, 16, 175-204.
- Sugden, R. (2003) "The Logic of Team Reasoning," *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, 6, 165-181.
- Sugden, R. (2008) "Nash Equilibrium, Team Reasoning and Cognitive Hierarchy Theory," *Acta Psychologica*, 128, 402-404.
- Sugden, R. (2011) "Mutual Advantage, Conventions and Team Reasoning," *International Review of Economics*, 58, 9-20.

Sugden, R. (2015) "Team Reasoning and Intentional Cooperation for Mutual Benefit," *Journal of Social Ontology*, 1, 143-166.

Tuomela, Raimo. (2009) "Collective Intentions and Game Theory," *Journal of Philosophy*, 106, 292-300.

Tuomela, R. (2006) "Joint Intention, We-Mode and I-Mode," *Midwest Studies In Philosophy*, 30, 35-58.