

PUBLIC REASON, SOCIAL CHOICE, AND COOPERATION*

Bruce Chapman
Faculty of Law, University of Toronto
bruce.chapman@utoronto.ca

Economic theory and legal theory can both claim to provide plausible accounts of rational decision-making. Yet, despite the growth of “law and economics” as a hugely successful area of interdisciplinary study, there is very little intellectual exchange between the rational choice theorist who attempts to explain economic behaviour on the one hand, and the more philosophically inclined theorist who seeks to comprehend legal reasoning and adjudication on the other. Thus, the claim that each sort of theorist makes to account for rational decision-making seems largely to go unanswered by the other, this despite the fact that the two disciplines are otherwise so interconnected.

While the two sorts of theory loosely understand the rationality of a set of decisions in the same way, namely as an “ordered particularity”, the notion of ordering is fundamentally different between the two. In economics, no matter how diverse the motivations for choice might appear to be, the idea of an ordering remains somewhat single-minded and “quantitative”, the sort of thing over which a chooser can *maximize*. Thus, in social choice theory, for example, the idea persists that the plurality of orderings (of either individuals or choice criteria) that are the inputs into the social choice function are to be reduced to, or reconciled into, the linearity of an “all things considered” social ordering of the alternatives available for choice. But in law the notion of an ordering is less quantitative and more “categorical”, the sort of thing that informs an *understanding*. We make sense of the many particular decisions that we observe in the legal world by rendering them coherent with one another under the aspect of general concepts and rules of thought that are accessible to shared reasons and amenable to public justification. One of the challenges for those who do research on rational decision-making is to comprehend the intellectual difference between these two accounts of rationality in a way that makes each accessible to the other, that is, that puts them within some common theoretical framework

* This paper was originally prepared for the Eighth Conference on *Theoretical Aspects of Rationality and Knowledge* (TARK VIII), held at the Department of Economics, University of Siena, July 2001.

Shared Understandings and Public Reason

The following legal example is suggestive of how a common framework might be developed and why it might be useful. Suppose that a panel of three judges has to decide whether a defendant should be held liable to a plaintiff for breach of contract. Judge A believes that there was a properly formed contract in this case, but that the defendant has not breached it. Thus, he would find in favour of the defendant. Judge B thinks that the defendant's behaviour does amount to breach of such a contract, but believes that the contract was not properly formed in this case. Therefore, she too would find in favour of the defendant. Finally, Judge C believes both that there was a contract and that the defendant has breached it. Judge C, therefore, would find in favour of the plaintiff. Thus, a majority of the judges, A and B, share the view that the defendant should win in this case and, absent an obligation to provide reasons, would choose that as their preferred result. Table 1 summarizes the three different views of these three judges.

Table 1

	1. Was there a contract?	2. Was there conduct constituting breach?	3. Was there a breach of contract?
1. Judge A	Yes	No	No
2. Judge B	No	Yes	No
3 Judge C	Yes	Yes	Yes
4. Majority	Yes (2:1)	Yes (2:1)	No (2:1)

The interesting question is whether an obligation to provide publicly accessible and comprehensible reasons for their shared view in column 3, the sort of obligation that characterizes judicial decision-making in the common law, makes it any more difficult for Judges A and B to reach their preferred result. It seems that it might. After all, while Judges A and B might have a *shared preference* for a particular legal outcome (as indicated in column 3), it is not at all clear that they have a *shared understanding* of *what it is* they are doing to reach that outcome. There are two legal issues underlying this case,

the breach issue and the contract formation issue. These are the issues that make the case rationally comprehensible to us, a proper object of our legal understanding. Yet, on each of these salient legal issues in the case, the two judges who form the majority in favour of the defendant have completely opposed views (as indicated in columns 1 and 2).

It seems less obvious, therefore, that there is a majority agreement between these two judges on *any matter in law*. Certainly, it would be a challenge for this majority, despite their shared preference for a particular legal outcome favouring the defendant, to articulate any common or coherent legal view supporting that result. We might say that their shared sense of an appropriate *doing* is hampered somewhat by their inability to offer a publicly comprehensible shared *saying* for what they do (Chapman 1998a). Indeed, to the extent that there is any majority agreement on the salient legal issues in this case, it is that a majority of the court believes both that there *is* a contract *and* that it has been breached (as summarized by the last row of columns 1 and 2). That is where the majority's shared reasons are, and they are in tension with the majority preference for an outcome that denies the plaintiff her remedy (last row of column 3). Because this is a tension about what one can say in support of what one wants to do (or, alternatively, what one can do given what one must say), it is sometimes referred to as the *discursive dilemma* (Pettit 2000, 2001).

Public Reason and Social Choice

Of course, in many collective choice situations it is possible that the individual members of a decisive majority may not have any reasons in common for what they most want to do. Yet we do not feel that this is any way problematic for them. For example, a purchasing consortium may decide by majority vote against purchasing a particular white sports car, where some members of the decisive majority have voted this way because they do not like the idea that this car is white and others that it is a sports car. Table 2 summarizes this possibility in a way that appears to be fully analogous to the earlier breach of contract example.

Table 2

	1. Do you like that this car is a sports car?	2. Do you like that this car is white?	3. Do you like this white sports car?
1. Individual A	Yes	No	No
2. Individual B	No	Yes	No
3. Individual C	Yes	Yes	Yes
4. Majority	Yes (2:1)	Yes (2:1)	No (2:1)

The fact that the members A and B in the majority in column 3 do not have “reasons in common” to support their shared preference for not buying this particular car is not thought to present them with any real difficulty. Nor is it thought to be rationally compelling that this purchasing consortium buy a white sports car just because a majority of the consortium likes both that it is a sports car and that it is white.

However, there are important structural differences between the two sorts of examples presented in Tables 1 and 2. In Table 1, the column 3 proposition that calls for judgement just *is* the compound of the two logically prior atomic propositions represented by columns 1 and 2. A judge charged with the issue of whether there has been a breach of contract simply has to assess both whether there has been a contract and whether there has been conduct amounting to its breach. With these judgements in hand, the proposition that there has been a breach of contract follows immediately, in the manner of *modus ponens*, and an individual judge, or judges as a group, should accept the implication as a matter of deductive closure (List and Pettit 2000). That the judges do not accept the implication *as a group* in Table 1 is what makes their conduct appear (in row 4) to be *collectively* irrational.

But the decision to purchase a car, even (more particularly) a white sports car, is not essentially decomposable into two prior atomic questions, Is it a sports car? Is it white? That underlying structure, while possibly a helpful guide to the purchasing decision, is not an essential part of the problem in the same way. Rather, the purchasing consortium is out to purchase a car, perhaps even the *best* car that is available to it *all*

things considered. But that judgement is ultimately made *of the car* and *on the whole*. Thus, it is a judgement that could have the consortium appealing to *modus tollens* as much as to *modus ponens*, that is, to the possibility that the more holistic column 3 judgement should determine what the (group or individual) judgements should be in columns 1 and/or 2, not vice versa. (One can imagine someone saying to the group or individual, for example, “Well, if you’re sure you don’t like the car, then there must be *something* about it, perhaps its colour or type, which you don’t like.”)¹

However, despite this structural difference in the examples, there might be something useful, even for what the economist seeks to accomplish by way of social choice, or what a purchasing consortium seeks to achieve in the market for cars, in insisting on the greater rationality requirement that is inherent in the legal idea that members of a group can act *sensibly* together only if they can organize what they would prefer to do under a common understanding, that is, if they can act together under a common set of categories or concepts. That this is sometimes difficult to do, and that it sometimes frustrates the achievement of shared preferences, might be precisely what is so useful about it.

To illustrate this point, suppose that the three individuals in our Table 2 purchasing consortium, considering the joint purchase of a car, originally had preferences over three alternative cars as follows:

Table 3

Individual A	Individual B	Individual C
Black sports car (BS)	White family car (WF)	White sports car (WS)
White sports car (WS)	Black sports car (BS)	White family car (WF)
White family car (WF)	White sports car (WS)	Black sports car (BS)

¹ Even *modus tollens* may not carry the day. One can imagine a group being very sure of its holistic judgement in column 3, but being quite unsure amongst themselves on what the underlying reasons are for that judgement. In such a case there may be no reason to think that the column 3 judgement should force a change in the judgements appearing in columns 1 and/or 2; a third reason, as yet still unidentified, might be the one that unites the group and brings collective rationality into row 4 of the Table. Some developing areas of jurisprudence, like discrimination law, might exemplify this; however, it seems unlikely that a more established area of law, like the law on breach of contract, is subject to the same sort of indeterminacy as to its underlying issues.

This is, of course, the preference profile that makes for the familiar majority voting paradox. A majority prefers BS to WS, WS to WF, and WF to BS. Thus, within the social choice framework, there is the danger here of a kind of *excess* of rational doing: for every alternative that one is tempted to choose, there is another that a majority would prefer to have instead. It is this excess of rational doing that gives rise to cycling and instability.

Now it is common for economists to point out that the problem here is that individual preferences are not “single peaked”, that is, that there is no general agreement (1) that all the alternatives are to be assessed according to some single decisive dimension and (2) that one of the alternatives is of intermediate value on that decisive dimension. If only that were so, the argument goes, then that alternative would never be a worst alternative for anyone and the majority voting paradox would be avoided.

This is, in effect, to insist that individuals organize their preferences in a single-minded way along one decisive dimension and to allow them only the limited scope of ordering the social alternatives according to how those alternatives vary *quantitatively* (more or less) along that decisive dimension. But, as the example suggests, and as multi-dimensional models show more generally (McKelvey 1976), individuals react, reasonably, to a broad range of *categorically* different dimensions or aspects of the social alternatives on offer. And so the question arises whether these different and plural dimensions of a social choice problem can be rationally organized in some way so that instability can be avoided.

The car example is suggestive. The majority coalition of A and B can *say* together (in support of what they might *do* together) “Given a sports car, we would prefer it to be black.” Likewise, the majority coalition of A and C might be able to say, “Given that it’s white, we prefer to have a sports car.” In this respect, these coalitions can make use of what are sometimes referred to as *generic* preferences (Doyle and Thomason 1999). But what would the majority coalition B and C say together? In some sense, of course, they have a shared preference over a pair of particular alternatives just like the other majority coalitions do. Indeed, as already intimated, that is what gives rise to the instability. But their shared preference for WF over BS lacks any of the generic structure that characterizes the shared preferences of the other two majority coalitions. Thus, it is harder for them to articulate their shared preference in any sort of categorical way, that is,

in a way that makes use of the generic preferences that are in play in the choice problem. In this respect they are rendered “speechless”, just like Judges A and B were in the Table 1 breach of contract example. But now we can see that there may be some stabilizing effect in using the discipline of public categorical reason to restrict the formation of this majority group. After all, without this additional discipline and structure, there is only a senseless (i.e., non-categorical, non-conceptual) aggregation of (merely particular) preferences and the cycling problem that this permits.

The discipline that is provided by public (categorical) reason can be related more generally to a particular form of “value restriction” (specifically, “not-between value restriction”) that Sen (1966) has shown is sufficient for avoiding the majority voting paradox. (Specifically, if all individuals agree that in any triple a given alternative is “not between” the other two, that is, is either best or worst of the three, then the majority voting paradox cannot occur.) For instance, in the car example, it is easy to see that an alternative way to express what A and B have in common is their view that WF is a “not-between” alternative for them, viz., the real issue between them is whether or not the purchased car should be a sports car (BS, WS) or not (WF). Individual A puts the pair (BS, WS) first and the single alternative (WF) last, whereas individual B has the reverse ordering of these two partitions. This is something that they could have decided first, before they went on to decide, if necessary, what was a secondary issue to them, viz., what colour the sports car should be.² Likewise, what the coalition of A and C has in common might have been expressed as an agreement over BS as a “not between” alternative, the sort of agreement that asks each to decide first whether the car chosen should be black and, second, if not, whether it should be a sports car or family car. But, again, the pair of individuals B and C would have some difficulty articulating their common understanding of the relevant issues. They agree between them that WS is a “not-between” alternative, but what, exactly, is the category or concept that embraces the partition (WF, BS) of alternatives that is the complement to that not-between alternative?

² It could be, of course, that A feels there is a great deal more at stake in the choice of colour than the choice of car type, viz., that the preferential distance between BS and WS is large compared to the preferential distance between WS and WF. But this cannot be a view that he has *in common* with B. For B the preferential distance between BS and WS is *contained within* the distance between WF and WS. So the search for a *public* reason for choice, at least one linking A and B, cannot be found here. As the text following this note suggests, this suggested interpretation is better for the pair of voters A and C.

The problem, again, is that it is hard to “make sense” of such a partition of the alternatives in terms of the categories or concepts (colour and type of car) that are in play in the example. We might say, as we can for all the other possible pairs of individuals, that individuals B and C agree at the level of preferences, but that they do not share any sort of categorical agreement about the sorts of issues that inform their choice and the order in which these issues might be considered.

Now it might be objected that the imposition of a categorical discipline on preferences still leaves too much unresolved to be helpful. After all, even those two groups of voters, AB and AC, which (unlike group BC) agree that the salient issues are the type and colour of the car to be purchased, disagree fundamentally on the *order* in which these two issues should be addressed. For AB the most salient issue is whether the car should be a sports car or not; only after considering that would this group turn its attention, if necessary, to colour. But for the coalition AC the most important issue is colour, and only if white was chosen would the coalition turn its attention to the type of car. Moreover, the order in which the issues are considered is likely to affect the outcome; in this respect the matter is similar to the problem of path dependent choice. For example, if type of car is considered first, then it seems less likely that WS will end up being chosen. Individual A will choose in favour of sports cars, and individual B against sports cars, in the first round. Whether sports cars are chosen categorically in that round depends a good deal on how individual C, whose preferences are not categorical in this way, i.e., they do *not* satisfy not-between value restriction on alternative WF, actually votes. But, in the event of a first round vote for a sports car, it does seem likely that BS will defeat WS in the vote on the issue of colour. An analogous argument would suggest that WF is the less likely choice if the issue of colour is decided first.

However, in some contexts, there is good reason to think that this sort of path dependence will be less of a problem for categorically sensitive choice (Chapman 1998b). In other words, the categories or concepts that make sense of certain *partitions* of the alternatives for choice will often make sense of certain *paths* (or sequences of those partitions) as well, at least if we want to continue to make use of the stabilizing effects of not-between value restriction. To see this, consider the example of a criminal trial, where the two issues to be decided are the verdict and the sentence for the accused. Again, one

could imagine a panel of judges considering the three possible outcomes, innocent (I), guilty with a severe sentence (GS), and guilty with a lenient sentence (GL). Again, a natural partition of the alternative outcomes might be into the two issues, Verdict “(I or (GS, GL)?” and Sentence (“GS or GL?”), a partitioning that would “make sense” in a way that the alternative partitions, “(GS) or (I, GL)?” or “(GL) or (I, GS)?” would not. (What single concept or category sensibly comprehends the partition (I, GS), for example?) But, still, it seems that one could take these issues, and the partitions to which they lend sense, in order of either “Sentence first, verdict afterwards” or “Verdict first, sentence afterwards”. The law adopts the second of the two possibilities (and the Queen at Alice’s trial in Wonderland adopts the first), but is there any reason to do so? One answer, of course, is simple economy: why bother attending to the issue of sentencing until we know that the verdict decision makes it necessary? But the analysis provided here suggests a different answer. While both sequences respect the *partition* of the alternatives that makes most sense, only the *path* that has us consider the verdict first, or the partition “(I or (GS, GL))”, imposes any sort of not-between value restriction on the panel of judges. Under the verdict first sequence, each judge must order his or her preferences around the salient legal categories, deciding whether to put the alternative I either better or worse than (but not between) the alternatives GS or GL. The sentence first sequence, on the other hand, while paying a kind of lip service to the same set of issues, does not require the judge to order his or her preferences around those issues. For example, a judge who preferred the three alternatives in the order first GL, then I, and then GS, that is, someone who might be saying “Whether or not I would find him guilty of the offence depends on the sentence he would receive”, would have no difficulty voting these preferences under the sentence first procedure even though these preferences do not seem to show a categorical commitment to the issues that are salient in the case. The verdict first sequence, on the other hand, does force this judge to ask a more categorical sort of question about the verdict, that is, to show the same sort of commitment to the issues in the case as does the law she personifies.³ Furthermore, under

³ Requiring this sort of structure can, of course, tempt the judge to “nullify” a possible guilty verdict for fear of risking the worst (for her) possible sentencing outcome GS. Verdict nullification has attracted a good deal of critical comment, particularly in the United States where in jury trials there is the possibility of the death penalty. Juries are said to be charged with the responsibility of reaching a verdict within the law

a verdict first procedure, we not only make sense of the issues in the case, we also impose a domain restriction on the preferences that legal decision-makers can bring to bear on legal decisions so that certain problems of instability are avoided.

The thrust of the paper so far has been, first, to motivate the idea that public reason can restrain choice in a way that preference cannot, and, second, to argue that the categories of public reason can be structured around some well known results in the theory of social choice that relate to restrictions on the domain of individual preferences. While the latter results are not in any way new for rational social choice, it is novel to motivate those results in, and connect them to, the more philosophical (and legal) idea of public reason. I now want to suggest the possible relevance of public reason for cooperation in the theory of games.

Public Reason and Co-operation

The idea of categorical reason, initially developed here in the context of social choice, can also be related to some recent discussion of the “we frame” in the theory of games. Bacharach (1997, 1998) has suggested that the players in a prisoners’ dilemma might approach that game with “variable frames”, that is, with different (not simultaneously available) conceptual representations or understandings of their situation. Bacharach contemplates in particular the idea that there might be both an “I/he frame” and a “we frame”. These different frames, which pose the quite different questions “What should *I* do?” and “What should *we* do?” call for quite different notions of reasoning. Specifically, the I/he frame accommodates the idea, familiar in game theory, that a player should ask what strategy is best for herself *given* what the other player might do, and allows that player, under common knowledge of such reasoning, to replicate that same sort of thinking in the other player as well. The we frame, on the other hand, encourages the player to think about what *profile S* of strategies should be adopted by the players as a

as explained by the trial judge; it is the task of the judge to determine the sentence. For members of the jury to worry about the sentence rather than the verdict, particularly if they think the accused has committed the offence in question, is thought by some to violate the rule of law. Whatever the merits of verdict nullification by juries, our analysis here, based on the stabilizing impact of imposing the categorical constraints of not-between value restriction, offers an independent reason for supporting the verdict first procedure.

group, and then identifies the rational strategy for each player as the one that simply (categorically, non-contingently) has that player “doing her part” s_i within that overall profile.⁴ Unlike for the I/he frame, a player who is in the we frame does not have to consider whether the other players are themselves doing their parts as components of this profile of strategies in order to justify her strategy choice. Rather, in response to any question about *why* she was doing what she was doing, she would only say “This is simply what *we* do when we do *S* (as best)” or, perhaps, “This is simply what *I* do when *we* do *S* (as best)” or even, most provocatively (because most categorical in tone), “This is simply *what it is* for us, you and me, to do *S* (as best).”

Bacharach argues not only that these two different frames might be available to each player, but also that it might be common knowledge that they are so available. And therein lies the advantage of the we frame. Common knowledge of the we frame can make it rational for players to achieve certain shared goals, for example, the Pareto-superior outcome in a co-ordination game, goals which would not be rationally achievable if the players were restricted to (their common knowledge of) the more conventional I/he frame. Under the I/he frame it is only rational for the first player to do “her part” in an attractive co-ordination equilibrium *if* the other player is also doing his part. But that other player (as the first player well knows under the common knowledge assumption) can only rationally do his part *if* the first player does hers. Thus, under common knowledge of the I/he frame, each player is caught in a problem of self-reference without being able to determine a unique strategy. Hence we have the co-ordination problem despite, perhaps, the salience of a Pareto-superior outcome. The we frame, by contrast, when it arises as a possibility under common knowledge, does not allow an individual player even to *conceive* of the possibility that it might not be rational for her to do her part under the best possible profile of strategies because the other player might not be doing his (perhaps, self-referentially, because he is waiting for her to do hers, and so on). Such individualistic I/he thinking (“What if *he* isn’t? What then should *I* do?”) is simply not available under the we frame. Where it might be said that the I/he frame and, more specifically, the players common knowledge of it, incorporates a form of *intersubjective* rationality, and leaves open those characteristically recursive iterations of rational rethinking across the individual players, the we frame is more *objective*

⁴ Robert Sugden’s “team reasoning” has a comparable structure; see Sugden (1993), (2000).

and, accordingly, more restrictive of each player's thinking (and rethinking) about possible strategic choices. Under the we frame each player asks only what is best for the group *as a group*, and then goes on (again, unthinkingly, unreflectively, *uncontingently*, at least in the I/he sense) simply to do *what it is* (categorically) that the best group strategy *S* requires of her. She does all this, in other words, without any immediate (this qualification is important; see note 5) and paralysing reconsideration of whether the other players in the group are also doing their individual parts in the overall scheme. Bacharach has argued effectively that an objectively restricted we frame, available at least as a possibility for the players, can be important not only for achieving Pareto optimality in co-ordination games, but also (sometimes) for achieving joint co-operation in the prisoners' dilemma game as well.

Now the shared understanding that comes with an effective communication between interlocutors must surely presuppose at least a common (or overlapping) conceptual scheme, that is, one that is intersubjectively accessible. Without that there would be no real possibility of communicating anything at all. However, in the case of actual *communication* (not mere existence) of this shared understanding there is likely to be much more than this. . For in the very use of language to articulate the shared meaning, there is the necessity to order one's thoughts, and more specifically one's thoughts about one another's strategic choices (Bacharach's "unit of agency"), under the aspect of a shared language which itself must *transcend* the momentary interaction of the two interlocutors. After all, it is not open to the parties to invent a completely private and local meaning for their particular interaction. At some point the words they use to capture the shared significance of their interaction for each other ("Exactly *what* are *we* doing within this overall profile of strategy choices? Is it *S* or something else?") are words that they must take as given and, ultimately, bring *to* that interaction from the outside. The words and their meaning cannot (or, at least, cannot all) arise *out* of the interaction, for example, by agreement, since that would beg the question as to how the meanings of the words that form the basis for this meaningful agreement were themselves agreed upon.

But then this must mean that the parties to such an interaction, at least when they become *interlocutors* and not merely *interactors*, are ultimately forced, at least in part, into the objectivity of a we frame. For when one player communicates an understanding of what she is doing by way of publicly accessible meanings, she must at some point order her

particular action s_i under the aspect of S , that is, as something that she does *categorically* for the other.⁵ We might say, again provocatively, that this simply is *what it is* for each player to do his or her part under the profile S of strategies where S is now understood, not merely intersubjectively (as would be the case if the players merely had to glimpse a private and momentary understanding of one another's behaviour), but objectively as well (where the words, and their meanings, which make up the communication must be brought *to* the interaction and, therefore, transcend its moment).

The greater discipline that is provided by public reason, that is, by the obligation to offer a publicly comprehensible articulation of the reasons that inform a concerted action, is what takes us, therefore, from the merely intersubjective (or Nash-like) thinking of the theory of games to the sort of thinking that is more objective and more conducive to co-operation and co-ordination under the aspect of a we frame. Of course, the obligation to provide reasons for what one does is what characterized our judges in the Table 1 breach of contract example and which disciplined the realization of the majority's preference there for an outcome where the defendant was not held liable. I hope to have provided some reason for thinking that the extra discipline that the obligations of public reason bring to collective action and preference can be exercised beyond that simple legal example, and be put to use, not only for the avoidance of cyclical instability in the theory of social choice, but also for the achievement of co-operation and co-ordination in the theory of games.

⁵ It should be emphasized that *categorical* action is categorically framed action, and action done *categorically for another* is action done under the aspect of categories which are publicly accessible or, at least, accessible to the other party. But categorical action is not action without exception or *absolute* action. Indeed, categorically framed "we" conduct, which would have each of us do our part s_i within some larger profile S , might only provide a *defeasible presumption* in favour of doing that part, a presumption that would sensibly give way to acting more within the I/he frame if others were not doing theirs. After all, the very idea of doing one's part presupposes the *whole* (and, therefore, others doing their parts) of which one's conduct *is* a part. In this respect the we frame can make sense of *conditional* cooperation, but in a way that avoids the problems of paralysing self-reference that plague mutually conditioned conditional cooperation in the I/he frame. For further discussion of a defeasible presumption in favour of cooperation in the context of voluntary contributions to public goods, see Chapman (2001). The dual aspect of defeasibility, viz., that there is, first, an initial presumption in favour of (the value of) some course of conduct and, second, the possibility of relaxing that presumption in the face of countervailing considerations, is also closely related, more generally, to the idea of a conceptually sequenced argument. This idea should be distinguished from both a model of tradeoffs (with no initial presumption) and a model of lexical priority (with an absolute, not merely presumptive, prior value); for discussion, see Chapman (1998b).

References

- Bacharach, M. (1997) "'We' Equilibria: A Variable Frame Theory of Cooperation", unpublished paper presented at the Seminar on Co-operative Reasoning, St. John's College, Oxford.
- Bacharach, M. (1998) "Players' Representations and the Theory of Games with Multiple Identity", unpublished paper (short version), Institute of Economics and Statistics, University of Oxford
- Chapman, B. (1998a) "More Easily Done Than Said: Rules, Reasons, and Rational Social Choice" 18 *Oxford Journal of Legal Studies* 293
- Chapman, B. (1998b) "Law, Incommensurability, and Conceptually Sequenced Argument" 146 *University of Pennsylvania Law Review* 1487
- Chapman, B. (2001) "Rational Voluntarism and the Charitable Sector" in J. Phillips, B. Chapman, and D. Steven eds. *Between State and Market: Essays on Charities Law and Policy in Canada* (Montreal: McGill-Queen's University Press)
- Doyle J. and Thomason R. (1999) "Background to Qualitative Decision Theory" *AI Magazine* 55
- List C. and Pettit, P. (2000) "The Aggregation of Reason", unpublished, Research School of Social Sciences, Australian National University, Canberra, Australia
- McKelvy, R. (1976) "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control" 12 *Journal of Economic Theory* 472
- Pettit, P. (2000) "The Discursive Dilemma and Social Ontology", unpublished, Research School of Social Sciences, Australian National University, Canberra, Australia
- Pettit, P. (2001) "Deliberative Democracy and the Discursive Dilemma" forthcoming 11 *Philosophical Issues* (Supp. *Nous*)
- Sen, A. (1966) "A Possibility Theorem on Majority Decisions" 34 *Econometrica* 491
- Sugden, R. (1993) "Thinking as a Team: Toward an Explanation of Nonselfish Behaviour" 10 *Social Philosophy and Public Policy* 69
- Sugden, R. (2000) "Team Preferences" 16 *Economics and Philosophy* 175